

Clonal Selection as an Inspiration for Adaptive and Distributed Information Processing

Jason Brownlee

2008

A dissertation presented for fulfillment of the requirements for the award of
Doctor of Philosophy

Abstract

The development of adaptive and intelligent computational methods is an important frontier in the field Artificial Intelligence given the fragility of top-down software solutions to complex problems involving incomplete information. This dissertation describes a systematic investigation of the Clonal Selection Theory of acquired immunity as a motivating information processing metaphor of a series of adaptive and distributed Computational Intelligence algorithms. The broader structure and function of the mammalian immune system is used to frame the cellular theory and classically inspired approaches, providing the additional distributed perspectives of a ‘host of tissues’ called the Tissue Paradigm and a ‘population of hosts’ called the Hosts Paradigm. This investigation was motivated by three open problems in the broader field of Artificial Immune Systems, specifically the perceived impasse in the development, identity, and application of the field, the promise of distributed information processing, and the need for a framework to motivate such work.

The state of Clonal Selection Algorithms is investigated in the context of immunological theory, and considered in the context of broader related machine learning fields and adaptive systems theory. A systematic approach is adopted in considering the adaptive qualities of clonal selection beyond a cellular perspective, involving the identification of the lymphatic system and lymphocyte migration as a motivating metaphor for intra-host distributed systems, and host immunisation and evolutionary immunology as a motivating metaphor for intra-population distributed system design. Relevant immunophysiology and theory was reviewed, abstracted to computational models and algorithms, and systematically assessed on model pattern recognition problems to demonstrate and verify expected information processing capabilities. The empirical investigation reveals a variety of tissue and host based clonal selection systems capable of acquiring distributed information via internal processes of controlled localisation and dissemination, in a decentralised information exposure environment.

The general capabilities of Clonal Selection Algorithms as a Computational Intelligence paradigm are defined in the context of a detailed assessment of the suitability of the approaches to the important problem primitives of Function Optimisation and Function Approximation. The findings highlight the general capabilities of the approaches as mutation-based parallel hill climbers for global optimisation and prototype quantisation approaches to function approximation. Finally, an Integrated Hierarchical Clonal Selection Framework demonstrates the subsumed relationship between the cellular, tissue, and host classes of algorithm, the dependent relationship with the complexity of the abstract antigenic environment addressed by the system, and a general scaffold for a broader class of distributed artificial immune systems.

Acknowledgements

I could not have completed this work without the help and assistance of a large number of people and organisations.

Professionally, I would like to thank:

Tim Hendtlass for your wisdom, enthusiasm, encouragement and seemingly unbounded patience and support through times easy and difficult. Your hands off approach has made me both more mature as a person, and more confident as a researcher. I am fortunate to have counted you as mentor and friend these past years. The *Faculty of Information and Communication Technologies, Swinburne University of Technology* for providing me with candidature and support throughout my PhD. The *Australian Government*, for providing me with the financial assistance of an Australian Postgraduate Award (APA) during my candidature.

Personally, I would like to thank:

Daniel Angus for his friendship, discussions, rants, debates, collaboration, competition, and ultimately for defining my PhD experience. My officemates, past and present: *Irene Moser, Nigel Fisher, Gerard Murray, Clinton Woodward, Hans Pierrot, James Montgomery*, and *Ben Prime* for providing much needed fun, drinking partners, and most importantly anchors for my sanity. *Friends and Housemates* for the good times, bad times, and the much needed balance you brought to my life. *Ying Liu* for your intimate understanding and patience for me whilst writing up my thesis. Finally, my parents *Lynette Brownlee* and *Derek Brownlee*, and my sister *Laura Brownlee* for your unconditional continued support of my decision to quit a perfectly good job and pursue my research interests.

Declaration

I declare that: This thesis contains no material which has been accepted for an award to myself for any other degree or diploma, except where due reference is made in the text of the thesis. This thesis, to the best of my knowledge contains no material previously published or written by another person except where due reference is made in the text of the thesis. Where the work is based on joint research or publications, this thesis discloses the relative contributions of the respective works or authors.

Jason Brownlee

Date:

Contents

Abstract	i
Acknowledgements	ii
1 Introduction	1
1.1 Motivation	1
1.2 Hypothesis	2
1.3 Goals	3
1.4 Contributions	4
1.5 Thesis Organisation	4
2 Background	8
2.1 Chapter Overview	8
2.2 Artificial and Computational Intelligence	8
2.3 Natural and Artificial Immune Systems	11
2.4 Three Schools of AIS	15
2.5 Distributed Artificial Immune Systems	19
2.6 Open Problems in AIS	23
2.7 Methodology	25
2.8 Chapter Summary	33
3 Clonal Selection Paradigm	35
3.1 Chapter Overview	35
3.2 Clonal Selection Theory	35
3.3 Clonal Selection Algorithms	43
3.4 Related Computational Paradigms	55
3.5 Clonal Selection Adaptive Strategy	64
3.6 Cellular Strategy and Beyond	72
3.7 Chapter Summary	75
4 Cellular Clonal Selection	76
4.1 Chapter Overview	76
4.2 Abstract Cellular Paradigm	76
4.3 Realised Cellular Paradigm	89
4.4 Cellular Clonal Selection	98
4.5 Spatial Repertoire	112
4.6 Response Mediation	118
4.7 Cell-Cell Recognition	129
4.8 Chapter Summary	143

5	Tissue Clonal Selection	145
5.1	Chapter Overview	145
5.2	Physiology of Lymphocyte Migration	145
5.3	Abstract Tissue Paradigm	153
5.4	Realised Tissue Paradigm	169
5.5	Lymphocyte Recirculation	180
5.6	Lymphocyte Homing	190
5.7	Tissue Inflammation	199
5.8	Chapter Summary	207
6	Host Clonal Selection	209
6.1	Chapter Overview	209
6.2	Populations of Interacting Immune Systems	209
6.3	Abstract Host Paradigm	214
6.4	Realised Host Paradigm	220
6.5	Population Elicited Immunity	228
6.6	Population Shared Immunity	239
6.7	Generational Maternal Immunity	246
6.8	Generational Evolved Immunity	253
6.9	Chapter Summary	259
7	Hierarchical Frameworks	261
7.1	Chapter Overview	261
7.2	Clonal Selection Framework	261
7.3	Antigenic Exposure Framework	268
7.4	Integrated Hierarchical Framework	273
7.5	Methodology and Hierarchical Framework	278
7.6	Chapter Summary	288
8	Suitability and Application	290
8.1	Chapter Overview	290
8.2	Suitability of Application	290
8.3	Function Optimisation	302
8.4	Function Approximation	311
8.5	Chapter Summary	318
9	Conclusions	320
9.1	Research Goals	320
9.2	Contributions	323
9.3	Limitations	324
9.4	Further Research	326
9.5	Final Words	331
A	Tables of Empirical Results	332
A.1	Tissue Clonal Selection	332
A.2	Host Clonal Selection	336
B	Works Arising from the Study	342
B.1	Peer-Reviewed Publications	342
B.2	Software	342
B.3	Unpublished Technical Reports	343
	Bibliography	349

List of Figures

2.1	Depiction of the conceptual framework for devising Biological Inspired Algorithms, taken from [373]	27
3.1	Depiction of the shape-space formalism taken from [322] (page 1225). X_i represents antibodies with corresponding complementary recognition volumes V_{ei} .	41
3.2	Theory of receptor editing providing long jumps in the affinity landscape, from [167].	42
4.1	Depiction of external stimulation resulting in internal activation.	80
4.2	Example exposure regimes demonstrating some archetypical pathogen virulence behaviours.	82
4.3	Flow diagram of the single and multiple exposure models.	83
4.4	Depiction of the trade-off between control of the environment and control of the system.	85
4.5	Box-and-whisker plot's from the Cellular Empirical Study.	101
4.6	Example plots of T_{rs} from the CCSA(N+N) at the end of the run on left section on both plots, for both ACSP-1 and ACSP-10, where the problem solution is represented on the right of each plot, and colour components within each CSP.	101
4.7	Box-and-whisker plot's from the Replacement Empirical Study.	105
4.8	Plots of the ordered state of the repertoire after the triggering of the stop condition (left of each plot) for the assessed CCSA and RCCSA algorithms on the ACSP-10, where the right of each plot provides the problem optima.	106
4.9	Depiction of a polyclonal activation and a polyclonal response.	108
4.10	Depiction of a one-dimensional toroidal lattice.	113
4.11	Depiction of clones competing in the neighbourhood of their progenitor.	114
4.12	Depiction of the explicit versus implicit partitioning of input signals and the effects on the repertoire.	115
4.13	Box-and-whisker plot's from the Spatial Repertoire Empirical Study.	117
4.14	Example plots of the spatial repertoire from two variations of SCCSA at the end of the run on ACSP-10 with BMC represented with circles.	117
4.15	Depiction of inter-repertoire structures based on the mapping of cells.	120
4.16	Depiction of the integration of feedback for supervised structure formation.	122
4.17	Depiction of the principle concerns of the MCCSA and the relationships in terms of the selected measures.	124
4.18	Example plots of the B- (top of each plot) and T-Cell (bottom of each plot) repertoires from the ECCSA at the end of the run on ACSP-10 for two different inter-repertoire mapping types.	125
4.19	Box-and-whisker plot's for the relationship configuration with ECCSA.	126

4.20	Example plots of the B- (top of each plot) and T-Cell (bottom of each plot) repertoires from the MCCA at the end of the run on ACSP-10 for all four inter-relationship types.	127
4.21	Examples of more complex relationships that may be formed given receptor-receptor recognition.	130
4.22	Depiction of the two-string representation scheme.	131
4.23	Example of a simple two-receptor promotion scheme.	132
4.24	Depiction of the example difference between a natural and remapping scheme.	132
4.25	Depiction of the relationship between remapped antigen and receptors (<i>ps</i> is primary and <i>ss</i> is secondary string).	133
4.26	Box-and-whisker plot's from the Recurrent NCCA Empirical Study.	136
4.27	Example plots from of the response mappings and final repertoire from R-NCCA-PS on ACSP-10.	136
4.28	Box-and-whisker plot of the Average Cell Error for the DE-NCCA schemes on ACSP-10.	140
4.29	Example repertoire plots of the four Dual Exposure NCCA configurations on ACSP-2 showing primary and secondary representations and antigen left to right.	141
5.1	Multiple tissue model and the situated spatial properties that may affect antigenic exposure.	156
5.2	Example of two different Tissue Exposure Regime with controlled selection of discrete repertoires on exposure events.	156
5.3	Depiction of the temporal memory effect in a single repertoire model over time, where the arrows at the top represent information flow into the system and the arrow along the bottom represents time.	157
5.4	Depiction of the repertoire specialisation that results in the Spatial Self-Organisation effect, where the arrows from above represent information flow into the system and the arrows below represent dissipation of acquired information.	158
5.5	Depiction of the unbiased dissemination of information that results in the Spatial Consistency of Response, where the arrow above represents information flow into the system and the arrow below represents information availability throughout the system.	159
5.6	Depiction of information sharing and a cross-reactive response, where the arrows above represent the flow of information into the system as well as an implicit anticipation for the application of acquired information.	160
5.7	Depiction of the trade-off between the general anticipation of exposure and organised anticipation of exposure.	161
5.8	Depiction of the relationship between primary and secondary lymphoid tissues, where the arrows show the sharing of information between primary tissues and the outward push of information from primary to secondary tissues.	161
5.9	Depiction of the relationship between secondary tissues and secondary and tertiary tissues, where the arrows show the sharing of information between secondary tissues and the outward push of information from secondary to tertiary tissues.	162
5.10	Depiction of the relationship between tertiary and secondary tissues and tertiary tissues and the environment, where the arrows depict the flow of information from the environment to the tertiary tissues and ultimately to the secondary tissues.	163

5.11	Depiction of the layered tissue architecture and information flows between the layers, the arrows show the flow of information from the environment into the system as well as the outward push of new detectors from the centre of the system.	164
5.12	Depiction of the information streaming providing by lymph collecting in lymph nodes.	164
5.13	Depiction of the courier-like behaviour of dendritic cell's.	165
5.14	Depiction of the layered information processing properties of the lymphoid tissue architecture.	166
5.15	Depiction of the simple filter model.	167
5.16	Depiction of the lymph node model.	167
5.17	Depiction of the host model.	168
5.18	Depictions of general and minimal configuration examples of the Infection Antigenic Exposure Problem (IAEP).	171
5.19	Depictions of the four different Tissue Exposure Regimes (TER).	172
5.20	Box-and-whisker plot of Host Diversity (HD) across all TER for the RTCSA study.	183
5.21	Box-and-whisker plot of Host Error (HE) across all TER for the RTCSA study.	184
5.22	Box-and-whisker plot of Average Tissue Diversity (ATD) across all TER for the RTCSA study.	184
5.23	Box-and-whisker plot of Average Tissue Error (ATE) across all TER for the RTCSA study.	185
5.24	Box-and-whisker plot of Host Diversity (HD) across all TER for the HTCSA study.	194
5.25	Box-and-whisker plot of Host Error (HE) across all TER for the HTCSA study.	195
5.26	Box-and-whisker plot of Average Tissue Diversity (ATD) across all TER for the HTCSA study.	195
5.27	Box-and-whisker plot of Average Tissue Error (ATE) across all TER for the HTCSA study.	196
5.28	Box-and-whisker plot of Host Diversity (HD) across all TER for the ITCSA study.	203
5.29	Box-and-whisker plot of Host Error (HE) across all TER for the ITCSA study.	204
5.30	Box-and-whisker plot of Average Tissue Diversity (ATD) across all TER for the ITCSA study.	204
5.31	Box-and-whisker plot of Average Tissue Error (ATE) across all TER for the ITCSA study.	205
6.1	Host Clonal Selection Model and the situated spatial properties that affect antigenic exposure, where the arrows represent interaction between the discrete components (hosts) in the system.	216
6.2	Elaboration on figure 5.18b to depict the Minimal Habitat Antigenic Exposure Problem (MAHEP)	222
6.3	Box-and-whisker plot of Population Diversity (PD) across all HER for the THCSA study.	234
6.4	Box-and-whisker plot of Population Error (PE) across all HER for the THCSA study.	235
6.5	Box-and-whisker plot of Average Host Diversity (AHD) across all HER for the THCSA study.	235

6.6	Box-and-whisker plot of Average Host Error (AHE) across all HER for the THCSA study.	236
6.7	Box-and-whisker plot of Population Diversity (PD) across all HER for the SI-HCSA study.	242
6.8	Box-and-whisker plot of Population Error (PE) across all HER for the SI-HCSA study.	242
6.9	Box-and-whisker plot of Average Host Diversity (AHD) across all HER for the SI-HCSA study.	243
6.10	Box-and-whisker plot of Average Host Error (AHE) across all HER for the SI-HCSA study.	243
6.11	Box-and-whisker plot of Population Diversity (PD) across all HER for the MI-HCSA study.	249
6.12	Box-and-whisker plot of Population Error (PE) across all HER for the MI-HCSA study.	250
6.13	Box-and-whisker plot of Average Host Diversity (AHD) across all HER for the MI-HCSA study.	250
6.14	Box-and-whisker plot of Average Host Error (AHE) across all HER for the MI-HCSA study.	251
6.15	Box-and-whisker plot of Population Diversity (PD) across all HER for the EI-HCSA study.	256
6.16	Box-and-whisker plot of Population Error (PE) across all HER for the EI-HCSA study.	256
6.17	Box-and-whisker plot of Average Host Diversity (AHD) across all HER for the EI-HCSA study.	257
6.18	Box-and-whisker plot of Average Host Error (AHE) across all HER for the EI-HCSA study.	257
7.1	Depiction of the focus on systems and the compression of the environment in the Hierarchical Clonal Selection Framework (HCSF).	262
7.2	Depiction of the compression of both the system and the environment when the HCSF is compressed to the primary objects of the adaptive system formalism.	263
7.3	Depiction of an expanded realisation of a hierarchical clonal selection system with all three tiers.	266
7.4	Depiction of the focus on environment and the compression of the systems in the Hierarchical Antigenic Exposure Framework (HAEF).	269
7.5	Depiction of the focus on system and environment in the Integrated Hierarchical Clonal Selection Framework (IHCSF).	273
7.6	Depiction of the integrated hierarchical clonal selection and hierarchical antigenic exposure frameworks, showing subsumed relationships and component cardinalities across the investigated paradigms.	276
8.1	Depiction that the suitability of application of a strategy to a problem is defined by the overlap in their features, as well as the comparability with other strategies, and transferability of results to other problems.	292
8.2	Depiction of the cellular algorithm approach to inductive learning with holistic, sub-, and partial-models of a given problem.	295
8.3	Depiction of the tissue algorithm approach to inductive learning with holistic, sub-, and partial-models of a given problem.	298
8.4	Depiction of the host algorithm approach to inductive learning with holistic, sub-, and partial-models of a given problem.	301

List of Tables

3.1	Clonal selection algorithm taxonomy overview.	46
3.2	CLONALG parameters.	47
3.3	BCA parameters.	49
3.4	SIA parameters.	51
3.5	Summary of primary objects in the adaptive systems formalism.	65
3.6	Summary of secondary objects in the adaptive systems formalism.	66
3.7	Fundamental questions regarding adaptive systems, taken from [216] (page 29).	67
4.1	List of goals for a colour-space domain.	87
4.2	Summary of terms for the colour space domain formalism.	88
4.3	Summary of the assessed configuration of CCSA for ACSP-1.	99
4.4	Summary of the assessed configuration of CCSA for ACSP-10.	99
4.5	Summary of results for CCSA on ACSP-1 and ACSP-10.	100
4.6	Summary of the assessed configuration for the CCSA-RCCSA empirical study.	103
4.7	Summary of results from the CCSA-RCCSA empirical study on ACSP-10.	104
4.8	Summary of results from the DCCSA empirical study on DCSP-10.	110
4.9	Summary of results for SCCSA on AEP 10.	116
4.10	Summary of the configuration of the ECCSA with a series of inter-repertoire selection relationships.	123
4.11	Summary of results for ECCSA with two different mapping schemes on ACSP-10.	125
4.12	Summary of results for ECCSA with four different inter-repertoire relationship schemes on ACSP-10.	125
4.13	Exposure-activation scheme for a minimal implementation.	132
4.14	Summary of the various replacement configurations for the mapped Recurrent Exposure Network Cellular Clonal Selection Algorithm (RE-NCCSA).	134
4.15	Summary of results for Recurrent Network Cellular Clonal Selection Algorithm (R-NCCAS) on ACSP-10	135
4.16	Summary of the various replacement configurations for the Dual Exposure Network Cellular Clonal Selection Algorithm (DE-NCCSA).	139
4.17	Summary of results for the Dual Exposure Network Cellular Clonal Selection Algorithm (DE-NCCSA) on ACSP-10.	140
5.1	Summary of information processing needs, and proposed algorithms that meet those needs.	165
5.2	Summary of the principle responsibility of each tier of the architecture.	166
5.3	Configuration of the architecture for the filter model.	167
5.4	Configuration of the architecture for the lymph node model.	168
5.5	Configuration of the architecture for the host model	169
5.6	Summary of the difference in effect, strategy and mechanism between the complementary concerns of tissue algorithms under discrete exposures.	177

5.7	Summary of the exposure and re-exposure behaviour of the five defined tissue exposure regimes and how they relate to system information availability and anticipated need.	178
5.8	Summary of general trends and their interpretation for Tissue Clonal Selection Algorithms.	178
5.9	Summary of the expected behaviours of the defined measures with regular and irregular exposure regimes on systems with mixing and isolated information dissemination strategies.	179
5.10	Summary of the results for MTCSA and RTCSA-LRG on ATER and RTER and the relative quantitative difference in measures.	185
5.11	Summary of confirmed behaviours of global and local information strategies and exposure regimes, highlighting where expectations were incorrect. . . .	186
5.12	Summary of results from an empirical study into the HTCSA comparing TER's with the HTCSA-L information management strategy.	197
6.1	Summary of the different types of immunity.	210
6.2	Summary of the differences between the innate and acquired immune systems, taken from [257] (page 565).	213
6.3	Summary of the difference in mechanism and strategy between the complementary concerns of host algorithms under discrete exposures.	226
6.4	Classification of Host Exposure Regimes in terms of general attributes. . . .	228
7.1	Mapping of the cellular, tissue, and host clonal selection perspectives onto the secondary objects of the adaptive systems formalism.	263
7.2	Summary of the component cardinality across the principle abstract and realised clonal selection algorithms, where * represents a wild-card for descendant algorithms of a given base algorithm.	265
7.3	Summary of the component interaction across the abstract and realised clonal selection algorithms, where * represents a wild-card for descendant algorithms of a given base algorithm.	267
7.4	Summary of the component cardinality across the principle abstract and realised antigenic exposure problems.	269
7.5	Summary of the bracketing of spatial exposure across the abstract and realised antigenic exposure problems.	271
7.6	Summary of the bracketing of temporal exposure across the principle abstract and realised antigenic exposure problems.	271
7.7	Summary of the suitability match between clonal selection and antigenic exposure components	274
7.8	Summary of the component cardinality for both the clonal selection and antigenic exposure models.	274
7.9	Summary of the Negative Selection AIS paradigm mapped onto the three tiers of the IHCSF.	281
7.10	Summary of the Immune Network AIS paradigm mapped onto the three tiers of the IHCSF.	282
7.11	Summary of an interpretation of the field of Evolutionary Computation in terms of the hierarchical reductionist methodology used to investigate clonal selection.	283
7.12	Summary of an interpretation of the field of Connectionist Intelligence in terms of the hierarchical reductionist methodology used to investigate clonal selection.	285

8.1	Summary of the projected applicability of Cellular Clonal Selection Algorithms to Function Optimisation Problems.	308
8.2	Summary of the projected applicability of Tissue Clonal Selection Algorithms to Function Optimisation Problems.	309
8.3	Summary of the projected applicability of Host Clonal Selection Algorithms to Function Optimisation Problems.	310
8.4	Summary of the projected applicability of Cellular Clonal Selection Algorithms to Function Approximation Problems.	316
8.5	Summary of the projected applicability of Tissue Clonal Selection Algorithms to Function Approximation Problems.	316
8.6	Summary of the projected applicability of Host Clonal Selection Algorithms to Function Approximation Problems.	317
A.1	Summary of results from an empirical study into the RTCSA compared to the MTCSA.	333
A.2	Summary of results from an empirical study into the RTCSA comparing TER's with the MTCSA local and RTCSA-LRG global information management strategies.	334
A.3	Summary of results from an empirical study into the HTCSA.	335
A.4	Summary of results from an empirical study into the ITCSA.	336
A.5	Summary of results from an empirical study into the THCSA compared to MP-HCSA.	337
A.6	Summary of the pathogen transmission and vaccination transmission results from an empirical study into the THCSA.	338
A.7	Summary of results from an empirical study into the SI-HCSA compared to MP-HCSA.	339
A.8	Summary of results from an empirical study into the MI-HCSA.	340
A.9	Summary of results from an empirical study into the EI-HCSA.	341

Chapter 1

Introduction

1.1 Motivation

Artificial Intelligence is concerned with the investigation of the information processing mechanisms that underlie intelligent behaviour in systems that act *rationally* or like *humans*. The traditional approach toward designing and investigating such systems is a rigorous top-down approach using a symbolic representation that provides clear indication *why* such systems work. A newer scruffy paradigm called Computational Intelligence considers intelligence from a bottom-up perspective of *strategies* and *outcomes*. Examples of systems of this type are prescriptive indicating *how* they work, although they are typically too complex to analyse to see *why* they work. The process for developing such strategies, like many other fields of science and engineering, is to begin by patterning them from biological and natural processes. Toward this end, the field of Artificial Immune Systems is concerned with abstracting and exploiting the computational patterns observed in the structure and function of the immune system [109].

The immune system is vast in scale and complexity, comprised of specialised organs and trillions of cells and molecules. The complexity of the system is commonly considered in the same order as the central nervous system. It is responsible for maintaining the homeostasis of a host organism and in particular it is responsible for identifying and eliminating invading micro-organisms called pathogen [81, 314]. The immune systems of vertebrates has a capacity to adapt so as to acquire immunity and better respond in the future to the same and similar infections. The immune system is both essential to a hosts survival and a powerfully efficient garbage disposal. The precision mechanisms that exist to keep a host alive can become corrupted and turn on and attack the host, an affliction that underlies many types of autoimmune disease. The acquired immune system is responsible for the manufacture of an enormous diversity of detector proteins called antibodies, such that collectively a repertoire of these molecules can potentially detect many millions or billions of different molecular structures. Interestingly the information required to prepare such a potentially diverse army of protective protein conformations is far greater than the number of genes encoded in a given vertebrates genome. This mismatch between the limited information that a genome can represent and what the acquired immune system is capable of detecting is achieved via an adaptive learning process described by the Clonal

Selection Theory [63]. This Darwinian inspired theory for antibody diversity begins with a repertoire of general-purpose detectors that can partially detect most infections. A process of *selection of the fittest* and cellular mitosis with mutations that affect what descendant cells can identify, results in the specialisation of a host organisms detection repertoire to its environment.

The learning, memory, and adaptive information processing properties of clonal selection are a ‘corner stone’ in the field of Artificial Immune Systems (AIS). The abstracted clonal selection strategy has been embodied in optimisation, pattern recognition, and classification algorithms and provides a foundational improvement procedure for those AIS that exploit the information processing attributes of other cell-centric immunological theories. Beyond adaptation, clonal selection describes a process by which the acquired immune system coordinates learning and memory of a lifetime of unknown infections without central control or organisation. This adaptive process occurs autonomously and concurrently throughout the distributed tissues of a host organism. Interestingly the distributed and decentralised attributes are an often cited advantage of patterning strategies from the immune system, although much of the research in the field has focused on monolithic systems and algorithms [417, 420]. There has been provoking work on the investigation of distributed signature based computer security applications [212, 255], and in autonomous navigation and group decision making in mobile robotics [238, 359, 268], although there exists no clear framework to guide the realisation of such systems and little work on investigating the clonal selection paradigm toward this end. Finally, there is some consensus among leaders in the field of a recent impasse in the general progress in the field of Artificial Immune Systems. The same authors suggested a trend of turning back to the biological system to abstract novel and more accurate models to replace the crude *first generation* approaches with so-called *second generation artificial immune systems* [24, 389, 388, 401].

The thesis of this work is that the clonal selection paradigm can be elaborated to exploit the distributed and decentralised information processing characteristics inherent in the inspiring metaphor, and that the structure and function of the immune system provide a path toward realising such computational models. Such information processing characteristics are beneficial for classes of problems that can be addressed using parallel or cooperative problem solving strategies, such as those problems partitioned along functional or information availability axes. This work addresses the open problem of considering a *second generation* of artificial immune systems, in this case distributed clonal selection algorithms, by following the trend of devising novel and more accurate abstractions and models from the inspiring biological system.

1.2 Hypothesis

The main hypothesis of this research is defined, as follows:

An in depth study of immunological structure and function, in particular the distributed nature of the immune system, provides effective computational inspiration for the extension and improvement of existing clonal selection algo-

rithms.

The aim of this research is to improve the current understanding of the clonal selection paradigm for adaptive information processing and more specifically the state of distributed and decentralised approaches. This aim was pursued by using an understanding of immunological structure and function to provide a conceptual framework for the investigation. To verify the effectiveness of the investigated framework and clonal selection algorithms the hypothesis was divided into a series of goals, defined in Section 1.3.

1.3 Goals

The research thesis is addressed in the context of five main research goals, as follows:

1. Identify a systematic methodology for realising a novel biological inspired computational framework and models. Before the investigation of a novel immunological inspired framework and resultant models and algorithms, it is essential to identify a systematic methodology to provide a set of well defined procedures as to how such a framework may be realised and how the effectiveness of the models and algorithms may be assessed.

2. Identify limitations with and elaborate upon the base clonal selection paradigm. The clonal selection paradigm is a core information processing pattern in cellular immunology and inspired computational models, therefore it is essential that both the capabilities and limitations of existing adaptive and distributed computational models are understood.

3. Identify immunological structures and/or functions which clearly exhibit distributed information processing. The structures and functions of mammalian immune system must be scrutinised through a lens of clonal selection for processes and architectures that constitute distributed information processing, and formulated into abstracted computational models.

4. Study the behaviours, capabilities, and limitations of devised computational models. In order to assess the effectiveness of the devised abstracted models they must be examined and understood with regard to their information processing capabilities, limitations, and general behaviours.

5. Propose an integrated clonal selection framework that unites the base and distributed variants. Finally, the adaptive clonal selection paradigm and realised distributed models must be integrated into a framework such that their benefits and limitations can be assessed given the success of its descriptive and predictive capabilities.

1.4 Contributions

This dissertation makes the following specific contributions:

1. The systematic elaboration and development of the existing clonal selection computational paradigm as an explicit adaptive strategy and suggestion of additional models based on the interaction of cells and molecules, called the *Cellular Clonal Selection Paradigm* (Chapter 3 and Chapter 4).
2. The design and investigation of novel *Spatial*, *Mediated*, and *Network* Cellular Clonal Selection Algorithms (Chapter 4).
3. The identification and abstraction of the lymphatic system and lymphocyte trafficking as a framework for distributed Clonal Selection Algorithms called the *Tissue Clonal Selection Paradigm* (Chapter 5).
4. The design and investigation of novel *Recirculation*, *Homing*, and *Inflammation* Tissue Clonal Selection Algorithms (Chapter 5).
5. The identification and abstraction of host immunisation and evolution as a framework for distributed Clonal Selection Algorithms called the *Host Clonal Selection Paradigm* (Chapter 6).
6. The design and investigation of novel *Transmission* and *Shared Immunity* Population Host Clonal Selection Algorithms, and the *Maternal Immunity* and *Inherited Immunity* Generational Host Clonal Selection Algorithms (Chapter 6).
7. The integration of the three perspectives of the computational properties of clonal selection (Cellular, Tissue, Host) into an *Integrated Hierarchical Clonal Selection Framework* with both explanatory and predictive properties (Chapter 7).
8. The elicitation of salient features of the Cellular, Tissue, and Host Clonal Selection Algorithms and systematic assessment of the suitability of application of the algorithms to the foundational problem domains of *Function Optimisation* and *Function Approximation* (Chapter 8).

1.5 Thesis Organisation

This section presents the organisation for this work, highlighting the objectives and setting the scene for the remainder of the dissertation.

Chapter 2 provides background documentation on the general fields of study in which this work is situated, specifically Computational Intelligence and Biologically Inspired Computation. A clear definition of Artificial Immune Systems (AIS) is outlined highlighting the computational and problem solving ends, and differentiating it from theoretical and computational immunology. This chapter supports the contention of a cellular-theoretic focus in the field of AIS, and the promise of patterns for distributed, decentralised and

autonomous systems. The three dominant paradigms of the field are reviewed as is the state of distributed AIS, highlighting the lack of a suitable framework and limited elaboration of the clonal selection paradigm toward this end. Finally, this section outlines a systematic research methodology for realising immune inspired algorithms, an experimental methodology to ensure systematic assessment of devised systems, and an integrated methodology that focus on decomposition and an economy of models.

Chapter 3 provides a review of clonal selection theory providing a context for the effective interpretation of the state of the art of inspired computational models. A taxonomy is presented of clonal selection algorithms highlighting the commonality, and more importantly open problems in the field. The placement and assessment of the field is considered in the broader context of computational intelligence demarcating related research that contributes to an improved understanding of the approach. The underlying computational strategy is considered in the context of relevant adaptive systems theory demonstrating the importance of the information environment of the strategy, interactions of components, and emergent behaviours. The cellular perspective of the field is addressed and an agenda is defined which involves the investigation and elaboration of the classical cellular clonal selection perspective, and the consideration of a ‘host of tissues’ and a ‘population of hosts’ as distributed cellular perspectives of the immunological theory.

Chapter 4 presents a conceptual framework for the clonal selection paradigm from a cellular perspective called the *Cellular Clonal Selection Paradigm*. The framework includes the principle components and base models that describe the current state of clonal selection algorithms referred to as *Quintessential Clonal Selection*. Symmetry is provided to the framework in the form of a complementary *Antigenic Exposure Paradigm* that encapsulates the domain information to which a system governed by clonal selection is triggered and to which it must respond. Exploiting the predictive properties of the cellular framework, three clonal selection approaches are defined and investigated: a *Spatial*, *Mediated*, and *Network Clonal Selection Algorithm*, that are empirically demonstrated on an antigenic exposure inspired pattern recognition problem domain called *Colour Space*.

Chapter 5 considers the clonal selection paradigm in the context of the entire host organism. The structure and function of the human immune system are reviewed, highlighting the important role of the lymphoid tissues and related systems in the trafficking of immune cells around the body. The tissue architecture and cell mobility information processing are abstracted into a *Tissue Clonal Selection Paradigm* that both describes the cellular clonal selection paradigm as a single constrained instance of a tissue-based system, and predicts a range of tissue-based clonal selection architectures and algorithms. Matching the symmetry provided in the cellular paradigm, an *Infection Exposure Paradigm* is defined to satisfy tissue-centric concerns of multiple concurrent points of exposure and information acquisition. A *Lymphocyte Recirculation*, *Lymphocyte Homing* and *Tissue Inflammation Clonal Selection Algorithm* are investigated and demonstrated on problem instances drawn from the colour space pattern recognition domain. Generally, the empiri-

cal results confirm the expected localisation and dissemination of acquired information as emergent behaviours from the decentralised cellular repertoires.

Chapter 6 considers the clonal selection paradigm in the context of a population of host organisms with interacting immune systems. Immunisation and host evolution are identified as the two primary facets of immune system interactions at the population level that provide a different perspective on information dissemination and organisation than the tissue paradigm. A *Population* and *Generational Host Clonal Selection Paradigm* are proposed that are descriptive at both considering an instance of the tissue paradigm as a constrained host system, and at integrating the related work on models of gene library evolution. A *Habitat Exposure Paradigm* is defined to facilitate the asymmetric properties of population-based exposure and transmission. The predictive qualities of the framework are exploited in the design and empirical investigation of four host-based algorithms on problem instances from the colour space domain. These include a *Transmission* and *Shared Immunity* algorithms from the Population paradigm, and the *Maternal Immunity* and *Evolved Immunity* algorithms from the Generational paradigm. Results confirm the expectation of the higher level localisation and dissemination of acquired information emergent behaviours from the decentralised systems using varied structures and functions to those used in the Tissue paradigm.

Chapter 7 provides a systematic integration firstly of the computational clonal selection models into a *Hierarchical Clonal Selection Framework*, and secondly of the exposure paradigm into a *Hierarchical Antigenic Exposure Framework*. Each framework provides a one-sided perspective for the development and investigation of clonal selection approaches and the mapping of complementary exposure problem domains. The integration of the frameworks highlight both the specific design decisions that bound and control the complexity of the investigated algorithms. In addition to explanatory power, the *Integrated Hierarchical Clonal Selection Framework* suggests at the potential for the development and investigation of additional models, algorithms, and even new paradigms inspired by the structure and function of the immune system that constrain the computational concerns of clonal selection. The generality of the immunologically-centric framework is considered as a model for the broader study of Artificial Immune Systems, and the methodology adopted to develop and investigate the framework is considered as a model for related Computational Intelligence fields of research.

Chapter 8 considers the applicability of the developed clonal selection algorithms. A feature-based methodology for addressing the suitability of an approach is proposed and adopted to elicit the salient features from the algorithms from across the three clonal selection paradigms. Two foundational problem domains from Computational Intelligence are considered and systematically reviewed: *Function Optimisation* and *Function Approximation*. The salient features of each domain are described and elaborated in the context of specific application examples that demonstrate information processing properties relevant to the clonal selection algorithms. The applicability of the algorithms from each clonal

selection paradigm is defined by the overlap in features between specific approaches and problem instances, suggesting at an abundance of ongoing application-centric research.

Chapter 9 provides a detailed assessment of the research hypothesis by highlighting the specific examples of how each research goal from Section 1.3 was addressed in this work. Immunological structure and function are demonstrated as providing an effective path for both the elaboration of the classical clonal selection computational paradigm, and as a path toward realising decentralised and distributed extensions of the paradigm. The contributions of the work are rephrased in terms of computational and conceptual artefacts, highlighting the discrete and standalone products that may be differentiated and drawn from this work. Three key methodological limitations of the project are considered as acceptable trade-off's in addressing the specific research question that highlight alternative and potentially future approaches of investigating the emerging sub-field of study. Finally, the hierarchical perspective of the work is used to organise and systematically suggested extensions to the project at three levels of abstraction including *Framework*, *Paradigm*, and *Algorithm*, outlining an ongoing research agenda for the study of Clonal Selection Algorithms.

Chapter 2

Background

2.1 Chapter Overview

This chapter provides background information and reviews the base motivations for the work. Section 2.2 generally positions the work within the sub fields of Biologically Inspired Computation and Computational Intelligence under the broader field of Artificial Intelligence. Section 2.3 clarifies this further by outlining the specific field of Artificial Immune Systems. Section 2.4 reviews the core paradigms in the field, highlighting the cellular focus that inspires the computational approaches, and the general fundamental role the adaptive clonal selection strategy has across the field. Section 2.5 motivates the specific research hypothesis of this work in the development of distributed Artificial Immune Systems, and how the this decentralised adaptive property is generally promoted in the field, although remains underdeveloped. Section 2.6 reviews open problems in the field, highlighting the general impasses and lack of satisfying progress and advocating the specific approach taken in this work of turning back to the motivating metaphor toward developing more sophisticated approaches. Finally, a systematic methodology is pieced together in Section 2.7 from best practices with regard to (1) biologically inspired algorithm development, (2) experimental investigation, and (3) the economy of modelling and integration of findings. This chapter sets the scene for the following chapter that outlines the specific elaborations and distributed realisations of clonal selection investigated throughout the remainder of the dissertation.

2.2 Artificial and Computational Intelligence

2.2.1 Artificial Intelligence

The general field of study is the multi-disciplinary field of Artificial Intelligence (AI). Russell and Norvig provide a perspective that defines Artificial Intelligence in four categories: (1) systems that think like humans, (2) systems that act like humans, (3) systems that think rationally, (4) systems that act rationally [342]. In the definition, acting like a human suggests that a system can do some specific things humans can do, this includes fields such as the Turing test, natural language processing, automated reasoning, knowl-

edge representation, machine learning, computer vision, and robotics. Thinking like a human suggests systems that model the cognitive information processing properties of humans, for example a general problem solver and systems that build internal models of their world. Thinking rationally suggests laws of rationalism and structured thought, such as syllogisms and formal logic. Finally, acting rationally suggests systems that do rational things such as expected utility maximisation and rational agents. Luger and Stubblefield suggest that AI is a sub-field of computer science concerned with the automation of intelligence, and like other sub-fields of computer science has both theoretical (*how and why do the systems work?*) and application (*where and when can the systems be used?*) concerns [276]. They suggest a strong empirical focus to research, such that although there may be a strong desire for mathematical analysis, the systems themselves defy analysis due to their complexity. The machines and software themselves are not black boxes, rather analysis proceeds by observing the systems interactions with their environment, followed by an internal assessment of the system to relate their structure back to their behaviour.

Artificial Intelligence is therefore concerned with investigating mechanisms that underlie intelligence and intelligence behaviour. The traditional approach toward designing and investigating AI (the so-called ‘good old fashioned’ AI) has been to employ a symbolic basis for these mechanisms. A newer approach historically referred to as *messy artificial intelligence* or *soft computing* does not use a symbolic basis, instead patterning these mechanisms after biological or natural processes. This represents a modern paradigm shift in interest from symbolic knowledge representations, to inference strategies for adaptation and learning, and has been referred to as *neat* versus *scruffy* approaches to AI. The neat philosophy is concerned with formal symbolic models of intelligence that can explain *why* they work, whereas the scruffy philosophy is concerned with intelligent strategies that explain *how* they work [364].

2.2.2 Computational Intelligence

A modern name for the sub-field of AI concerned with sub-symbolic (messy, scruffy, soft) mechanisms is *Computational Intelligence*. This name provides a banner which groups four principle approaches: Fuzzy Intelligence, Connectionist Intelligence, Evolutionary Intelligence, and Swarm Intelligence [135, 317]. A second popular and general name for the *strategy-outcome* perspective of scruffy AI is *Metaheuristics*, which evolved from the neater field of heuristics methods applied in Operations Research. A metaheuristic is defined as a general algorithmic framework which can be applied to different optimisation problems with relative few modifications to make them adapted to a specific problem [48]. Another important perspective is that provided by the field of *Machine Learning* that focuses on the learning properties of Artificial Intelligence. The term is commonly used to describe inductive model building techniques (that generalise from specific observations) that are applied to ‘learn’ relationships in data sets (the application of which is referred to as Data Mining [433]), with or without supervision (corrective behaviour) [293]. The following provides a summary of the four primary areas of study in Computational Intelligence:

Evolutionary Computation A paradigm that is concerned with the investigation of systems inspired by the neo-Darwinian theory of evolution by means of natural selection. Popular evolutionary algorithms include the Genetic Algorithm, Evolution Strategy, Genetic and Evolutionary Programming, and Differential Evolution. The evolutionary process is considered an adaptive strategy and is typically applied to search and optimisation domains.

Swarm Intelligence A paradigm that considers collective intelligence emerges through the interaction and cooperation of large numbers of lesser intelligent agents. The paradigm consists of two dominant sub-fields (1) *Ant Colony Optimisation* that investigates probabilistic algorithms inspired by the stigmergy and foraging behaviour of ants, and (2) *Particle Swarm Optimisation* that investigates probabilistic algorithms inspired by the flocking and foraging behaviour of birds and fish. Like evolutionary computation, swarm intelligences are considered adaptive strategies and are typically applied to search and optimisation domains.

Connectionist Intelligence An approach that is concerned with the investigation of architectures and learning strategies inspired by the modelling of neurons in the brain called *Artificial Neural Networks*. Learning strategies are typically divided into supervised and unsupervised which manage environmental feedback in different ways. Neural network learning processes are considered adaptive learning and are typically applied to pattern recognition domains.

Fuzzy Intelligence An approach that is concerned with the investigation of fuzzy logic which is a form of logic that is not constrained to true and false, but rather functions which define degree's of truth. Fuzzy logic is reasoning strategy and is typically applied to expert system and control system domains.

2.2.3 Natural Computation

An important perspective on scruffy Artificial Intelligence is the motivation and inspiration for the core information processing strategy of a given technique. Computers can only do what they are instructed, therefore a consideration of Computational Intelligence is to distil information principles and strategies from other fields of study, such as biology. The study of biologically motivated Computational Intelligence may be called *Biologically Inspired Computing* [118], and is one of three related fields of *Natural Computing* [145, 146, 313]. Natural Computing is an interdisciplinary field concerned with the relationship of computation and biology, which in addition to Biologically Inspired Computing also comprises of *Computationally Motivated Biology* and *Computing with Biology* [316, 281, 2]. Therefore, the field of Artificial Intelligence, specifically the scruffy variety of Computational Intelligence motivates this work from the perspective of an intelligent problem-solving strategy in computer science, whereas the field of Natural Computing,

specifically the Biological Inspired Computation variety motivates the actual principles and information processing capabilities of the strategy.

Biologically Inspired Computation (*Computation inspired by biological metaphor*)

The intent is to devise mathematical or engineering tools to address problem domains. Biologically Inspired Computation fits into this category, as do other non-computational areas of problem solving not discussed. At its simplest, its using solutions (a procedure for finding solutions is considered a solution) found in the biological environment.

Computationally Motivated Biology (*Biology with digital computers*)

The intent of this area is to use information sciences and simulation to model biological systems in digital computers with the aim to replicate and better understand behaviours in biological systems. The field facilitates the ability to better understand life-as-it-is and investigate life-as-it-could-be. Typically, work in this sub-field is not concerned with the construction of mathematical and engineering tools, rather it is focused on simulating natural phenomena. Common examples include Artificial Life, Fractal Geometry (L-systems, Iterative Function Systems, Particle Systems, Brownian motion), and Cellular Automata.

Computing with Biology The investigation of substrates other than silicon in which to implement computation. Common examples include molecular or DNA Computing and Quantum Computing.

2.3 Natural and Artificial Immune Systems

This section reviews the biological immune system which motivates the biologically inspired field of Artificial Immune Systems. The varied approaches in this field have led to some debate as to how it relates to sibling Computational Intelligence sub-fields, comprising features from Evolutionary Computation, Swarm Intelligence, and Connectionist Intelligence. These relationships are further considered in Section 3.4.

2.3.1 Overview of the Immune System

This review of the immune system with a focus on the general problem to which the biological system is attributed to addressing, which is to protect the host organism from the threats posed to it from pathogens and toxic substances. Pathogens encompass a range of micro-organisms such as bacteria, virus, parasites and pollen. This review provides sufficient immunology such that a computer scientist may grasp the core architecture, processes and general problem of addressed by the immune system. The focus is the vertebrate immune system, specifically that of mammals given that disproportionately more information is known about the system. For more general information, the reader is directed to two addition reviews of immunology by and for Computer Scientists, de Castro and Timmis [110] (Chapter 2), and Hofmeyr [213].

The traditional perspective regarding the role of the immune system is divided into two primary tasks: the *detection* and *elimination* of pathogen¹. This behaviour is typically referred to as the differentiation of self (molecules and cells that belong to the host organisms) from potentially harmful non-self. The architecture of the immune system is such that a series of defensive layers protect the host. The most basic of these layers provide *physical barriers* to prevent pathogen from entering the host such as skin, and mucous membranes that trap and filter out pathogen. The next level is physiological that provides inhospitable conditions for pathogen at entry points to host such as high pH acidity in the digestive tract and varied temperature to inhibit growth of micro-organisms. Once a pathogen makes it inside the host, it must contend with the *innate* and *acquired* immune system. These interrelated immunological sub-systems are comprised of many types of cells and molecules produced by specialised organs and processes to address the self-nonself problem at the lowest level using chemical bonding, where the surfaces of cells and molecules interact with the surfaces of pathogen.

Innate Immunity

The innate immune system is named such because the host organism is born with a set of fast acting detection and elimination mechanisms that do not change over the hosts lifetime. In this regard, the mechanisms are reflexive providing a *first line of defence* against pathogen. Innate mechanisms are the most primitive found in all classes of plant and animal life improving in a species over evolutionary time. The complement system refers to a mechanism where associated molecules bind to and coat specific types of bacteria. The activation of this mechanism (a complement cascade) can lead to a rupturing of the bacteria called *lysis*, or detection and subsequent killing of the bacteria by macrophage cells called *opsonisation*. The process is fast acting, typically occurring within a the first few hours of an infection. Macrophages are scavenger cells responsible for cleaning up cellular debris, some bacteria, and micro-organisms that have been identified by the complement system. These cells engulf material and neutralise it using a digestive process. When a virus infects a cell, the cell may release chemicals called cytokines that provide an indication to other cells of the infection. *Natural Killer* (NK) cells are a specialised class of immune cell responsible for identifying virally infected cells via their chemical signals and releasing chemicals that trigger apoptosis (programmed cell death) of these cells. These NK cells are also responsible for identifying and eliminating tumour cells.

Adaptive Immunity

The adaptive immune system, also referred to as the acquired immune system, is named such because it is responsible for specialising a defence for the host organism based on the *specific* pathogen to which it is exposed. Unlike the innate immune system, the acquired immune system is present only in vertebrates (animals with a spinal column). The system retains a *memory* of exposures which it has encountered. This memory is

¹More recent perspectives on the role of the system include a maintenance system [81], and a cognitive system [403].

recalled on reinfection exhibiting a *learned* pathogen identification. This learning process may be divided into two types of response. The first or *primary response* occurs when the system encounters a novel pathogen. The system is slow to respond, potentially taking a number of weeks to clear the infection. On re-encountering the same pathogen again, the system exhibits a *secondary response*, applying what was learned in the primary response and clearing up the infection rapidly. The *memory* the system acquires in the primary response is typically long lasting, providing pathogenic immunity for the lifetime of the host, two common examples of which are the chickenpox and measles. White blood cells called lymphocytes (or leukocytes) are the most important cell in the acquired immune system. Lymphocytes are involved in both the identification and elimination of pathogen, and recirculate within the host organisms body in the blood and lymph (the fluid that permeates tissue). The behaviour of the adaptive immune system may be described with regard to two general sub-systems: *Humoral Immunity* for addressing free-floating pathogen (for example in blood), and *Cellular Immunity* for addressing virally infected cells.

Humoral Immunity B-lymphocyte cells are a specialised class of white blood cell that are created in the bone marrow and are predominantly responsible for the secretion of antibodies. These are protein molecules that bind to parts or specialised features of pathogen called pathogenic determinants. B-cells are activated by identifying pathogen via their surface bound receptors which provide a complementary shape for pathogen surface features, fitting like a lock and a key. Once activated a B-cell will produce vast quantities of antibody which enter the blood stream and recirculate around the host organism to clear up the infection.

Cellular Immunity T-lymphocyte cells are another class of white blood cell that are matured in the thymus (an immune organ located in the chest). T-cells perform specific roles both mediating the activation of B-cells (Helper T-cells) and in the identification and neutralisation of virally infected cells (Killer T-cells). In this latter role, T-cells identify processed (chopped up) pieces of pathogen on the surfaces of infected cells called histocompatibility complex (commonly referred to by its abbreviation: MHC). Cells responsible for processing pathogen and presenting MHC are appropriately called antigen-presenting cells, examples of which include some B-cells, macrophages and dendritic cells.

2.3.2 Artificial Immune Systems

History

Artificial Immune Systems (AIS) is a sub-field Computational Intelligence motivated by immunology (primarily mammalian immunology) that emerged in the early 1990's (for example [40, 227]), based on the proposal in the late 1980's to apply theoretical immunological models to machine learning and automated problem solving (such as [210, 138]). The early works in the field were inspired by exotic theoretical models (immune network theory) and were applied to machine learning, control and optimisation problems. The

approaches were reminiscent of paradigms such as Artificial Neural Networks, Genetic Algorithms, Reinforcement Learning, and Learning Classifier Systems. The most formative works in giving the field an identity were those that proposed the immune system as an analogy for information protection systems in the field of computer security. The classical examples include Forrest, et al.'s Computer Immunity [151, 148] and Kephart's Immune Anti-Virus [248, 249]. These works were formative for the field because they provided an intuitive application domain that captivated a broader audience and assisted in differentiating the work as an independent sub-field.

Motivation

The motivation for the field of AIS is the immune system, specifically the architecture, mechanisms, principles, and theories used to explain observed immunological function. The authors de Castro and Timmis defend the immune system as an inspiration for computational intelligence by providing a comprehensive listing of abstracted information processing principles [110] (pages 55-56). This listing contains nineteen desirable computational attributes, the following of which specifically motivate this research:

- *Uniqueness*: Each individual with an immune system has distinct and different capabilities and vulnerabilities.
- *Disposability*: There is no single cell or molecule that is essential to the functioning of the immune system.
- *Autonomy*: There is no central or coordinating organ controlling the immune system
- *Distributivity*: The immune cells, molecules, and organs are distributed all over the body and are not subject to centralised control.
- *Fault Tolerance*: The complementary roles performed by several immune components allow the re-allocation of resources in the event of failure.
- *Self-Organisation*: There is no specific information as to how to respond to a given pathogen, the system responds locally providing a global effect of defence.

Forrest and Hofmeyr take a similar approach in considering and abstracting the information processing principles of the immune system [147]. In their work, the authors focus on three specific information-processing principles, although they highlight a number of general design principles that have much overlap with the de Castro-Timmis listing. Perhaps the more important features that strongly entrench the field of Computational Intelligence are *Immune Learning* (acquired information through interaction with the environment) and *Immune Memory* (persistence and ongoing application of acquired information for short and long durations), both of which fit into a broad notion of intelligence. Using this general description, the mammalian acquired immune system may be considered rudimentary intelligent.

Definitions

Although the motivations of the field are easy to relate, it is important to clarify those motivations that do not contribute to the specific field. This section focuses on some standard definitions of what is and is not an Artificial Immune System. The authors de Castro and Von Zuben [113] define an AIS as “*a computational system based upon metaphors of the biological immune system*”. They continue by defining *Immune Engineering* as the application of an AIS in “*a meta-synthesis process that uses the information contained in the problem itself to define the solution tool to a given problem, and then apply it to obtain the problem solution*”. The distinction is that Immune Engineering is the engineering process, differentiated from conventional engineering process via its biological (immunological) motivation. Immune Engineering was later formalised into a framework by de Castro and Timmis as a set of three principles of what an Artificial Immune System should contain [109], as follows: (1) A representation for the components of the system, (2) A set of mechanisms to evaluate the interaction of individuals with the environment and each other, and (3) Procedures of adaptation that govern the dynamics of the system.

A further, more concise definition of AIS is provided in the same work as “*adaptive systems, inspired by theoretical immunology and observed immunological functions, principles and models, which are applied to problem solving*”. This definition clearly highlights the role of theoretical immunology and immunological observations as an inspiration and not the pursuit of research in AIS, and is the definition adopted in this dissertation. In addition to the clarification it provides, the authors propose that a system to be referred to as an AIS must incorporate a minimum level of immunology, such as an immune component (for example cell, molecule, and organ), it has to be designed by incorporating ideas from theoretical and/or experimental immunology, and it has to be directed toward problem solving. They stipulate that attributing immunological terminology is insufficient to call a system an artificial immune system.

Another perspective of AIS is provided by one of the pioneers in the field Ishida who defines an Immunity-Based System (IMBS) as a “*self-maintenance systems learned from and inspired by the immune system*” [229]. His IMBS is described as (1) a *self-maintenance system* with monitoring of self and nonself, (2) a *distributed system* with autonomous components capable of mutual evaluation, and (3) an *adaptive system* with diversity and selection. The three properties of Ishida’s AIS complement the de Castro-Timmis definition by highlighting core information processing characteristics as a baseline when abstracting from the biological system that may be distilled to (*autonomy, distributed, adaptive*).

2.4 Three Schools of AIS

This section provides a summary of the field of Artificial Immune Systems. The review presents a taxonomy that covers the three core paradigms employed that collectively cover the majority of the work in the field: *Clonal Selection*, *Negative Selection*, and *Immune Network*. Each paradigm is presented from the perspective of immunological theory of immune cell interactions, the abstracted information processing principles, and the al-

gorithms and applications that have been proposed. As such, the review highlights the central role the adaptive clonal selection theory plays across the principles paradigms in the field.

2.4.1 Clonal Selection Paradigm

The clonal selection theory credited to Frank Macfarlane Burnet was proposed to account for the behaviour and capabilities of antibodies in the acquired immune system [63, 64]. Inspired by the principles of Darwinian natural selection theory of evolution, the theory proposes that antigens select-for lymphocytes (both B and T-cells). When a lymphocyte is selected and binds to an antigenic determinant, the cell proliferates making many thousands more copies and differentiates into different cell types (plasma and memory cells). Plasma cells have a short lifespan and produce vast quantities of antibody molecules, whereas memory cells live for an extended period in the host anticipating future recognition of the same determinant. The important feature of the theory is that when a cell is selected and proliferates, it is subjected to small copying errors (changes to the genome called somatic hypermutation) that change the shape of the expressed receptors and subsequent determinant recognition capabilities of both the antibodies bound to the lymphocytes cells surface, and the antibodies that plasma cells produce.

The theory is interesting from an information processing perspective. It suggests that starting with an initial repertoire of general immune cells, the system is able to change itself (the compositions and densities of cells and their receptors) in response to experience with the environment. Through a blind process of selection and accumulated variation on the large scale of many billions of cells, the acquired immune system is capable of acquiring the necessary information to protect the host organism from the specific pathogenic dangers of the environment. It also suggests that the system must anticipate (guess) at the pathogen to which it will be exposed, and requires exposure to pathogen that may harm the host before it can acquire the necessary information to provide a defence.

The information processing principles of the clonal selection theory describe a general learning strategy. The theory has inspired a suite of algorithms for optimisation, classification, and other rudimentary machine learning problem domains. In each algorithm, a population of adaptive information units (each representing a problem-solution or component thereof) is subjected to a competition processes for selection, which together with the resultant duplication and variation, ultimately improves the information units, and solves or approximately solves a problem. The seminal algorithm was proposed by de Castro called the CLONal selection ALGORITHM (CLONALG) [114, 117] and was applied to Function Optimisation, the Travelling Salesman Problem (a combinatorial optimisation problem) and Binary Character Pattern Recognition. Other algorithms include the B-cell Algorithm (BCA) [247, 246], and the Multi-objective Immune System Algorithm (MISA) [79, 83] for optimisation, and the Artificial Immune Recognition System (AIRS) by Watkins [418, 421] for supervised classification. This paradigm is elaborated in Chapter 3.

2.4.2 Negative Selection Paradigm

An important consideration for Burnet in developing the clonal selection theory was the integration of the fact that such a powerfully destructive process as acquired immunity does not immediately turn against the host organism in the generation of self-antibodies (autoimmune response). This ability of the clonal response to mature lymphocyte cells for pathogen and not self-tissues became known as *self-nonsel self discrimination* [86, 305], and in conjunction with clonal selection forms paradigms by which the acquired immune systems is understood in the field of immunology. Clonal selection accounts for this observation by proposing both that (1) the initial repertoire of immune cells are prepared in such a way that none are autoimmune, and (2) that during the maturation process in the ongoing interaction with pathogen, that autoimmune clones are destroyed. Immune cells are continually being destroyed, therefore there is a constant stream of new immune cells that must be created and integrated into the repertoire. This theory was confirmed by the discovery of T-lymphocyte cells that are matured using both a positive and negative selection process in the thymus.

Self-nonsel self discrimination primed with a negative selection process has interesting information processing properties. The theory suggests that the anticipatory guesses made in clonal selection are filtered by regions of infeasibility (protein conformations that bind to self-tissues). Further, the self-nonsel self immunological paradigm proposes the modelling of the unknown domain (encountered pathogen) by modelling the complement of what is known. This is unintuitive as the natural inclination is to categorise unknown information by what is different from what is known rather than guessing at the unknown information and filtering guesses by what is known.

The information processing principles of the self-nonsel self discrimination via negative selection are that of a *anomaly* and *change* detection system that model the anticipation of variation from what is known. The seminal negative selection algorithm was proposed by Forrest, et al. [151] in which a population of detectors are prepared in the presence of known information, where those randomly generated detectors that match against known data are discarded. The population of pattern guesses in the unknown space then monitors the corpus of known information for changes. The algorithm was applied to the monitoring of files for changes (corruptions and infections by computer viruses), and later formalised as a change detection algorithm [129, 130]. It was applied to monitoring changes in the execution behaviour of Unix processes [149, 215], and to monitor the changes in remote connections of a network computer (intrusion detection) [211, 212]. The application of the algorithm has been predominantly to virus host intrusion detection and their abstracted problems of classification (two-class) and anomaly detection. More recently, the validity of the application of negative selection algorithms in high-dimensional spaces has been questioned, specifically given the scalability of the approach in the face of the exponential increase in volume within the problem space [374].

2.4.3 Immune Network Paradigm

A concern of the clonal selection theory is that it presumes that the repertoire of reactive cells remains idle when there are no pathogen to which to respond. Considering this problem, Niels Jerne proposed an Immune Network Theory (Idiotypic Networks) where immune cells are not at rest in the absence of pathogen, instead antibody and immune cells recognise and respond to each other [233, 234, 235]. The theory proposes that antibody (both free floating and surface bound) possess idiotopes (surface features) to which the receptors of other antibody can bind. As a result of receptor interactions, the repertoire becomes dynamic, where receptors continually both inhibit and excite each other in complex regulatory networks (chains of receptors). The network theory suggests that the clonal selection process may be triggered by the idiotopes of other immune cells and molecules in addition to the surface characteristics of pathogen, and that the maturation process applies both to the receptors themselves the idiotopes which they expose.

The immune network theory has interesting resource maintenance and signalling information processing properties. The clonal selection and negative selection paradigms integrate the accumulative and filtered learning of the acquired immune system, whereas the immune network theory proposes an additional order of complexity between the cells and molecules under selection. In addition to cells that interact directly with pathogen, there are cells that interact with those reactive cells and with pathogen indirectly, in successive layers such that networks of activity for higher-order structures such as internal images of pathogen (promotion), and regulatory networks (so-called anti-idiotopes and anti-anti-idiotopes).

Early works, such as Farmer, et al. [138] suggested at the exploitation of the information processing properties of network theory for machine learning. A seminal network theory based algorithm was proposed by Timmis, et al. called the Artificial Immune Network (AIN) [392] (later extended and renamed the Resource Limited Artificial Immune System [393, 391] and Artificial Immune Network [258]). The algorithm maintains a population of adaptive units that operate under a clonal selection process. Units are connected to each other based on distance measures in the information space (for example Euclidean distance if the units are real-valued vectors), such that adaptive units that are *close* in the information space suppress the activation of each other. Therefore, the more *connections* a given unit has with other nodes influences the units potential for being selected and cloned under clonal selection. Timmis' system was applied to feature extraction and data clustering. Another seminal network based algorithm was proposed by de Castro and Von Zuben called Artificial Immune Network (aiNet) which extended the principles of AIN and CLONALG and was applied to clustering [115, 116]. The aiNet algorithm was further extended to optimisation domains and renamed opt-aiNet [108].

2.4.4 Summary

The clonal selection, negative selection, and immune network paradigms represent the majority of the influences for the information processing properties from immunology exploited in the field of artificial immune systems. The proposed taxonomy is not complete,

for example an additional influence that is not discussed is the Danger Theory of Polly Matzinger [287, 288] which proposes that the acquired immune system responds to signals of damage, which opposes the fundamental self-nonsel paradigm. This coarse taxonomy of the field of Artificial Immune Systems clearly highlights the central role clonal selection to both the cognitive properties assigned to the immune system (learning and memory), and the three principle paradigms that underlie the field. This centrality and even reliance on clonal selection suggests that developments and improvements to the understanding of the paradigm will have follow-on effects to the other cell-based paradigms in the field.

2.5 Distributed Artificial Immune Systems

One of the often cited propositions of using the immune system as an inspiration for computational systems is that it affords decentralised and distributed information processing. This section reviews this promise of the field of Artificial Immune Systems, and reviews systems and algorithms that have made progress toward this end.

2.5.1 Promise of Distributed Information Processing

The terms *autonomous*, *decentralised*, and *distributed* are used with abandon in the field of Artificial Immune Systems to both describe information processing abstractions of the immune system and the desired characteristics of inspired systems. For example, the motivations and definitions for the in Section 2.3.2 both feature the distributed aspects of the immune system. Watkins, et al. described the “*inherent distributedness*” of the biological immune system, and the “*almost embarrassingly parallel nature*” of the clonal selection algorithm [417]. Watkins and Timmis further pushed this perspective and commented that “*Among the oft-cited reasons for exploring the mammalian immune system as a source of inspiration for computational problem solving include the observations that the immune system is inherently parallel and distributed*” [420]. Therefore, it is not unreasonable to consider that a promise of the field of Artificial Immune Systems is a source of distributed information processing paradigms, systems, and/or algorithms, imbued with notions of autonomy, decentralisation, parallelism, concurrency.

In this work, the terms are used with explicit meaning. *Autonomy* refers to the self-governance of discrete facets of the system such as the molecules and cells and to the processes in which they are engaged for example the identification and neutralisation of pathogen. *Decentralised* refers to the lack of centralised control, specifically the localisation of the self-governance such as at the spatial and temporal characteristics of a pathogenic encounter. Finally, *distributed* is used in the systems sense of the integration and coordination of the localised self-governance of resources towards a holistic end, such as homoeostasis and general organism defence from infection. The definitions used align with the pursuit of Computational Intelligence systems that exploit the distributed information processing with multiple interacting and intelligent agents referred to as Distributed Artificial Intelligence and Multi-Agent Systems [425]. From Weiss’ definition, the *autonomy*, *emergent cooperation* and *intelligence* are preserved although the constraints of

rationality, local determinism, and goal or task-directed behaviour are relaxed in favour of the biologically inspired bottom-up and stochastic mechanisms of Computational Intelligence. Therefore, a distinction is made between centralised, distributed, and distributable, where a system may be designed or re-engineered to be *distributed* although the benefits of such a process are not limited to concurrent, parallel, or distributed implementation in the computer hardware sense.

2.5.2 Algorithms, Systems, and Applications

This section reviews the efforts towards the realisation of distributed Artificial Immune Systems. Considered in this review are the seminal works and main efforts toward such realisation which are categorised as (1) distributed negative selection approaches for computer security, (2) efforts toward autonomous and decentralised control systems and robotics, (3) network systems and application domains, (4) parallel extensions of existent algorithms, and (5) miscellaneous works.

Distributed Detectors for Computer Security

In their proposal of a distributed detection system based on the negative selection algorithm Forrest, et al. comment on the characteristics of the acquired immune systems pathogen identification and memory capabilities, highlighting that they are “*highly distributed*”, and that responses are decentralised and “*highly localized*” [148, 212]. Hofmeyr, et al. extended the approach and applied it to the domain of computer network intrusion detection [211]. In their system, a central negative selection process (central tolerisation) prepares a private set of detectors for each machine on a computer network such that the diversity of each machines detector set contributes to the protection of the entire network. The system was simulated demonstrating the principle, and was implemented although not in a distributed manner. The research was abstracted to a general framework for distributed adaptive systems called the *ARTificial Immune System* (ARTIS), with the application to intrusion detection renamed the *Lightweight Intrusion Detection System* (LISYS) [214]. The detectors of LISYS involve a life-cycle that extends the negative selection algorithm including a naïve state for untested detectors, a memory state for proven detectors, and an activation threshold used to manage the sensitivity of raised intrusion signals. Three specific distributed characteristics of acquired immunity are not realised in LISYS, specifically (1) the detectors are not mobile although memory detectors are shared with neighbouring systems, (2) the detectors do not undergo a clonal selection process when activated, and (3) the system is not autonomous requiring the intervention of a human operator in the promotion of activated detector to memory detector.

The Computer Defence Immune System (CDIS) proposed an elaborate multi-agent system with much commonality to ARTIS and LISYS [432, 197, 17, 18]. Among the many enhancements to LISYS, the CDIS intrusion detection system facilitated the hierarchical dissemination of acquired information, specifically the signatures for detectors of verified intrusion attempts. Kim, et al. proposed an intricate and integrated system of artificial immune algorithms for intrusion detection [252, 253, 254, 255, 251]. The work

approached the problem of network intrusion detection in the same manner as LISYS and CDIS, although it employed gene library evolution to improve detector generation during centralised negative selection, and a dynamic clonal selection algorithm to improve detectors successful at identifying anomalies. Although the system was designed for distributed intrusion detection, it was not simulated or implemented in a distributed environment. Finally, Dasgupta and Brian proposed a multi-agent system architecture for intrusion detection with mobile agents on a network where anomaly detection was achieved through a loose analogy of distributed immune cells [107].

Control, Robots, and Autonomous Navigation

Early contributions to the field of AIS were made by Ishida, et al. by phrasing distributed fault diagnosis problems in the connectionist paradigm of Parallel Distributed Processing, and exploring the capabilities of Jerne’s network theory on such problems [227]. The models were graphs of sensor nodes that performed distributed information processing (sensor fault diagnosis) without a central point of control [295]. Ishida expanded upon the work to exploit the systems autonomous cooperative decision making and self-organising capabilities in domains such as coin balancing and flow monitoring [228].

Lee and Jun, et al. proposed a Distributed Autonomous Robotic System that achieved cooperative control of robots through emergent swarm behaviour for the selection of group behaviour strategies in autonomous robots [270, 238]. The system was modelled such that each robot is a B-cell that expresses behaviour strategies antibodies in an environment of conditions (pathogens), with control parameters managed as T-cells. Robots select their own behaviour strategies which are communicated with each other, propagating (clonal selection proliferation) if useful, or suppressed if not, in a simulated immune network. The robots were simulated in a dynamic environment to perform group tasks collectively selecting action strategies. A similar approach inspired by Lee and Jun, et al. was employed by Meshref and VanLandingham and applied to the benchmark dynamic control dog-and-sheep problem, where dogs (B-cells) coordinate their behaviour to get sheep (pathogen) into a pen [291]. Singh and Thayer proposed an architecture called the Immunology-derived Distributed Autonomous Robotics Architecture for the large-scale coordination and control of robot teams (many thousands of robots) simulated on a dynamic mobile mine clearing domain [359]. Tasks in the domain were implements as pathogen, whereas robot action strategies of varying generality were represented as immune cells that are selected for a given robots current needs. A uni-directional beacon broadcast inter-robot communication was used to coordinate behaviour inspired by a histamine-based recruitment mechanism. The approach resulted in emergent coordination that the authors claimed did not become apparent without large colonies of robots. The framework was further applied to simulations of an exploration and mapping domain [384, 361], and a search-and-rescue domain [360].

Lau, et al. proposed a multi-agent control system for fleets of autonomous guided vehicles where vehicles are antibodies, and unassigned tasks they must perform are pathogen to which they specialise [267, 266, 268]. Michelan and Von Zuben propose a decentralised

autonomous navigation system for a robot for garbage collection (locate, collect, obstacle avoidance, limited resources), that although not distributed, provided an archetype AIS application in robotics toward complex autonomous behaviour [294]. Ishiguro, et al. also proposed a decentralised robotic application with behaviour arbitration for a mobile robot that although not implemented in multiple robots, demonstrated complex autonomous behaviour [230, 231].

Network Systems

Boudec and Sarafijanovic proposed an AIS for the detection of irregular dynamic routing behaviour in mobile ad hoc networks [52]. Their negative selection based system identified misbehaving relays in a simulated network test bed by observing routing events with detectors prepared off-line. Their approach was extended through the addition of a virtual thymus used to maintain a dynamic representation of normal and abnormal relay behaviour, distributed danger signals to govern the flow of information into the virtual thymus, and memory detectors for improved secondary response [345, 346]. Ganguly, et al. proposed an *ImmuneSearch* algorithm for a peer-to-peer network topology that was decentralised using affinity-based proliferation and mutation to control message movement to avoid flooding the network with messages [160]. Kephart proposes a vision for a scaled up automated and adaptive virus signature system providing protection across the Internet [250]. Mohr, et al. proposed exploiting information processing properties of the immune system such as continual adaptation, distributedness, and context-dependant response for a context-aware ubiquitous computing platform (also known as pervasive computing and ambient intelligence) suitable for integration of hand-held devices [296]. Ishida proposed the immune system from the perspective of the immune network theory provides the foundational inspiration for developing an Autonomous Distributed Systems (ADS), which is a general class of system of autonomous decision making agents organised in dynamic structure that cannot be designed beforehand [228]. Ishida highlights examples of adaptive, autonomous, and decentralised artificial immune systems, although does not provide any specific information as to how an ADS may be realised from an immunological basis.

Parallel Algorithms

The work by Watkins, et al. addressed the so-called unexploited *inherent parallelism* of the CLONALG algorithm [417]. Their effort extended the algorithm such that pattern recognition tasks are partitioned and executed concurrently (explicit divide and conquer), and reassembled after execution. The same parallel strategy was applied to the AIRS algorithm by Timmis and Watkins [420]. Coello Coello, et al. proposed a parallel implementation of their constraint handling immune algorithm as a strategy for assessing infeasible solutions [79, 78]. Their approach was implemented as an extension to the standard genetic algorithm implemented in an multiple island population configuration.

Other

Hart and Ross explored the analogy between acquired immunity to addressable memory, specifically as Sparse Distributed Memories (SDM). These structures are a robust content-addressable memory which are efficient at storing vast amounts of binary data with a small number of physical addresses [365]. A hybrid method was applied to clustering dynamic datasets that combined SDM, antigen-antibody matching for addressing, and a co-evolutionary algorithm [199, 200]. Newborough and Stepney proposed a generic framework for population based algorithms based on a general biologically-inspired conceptual framework [304] (considered further in Section 2.7.1). Their framework provided abstractions of common optimisation algorithms (clonal and negative selection, genetic algorithm, swarm algorithms such as particle swarm and Ant Colony Optimisation), and a proof-of-concept implementation on Field Programmable Gate Array demonstrated both hardware parallelism and distribution. Toma, et al. proposed an Immune Distributed Competitive Problem Solver that employed aspects of MHC and immune networks in a multi-agent system to address division of labour problems such as n -Travelling Salesman Problem [395].

2.5.3 Summary

The clonal selection, negative selection and immune network cellular immunological theories describe decentralised processes capable of general autonomous (per-cell and per-pathogen) decision making, suitable for distributed and parallel information processing. The exploitation of these information processing attributes imposed on the biological system is the promise of distributed Artificial Immune Systems, and is one of the principle areas of research in the field of AIS, enshrined in definition. The three schools of AIS demonstrated that the first-order realisation of Artificial Immune Systems is typically simplistic and monolithic. The work on distributed detector-based computer security driven by negative selection, and emergent cooperation in autonomous mobile robots are the two most promising thrusts of research toward realising this goal. The work towards parallelising existent monolithic immunological algorithms and network applications may be fruitful areas, although these are still immature and emergent areas of research. Distributed and decentralised realisations of artificial immune systems represent the overarching ambition of the field in general. Such information processing characteristics can be directly harnessed for classes of function optimization and function approximation problem domains suited to parallel and cooperative solving strategies. What is lacking in these efforts is a clear framework to guide the development of a distributed AIS, although Kephart's vision for cyberspace anti-virus, Mohr ubiquitous AIS, and Ishida's vision for an immunological ADS are thought provoking.

2.6 Open Problems in AIS

This section considers the open problems in the field of artificial immune systems, and the current trends to address these problems. The review highlights the perceived impasse

in AIS and the trends of drawing on novel and more accurate models from the inspiring biological system.

2.6.1 Open Problems

The authors de Castro and Timmis comment on the lack of and need for a adequate framework for AIS design, interpretation, and application [111]. They propose a general AIS framework for such purposes that includes three core facets, which are presented formally in [109] (pages 60-61) as an AIS Engineering Framework, they are: (1) *Representation* for the components of the system, (2) *Affinity Measures* and mechanisms to evaluate the components of the system in the context of the domain (such as fitness scoring), and (3) *Algorithms* and procedures for adaptation that govern the dynamics of the system. They highlight both the need for more suitable real-world applications of AIS, and the need to consider the accuracy of the inspiring metaphor, specifically the importance for computer scientists to grasp the more subtle aspects of immunology [109] (page 310). They outline four general areas of research as future trends in the field: (1) improvement and augmentation of AIS, (2) new artificial immune systems, (3) application areas, and (2) extending their general AIS framework.

Freitas and Timmis highlight the increasing trend in the field of applying generic AIS algorithms to specific and well-defined benchmark domains without the required specialisation of such algorithms [155]. The authors consider the classification domain and negative selection algorithms specifically, highlighting the need to consider standard machine learning principles such as over fitting and inductive bias which are not addressed in such applications. Andrews and Timmis consider a path to *next generation* artificial immune systems, commenting that an assessment of the current state of the field reveals the majority of AIS are almost exclusively inspired by the theories of Jerne's network theory and Burnet's clonal selection theory [24]. They propose to consider other and potentially controversial immunological theories such as the Cognitive Theory proposed by Cohen, and advocate the adoption of a guiding conceptual framework.

Timmis, et al. proposed that the *first generation* of biologically motivated systems are crude analogies of biological systems, proposing that an inter-disciplinary approach (computer scientists and biologists) is needed to define the *second generation* of such systems [389]. They highlight both the importance of understanding the biological system being exploited, and that high level abstractions of underlying biology are no longer sufficient requiring a deeper level of understanding. They propose the path to addressing these concerns is an inter-disciplinary approach. Their argument is made in the context of a number of computational intelligence paradigms, including AIS. They comment on the crudeness of AIS algorithms such as CLONALG that "*whilst intuitively appealing, lacks any notion of interaction of B-cells with T-cells, MHC, or cytokines*". Garrett proposes a framework for assessing usefulness based on the *distinctiveness* and *effectiveness* of AIS and applies the criteria to the three schools of AIS, showing that there is indeed some useful contributions [164].

Timmis suggests that the field of AIS has reached an impasse [388], specifically with the

lack of theoretical advancement, limited application, and the adoption of naïve biological models. He proposes a number of challenges in the field to address the impasse: (1) the development of novel and accurate metaphors that provide benefit to immunology, (2) improved application of AIS models work toward a *killer application* for AIS, (3) to develop a theoretical basis for AIS, and (4) to consider the integration of the immune system with other biological systems. Hart and Timmis assess the field from an application centric perspective, highlighting distinct lack and narrow focus of real-world applications [201]. They consider the so called *scatter gun* approach to applications is caused by poor foundational biological metaphors, suggesting three areas that may provide a fertile future for the field: (1) the innate immune system, (2) embodied immune system working towards host homeostasis, and (3) life-long learning of the system. Twycross and Aickelin consider Timmis’ proposed impasse in AIS, highlighting its causes as limited application to challenging problems, lack of theoretical advances, and use of naïve biological metaphors [401]. They proposed to address this impasse by considering two addition unexploited biological metaphors that are suggested as being *less complex* than those of the human immune system: plant immunity and invertebrate immune systems.

2.6.2 Summary

The general open problems and trends to address them are common to the field of Computational Intelligence, specifically the need for a theoretical underpinning for the the field and suitability and specialisation of algorithms to application domains. Beyond these general concerns, there is a specific desire for a field-defining (so-called *killer*) application domains, and the trend to continue to assimilate more exotic immunological theory and sub-fields towards this end. Further, there is a concerted general intent to devise and employ general conceptual frameworks to unify existing algorithms, and in the investigation of novel Artificial Immune Systems. This work proposes to address the open problem of a perceived impasse in the field of Artificial Immune Systems by following the specific emerging trends in the field of (1) exploiting novel and more accurate immunological metaphor, (2) devising and exploiting a general framework that unites the facets of the novel metaphor, and (3) extending the existing AIS frameworks via integration. This work does not address the open problems of an improved theoretical underpinning of the field or killer applications, although it is believed that the work will provide insight toward these questions.

2.7 Methodology

This section specifically address the research goal of identifying a systematic methodology for the development and investigation of immunologically inspired algorithms. The methodology proposed is comprised of the best practices regarding (1) the abstraction of immunological properties and development of inspired algorithms in Section 2.7.1, (2) experimental methodology for the qualitative empirical investigation of developed algorithms in Section 2.7.2, and (3) the important considerations regarding the ‘economy of

modelling’ of developed algorithms and the need for *patch-quilt* integration of findings into broader frameworks in Section 2.7.3.

2.7.1 Conceptual Framework

Conceptual Framework for Bio-Inspired Algorithms

Although a progression from inspiring biological system to inspired computation system may appear to be an intuitive process, it can involve problems of standardisation of nomenclature, effective abstraction and departure from biology, and rigour. Stepney, et al. caution that by following a process that lacks the detail of *modelling*, one may fall into the trap of *reasoning by metaphor* [400, 372, 373]. Besides the lack of rigour, the trap suggests that such reasoning and lack of objective analysis limits and biases the suitability and applicability of resultant algorithms. They propose that many algorithms in the field of Artificial Immune Systems have succumbed to this trap. This observation resulted in the development and application of a conceptual framework to provide a general process that may be applied in the field of Biological Inspired Computation toward realising Biological Inspired Computational Intelligence systems. The conceptual framework is comprised of the following actors and steps:

1. *Biological System*: The driving motivation for the work that possess some innate information processing qualities.
2. *Probes*: Observations and experiments that provide a partial or noisy perspective of the biological system.
3. *Models*: From the probes abstract and simplified models of the information processing qualities of the system are build and validated.
4. *Framework*: Build and validate analytical computational frameworks. Validation may use mathematical analysis, benchmark problems and engineering demonstration.
5. *Algorithms*: The framework provides the principles for designing and analysing algorithms that may be general and applicable to domains unrelated to the biological motivation.

Immunology as Information Processing

Forrest and Hofmeyr summarised their AIS research efforts at the University of New Mexico and the Santa Fe Institute as “*immunology as information processing*” [147]. They define information as spatio-temporal patterns that can be abstracted and described independent of the biological system, and information processing as computation with these patterns. They proposed that such patterns are encoded in the proteins and other molecules of the immune system, and that they govern the behaviour of the biological system. They suggest that their information processing perspective can be contrasted

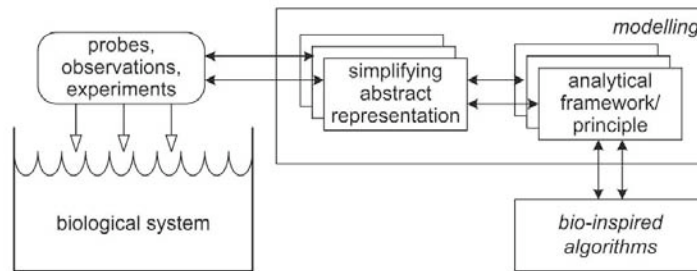


Figure 2.1: Depiction of the conceptual framework for devising Biological Inspired Algorithms, taken from [373]

with the conventional structural perspective of cellular interactions as mechanical devices. They consider a simple four-step procedure for the investigation of *immunology as information processing*, transitioning from the biological system to a usable computational tool:

1. Identify a specific mechanism that appears to be interesting computationally.
2. Write a computer program that implements or models the mechanism.
3. Study its properties through simulation and mathematical analysis.
4. Demonstrate capabilities either by applying the model to a biological question of interest or by showing how it can be used profitably in a computer science setting.

The procedure is similar to the outlined in the conceptual framework for Biologically Inspired Algorithms in that in addition to identifying biological mechanisms (input) and demonstrating a resultant algorithms (output), the procedure (1) highlights the need for abstraction involving modelling the identified mechanism, and (2) highlights the need to analyse the models and abstractions. The procedure of Forrest and Hofmeyr can be used to specialise Stepney, et al. conceptual framework by clearly specifying the immunological information processing focus.

2.7.2 Experimental Methodology

Optimization is a generalized class of problem involving improvement of a structure under a cost function. It is a one of a few standard perspectives for the investigation and application for computational intelligence systems, specifically including clonal selection algorithms. As such, this section considers the state of optimization algorithm experimental methodology, the concerns of which are applicable to the investigation of computational intelligence algorithms in general.

When it comes to evaluating an optimisation algorithm, every researcher has their own thoughts on the way in which to proceed. Unfortunately, many empirical evaluations of optimisation algorithms are performed and reported without addressing basic experimental design considerations. Perhaps before an experimental methodology can be adopted, a

researcher or practitioner may be paralysed by the perceived pessimism of the *no free lunch* theorem that contends the futility of the benchmarking exercise. A pervasive problem in the field of Computational Intelligence algorithms is the lack of meaningful and consistent algorithm experimental and benchmarking methodology. This includes but is not limited to issues of the selection of problem instances, the selection of algorithm specifications, the algorithm configuration parameters, and interpretation of results. The intention of this section is to summarise the general concerns of experimental methodology, with a focus on benchmarking Computational Intelligence algorithms in the face of the ‘no free lunch’ theorem.

No Free Lunch

Wolpert and Macready’s *No Free Lunch Theorem* of search and optimisation has caused a lot of pessimism and misunderstanding, particularly in related to the evaluation and comparison of computational intelligence algorithms [436, 435]. In simplest terms, the theory indicates that when searching for an extremum of a cost function, *all algorithms perform the same when averaged over all possible cost functions*. The implication is that the often perused general-purpose optimisation algorithm is theoretically impossible. The theory applies to stochastic and deterministic optimisation algorithms, and to algorithms that learn and adjust their search strategy over time. It is invariant to the performance measure used as well as the representation selected. Perhaps the catalyst for benchmarking cynicism is a comment accompanying the proof suggesting that: “... *comparisons reporting the performance of a particular algorithm with a particular parameter setting on a few sample problems are of limited utility*” [436]. The theorem is an important contribution to computer science, although its implications are theoretical. The original paper was produced at a time when grandiose generalisations were being made as to algorithm, representation, or configuration superiority. The practical impact of the theory is to *bound claims of applicability*. Wolpert and Macready encouraged effort be put into devising practical problem classes and the matching of suitable algorithms to problem classes. Further they compelled practitioners to exploit domain knowledge in optimisation algorithm application, now an axiom in the field.

Issues of Benchmarking Methodology

Empirically comparing the performance of algorithms on problem instances is a staple for the fields of heuristics and Computational Intelligence, and the problems of effective comparison methodology have been discussed since the inception of these fields. Johnson suggested that the coding of an algorithm is the easy part of the process, that the difficult work is getting meaningful and publishable results [236]. He provides a very thorough list of questions to consider before *racing* algorithms, as well as what he describes as his *pet peeves* within the field of empirical algorithm research. Hooker (among others) condemns what he refers to as *competitive testing* of heuristic algorithms, calling it *fundamentally anti-intellectual* [221]. Hooker continues by strongly encouraging a rigorous methodology of what he refers to as *scientific testing* where the aim is to investigate algorithmic be-

haviours. Barr, et al. list a number of properties worthy of a heuristic method making a contribution, which can be paraphrased as; efficiency, efficacy, robustness, complexity, impact, generalisability and innovation [33]. Barr, Golden et al. specify a loose experimental set-up methodology with the following intuitive steps; 1) define the goals of the experiment, 2) select measure of performance and factors to explore, 3) design and execute the experiment, 4) Analyse the data and draw conclusions, and finally 5) report the experimental results. They then suggest eight guidelines for reporting results, in summary they are; reproducibility, specify all influential factors (for example code and computing environment), be precise regarding measures, specify parameters, use statistical experimental design, compare with other methods, reduce variability of results, ensure results are comprehensive.

Peer, et al. summarise the problems of algorithm benchmarking (with a bias toward Particle Swarm Optimisation) to the following points; duplication of effort, insufficient testing, failure to test against state-of-the-art, poor choice of parameters, conflicting results, and invalid statistical inference [318]. Eiben and Jelasity site four problems with the state of benchmarking evolutionary algorithms; (1) test instances are chosen ad hoc from the literature, (2) results are provided without regard to research objectives, (3) scope of generalised performance is generally too broad, and (4) results are hard to reproduce [132]. The general problems with benchmarking methodology may be distilled into the following sub-problems:

- *Parameter Selection*: Computational Intelligence algorithms are typically parameterised, although the mapping of parameter values to predictable effects on problem domains is generally poorly understood. This is because typically unknown and non-linear dependencies commonly exist between the variables resulting in the algorithm being consider a complex system. Francois and Lavergne discuss the deficiencies of the traditional *trial-and-error* and the *experienced-practitioner* approaches to parameter tuning, further suggesting that seeking general rules for parameterisation will lead to optimisation algorithms that offer neither convergent or efficient behaviours [153]. There are many solutions not limited to the following: consulting the literature although often ignored [132]), generalise from large numbers of experiments [347], self-adaptive parameters, meta-algorithms for searching for good parameters values, and statistical sensitivity analysis over parameter ranges [74, 45].
- *Problem Selection*: Problem instances should be selected to demonstrate specific behaviours of the systems under study. Typically benchmarks are performed on ad hoc lists of standard benchmark instances, that although offering the benefit of comparability, typically devolve into racing contest. Eiben and Jelasity support the division of problem instances into categories and encourage the evaluation of optimisation algorithm over a large number of test instances [132]. In their paper on understanding the interactions of Genetic Algorithm parameters Deb and Agrawal propose four structural properties of problems for testing genetic algorithms; multimodality, deception, isolation, and collateral noise [125]. Yao, et al. divide their large test dataset into the categories of uni-modal, multi-modal with many local

optima, and multi-modal with few local optima [441]. Whitley, et al. provide a detailed study on the problems of selecting test instances for genetic algorithms, and suggest that difficult problem instances should be non-linear, non-separable, and non-symmetric [431].

- *Measure Selection*: There are many ways to measure the performance of an algorithm for a problem instance, although the most common involves measures of *efficacy* (problem specific costs) and *efficiency* (computation space and/or time). Most Biologically Inspired Algorithms have a stochastic element, typically in their random starting position(s) and in the probabilistic decisions made during sampling of the domain. Therefore, the measuring of performance must be repeated a number of times to account for the stochastic variance², which also could be a measure of comparison between algorithms. The most critical concern in the selection of measures is that of standardisation for the purposes of comparison [44, 33].
- *Statistical Significance*: The scientific knowledge of Computational Intelligence is predominantly accrued as empirical observations and conclusions. Therefore, there is a need for methods that permit interpretation and conclusions to be drawn beyond the standard reporting of results. Statistical hypothesis testing and related methods provides such tools, although unfortunately are rarely employed. Peer, et al. [318] and Birattari and Dorigo [43] provide a basic introduction (suitable for an algorithm-practitioner) into the appropriateness of various statistical tests for algorithm comparisons. Parametric statistical methods are used for interval and ratio data (like a real-valued performance measure), and non-parametric methods are used for ordinal, categorical and rank-based data. Interval data is typically converted to ordinal data when salient constraints of desired parametric tests (such as assumed normality of distribution) are broken such that the less powerful non-parametric tests can be used. The use of non-parametric statistical tests maybe preferred as some authors (for example [75]) claim the distribution of cost values are very asymmetric and/or not Normal.

2.7.3 Designing and Investigating Adaptive Systems

Complex and adaptive systems are difficult to design and investigate, therefore the need for a coherent and proven methodology for such activities is paramount. Inspired by an interpretation of the achievements of the Wright brothers in achieving the first powered flight, this section provides a summary of a methodology for investigating and designing *conceptual machines* proposed by David Goldberg [177].

For the purposes of clarity, a *model* is a conceptualisation that provides a description of a system that accounts for its properties and may be used to study the characteristics of the system. For the remainder of this thesis, conceptual models are instantiated as *algorithms* for the purposes of computational implementation, simulation, and study. The

²Typically ≥ 30 according to the central limit theorem such that the underlying distribution can be meaningfully summarised.

algorithms provide procedures for problem-solving when the behaviour of the instantiated model suitably match the characteristics of a problem.

Engineers and Mathematicians

Goldberg describes the airplane and other products of engineering as *material machines*, and distinguishes them from the engineering of genetic algorithms and other adaptive systems as *conceptual machines*. He argues the methodological distinction between the two is counter-productive and harmful from the perspective of conceptual machines, specifically that the methodology of the material is equally applicable to that of the conceptual [174]. The obsession of mathematical rigour in computer science, although extremely valuable, is not effective in the investigation of adaptive systems given their complexity. Goldberg sights the airplane as an example where the engineering invention is used and trusted without a formal proof that the invention works (that an airplane can fly³). This defence leads to what Goldberg refers to the *economy of design* which is demonstrated with a trade-off that distinguishes ‘model description’ (mathematician-scientists) that is concerned with model fidelity, and model prescription (engineer-inventor) that is concerned with a working product. In descriptive modelling *the model is the thing* (of interest) whereas in ‘prescriptive modelling’, *the object is the thing* (of interest). In the latter, the model (and thus its utility) serves the object, in the former model accuracy may be of primary concern. This economy of modelling provides a perspective that distinguishes the needs of the prescriptive and descriptive fields of investigation.

The mathematician-scientist is interested in increasing model accuracy at the expense of the speed (slow), where as the engineer may require a marginally predictive (inaccurate) model relatively quickly. This trade-off between high-cost high-accuracy models and low-cost low-fidelity models is what may be referred to as the *modelling spectrum* that assists in selecting an appropriate level of modelling. Goldberg proposes that the field of genetic algorithms expend too much effort at either ends of this spectrum. There is much work where there is an obsession with blind-prototyping many different tweaks in the hope of striking it lucky with the *right* mechanism, operator, or parameter. Alternatively, there is also an obsession with detailed mathematical models such as full-blown differential equations and Markov chains. The middle ground of the spectrum, what Goldberg refers to as *little models* is a valuable economic modelling consideration for the investigation of conceptual machines to *do good science through good engineering*.

Methodology

The methodology has been referred to as post-modern systems engineering and is referred to by Goldberg as a methodology of innovation [175]. The core principles of the process are as follows:

1. *Decomposition*: Decompose the large problem approximately and intuitively, breaking into quasi-separate sub-problems.

³Goldberg is quick to point out that sets of equations do exist for various aspects of flight, although no integrated mathematical proof for airplane flight exists.

2. *Modelling*: Investigate each sub problem separately (or as separate as possible) using empirical testing coupled with adequately predictive, low-cost models.
3. *Integration*: Assemble the sub-solutions and test the overall invention, paying attention to unforeseen interactions between the sub-problems.

Decomposition Problem decomposition and decomposition design is an axiom of reductionism and is at the very heart of problem solving in computer science. Therefore, it is not worth dwelling on the topic other than to comment as to its meaning within the context of adaptive systems. One may consider the base or medium on which the system is performing its computation mechanisms, the so-called building blocks of information processing. A structural decomposition may involve the architecture and data structures of the system. Additionally, one may also consider a functional breakdown of mechanisms such as the operators applied at each discrete step of an algorithmic process or mechanisms. The reductions achieved provide the basis of investigation and modelling.

Small Models Given the principle of the economy of modelling presented as a spectrum, one may extend the description of each of the five presented model types. *Small Models* refers to the middle of the spectrum, specifically to the application of dimensional and facet-wise models. These are mid-range quantitative models that make accurate prediction over a limited range of states at moderate cost. Once derived, this class of models generally requires a small amount of formal manipulation and large amounts of data for calibration and verification. The following summarises the modelling spectrum:

- *Unarticulated Wisdom*: (low-cost, high-error) Intuition, what is used when there is nothing else.
- *Articulated Qualitative Models*: Descriptions of mechanisms, graphical representations of processes and/or relationships, empirical observation or statistical data collection and analysis.
- *Dimensional Models*: Investigate dimensionless parameters of the system.
- *Facet-wise Models*: Investigation of a decomposition element of a model in relative isolation.
- *Equations of Motion*: (high-cost, low-error) Differential equations and Markov chains.

Facet-wise models are an exercise in simple mathematics that may be used to investigate a decomposition element of a model in relative isolation. They are based on the idea of *bracketing high-order phenomena* by simplifying or making assumptions about the state of the system. An example used by Goldberg from fluid mechanics is a series of equations that simplify the model by assuming that a fluid or gas has no viscosity, which matches no known substance. A common criticism of this modelling approach is “*system X doesn’t work like that, the model is unrealistic*”. The source of such concerns with adaptive systems is that their interactions are typically high-dimensional and non-linear.

Goldberg’s response is that for a given poorly understood area of research, any useful model is better than no model. Dimensional analysis or the so-called dimensional reasoning and scaling laws are another common conceptual tool in engineering and the sciences. Such models may be used to investigate dimensionless parameters of the system, which may be considered the formalisation of the systemic behaviours.

Integration Integration is a unification process of combining the findings of various models together to form a *patch-quilt* coherent theory of the system. Integration obviously is not limited to holistic unification, one may address specific hypothesis regarding the system resulting in conclusions about existing systems, and design decisions pertaining to the next generation of systems.

Application In addition to elucidating the methodology, Goldberg specifies a series of five useful heuristics for the application of the methodology as follows (taken from [174], page 8):

1. *Keep the goal of a working conceptual machine in mind.* Experiments commonly get side tracked by experimental design and statistical verification; theoreticians get side tracked with notions of mathematical rigour and model fidelity.
2. *Decompose the design ruthlessly.* One cannot address the analytical analysis of a system like a genetic algorithm in one big ‘gulp’.
3. *Use facet-wise models with almost reckless abandon.* One should build easy models that can be solved by bracketing everything that gets in the way.
4. *Integrate facet-wise models using dimensional arguments.* One can combine many small models together in a patch-quilt manner and defend the results of such models using dimensional analysis.
5. *Build high-order models when small models become inadequate.* Add complexity to models as complexity is needed (economy of modelling).

2.8 Chapter Summary

This Chapter positioned the investigation broadly in the field of Biologically Inspired Computational Intelligence, a sub-field of Artificial Intelligence, and specifically in the field of Artificial Immune Systems. Clonal Selection was highlighted as an adaptive paradigm in the field of Artificial Immune Systems that underlies much of the general field. Distributed Artificial Immune Systems including distributed Clonal Selection Algorithms is an over promised and under delivered, although potentially beneficial field of study. An effective realisation of distributed clonal selection algorithms is expected to facilitate the broader application of such approaches to classes of problem that may benefit from parallel and cooperative problem solving strategies, such as problems with functional or information availability decompositions. Turning back to the structure and function of the immune

system to develop new AIS approaches including distributed approaches is currently an advocated approach in the field given the apparent impasse experienced in the progress of such approaches. A systematic framework was adopted for the development of new immunologically inspired approaches, with careful empirical investigation, and awareness of the economy of models and need for the re-integration of findings. This chapter both positioned and motivated the research hypothesis outlined in Chapter 1, and defined the systematic methodology adopted for the development and investigation of adaptive and distributed clonal selection algorithms throughout the remainder of the dissertation. Toward this end, Chapter 3 provides an in depth investigation into the clonal selection paradigm, both reviewing and elaborating the cellular perspective of such approaches which is later pursued in Chapter 4, and outlining the agenda for realising distributed approaches in Chapters 5 and 6.

Chapter 3

Clonal Selection Paradigm

3.1 Chapter Overview

This chapter provides an in depth review of clonal selection theory and inspired computational algorithms, highlighting the open problems in the state of the field that specifically motivate the work. Section 3.2 reviews the clonal selection theory and relevant immunology and theoretical immunology providing a detailed foundation to effectively interpret the motivations for the state of clonal selection algorithm research. The field of clonal selection algorithms is reviewed in Section 3.3 providing a compressive taxonomy of approaches and focusing on the abstracted immunological principles and application domains for resultant systems. The review identifies three limitations that motivate the remainder of the chapter, specifically (1) the lack of placement of clonal selection approaches in a broader context, (2) the lack of explicit investigation of clonal selection as an adaptive system, and (3) the cellular focus of abstracted computational principles. These general open problems are investigated in turn. Section 3.4 reinforces the relationship of the paradigm with evolutionary algorithms and instance based learning, and highlights the relationships with binary hill climbing and competitive learning paradigms. Section 3.5 investigates clonal selection in the context of adaptive systems theory, highlighting the need to focus on the system-environment relationship of the paradigm, and the general short-sightedness of selection-based adaptive systems. Finally, Section 3.6 outlines an agenda for investigating the clonal selection paradigm for the remainder of the dissertation, firstly at the classical adaptive cellular level, as well as the inherently distributed ‘host of tissues’ and ‘population of hosts’ perspectives of immunology.

3.2 Clonal Selection Theory

3.2.1 Theory of Acquired Immunity

This section provides an overview of the theory including its inception, main processes, and the mechanisms employed for diversity. This discussion of clonal selection is limited to the discussion of B-lymphocytes and humoral immunity given that the development and much of the experimental and theoretical work on clonal selection has been completed in

this context.

Theory Development

Paul Ehrlich observed the exponential generation of antibodies triggered by the primary exposure to antigen in blood. He proposed a *side-chain theory* (circa 1900) in which blood cells naturally made side chains (now known as receptors) capable of binding to all antigens. His theory proposed that the information and capability to detect all antigen was already possessed by the host (for example in the genome) and expressed by the antibodies. This model was ultimately rejected because it was demonstrated that antibodies could be produced in response to synthetic substances of which the antibodies could have no prior knowledge. Further, the theory could not explain the improvement in the specificity of the response in the secondary and subsequent exposures to the same antigen. The *antigen-template theory* was proposed and suggested that antigen direct the response [315]. In the theory antibodies use the antigen as a template or model and modify themselves to provide a more effective complement. The emerging field of genetics suggested the invalidity of this theory as the specificity of an antibody could not be changed after its creation given that the process of *gene to amino acids to protein conformation* is one-way. Niles Jerne had great difficulty with the template theory and listing a number of observations to which the theory could not account [232]. The template theory proposed a one-to-one relationship between antibodies and antigen, although experimental observations suggested that far more antibodies were produced in the primary response than there were antigen. Further, the template theory suggested the role of antibodies ended after neutralising the antigen, whereas observations indicated that antibodies have a short lifespan, therefore failing to account for observed immunological memory exhibited in the improved specificity of the secondary response.

Jerne proposed an alternative called the *natural selection hypotheses* of antibody diversity. His theory proposed that the host contains a small pre-existing collection of randomly generated antibodies. In generating this pre-existing pool, self-antibodies were proposed to be *somehow* neutralised before causing harm to the host. The theory proposed that antigen *naturally select* antibodies of high affinity (a good match) from which more antibodies of the same affinity are produced. The result of the antigen-to-antibody selection mechanism the production of antibodies specific to the antigen in large numbers, which presented an improved response of the host to the antigen. Finally, he suggested that copying errors during the triggered production of antibodies may result in an improved fit with the antigen and account for the observed immunological memory effect. In his work on antibodies and allergy, David Talmage extended upon the natural selection hypothesis suggesting that Jerne's theory was an extension of Ehrlich's *side-chain theory*, and that replicating cells, not antibodies, were the main feature of the immune response [380]. Talmage suggested the notion that immune cells must specialise in producing antibodies with the same specificity as the receptors of the surface of the cell that trigger its proliferation.

Frank Macfarlane Burnet also extended Jerne's theory at the same time as Talmage (and subsequently received priority) focusing on a *clone* perspective and the self-replicating

cellular mechanisms during the immune response [63]. He called his theory the *clonal selection hypothesis of antibody diversity*. The ‘missing link’ for Burnet was the pre-existing naïve antibody population that was proposed in Jerne’s theory. As with Talmage’s theory, antigen select the best matching cells, that results in their subsequent replication. He suggested that for the selectionist theory to hold, that each cell must produce antibody of one type (all with the same receptor configuration). His theory attributed the exponential rise in the number of antibodies after infection to the exponential rise in the number of antibody producing cells due to clonal expansion of selected cells. It also explained the improved secondary response given the specialisation of receptors during expansion, although the mechanisms by which receptors specialised was only speculated. The following provides a summary of the characteristics of Burnet’s clonal selection, from [65]:

1. *Initial Repertoire*: Randomly generated initial population of antibody producing immune cells.
2. *One-to-one*: An immune cell can detect and release antibody of one specificity, which is passed on to descendent cells.
3. *Elimination*: Those initial cells that are reactive to self-tissues are selected against (killed).
4. *Proliferation*: Generation of antibody and the proliferators of the cell after contact (selection) with antigen.
5. *Forbidden*: The elimination of progeny that can select for self tissues.

Talmage and Burnet both followed up two years later with monographs elaborating on the theory, providing a more detailed account of the clonal selection process [381, 64]. Burnet continued his work on the study of immunological tolerance, to which his theory supported, and with Peter Medawar won the Nobel Prize in 1960 for work that provided the foundation for viable tissue and organ transplant. It is also interesting to note that it was Burnet who conceptualised the *self-nonsel* abstraction of the immune system in his work on immunological tolerance, which like his clonal selection theory has remained a foundational concept of modern immunology¹. Nossal and Lederberg supported the one-cell-one-antibody assertion of clonal selection providing first test and confirmatory evidence for the theory [306]. As mentioned, Burnet speculated as to the mechanisms for antibody diversification and specialisation in his theory. His prediction of the existence of such a mechanism was confirmed approximately 20 years later by Tonegawa et al. who investigated the fine tuning of antibody receptors called *somatic hypermutation* for which he later won the Nobel Prize [223, 397].

Theory Detail

An antigen is a molecule which can elicit an immune response, an antibody (immunoglobulin) is a molecule which can bind to and neutralise an antigen, and a B-lymphocyte is an

¹For an overview of the history of self-nonsel paradigm see Crist [86]

immune cell which can bind to antigen as well as produce antibodies. The physical surface of an antigen has a number of different structural features called antigenic determinants. An antibody and B-cell have receptors which have a specificity (ability to bind) to only one surface feature of an antigen, although the feature an antibody binds to may be represented on more than one different antigen. A host possesses a base pool of randomly generated lymphocytes that define the hosts natural immunity, and are prepared during the embryonic stage of development. During this time, the immune system learns a tolerance for the tissues of the host organism and those lymphocytes that are self-reactive are removed from the pool. The repertoire of natural lymphocytes possesses the ability to bind to any antigenic stimulus (natural or synthetic) with some general affinity. The introduction of an antigen triggers an immune response. The lymphocytes that are most suited (have the best complementary shape) for the antigen bind to it triggering cell division and differentiation. Therefore, it is the features of the antigen that select the best matching lymphocytes from the hosts cellular repertoire. After a lymphocyte binds to an antigen it moves to lymphatic tissue (germinal centre) to begin the differentiation process. The function of this process is the preferential proliferation to replicate a clone of cells to produce antibody capable of neutralising the triggering antigen. The lymphocyte proliferates and differentiates into two different cell types: *plasma cells* that continue to replicate and which release large numbers of antibody, and *memory cells* that like the parent cell act as a catalyst for response to subsequent infections.

During the proliferation and differentiation process copying errors (somatic mutations) may occur with a probability of mutation much higher than the genetic mutations that occur in typical cell divisions. The result is a clone of lymphocytes that have receptors similar to the parent cell, although vary slightly in their specificity to the triggering antigen. This means that on subsequent exposures to the antigen, the host possesses a larger repertoire of lymphocytes and antibodies capable of neutralising the antigen. Given the somatic mutations introduced during replication, some of the lymphocytes will contain an improved match for the antigen, improving their chance of being selected. An antibody producing lymphocyte will only produce antibodies of the same affinity. This means that after the cell has formed (after mitosis) it may be considered an antibody factory where all antibodies produced have the exact same complementary shape for a specific antigen surface feature. The fact that a surface feature may be represented by more than one antigen provides the basis for vaccinations. This commonality of shape facilitates the capability for generalisation, where lymphocytes and antibody may be raised on one harmless antigen and be effective on other more harmful antigens that possess the same structures. The memory cells generated during the differentiation process may live for months and years which is much longer than typical lymphocytes that lives for hours or days, providing long term memory of the learned antigenic defence.

The problem of antibody diversity was considered for at least 100 years before a coherent theory was proposed, although it is important to highlight that the theory in its original inception was not perfect. Specific details, such as the mechanisms of diversity were not known for many years after the theory's acceptance. Silverstein suggested that

with regard to the clonal selection theory, like Darwin's theory of evolution by *natural* selection, much of the specifics have been out-dated given scientific progress [358]. Rather than abandon the core principles, Silverstein emphasises the need to embrace the continually updated version of the theory. A pioneer himself in immunology, Lederberg suggests similar sentiments: “*Don't let conflicting and awkward 'facts' stand in the way of an esthetically satisfying theory whose fundamentals are consistent with the world model and with one another! And be suspicious of 'facts' that seem in the way of any coherent theory*” [269] (page 181).

Finally, the clonal selection process of acquired immunity is not without problems. An observed effect known as *Original Antigenic Sin*² refers to the commitment of the immune response and subsequent memory to specific surface characteristics of a pathogen, such as a virus, then having the virus mutate and change those surface characteristics [152, 122]. The result of the commitment is that the previously high-affinity lymphocytes and antibody become low-affinity within the context of the change to the virus. Given the clonal selection and expansion of lymphocytes in the first exposure, the now low affinity receptors dominate the subsequent exposures, effectively out-competing other lymphocytes that may be better suited to the changed virus.

Diversity Mechanisms

The genetics and chemistry of antibody diversity is beyond the scope of this work. This section provides a summary of the main mechanisms that lead to the diversity of the B-lymphocyte repertoire, which includes gene recombinations, somatic mutations, receptor editing, class switching, and gene conversion.

- *Genetic Recombination*: The process that accounts for the initial and on-going high diversity of new lymphocytes is called *secondary V(D)J rearrangements*, which refers to the recombination of the genetic code during the development of B-lymphocytes. This genetic recombination process accounts for the diversity of the initial lymphocyte repertoire and in the creation of new (known as naïve) B-lymphocytes [396, 397].
- *Somatic Hypermutation*: Along with genetic recombination was the first discovery made regarding the genetic component of antibody diversity [396]. It refers to the point mutations of the genome that occur during the clonal expansion phase when the cell moves to the germinal centre for differentiation (for more information regarding cell maturation see [37, 38]). The mutations are targeted at the genes in the genome that code for the receptors and occur at a rate about one million times higher than the mutation rates in other genes. Mutations include substitutions, additions, and deletions of genetic code. Specific sequences (motifs) within the receptor genes are targeted with high frequency, referred to as ‘hot spots’.
- *Receptor Editing*: Clonal selection proposes that those cells with receptors that bind to self-tissues (auto-reactive) are clonally deleted before they can cause damage.

²“Original Sin” is a term from Christian theology describing the condition of sin that marks all humans as a result of Adam's first act of disobedience.

Receptor editing is an additional process for re-programming auto-reactive receptors by recombining their genes [332, 307]. Those cells that have receptors that are still auto-reactive after this *editing* process are eliminated from the repertoire. Receptor editing was proposed to occur before somatic mutation and in response to contact with antigen (similar to template theory). It has also been proposed that receptor editing may provide an additional general receptor diversity mechanism for so-called ‘long jumps’ (large changes) in receptor affinity [167].

- *Class Switch Recombination*: The genetic recombination that results in antibodies changing their behaviour by *becoming* different immunoglobulin types, although maintaining the same specificity for antigen [198, 273]. Gene conversion is another genetic recombination process and the primary mechanism for diversification of immune receptors in chickens (and some other species) that is employed instead of somatic mutation used in mammals [438, 273].

3.2.2 Shape Space and Affinity Landscape

The *shape-space* formalism and its sibling the *affinity landscape* are two common geometric paradigms from theoretical immunology, both of which have had a recent revival and reapplication in the related field of Artificial Immune Systems. This section briefly surveys the history of these terms in the context of computational immunology, and consider their pervasive application in AIS.

Shape-Space

Perelson and Oster introduced the shape space formalism in their theoretical investigation of antibody repertoire size and reliable recognition [321]. Given antibody ab and antigen ag , the authors discretised the physical aspects of the antibodies combining site and the antigenic determinant into a vector of n so called *shape* parameters. The parameters included the set of intermolecular forces important to the $ab - ag$ recognition. The abstraction of a molecule was later referred to as its *generalised shape*, where the shape parameters for a ag and ab molecule may be defined as points in an n -dimensional Euclidean space called *shape-space* S [320]. The authors suggested that given an adequate method of defining molecule shape, the dimensionality of such a space could be relatively small (as low as five dimensions). Further, the formalism simplified the complementarity nature of $ab - ag$ interactions, thus rather than the matching convention of $ab \neq ag$, the formalism considers $ab = ag$. Given this simplification, the formalism defines distances between ab and ag in the Euclidean shape space as the *degree of complementarity* or *affinity* of the interaction. A generalised affinity function may be defined as $c_{ij} = f(x_i, x_j)$, where c_{ij} is the affinity between the molecules i and j , and x_i and x_j are vectors representing the molecules in shape space, f is an appropriate function for the selected representation. The shape space is defined as a finite hypercube, with a volume V of uniform density. Each antibody was defined with a local recognition region in the shape space ϵ , with a Gaussian falloff. An $ab-ag$ interaction only occurs if the ag is above an affinity threshold, within the

triggering hyper-sphere of the *ab*. In their application of the model, random distributions of *ab* and *ag* were used and self-antigens were ignored.

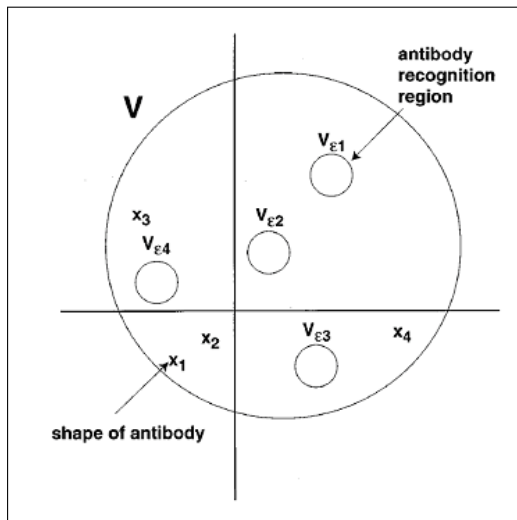


Figure 3.1: Depiction of the shape-space formalism taken from [322] (page 1225). X_i represents antibodies with corresponding complementary recognition volumes $V_{\epsilon i}$.

The shape-space formalism provides a geometric way to consider molecule interactions in the context of the immune system, it was not the first of such a geometric abstraction of molecular interaction. For example, Write considered a *genotypic-space* for possible gene combinations and a mapping onto the now ubiquitous *fitness landscape* in theoretical genetics [437]. This fundamental *ag – ab* interaction paradigm was used throughout the late 1980’s and 1990’s predominantly in theoretical works modelling clonal selection and various aspects of Jerne’s idiotypic network theory (for example work on a one-dimensional Euclidean shape-space by Segel [355], and an investigating the stability of the idiotypic model by Perelson [320]). Carneiro and Stewart criticised the shape-space formalism, focusing on the simplicity of the abstracted space and the shortcomings of the simple affinity functions [70]. They cautioned against the extrapolation of results from the abstraction to the real shape-space, highlighting that the dimensionality of the space should be much higher (at least 10, and as high as 20), and that the affinity mapping function would have to be irregular and discontinuous. They support their claims with observations from simulated docking of molecules based on crystallographic structures. They propose a splitting of the formalism into a *realistic shape-space* in which real molecules can be evaluated, and an abstracted *inversion space* as a tool for modelling.

Affinity Landscape

It is typical in physics for systems to minimise some utility in the context of a response surface, whereas in biology it is common for systems to maximise some utility. This is highlighted by the already mentioned evolutionary fitness landscape paradigm introduced by Wright in conceptualising gene combinations [437]. In the immune system, one may consider the utility of an antibody as its affinity in binding to a specific antigen. A related concern called *avidity*, refers to strength of the binding between antigen and antibody.

Among the first to employ the conceptualisation of an *affinity landscape* were Kauffman and Weinberger who used this geometric paradigm in the field of theoretical immunology to investigate and describe the theoretical effects of affinity maturation [242, 243]. They investigated adaptive walks using Kauffman’s NK³ model on the affinity landscape based on antibody sequences or so called *antibody-space* in response to a given antigen. Kauffman and others used the NK model to investigate various molecular sequence spaces and the corresponding response surfaces, for example [137, 240, 241]. One may define an affinity landscape as a degree of complementarity response surface for a given collection of antigen detecting agents (lymphocytes and antibody), for a given *specific* antigenic determinant. Abstractions of this surface considered in theoretical immunology are typically continuous, the topology of which consists of many local optima. Detecting agents may navigate this response surface through the process of affinity maturation, which involves manipulations to genotypic sequence information (such as the clonal selection diversity mechanisms discussed in Section 3.2.1).

The affinity landscape is a simple yet important conception, and although Kauffman et al. were perhaps the first to apply such a name, the geometric formalism was no doubt in the mind of Perelson and others in the early days of theoretical and computational immunology. The formalism has played an important role in conceptualising and formalising *affinity maturation* of the immune response by hypermutation (for example see [325, 356, 385, 386]), which as has been demonstrated is a critical aspect of the clonal selection theory of antibody diversity. George and Gray used the affinity landscape as a crutch in describing their theories on receptor editing, which were generalised very effectively on a simple two-dimensional affinity graph in Figure 3.2 [167, 168].

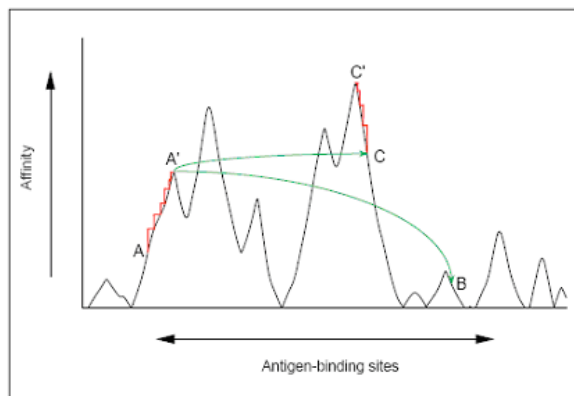


Figure 3.2: Theory of receptor editing providing long jumps in the affinity landscape, from [167].

Artificial Immune Systems

The shape-space and affinity landscape paradigms have been widely used in the field of Artificial Immune Systems. Some examples include the quintessential binary shape-space used in negative selection (for example [375]) and immune network algorithms, the recogni-

³Not to be confused with Natural Killer cells, N and K are parameters of the Kauffman’s model.

tion region and threshold used in data mining and classification algorithms [422, 421, 259], and the affinity landscape cost functions navigated by optimisation algorithms [427, 117]. The usage of these paradigms is unified in the framework for AIS of Engineering de Castro and Timmis, and listed in Section 2.6.1. The authors presented abstract models of immune cells, molecules and their interactions including the affinity landscape and shape-space paradigms in an effort to elaborate the three facets of their framework (representation, affinity measures, and immune algorithms). One may distil the discussion of the AIS Engineering Framework and related abstraction by de Castro and Timmis, to a triad relationship between *representation* in a shape-space, a *mapping* function between points in that space, and an *affinity landscape* of scores assigned by the mapping function, as follows:

- *Representation*: Extending the shape-space paradigm, the sequence, or search space is considered essentially a representation domain. Examples provided include real-valued, binary, integer and symbolic.
- *Mapping*: A function that translates from the representation shape-space to an affinity landscape surface. Distance and matching within the shape-space are the mapping functions discussed with a number of appropriate suggestions for the various shape-spaces listed. A point in one of these spaces may have a *recognition region*, and a *cross-reactive threshold*. A mapping function may be complemented with an additional threshold binding function, perhaps to convert a scalar affinity into a decision variable.
- *Affinity*: Affinity may be further abstracted from the degree of complementarity in an antibody-antigen interaction, to a more general scoring assigned to the mapping between points in the shape-space. Affinity may be a measure of quality of an agent in a specific environment or in the context of a specific requirement.

The proposed triad provides a geometric interpretation and relation of *shape-space* to *affinity landscape* paradigms and provides an open framework for investigating immunological principles in domains, not limited to the pattern recognition problem. One may consider further geometric conceptions related to affinity maturation in immune system that lack a suitable formalism, let alone an abstraction for use in Artificial Immune Systems. Two examples may include: (1) the aggregation of multiple affinity landscapes, that is the affinity of a repertoire in the context of an antigenic environment not limited to one antigen, and (2) the temporal and spatial aspects to affinity throughout the host organism in the context of lymphocyte and antibody mobility.

3.3 Clonal Selection Algorithms

The previous section outlined in detail the abstraction, immunochemistry, and theoretical immunology that inspired the field of clonal selection algorithms. The review highlighted that the main actors in clonal selection are cells, antibodies, and antigen, and the main

processes are selection, proliferation, and genetic mutations. Further, it showed that the common measures used to describe the relationship between antibody and antigen are affinity and avidity, and that the main conceptual formalisms used to describe the constraints and effects of such interactions are shape-space and the affinity landscape. This section reviews the so called *clonal selection principle* which is the abstraction that motivates the field of clonal selection algorithms. A taxonomy of such algorithms is presented, highlighting the open problems in the state of the art.

3.3.1 From Theory to Principle

In the glossary of their seminal book on Artificial Immune Systems, de Castro and Timmis provided a definition of the clonal ‘selection principle’ as “*the prevalent theory stating that the specificity and diversity of an immune response are the result of selection by antigen of specifically reactive clones from a large repertoire of preformed lymphocytes, each with individual specificities*”, not distinguishing between the theory and information processing abstraction [109] (page 322). In their treatment of Burnet’s theory, the authors comment that lymphocytes may be considered to undergo a process similar to natural selection and that only those cells that contact an antigen may be selected for proliferation, strongly highlighting the Darwinian connection made by Jerne and Burnet. The authors focused on the clonal expansion and variation aspects of the process, referring to the abstraction of the theory as **affinity maturation of the immune response**. They propose two important computational features of the theory in their abstraction (page 80): (1) An antigen selects multiple cells to proliferate where each cell has an individual clonal expansion rate proportional to its affinity to the antigen (*higher affinity equals more clones*), and (2) the mutation of each cell during clonal expansion is inversely proportional to the affinity of cell to the antigen (*higher affinity equals less point mutations*). The affinity proportional rates imposed on clonal expansion and mutation in their general clonal selection abstraction may be traced back to an earlier work of de Castro and Von Zuben on the seminal clonal selection algorithm CLONALG [114] (discussed in Section 3.3.2), in which they cited the work of Berek and Ziegner [38], commenting that a controlled cloning and mutation process could improve the efficiency of the process, and *may* be employed by the immune system.

In their treatment of the clonal selection and inspired algorithms Cutello and Nicosia proposed that clonal selection allows “*the learning of patterns during the primary response, and retrieval of previous knowledge during the secondary response or the cross-reactivity process*” [97] (page 112). They defined the abstraction of the theory called the **clonal selection principle** as involving both a learning or training phase followed by an applied or testing phase that may continue for the lifetime of the system, an approach which is common to the field of machine learning and pattern recognition. They proposed two key features of the clonal selection theory be taken into account in realising such a principle: (1) the *hypermutation mechanism* which they consider a local search procedure that leads to a fast maturation during the learning phase, and (2) *clonal expansion mechanism* that triggers the growth of a new population of high-value (affinity) cells. They consider all

algorithms based on the clonal selection principle to be population based, where each member of the population represents a candidate solution belonging to the combinatorial fitness landscape of a given computational problem.

What is distinctly common between these two sets of principles is (1) the absence of concerns of cell and antibody preparation for the task of pathogen identification associated with the negative selection paradigm, and (2) the absence of the concerns of cell-cell (inter-repertoire) interactions such as those associated with the immune network paradigm. Neither sets of principles explicitly exclude these concerns, rather they are not the focus of the information processing. The Cutello-Nicosia and the de Castro-Timmis abstractions of the clonal selection theory are the two main sets of principles inspiring the IA and CLON-ALG families of clonal selection algorithms respectively (Section 3.3.2). Both abstractions consider clonal selection to be an **adaptive process** that operates on a finite set of discrete units, the composition of which is changed through the iterative application of selection and clonal expansion with variation. Both abstractions also value the diversity provided through bind mutation that results in pre-committed configurations, some of which offer a relative selective advantage. The Cutello-Nicosia abstraction considers distinct training (change) and test (fixed) phases of adaptation, whereas the de Castro-Timmis abstraction considers affinity-biased controls on a single adaptive phase.

3.3.2 Algorithm Taxonomy

This section presents a taxonomy of clonal selection algorithms that divides the scope of such algorithms into a genealogical tree of five groups, with a sixth miscellaneous group⁴. The presented taxonomy only considers those algorithms that explicitly exploit (or claim to exploit) the clonal selection principle, which as outlined above requires the integration of the abstract actors and processes of the the clonal selection theory. Also listed in the miscellaneous category are examples of those AIS that although do not fit the adopted definition of a clonal selection algorithm, exploit considerable aspects of the clonal selection principle. Each lineage is reviewed with regard to the core principles, exemplar algorithms, and summary of general application. Table 3.1 provides a summary of each lineage and classified algorithms.

CLONALG

Hidden at the back of a technical report surveying the applications of Artificial Immune Systems, de Castro and Von Zuben proposed the Clonal Selection Algorithm (CSA) as a computational realisation of the clonal selection principle for pattern matching and optimisation [112]. This algorithm which has become perhaps the most popular in the field of AIS, was later published and represented [114], and again [117] where it was renamed to CLONALG (CLONal selection ALGORITHM). The general CLONALG model involves the selection of antibodies (candidate solutions) based on affinity either by matching against

⁴This is the scope of clonal selection algorithms to the author's knowledge. It is not expected to be complete given the vastness and disparate nature of related publications, nor is it required to be complete for the purposes of this dissertation.

Lineage	Algorithms	Primary Application
<i>CLONALG</i>	CSA, CLONALG, CLONALG (1,2), ACS, CLONCLAS, RCSA, MOCSA, IMCSA, AISMM, SACSA, ECA	Optimisation
<i>AIRS</i>	AIRS, AIRS2, Parallel AIRS	Classification
<i>BCA</i>	BCA	Optimisation
<i>IA</i>	IA, SIA, I-PAES, CLIGA, CLIGA+, NC-IA, READ-Alg, opt-IA, opt-IMMALG, Par-IA, Dyn-IMMALG	Optimisation
<i>MISA</i>	MISA	Multi-Objective Optimisation
<i>Other</i>	Too large and varied to classify	Varied

Table 3.1: Clonal selection algorithm taxonomy overview.

an antigen pattern or via evaluation of a pattern by a cost function. Selected antibodies are subjected to cloning proportional to affinity, and the hypermutation of clones inversely-proportional to clone affinity. The resultant clonal-set competes with the existent antibody population for membership in the next generation. In addition, low-affinity population members are replaced by randomly generated antibodies. The pattern recognition variation of the algorithm includes the maintenance of a memory solution set which in its entirety represents a solution to the problem. A binary-encoding scheme is employed for the binary-pattern recognition and continuous function optimisation examples, and an integer permutation scheme is employed for the Travelling Salesman Problem (TSP). Table 3.2 summarises the algorithms parameters, and Algorithm 3.1 presents a general formulation of the CLONALG.

The work by Watkins, et al was proposed to exploit the *inherent distributedness* of the CLONALG [417], as discussed in Section 2.5.2. In the work, the pattern recognition variation of the CLONALG was modified such that each memory cell is partitioned to different processes and evolved independently or in small groups, the results from which are collated at the end of the algorithm run and returned as the algorithm result. White and Garret also investigated the pattern recognition version of CLONALG and generalised the approach for the task of binary pattern classification renaming it Clonal Classification (CLONCLAS) where their approach was compared to a number of simple Hamming distance based heuristics [427]. Walker and Garrett investigated CLONALG and Evolution Strategies (ES) on dynamic function optimisation, showing that although CLONALG can achieve better results faster than ES on low dimensional dynamic functions, ES consistently outperforms CLONALG on the two high-dimensional problems tested [415]. In an attempt to address concerns of algorithm efficiency, parameterisation, and representation selection for continuous function optimisation Garrett proposed an updated version of CLONALG called Adaptive Clonal Selection (ACS). The mutation parameter, the number of antibodies selected for cloning, and the number of clones produced for each antibody were changed to automatic parameters, controlled in a similar way to those in ES [163]. CLONALG is perhaps the most popular basis for elaboration and application with regard

Parameter	Description
P	Repertoire of antibodies.
N	The fixed antibody repertoire size.
n	The number of antibodies to select for cloning.
L	Bitstring length for each antibody.
N_c	Number of clones created by each selected antibody. Originally expressed as a function of the repertoire size (for optimisation) $N_c = \text{round}(\beta \cdot N)$ (where β is a user parameter), although a direct integer specification of N_c is simpler. A rank-based (affinity proportionate) variation of the parameter is presented for pattern recognition.
d	Number of random antibodies to insert at the end of each generation. Random antibodies replace the d lowest affinity antibodies in the repertoire.
<i>Stop condition</i>	Typically a specified number of generations or function evaluations or epochs of exposure to patterns.
<i>affinity</i>	Solution evaluation, typically the solution is decoded into a domain specific representation and assigned a quality scoring.
<i>clone</i>	Duplication of a bitstring.
<i>hypermutate</i>	Modification of a bit string where the flipping of a bit is governed by an affinity proportionate probability distribution. Originally $p = \exp(-\rho \cdot f)$, although the opt-aiNET variant is also popular $p = (\frac{1}{\rho}) \cdot \exp(-f)$ (where ρ is a user parameter and f is the normalised affinity scoring).
<i>replace</i>	The set of the d lowest affinity clones in the population (P) are replaced with randomly generated solutions.

Table 3.2: CLONALG parameters.

to clonal selection algorithms, specifically for function optimisation problem instances, the extent of which is excluded in the interest of brevity.

Artificial Immune Recognition System

After CLONALG, the Artificial Immune Recognition System (AIRS) algorithm is perhaps the second most popular clonal selection algorithm, although the approach was designed for and has only been applied to the supervised classification problem domains. The earliest work on AIRS was in Watkins Masters work [422] that was later published [423]. The approach is a supervised learning algorithm for classification that uses the idea of an Artificial Recognition Ball (ARB) (introduced in earlier works on the Artificial Immune Network algorithm) to represent clones (groups) of identical B-cells. The AIRS procedure involves cloning and somatic hypermutation for preparing a set of real-valued exemplar vectors suitable for classifying unobserved cases, using a single iteration (epoch is the machine learning term) over a set of training data. Watkins and Boggess applied the AIRS to a suite of benchmark classification problems [418], and Goodman and Boggess problems did the same, comparing to a conceptually similar approach called Learning Vector Quantization (LVQ) [178].

Given the rapid popularity of the approach Marwah and Boggess investigated the algorithm seeking issues that affect the algorithms performance [286]. They compared various

Algorithm 3.1: CLONALG for Function Optimisation.

Input: N, n, L, d, β, ρ
Output: P

```
1  $P \leftarrow \text{CreateRandomCells}(N, L);$ 
2 while  $\neg \text{StopCondition}()$  do
3   foreach  $p_i \in P$  do
4      $\text{Affinity}(p_i);$ 
5   end
6    $P_{\text{select}} \leftarrow \text{Select}(P, n);$ 
7    $P_{\text{clones}} \leftarrow 0;$ 
8   foreach  $p_i \in P_{\text{select}}$  do
9      $P_{\text{clones}} \leftarrow \text{Clone}(p_i, \beta);$ 
10  end
11  foreach  $p_i \in P_{\text{clones}}$  do
12     $\text{Hypermutate}(p_i, \rho);$ 
13     $\text{Affinity}(p_i);$ 
14  end
15   $P \leftarrow \text{Select}(P, P_{\text{clones}}, N);$ 
16   $P_{\text{rand}} \leftarrow \text{CreateRandomCells}(d, L);$ 
17   $\text{Replace}(P, P_{\text{rand}});$ 
18 end
19 return  $P;$ 
```

variations of the algorithm with modified resource allocation schemes, tie-handling within the ARB pool, and ARB pool organisation. AIRS was again raced against LVQ by Boggess and Hamaker on datasets that contained irrelevant features to assess the algorithms ability to handle noise [49]. Greensmith and Cayzer applied AIRS to hierarchical document classification [186], that culminated in Greensmith’s masters work [185]. Watkins and Timmis proposed a new version of the algorithm called AIRS2 which became the replacement for AIRS1 [419]. The updates reduced the complexity of the approach while maintaining the accuracy of the results. An investigation by Goodman, et al. into the so called ‘*source of power*’ in AIRS indicated that perhaps the memory cell maintenance procedures played an important role [179]. A follow-up empirical investigation by Goodman and Boggess supported the original finding indicating that the process by which new memory cells are admitted into the ARB pool is critical to the success of the approach [180]. The work by Watkins, et al., already mentioned briefly in Section 2.5.2 proposed a parallel version of AIRS permitting the division of training patterns and memory pool suitable to exploit parallel hardware [420]. The extent of further elaborations and applications of the approach are excluded in the interest of brevity.

B-Cell Algorithm

Kelsey and Timmis proposed the B-Cell Algorithm (BCA) as an AIS designed for continuous function optimisation [246]. The algorithm maintains a pool of B-cells (binary-encoded candidate solutions) that are subjected to cloning and mutation. An elitist replacement population maintenance scheme is applied that ensures only improved cells are admitted

into the pool. The mutation operator, which is called *contiguous somatic hypermutation* selects a random sub-string of a solution to vary in a probabilistic manner, in what the authors claim as ‘hot-spot’ mutation. Kelsey, et al. applied the BCA to multi-modal dynamic chaotic test functions [247]. Empirical algorithm tuning by the authors revealed that small population sizes (3-5) show better results. In an investigation of AIS applied to optimisation, Hone and Kelsey provide a case study investigation of the BCA and showed fractal structures on the complex plane suggesting the potential usefulness of studying AIS as non-linear dynamical systems [219]. In a further empirical study, Timmis, et al. compared the BCA to opt-aiNET⁵ and a Hybrid Genetic Algorithm attributing the partial success of BCA to the mutation scheme, speculating it results in the escaping of local-optima search behaviour [390]. The parameters of BCA are summarised in Table 3.3, and the general procedure for the approach is listed in Algorithm 3.2.

Parameter	Description
P	Repertoire of antibodies.
N	Antibody population size.
N_c	Number of clones to create of each antibody.
N_r	Number of random antibodies to create and insert each generation.
<i>Stop Condition</i>	Typically if no progress is made for a number of generations.
<i>hypermutation</i>	Uses a processes called contiguous hypermutation, a random location in the bit string is selected, and a random bitstring length is selected. Each bit in the sub-string is flipped with the probability ρ .
<i>replace</i>	A parent is replaced only if a member of its clone has a higher affinity (greedy replacement).

Table 3.3: BCA parameters.

A proof of convergence for the BCA was proposed by Clark, et al. using a Markov Chain model [76]. The proof simplifies the algorithm to an elitist search with a single population member suggesting that the members of the population can be treated independently given the lack of interaction during the optimisation procedure. Further, they speculate that the introduction of inter-solution interactions in the BCA will have a detrimental effect on results of a search. Finally, in an empirical study Bull, et al. applied the BCA to what they refer to as less-smooth test problem instances (Diophantine equations) seeking empirical convergence heuristics [62]. Four variants of the algorithm were compared: an approach that used an elitist selection mechanism to introduce inter-solution interactions, and three of what the authors refer to as *mega-mutation* schemes that attempt to introduce further diversity into the search. These less-greedy modifications of the approach achieved a better final result compared to the classical BCA on the test functions chosen, perhaps suggesting the use of diversity introduction approaches in further BCA applications.

⁵(opt-aiNET) is an immune network algorithm for optimisation specified in [108]

Algorithm 3.2: B-Cell Algorithm (BCA).

Input: N, N_c, N_r, ρ, L
Output: P

```
1  $P \leftarrow \text{CreateRandomCells}(N, L);$ 
2 while  $\neg \text{StopCondition}()$  do
3   foreach  $p_i \in P$  do
4      $\text{Affinity}(p_i);$ 
5   end
6   foreach  $p_i \in P$  do
7      $P_{clones} \leftarrow \text{Clone}(p_i, N_c);$ 
8      $P_{clones} \leftarrow \text{CreateRandomCells}(N_r, L);$ 
9     foreach  $p_i \in P_{clones}$  do
10       $\text{Hypermutate}(p_i, \rho);$ 
11       $\text{Affinity}(p_i);$ 
12    end
13     $p_{i'} \leftarrow \text{SelectBest}(P_{clones});$ 
14    if  $p_{i'}.affinity \leq p_i.affinity$  then
15       $p_i \leftarrow p_{i'};$ 
16    end
17  end
18 end
19 return  $P;$ 
```

Immunological Algorithm Family

A simple clonal selection inspired algorithm was proposed by Cutello and Nicosia called Immunological Algorithm (IA) [95, 96]⁶ later renamed to Simple Immunological Algorithm (SIA) [97]. The algorithm maintains a population of B-cells that are exposed to a clonal expansion process each iteration. This expansion process involves the cloning of cells and the application of a hypermutation operator, and was demonstrated on the Minimum Hitting Set Problem (MHSP) and the 3-Satisfiability Problem (3-Sat). The SIA was extended and applied to the Graph Colouring Problem (GCP) [100]. The extensions involved the introduction of a local-search procedure that operated upon each B-cell after the clonal expansion phase. In addition, rather than an elitist selection method of maintain the population size after each expansion, an aging operator was introduced for each B-cell. B-cells are probabilistic deleted from the population using an equation inspired by the biological literature. Two variations of the aging operator were applied, an elitist version that ensured the best B-cell's were not deleted, and a pure strategy that probabilistically deleted irrespective of the elitist concerns. A birthing operator was also added to 'top-up' the population to the configured size, and an information gain (a stabilisation in the measure of information discovered by the algorithm) measure was used as the termination criteria for the algorithm. Given the enthusiastic renaming and tweaking of this family of approaches, the SIA (the simplest variation of the family) is presented in Table 3.4 and

⁶The Immunological Algorithm (IA) is renamed and represented many times by its authors. Other names include Simple Immune Algorithm (SIA), Cloning Information Gain Aging (CLIGA), and Optimization Immune Algorithm (opt-IA, opt-IMMALG). This lineage is named the IA family for simplicity.

Algorithm 3.3.

Parameter	Description
P	Antibody population.
L	Length of binary string representation.
d	Population (repertoire) size.
dup	The number of clones created for each antibody.
$clone$	Duplication of the bitstring.
$hypermutation$	Probabilistic modification of a bit string (bit flipping), requires the specification (ρ) of the probability of flipping each bit.

Table 3.4: SIA parameters.

Algorithm 3.3: Simple Immune Algorithm (SIA).

Input: d, dup, ρ, L
Output: P

- 1 $P \leftarrow \text{CreateRandomCells}(d, L);$
- 2 **foreach** $p_i \in P$ **do**
- 3 | $\text{Affinity}(p_i);$
- 4 **end**
- 5 **while** $\neg \text{StopCondition}()$ **do**
- 6 | $P_{clones} \leftarrow 0;$
- 7 | **foreach** $p_i \in P$ **do**
- 8 | $P_{clones} \leftarrow \text{Clone}(p_i, dup);$
- 9 | **end**
- 10 | **foreach** $p_i \in P_{clones}$ **do**
- 11 | $\text{Hypermutate}(p_i, \rho);$
- 12 | $\text{Affinity}(p_i);$
- 13 | **end**
- 14 | $P \leftarrow \text{SelectBest}(P, P_{clones}, d)$
- 15 **end**
- 16 **return** $P;$

The probabilistic aging operator was replaced with a simplified generational aging operator by Cutello, et al. in an application to the 2DHP protein folding problem [102]. In a more detailed study on different varieties of the same protein folding domain, the aging operator was further tweaked to facilitate longer life spans on some B-cells deemed useful to the search process [104]. The clonal expansion aspect of the algorithm (cloning and hypermutation) was grafted into to an existing evolutionary multiple-object optimisation technique and called I-PAES [91]. The transformed algorithm (generational aging, information gain stopping criteria) was reviewed and renamed to the ‘Cloning, Information Gain, Aging’ (CLIGA) algorithm [97]. A modified version called CLIGA+ was proposed in which each B-cell contains more than one receptor (pattern), permitting application of the algorithm to pattern recognition tasks. Also proposed in this work is a Noisy Channel variation of SIA (NC-IA), and a Reaction-Diffusion variation of SIA (READI-Alg) both of which were applied to instances of the GCP. Cutello, et al. again renamed the approach

to Optimisation Immunological Algorithm (opt-IA) and applied the approach to instances of binary trap functions [94]. An additional fitness inversely-proportional hypermutation referred to as *hypermacromutation* was proposed and compared to the traditional static approach. This algorithm was evaluated again in a larger study involving a number of machine learning domains [92], and again on a large number of continuous function optimisation instances [93]. Cutello, et al. investigated the hypermutation operators of opt-IA [101]. Cutello, et al. applied opt-IA to the 3DHP protein folding problem [90]. Cutello and Nicosia applied opt-IA to graph colouring, MHSP, and satisfiability [98]. Work by Anile, et al. hybridised opt-IA with a direct search method [27]. An extension of opt-IA called *Aligner* was proposed by Cutello, et al. and applied to multiple sequence alignment of DNA [89].

An extension to opt-IA was proposed by Cutello, et al. called the parallel immune algorithm (Par-IA) which is a master-slave version of the algorithm applied to numerical function optimisation [99]. Cutello, et al. renamed the approach to Optimisation Immune Algorithm (Opt-IMMALG), applying the approach to continuous function optimisation using a real-valued representation as opposed to the binary representation used in previous works [103]. Also stated in this work was the use of fitness inversely proportional hypermutation as the standard mutation operator for the approach. This algorithm was extended and renamed dynamic immune algorithm (dyn-IMMALG) by Cutello, et al. who propose a dynamic rather than static clonal operator [88]. The approach was applied to binary trap functions and compared to opt-IA and variations of CLONALG.

Multi-objective Immune System Algorithm

Coello Coello and Cruz Cortes introduced an AIS called the Multi-objective Immune System Algorithm (MISA), and as its name suggests was designed as a population-based approach for constrained and unconstrained multi-objective optimisation [77]. In the approach a repertoire of solutions is split into antigens (Pareto non-dominated and feasible solutions) and antibodies (Pareto dominated and infeasible solutions). A bit-string representation is used and antigens are selected at random and matched against antibodies using Hamming distance. After the selection step, antibodies are cloned, mutated the population is unioned and reduced back to the configured size, culling the lower quality solutions. An external (elitist) memory repertoire is maintained of non-dominated feasible solutions. Solutions are added to the memory set if they are non-dominated by the current memory set population and sufficiently diverse as determined by a grid-based maintenance structure. MISA was extended by the same authors and further compared to state of the art evolutionary approaches for multi-objective optimisation [83]. An EC-AIS hybrid terminology was adopted and an EC-based crossover mechanism was adopted within the memory set. The main algorithm was simplified such that all population members were consider antibodies, and only the lower score solutions was selected for cloning and hypermutation. The modified algorithm was shown effective, although it demonstrated rapid convergence behaviours on benchmark problem instances. Finally, Villalobos-Arias, et al. proposed a convergence proof for the update MISA showing that the elitism within the

algorithm was needed to guarantee convergence [408].

Other

Research into the five core lineages of clonal selection algorithms revealed a large number (>70 works) on new or extended clonal selection algorithms, predominantly in non-western conference proceedings. The vast majority of these works involved the application of a CLONALG, SIA, or derivative algorithms to benchmark and engineering problem domains. The extent of these works are omitted here for brevity, although the dominant applications are listed as follows: feature selection for models, parameter tuning for models or controllers, parameter tuning for a PID controllers, anomaly and/or intrusion detection, pattern recognition, multi-objective optimisation, function optimisation, general optimisation, multi-user detection, and finally hybridisation with other algorithms.

As discussed in Section 2.4, the clonal selection principle underlies the adaptive qualities of all three of the major paradigms of Artificial Immune Systems. As such, many works investigating the negative selection and immune network employ B-cells and antibodies that are modified using processes of selection and expansion with variation. Given that these approaches do exploit the clonal selection principle they may be considered clonal selection algorithms, although do not fit into the presented taxonomy as the clonal selection principle is not the primary focus of the works. This section provides some examples of such research. Weinand propose a dynamical systems computational model of somatic mutation of B cells to evaluate the effect of somatic mutation of affinity maturation of the immune response [424]. Zhang and Hou propose a Niching Clonal Selection Algorithm (NCSA) that combines negative selection and refinement applied to the pattern matching problem of anomaly detection [445]. Yu and Hou proposed an extension of CLONALG called CsAL (Clonal Selection Algorithm) to investigate the negative selection approach to virus detection [443]. In the body of work on an integrated intrusion detection system discussed in Section 2.5.2, the authors propose a static and dynamic clonal selection algorithms (DynamICS) which are an augmentation to T-cell inspired negative selection system that uses clonal selection mechanisms to improve detectors.

3.3.3 Criticisms and Open Problems

The review of the clonal selection principle demonstrated that abstractions of the theory are generally simplistic, considering the information processing properties and mechanistic procedure at a high-level, correlating the general AIS observations made in Section 2.6. The authors de Castro and Timmis reduced clonal selection to a single antigen optimisation system using affinity maturation with controlled affinity-proportionate cloning and inversely affinity proportional mutation. Cutello and Nicosia reduced clonal selection to the more general life-long acquisition and retrieval for application of patterns. The common features to clonal selection abstractions are the properties of (1) population based information processing (2) clonal expansion of selected high affinity population members, and (3) affinity maturation via mutation of clones. The commonly executed concerns of the abstractions include (1) concerns of self and nonself discrimination (cell preparation),

(2) distinction between cell types and/or antibody molecules, and (3) cell-cell interactions.

The realised clonal selection algorithms are monolithic, considering antibody conformations in binary and real value shape-spaces, mostly against a single affinity landscape. As such, the dominant application domains are optimisation (function optimisation), and supervised classification with procedures that closely resemble genetic algorithms (CLONALG, IA, BCA, MISA) and vector quantisation procedures (AIRS), and the dominant conceptual and quantitative comparisons have been made with evolutionary algorithms. The review demonstrated that the trend with these approaches have been to (1) enhance the algorithm procedure through domain specific assumptions (limited resources in the case of ARB for classification, and affinity proportionate descent with modification in CLONALG for optimisation), (2) tweaking and parameter tuning (numerous additional and alternative operators in the IA family, sensitivity analysis of CLONALG, BCA and AIRS), and (3) application, benchmarking and racing (the majority of the published research). Based on the detailed review of the state of the art of clonal selection algorithms in the context of clonal selection and related theory, this section identifies three specific limitations that represent open problems in the field, as follows:

1. *Related Approaches*: Clonal Selection Algorithms have generally been related and compared to Evolutionary Algorithms including the Genetic Algorithm and the Evolutionary Strategy, although broader connections with related computational intelligence fields are typically brief and/or non existent. Systematically placing and contrasting Clonal Selection as a Computational Intelligence metaphor with related paradigms provides both a context for comparison of behaviour and capability, and suggests at the specific overlapping features and findings that may be exploited as well as those features that differentiate the approach and should be pursued.
2. *Implicit Adaptive System*: Given the adaptive qualities of the clonal selection theory, computational models are implicitly assumed to be instances of adaptive systems. Whether or not a CSA is an adaptive system is not in question, rather the more important concerns of what it means for computational models of clonal selection to be classified as adaptive systems. Unlike the Genetic Algorithm that emerged from Holland's investigation and formalisation of adaptive systems, the inception and investigation of clonal selection algorithms has been focused on addressing benchmark instances of difficulty computational problems. The consideration of the clonal selection as a strategy in the context of an adaptive systems formalism will provide a stronger foundation of understanding, and insights into the broader and unrealised information processing capabilities of the approach.
3. *Cellular Focus*: The scope of clonal selection algorithms are concerned with abstractions of lymphocyte interactions with antigen, specifically with little or not detail of the molecular basis of such interactions, nor the broader consideration of where such interactions occur or the effect they have on the host organism. This cellular focus is reasonable for the preliminary realisation of computational models inspired by the theory embodied in the present state of the field, although elaboration and

development of so-called ‘second generation’ clonal selection algorithms requires a holistic consideration of the acquired immune system and the restrictions and constraints such a broader perspective imposes on improvement-centric adaptive clonal selection strategies.

3.3.4 Summary

The detailed treatment of the Clonal Selection Theory in Section 3.2 provided the context for effectively appraising the state of clonal selection algorithms, specifically with regard to the task of identifying limitations which may be exploited toward elaborating the general approach. The taxonomy presented in Section 3.3.2 unified disparate works by general algorithm lineages, that ultimately highlighted both the deep commonality between the lineages and the algorithms that belong to them, and the simplicity of the overarching computational abstraction that underlies such approaches. Importantly this insight provided the context to identify three limitations and open problems with the clonal selection paradigm that motivate the remainder of this Chapter. Specifically, Section 3.4 reviews machine learning approaches to clonal selection highlighting the findings that may be exploited and the features that differentiate the approach. Section 3.5 considers the clonal selection strategy in the context of adaptive systems theory, highlighting critical features of the theory excluded and underrated by existing computational abstractions. Finally, Section 3.6 address the cellular focus of the paradigm, and outlines an agenda for extending the approach beyond a discrete repertoire of cells toward distributed clonal selection algorithms.

3.4 Related Computational Paradigms

This section considers some related machine learning paradigms and their relationship with the clonal selection principle and inspired algorithms. This review includes both related paradigms that have been previously considered such as Evolutionary Computation and Reinforcement Learning, and additional newly considered paradigms such as Lazy Learning, Competitive Learning and Hill Climbing. The intent of this investigation is to highlight findings from related computational paradigms that may provide insight into the behaviour, function, and/or capability of the current state and potential elaborations of clonal selection algorithms. An interesting aside is the work by Newborough, et al. [304] provide a framework that suggested that all population algorithms (including clonal selection) may be treated in a similar manner.

3.4.1 Evolutionary Computation

Evolutionary Computation (EC) is a field of study of Evolutionary Algorithms (EA’s) that has much in common with AIS, although draws its inspiration for computation from *Darwinian* (theory of natural selection) and *neo-Darwinian* (findings of modern genetics also referred to as the new synthesis) theories of evolution. Clonal selection algorithms have a superficial similarity to some EC algorithms such as the Genetic Algorithm (GA),

and to the properties of modern Evolution Strategies (ES). See [142, 173] for a classical treatment of EC and [31, 32] for a modern treatment of the field of EC.

In their work on CLONALG, de Castro and Von Zuben consider the similarity between a GA and their approach, particularly in regard to the binary representation used and the stochastic-Darwinian processes employed by both algorithms [112]. They suggest the differences between the algorithms, include the vocabulary used (genetics and evolution verses the shape-space formalism and antibody-antigen cellular interactions), and the somatic mutation and receptor editing used to explore the shape space. The authors also claim that CLONALG can be categorised as an *evolutionary-like algorithm*, although they maintain the same arguments of inspiration, vocabulary, and formalism and the primary differences [117]. In their book de Castro and Timmis again acknowledge the similarity of CLONALG to an EA [109]. They are quick to point out that a major difference between the inspirations of the two approaches is that mutation in evolution is random, whereas the hypermutation process of clonal selection is controlled and directed, proportional to the receptors affinity with the triggering antigen (specific to their abstraction). They suggest that work on EA's can be leveraged by CSA's indicating that research on selection operators (e.g. tournament and roulette wheel selection) may be exploited. The shape-space formalism is presented as a CSA representation abstraction where non-binary shape-space schemes and corresponding mutation mechanisms are discussed, also leveraging from research from representation and mutation in EA's. The authors provide a treatment of evolution and the clonal selection of the acquired immune system. They suggest an important difference between the two theories is the fact that in the clonal process expansion occurs through cell cloning, that there is no sex or genetic recombination, rather only affinity inversely-proportionate somatic hypermutation.

In work on the MISA, Coello Coello and Cruz Cortes claim that their approach is not a genetic algorithm because it does not use recombination (crossover operator) [77], they later adopt a crossover procedure in their approach, as well as adopt a hybrid of EC and AIS terminology [83]. In their work evaluating the BCA on function optimisation Timmis, et al. suggest that BCA is not a GA, based on the empirical performance of the approach on a small suite of test problem instances, although they are very quick to point out the limitations of their small study and the requirement for further research [390]. Forrest, et al. investigated the pattern recognition properties of the immune system [150]. They used a binary coded GA to model antibody-antigen matching in the immune system, which included the clonal selection mechanism, claiming "*The GA without crossover is a reasonable model of clonal selection, while the GA with crossover models genetic evolution*". Hightower, et al. use a Binary GA model of somatic hypermutation of clonal selection to investigate the Baldwin Effect and evolution [207]. Fukuda, et al. [157] and Mori, et al. [297] used a GA to investigate clonal selection properties and immune network algorithms for scheduling and resource allocation. The niching-like properties (an EA property inspired by theories of population genetics and ecology) were observed by de Castro and Von Zuben with CLONALG on multi-modal function optimisation [117], and empirically compared to the fitness sharing approach of Goldberg, et al. [173, 176, 126]. The nich-

ing search properties were proposed to occur given the hill-climbing like behaviour of the independent and semi-independent evolution of B-cells of the various clonal selection algorithms.

The principle similarity between evolutionary algorithms like the genetic algorithm and clonal selection algorithms is the Darwinian inspired theories that inspire both fields (selection and descent with modification). Natural selection accounts for the diversity of species, whereas clonal selection accounts for the diversity of antibodies. As such, both processes are imperfect, stochastic process that operate on pre-committed discrete individuals with short terms goals (relative selective advantage).

The differences between the sub-fields are subtle, related to the functional specifics of the inspired theories. The theory of natural selection encompasses that of clonal selection, as the genetics of the organism define the fitness of that organism, one aspect of which is the structural and function aspects of the organisms adaptive immune system. Related to this point is that of the time-scales involved, evolution of species such as mammals involves the accumulation of small changes over geological time, where as the accumulation of changes in clonal selection are exploited by a single host over its lifetime. Although in both cases selection, decent, and modification occur on genetic material, the hereditary of evolution is the entire genome, where as the hereditary of the B-cell lines are those genes which affect antibody structure. The time-scales also highlight the properties of learning within the host, evolution involves learning at the species level with a single host representing a pre-committed unit of adaptation, whereas acquired immunity involves learning within the host, like the learning that occurs within the brain, so called lifetime or meta-ontogenetic and meta-epigenetic (non-inheritable changes to DNA) adaptation. The differences go beyond nomenclature to the specifics of the inspired processes. Selection in the immune system as a discrete event that triggers (or does not trigger) an immune response resulting in decent with modification. Selection or survival of the fittest is an abstraction that involves relative differences in reproductive success not contingent on anticipated triggered events. The clonal expansion process involves mass replication via cell division where modification is introduced through high-probability point mutations in related genes. Darwinian evolution employs both sexual and asexual reproduction, although evolutionary algorithms typically exploit the crossover of genetic material and point mutations to introduce variation.

3.4.2 Binary Hill Climbing

A realisation of the clonal selection principle includes the choice of strings of bits as a first-order representation. This choice of representation is (1) amenable to analysis and (2) amenable to remapping, and (3) provides a genetic metaphor facilitating genetic-like operations such as copying errors during cell division. The use of binary strings as a first-order representation is pervasive in computational intelligence, particularly in the field of evolutionary computation. Section 3.4.1 demonstrated that there has been much consideration of the relationship between EC and CSA, although little attention has been given to strong relationship between optimisation-driven CSA's and binary hill climbing.

This section considers this relationship by providing a brief overview of the state of the field.

The genetic adaptive plan which became the genetic algorithm operates on a population of bit-strings. This representation facilitated the genetic operators of the approach, and the analysis of the behaviour of the system as processing building blocks (schemata theorem). One of the generalised genetic operators is mutation, the probabilistic flipping of bits in the bitstring during genetic replication. Holland proposed that mutation is determined “*by a random process where each position has a small probability of undergoing mutation, independently of what happens in other positions*” [216] (page 109). In using random mutation alone Holland suggested the use of small mutation probabilities such that the process provides a high probability of maintaining history, which is the dependence of the approach. Holland pointed out the use of high mutation values results in little dependence between successive generations of individuals, calling such an approach *essentially enumerative*. Low probability mutation facilitates the dependence between observations and new trials, although is suggested to be *very unsophisticated* compared to genetic crossover. He proposed the use of low probability mutation as an improvement operator for existing solutions that rather than used in the construction of new trials. He proposed the operator should be used to complement the primary genetic operator by restoring lost genetic material. Further, the application of the mutation operator was described to be disrupting to the building blocks moved around by the crossover operator, further reinforcing the need for a low probability of occurrence. Holland called mutation a background operator, and thus the effects of the operator were largely ignored, and the probability of point-mutations was conventionally kept very small (for example $P_m \in [0.001, 0.01]$) [173, 31]. The probability of a bitstring being mutated is $P_m = 1 - (1 - p_m)^L$ where P_m is the mutation rate and L is the string length. Goldberg differentiated between mutation and crossover by suggesting that *selection+mutation* provides continual improvement with limited scope, whereas *selection+crossover* provides innovation via jumping and remixing of components [177].

Eiben and Schippers summarised the role of the mutation operator as a unary (or asexual) and unbiased operator, highlighting its role as a main operator in evolutionary strategies (traditionally real numbers) and the primary operator in evolution programming (traditionally operates on finite state machines) [133]. Spears highlighted the complementary nature of the two primary genetic algorithm operators suggesting that mutation the focus if optimality is important and the crossover is useful accumulated payoff is important [367]. He suggested that crossover and mutation are two versions of the same general operator that perturbs genetic representations based on available information. In his dissertation, Spears proposes that mutation has a high exploratory power (in that it is not dependent on the composition of the population) and no positional dependence (in that it has the potential to transform a given string into any other string in the search space) [368]. Work by Schaffer, et al. in the late 1980’s and early 1990’s highlighted the potential of the mutation operator, suggesting the operators power had been underestimated in genetic algorithms [348, 347]. These works attempted to formalise an optimal mutation rate

for genetic algorithms, work that was extended experimentally in [347] and theoretically [206]. Later work suggested the use of an initially large mutation rate that exponentially decreased with generational running time of the algorithm. A dynamic and decreasing mutation rate was further supported by others investigating the potential of the operator (not limited to [298, 141]). The reason for this is “*the probability that a mutation will give a better string decreases with the number of bits which are correct*” [30]. A simplified genetic algorithm was proposed for investigating the strengths and limitations of the mutation operator, which became known as a mutation hill-climber, $(1+1, m)$ [298] or simply $(1+1)$ [30]. The algorithm has a static mutation probability (P_m) applied uniformly and independently to each bit in the bitstring. The algorithm was called a hill climber because without crossover, there is no inter-population interaction of genetic information, thus if a population is used, it is a collection of unrelated parallel hill climbers. The $(1+1)$ is taken from the Evolution Strategy naming convention $ES(\mu+\lambda)$ where μ parents create λ offspring, and the best from both ‘+’ are selected. The scheme was reduced to a single parent and offspring for ease of analysis. The algorithm has the interesting property as a hill climber that there is no well-defined neighbourhood for a given string. Rather, a probabilistic neighbourhood is defined by the mutation probability. This neighbourhood covers the entire search space such that it is possible to move from a given string to any other string in the space, a probability that decreases with the increase in Hamming distance between strings and the current string. Mühlenbein analysed the algorithm and proposed an approximate optimal mutation rate (the standard mutation rate) as $P_m = \frac{1}{L}$.

The CLONALG, BCA and the basis of the IA family (SIA for example) are elaborations of either a parallel Random Mutation Hill Climbing Algorithm (RMHCA) or the $(1+1, m)$ -Algorithm. As such, it is reasonable to consider the state of (optimisation-based) clonal selection algorithms as investigations into parallel hill climbers (GA’s without crossover as commented by Forrest, et al. [150]). A fair assessment of algorithms such as CLONALG, BCA, and the IA family requires the consideration of the decades of empirical and theoretical findings from this related field, such as mutation rates, test problems, and existing hill climbing and mutation-based hill climbing algorithms for comparison. This perspective may also highlight a severe weakness in the state of the art of optimisation-centric clonal selection algorithms.

3.4.3 Lazy Learning

Instance-based learning is a non-parametric and supervised machine-learning paradigm where a model is constructed from domain data instances at query time. The complexity increases with the quantity of data, and the more data available, the more specific the model [8, 342]. Instance-based learning is typically referred to as Lazy Learning, which is a machine learning paradigm where generalisation occurs as required. Lazy learning may be contrasted with Eager Learning in which generalisation occurs before a query is received. In addition to the deferral of generalisation, lazy learning is also typified by the combination of stored data used to reply to requests, and the fact that the response is discarded after it is delivered [6, 7, 426]. The three principle characteristics of lazy

learning are: *Defer*: Store all training data and defer processing until queries are given that require replies, *Demand-Driven*: Queries are answered by combining training data, using a local learning (neighbourhood) approach. Instance are (1) identified in as points space, (2) a similarity measure is used to define the neighbourhood, (3) prediction function uses information to answer a query. *Discard* the answer and any intermediate results after answering the query.

Instance-based learning may be considered a type of local learning algorithm, where reasoning is based on models (kernels) constructed from instances similar to the input instance. A typical concern with such models is the capability of the system to generalise. This represents a trade-off between the capacity (size) of the model and the number of samples to consider in the model. Bottou and Vapnik rephrased this local learning trade-off to be between the capacity of the model and the locality of the samples, where locality is the shape and size of the instance selection region (neighbourhood) [51]. This locality property of model construction defines how local an algorithm is, such that an instance-based approach such as k -Nearest Neighbour is highly-local with a low capacity, and a connectionist approach such as a neural network is non-local with high capacity. Case-based Reasoning (CBR) (or memory-based reasoning), is a generalised application of instance-based learning applied to reasoning such as expert systems [5]. Aamodt and Plaza provide a concise description of the various different identities of instance based learning from a case-based reasoning perspective [1]. Their taxonomy highlights the perspective each identity places on this style of learning which covers exemplar, instance, memory, case, and analogy based learning and reasoning. Further, Aamodt and Plaza propose a descriptive framework for CBR, which includes a general algorithmic cycle for the application of the approach, as follows: *Retrieve* the most similar case or cases, *Reuse* the information and knowledge retrieved to solve the problem, *Revise* the proposed solution, and *Retain* the parts of the experience that may be useful for future problem solving.

There are two primary classes of lazy learning approaches: Nearest Neighbour Methods, and Kernel Methods. Nearest neighbour methods are function approximation algorithms that search through a database for similar instances (cases) to a given input instance and make a prediction [84]. The similar points to consider in the prediction are referred to as the neighbourhood, and the number of points is typically denoted k , thus the scheme may be referred to as k -nearest neighbour or k NN. A natural extension to nearest neighbour approaches is that each case generates a local density function or kernel. Predictions may be made by summing together densities functions. The kernel function depends upon the distance function between cases, where the Gaussian is a typical function employed. Some techniques include distance weighted regression such as Linear Weighted Regression methods [29], and a neural-network like approach called Radial Basis Function [59, 61, 328]. Much of the work on instance based learning algorithms is focused on reduction of storage requirements (thus increasing the efficiency of the approach), and improving the algorithms robustness to with regard to attribute noise and irrelevant attributes. Perhaps the most popular concern of nearest neighbour based approaches is what has been referred to as the *curse of dimensionality*. This refers to the

fact that in common datasets, the nearest neighbour is approximately the same distance away from a given point as the furthest neighbour. This effect is observed in dimensions as low as 10-15 [41]. The problem is caused by the exponential increase in volume of a space by the adding of additional dimensions⁷.

The clonal selection principle does share some commonality with the instance based learning, particularly in the context of the continual adaptation to antigen and Aamodt and Plaza's four R's of CBR. The specific knowledge about the antigen is discarded after the infection is addressed, although is retained implicitly in the selected cognate receptors. The AIRS is a realisation of clonal selection that considers the the repertoire as a model of cases that are compressed (quantised) to a set of exemplars. For example, a recent work by Seeker and Freitas [353] make the argument that AIRS is nothing more than an instance-creating algorithm for use by k NN. Exemplar-based modelling is used in a variety approaches for function approximation problems (such as the LVQ algorithm), which represent an important application perspective on clonal selection algorithms (considered further in Section 8.4). A similar approach was taken on the pattern recognition variation of CLONALG, although the number of patterns was known and not compressed as in the case of AIRS.

3.4.4 Competitive Learning

Intrator and Edelman commented that competition between functional units is a widespread biological phenomenon, particularly in the vision systems, the brain, and other sensory systems [226]. Competitive learning systems are based on the idea of competition for activation, and lateral (same level) inhibition. Without inhibition, the models result in an averaged or muddled converged state. They define competitive learning as a dynamic redistribution of responsibilities of various units over parts of the representation space and propose a global, local, and a hierarchical perspective. The authors suggest that resource allocation may be addressed through selective inhibition, such that those units or modules that have not been activated for some time are inhibited less. Further, they suggest investigation into two types of competitive learning: (1) competitive learning via lateral inhibition, (2) competitive learning via the top-down separation of flows of information. Three perspectives of competitive learning are as follows [226]:

- *Global Competition* Competition over the entire representation space, with a strong requirement for inhibition via lateral connections between neurons and the distinction between hard (single winner) and soft (multiple winner) competition
- *Local Competition* The division of the representation space such that local experts (modules of units) compete (known as mixture of experts). Assumes that different processes generate different parts of the representational space, thus hard competition assume a one-to-one matching between modules and hidden generator processes that they model.

⁷It should be noted that the curse of dimensionality has been considered in the context of negative selection algorithms by Stibor, for example [374]

- *Hierarchical Competition* A mixture of global and local competition, where the representation space is partitioned or split into a hierarchical tree structure.

Competitive learning is a connectionist machine-learning paradigm where an input pattern is matched to the node with the most similar input weights, and the weights are adjusted to better resemble the input pattern. This is called the *Winner-Take-All* (WTA or maximum activation) unsupervised learning method where the input pattern is compared to all nodes based on similarity. The nodes compete for selection (or stimulation) and ultimately adjustment (or learning) [276]. Some popular competitive learning algorithms include Kohonen's Self-Organising Map (SOM), Adaptive Resonance Theory (ART), and Counterpropagation Networks, and Neural Gas. SOM is a seminal achievement in the field of competitive learning inspired by the spatial order and organisation of brain functions [261]. The Self-Organizing Map is an unsupervised learning algorithm (clustering) that embodies the WTA principle and facilitates topological preservation in the face of dimensionality reduction. The Learning Vector Quantization (LVQ) is a supervised learning algorithm (classification) and is a zero-order SOM with no connectivity between nodes. Without the connectivity, there is no geometry or topological neighbourhood effects, although the unstructured collection of exemplars compete for stimulation using the winner-take-all principle, and winners that miss-classify during the training process are suppressed. A problem with the winner-take-all learning principle is that units may win too much, and thus dominate selection and response. The result is that some units may never win and the units are considered dead (resources are underutilised) [341]. A relatively simple solution is to dampen the competitiveness of the principle by keeping track of how often units win competitions and using this information to reduce the chance of winning in the future. This may be achieved by introducing a conscience to the units [128], or a similar method called frequency-sensitive competitive learning [9]. This effect may also be achieved by using a generalised version of the suppression mechanism of LVQ2 called rival-penalised competitive learning, where the second winner (rival) in each competition is suppressed [439].

Grossberg with his Adaptive Resonance Theory (ART) [189, 188, 190], along with von der Malsburg [412] were among the first to propose theories of self-organisation and competitive learning of neural cells. Grossberg exploited ART as the basis of a network architecture and unsupervised competitive learning model [71, 191]. The model addressed the instability of unsupervised competitive learning in the face of unexpected irregularities in the input signals. The ART models (ART1, ART2, and others) are in effect self-regulating control structures for autonomous learning and recognition. Hecht-Nielsen combined Kohonen's feature map with Grossberg's outstar network structure calling it a counterpropagation network. This network produces a mapping (like backpropagation) that functions as a statically optimal self-programmable look-up table [202, 203]. Competitive Hebbian Learning is a combination of competitive learning and Hebbian-style learning in which nodes specialise their response to inputs [428, 430, 429]. Martinetz employed competitive Hebbian-learning rule to construct topology preserving graphs by inserting edges between nearest-neighbour nodes distributed across feature space based on input signals

from the domain [282]. The Neural Gas algorithm of Martinetz and Schulten, like SOM is another self-organising and unsupervised competitive learning algorithm [285, 283, 284]. Unlike SOM (and more like LVQ), the nodes are not organised into a lower-dimensional structure, instead the competitive Hebbian-learning like rule is applied to connect, order, and adapt nodes in feature space. Martinetz calls this the *winner-take-most* learning rule distinct from Kohonen's *winner-take-all* rule. The result of the Neural Gas algorithm is a set of points that are distributed across feature space in relatively proportion to the input signal density, with a graph topology that preserves structures in the feature space.

Lymphocytes compete with each other for binding to an antigen. This process has stochastic aspects (the units are mixed in a diffuse substrate), and deterministic aspects (physics of the chemical bonding). Generally, the population is scanned and the highest affinity receptor wins (is proliferated and adapted). Clonal selection is an application of the winner-take-all principle which is manifest in the clonal dominance of receptors for antigen. This effect may also account for the antigenic sin, where the greedy-nature of the winners (winners keep winning) results in lateral inhibition of cells that may be better suited for an antigen. A second interesting point is that the inhibitory effect of clonal selection is attrition and dominance not only of the competition, but also with regard to resource allocation. Cells have a finite lifetime before they are discarded, and the cell population is continually turned-over. Therefore, those receptors that are not used, are not preserved in memory (memory cells) and are lost.

3.4.5 Reinforcement Learning

Reinforcement Learning is an area of study in the field of machine learning characterised by system (or agent) learning to perform a task through a process of trial-and-error in an environment that provides minimal feedback in the form of delayed rewards. Sutton and Barto provide a seminal treatment of reinforcement learning, commenting that problems and not learning methods characterise the field [376].

In the proposal of the CLONALG, de Castro and Von Zuben commented on the commonality between the clonal selection process and reinforcement learning. The learning task of the immune system may be the identification and neutralisation of antigen with the goal of host survival, where the clonal selection process reallocates resources to antibodies (B-cell lines) that are demonstrated as effective. They further comment that feedback is received by the process in what may be interpreted as an unsupervised manner through the re-introduction of previously encounter antigen. They highlight the desire to efficiently allocate limited resources, and the importance of generalisation via the cross-reactive response. The de Castro-Von Zuben interpretation is simplified from the conventional Sutton-Barto definition in which a delayed numerical scoring is given to the entire system given the systems overall performance on the whole uncertain environment. This simplification highlights the question as to how to concisely phrase clonal selection, acquired and even the immune system as a reinforcement learning problem. Sutton-Barto point out that in such a phrasing that a system must perceive the state of its environment, take actions to effect this state, and have internal goals that relate to this state (sensa-

tions, actions, goals). An adjustment to de Castro and Von Zuben’s description of clonal selection better relating it to reinforcement learning is as follows: *Sensations* collectively are the summation of all antigen perceived by the immune system (for example the exogenous pathogenic, and/or internal antigenic environment), *Actions* collectively are the mechanisms of the clonal selection process applied discretely to each antigenic exposure, and *Goals* collectively is the systems efficiency and effectiveness at pathogen identification and neutralisation toward host survival, the feedback of which may be perceived through injury to host tissues from infection or the holistic performance of the system for antigen identification. A reinforcement learning approach that combines aspects of pattern matching and adaptation using a genetic algorithm to evolve a set of binary encoded rules, called Learning Classifier Systems. The relationship of this approach has been made between negative selection algorithms (such as ARTIS), which is discussed further in Section 3.5.3.

3.4.6 Summary

This investigation reinforced the important association of clonal selection algorithms with the field of Evolutionary Computation, specifically with genetic algorithms as they are generally both stochastic adaptive process that operate on pre-committed discrete units, motivated by short term relative goals. The review of the related field of binary hill climbing algorithms demonstrated that the bit-string based CLONALG, BCA and much of the IA family may be considered population-based elaborations of the Random Mutation Hill Climbing Algorithm (RMHCA) or variations of the greedy (1+1,m)-Algorithm. The clonal selection principle and the exemplar-based AIRS share common properties with the field of lazy learning, specifically the deferred and demand driven properties of instance-based learning and the retrieve, revise, retain principles of case-based reasoning. Finally, the principle and algorithms exhibit the competition for selection and winner-take-all properties of competitive learning, where clonal dominance and proportional allocation of resources results in an implicit lateral inhibition of less competitive cells.

3.5 Clonal Selection Adaptive Strategy

This section explicitly considers the computational abstraction of clonal selection in the context of adaptive systems theory. An adaptive systems formalism is reviewed and used to phrase a clonal selection adaptive plan, that is generally similar to Holland’s classical reproductive and genetic plans. A modern complex adaptive system approach is considered, providing insights into both the acquired immune system and into assessing the emergent properties of clonal selection based approaches.

3.5.1 Adaptive Systems

Adaptive systems are difficult to define, compare, and evaluate. The use of formal tools is necessary for the investigation of systems of this type. Holland proposed a formalism in his seminal work on adaptive systems that provides a general manner in which to define an adaptive system [216]. Phrasing systems in this way provides a framework under which

adaptive systems may be evaluated and compared relative to each other, the difficulties and obstacles of investigating specific adaptive systems are exposed, and the abstracted principles of different system types may be distilled. This section provides a summary of the Holland’s seminal adaptive systems formalism and considers clonal selection as an example of an adaptive plan.

Adaptive Systems Formalism

This section presents a brief review of Holland’s adaptive systems formalism taken from [216] (Chapter 2). This presentation focuses particularly on the terms and their description, and has been hybridised with the concise presentation of the formalism by De Jong [120] (page 6). The formalism is divided in two sections: (1) Primary Objects summarised in Table 3.5 and (2) Secondary Objects summarised in Table 3.6. *Primary Objects* are the conventional objects of an adaptive system: the environment e , the strategy or adaptive plan that creates solutions in the environment s , and the utility assigned to created solutions U .

Term	Object	Description
e	Environment	The environment of the system undergoing adaptation.
s	Strategy	The adaptive plan which determines successive structural modifications in response to the environment.
U	Utility	A measure of performance or payoff of different structures in the environment. Maps a given solution (A) to a real number evaluation.

Table 3.5: Summary of primary objects in the adaptive systems formalism.

Secondary Objects extend beyond the primary objects providing the detail of the formalism. These objects suggest a broader context than that of the instance specific primary objects, permitting the evaluation and comparison of sets of objects such as plans (S), environments (E), search spaces (A), and operators (O).

A given adaptive plan acts in discrete time t which is a useful simplification for analysis and computer simulation. A framework for a given adaptive system requires the definition of a set of strategies S , a set of environments E , and criterion for ranking strategies X . A given adaptive plan is specified within this framework given the following set of objects: a search space A , a set of operators O , and feedback from the environment I . Holland proposed a series of fundamental questions when considering the definition for an adaptive system, which he rephrases within the context of the formalism (see Table 3.7).

Some Examples

This section provides a few examples of phrasing adaptive systems using the formalism and examples as to how the formalism may be used as a framework in the investigation of adaptive systems. Holland provides a series of illustrations rephrasing common adaptive systems in the context of the formalism [216] (pages 35-36). Examples include: genetics, economics, game playing, pattern recognition, control, function optimisation and the

Term	Object	Description
A	Search Space	The set of attainable structures, solutions, and the domain of action for an adaptive plan.
E	Environments	The range of different environments, where e is an instance. It may also represent the unknowns of the strategy about the environment.
O	Operators	Set of operators applied to an instance of A at time t (A_t) to transform it into A_{t+1} .
S	Strategies	Set of plans applicable for a given environment (where s is an instance), that use operators from the set O .
X	Criterion	Used to compare strategies (in the set S), under the set of environments (E). Takes into account the efficiency of a plan in different environments.
I	Feedback	Set of possible environmental inputs and signals providing dynamic information to the system about the performance of a particular solution A in a particular environment E .
M	Memory	The memory or retained parts of the input history (I) for a solution (A).

Table 3.6: Summary of secondary objects in the adaptive systems formalism.

central nervous system. Holland applies the formalism to investigate his schemata theorem, reproductive plans, and genetic plans, the foundational models became the field of Evolutionary Computation. From working within the formalism, Holland makes six observations regarding obstacles that may be encountered whilst investigating adaptive systems [216] (pages 159-160): (1) *High cardinality of A* : makes searches long and storage of relevant data difficult. (2) *Appropriateness of credit*: knowledge of the properties about ‘successful’ structures is incomplete, making it hard to predict good future structures from past structures. (3) *High dimensionality of U on an e* : performance is a function of a large number of variables which is difficult for classical optimisation methods. (4) *Non-linearity of U on an e* : many false optima or false peaks, resulting in the potential for a lot of wasted computation. (5) *Mutual interference of search and exploitation*: the exploration (acquisition of new information), exploitation (application of known information) trade-off. (6) *Relevant non-payoff information*: the environment may provide a lot more information in addition to payoff, some of which may be relevant to improved performance.

Cavicchio provides perhaps one of the first applications of the formalism (after Holland) in his dissertation investigating Holland’s reproductive plans [72] (and to a lesser extent in [73]). The work summarises the formalism, presenting essentially the same framework, although he provides a specialisation of the search space A . The search space is broken down into a representation (codes), devices (solutions), and a mapping function from codes to devices. The variation highlights the restriction the representation and mapping have on the designs available to the adaptive plan. Further, such mappings may not be one-to-one, there may be many instances in the representation space that map to the same solution (or the reverse). Although not explicitly defined, Holland’s specification of structures A is clear in pointing out that the structures are not bound to a level of abstraction, that his definition covers structures at all levels. Nevertheless, Cavicchio’s

Question	Formal
To what parts of its environment is the organism (system, organisation) adapting?	What is E ?
How does the environment act upon the adapting organism (system, organisation)?	What is I ?
What structures are undergoing adaptation?	What is A ?
What are the mechanisms of adaptation?	What is O ?
What part of the history of its interaction with the environment does the organism (system, organisation) retain in addition to that summarised in the structure tested?	What is M ?
What limits are there to the adaptive process?	What is S ?
How are different (hypotheses about) adaptive processes to be compared?	What is X ?

Table 3.7: Fundamental questions regarding adaptive systems, taken from [216] (page 29).

specialisation for a representation-solution mapping was demonstrated to be useful in his exploration of reproductive plans (early genetic algorithms). He proposed that an adaptive system is *first order* if the utility function U for structures on an environment encompasses feedback I . Cavicchio described the potential independence (component-wise) and linearity of the utility function with respect to the representation used. De Jong also employed the formalism to investigate reproductive plans in his dissertation research [120]. He indicated that the formalism covers the essential characteristics of adaptation, where the performance of a solution is a function of its characteristics and its environment. Adaptation is defined as a strategy for generating better-performing solutions to a problem by reducing initial uncertainty about the environment via feedback from the evaluation of individual solutions. De Jong used the formalism to define a series of genetic reproductive plans, which he investigated in the context of function optimisation.

3.5.2 Clonal Selection Adaptive Plan

Holland's work on adaptive systems, along with collaboration with colleagues at the Santa Fe Institute evolved into the field (and formalism) of Complex Adaptive Systems (considered in the next Section). Nevertheless, the adaptive systems formalism Holland proposed in 1975 is distinct and remains a useful tool, particularly in the phrasing and investigation of the information processing properties of biologically inspired adaptive plans. This section considers clonal selection as an adaptive plan, similar to Holland's reproductive and genetic adaptive plans.

Primary Objects A primary objects interpretation of clonal selection reveals an adaptive plan that encompasses the extent of clonal selection algorithms. The environment e is comprised of one or a set of antigen, the strategy s is the successive structure modification of a repertoire of antibody using antigen-driven selection and expansion with variation. The utility U is the antigen-wise affinity. This mapping provides little insight, other than the explicit identification of the environment as an actor in the process. As was demon-

strated in the review of the abstraction of the theory and resultant algorithms, the focus lies with the clonal selection principle itself, where as the environment, which defines the scope of the required adaptation, is ignored, and/or simplified to a single antigen (cost function in the case of optimisation).

Secondary Objects This section considers the detail of the formalism, the secondary objects, by addressing the questions proposed by Holland (listed in Table 3.7).

Environments (E) The acquired immune system adapts to antigen. Antigen may be endogenous, such as self-tissues and misbehaving cells (cancer), and exogenous such as pathogen. The tradition with clonal selection abstractions is to ignore the self-nonself discrimination problem which is delegated to the negative selection paradigm, thus one may either consider a general antigenic environment (exogenous and endogenous), or just a pathogenic environment (exogenous). Further, each host has a distinct antigenic environment given the unique internal composition and exposure with the environment.

Feedbacks (I) The environment acts upon the adapting system by explicitly selecting those structures to adapt. The environment is exposed discretely to parts of the system, where the internal repertoire of cells and molecules is compartmentalised and distributed throughout the host, and antigen of various sources have discrete entry and contact points with the repertoire within the host organism. The antigenic-wise feedback of affinity and avidity is relative to the space and time of the specific antigen and the specific sub-repertoire. This is an elaboration on the simplified clonal selection models in which antigen-antibody interaction is holistic with regard to both sets (all antigen interact with all antibody). In addition, different antigen may have different effects on the host, such as pathogen virulence or the speed of cancer metastasis. These antigen behaviour may effect the amount of damage or potential damage a given antigen may cause to the host organism, perhaps perceived by the acquired immune system as danger signals (Danger Theory), or through the release of specialised chemicals such as cytokines. Finally, environmental feedback may be assessed implicitly where the receptor affinity or avidity for given antigen may be a measure of the responsiveness, and thus the utility of the system.

Search Space (A) The search space has been well considered in the abstractions of the theory. The structures undergoing adaptation are the cell receptors, which are defined by receptor genes of the genome within each cell, and are manifest in protein conformations. As discussed, the common abstraction for this space is combinatorial shape-space formalism. One may consider the breakdown of A proposed by Cavicchio into that of representation, devices, and the mapping which clearly correlates with the framework proposed by de Castro and Timmis, discussed in Section 3.2.2.

Operators (O) The mechanisms of adaptation have also been well considered in abstractions of the theory. Beyond the generic selection, expansion, hypermutation, there

is an array of genetic diversity mechanisms outlined in Section 3.2.1, there are also inter-cell interactions such as Helper T-cells, and receptor-receptor recognition described in the immune network theory. Although the scope of the simple clonal selection operators is well considered, those operators which may influence such first-order operators are employed have not (such as cellular interactions during clonal selection).

Memory (M) The history of the interaction between receptors and antigen is recorded using two memory mechanisms. Implicit memory occurs through the change in a given receptors density given the exponential increase in cells that occurs after an exposure. Explicit memory, which is exploited by some clonal selection algorithms (CLONALG for example) is achieved through creation and retention of long-lived memory cells.

Strategies (S) The limits to the adaptive process have not been investigated. The adaptive process may be mediated by helper cells or un-mediated, it may be triggered by an exogenous or endogenous antigen. Beyond combinations of the operator set O , the clonal selection is localised to a space and time within the repertoire. Such constraints on antigen-centric selection have not been considered. Therefore, the consideration that influences the process suggested in the operators set provide much opportunity for different strategies beyond simple reflexive clonal selection.

Criterion (X) The adaptive processes may be compared by their general capabilities of identifying and neutralising antigen in an environment. Some adaptive plans may be suited to general or specific identification, others may be suited to endogenous or exogenous antigenic environments. Given that S considers that there are multiple clonal selection strategies, and E considers that there are multiple of such environments for strategies, the consideration of general clonal selection centric assessment criterion has not previously been considered.

Summary The primary objects consideration reveals, a *first-order* adaptive system where a simple abstraction of the clonal selection procedure considers a payoff that encompasses the extent of the interaction with the environment, where typically (in the case of optimisation) an environment is composed with one antigen and one affinity landscape. The secondary objects consideration of clonal selection provides an explicit decomposition and elaboration of the procedure, highlighting a number of weak points not considered, including: the environment as an actor in the process, a range of clonal selection environments, numerous explicit and implicit feedbacks beyond a single affinity landscape, an implicit memory mechanism, the potential for a class of clonal selection adaptive plans, and clonal-selection centric criterion for comparing such plans. In addition, the elaboration highlighted those aspects that have been considered in existing abstractions including: the search space formalism, the operators used in the procedure, and the core (simplest) arrangement of the operators. Therefore, the explicit phrasing of clonal selection as a Hollandian adaptive system provided a clear framework for communicating (1) the sub-elements of the strategy, and (2) the extent to which those sub-elements have or have not

been investigated.

3.5.3 Complex Adaptive Systems

Adaptive strategies are typically complex because they result in irreducible emergent behaviour that occur as a result of the non-linear interactions of systems components. The study of Complex Adaptive Systems is the study of high-level abstractions of natural and artificial systems that are generally impervious to traditional analysis techniques. Macroscopic patterns emerge from the dynamic and non-linear interactions of the systems low-level (microscopic) adaptive agents. The emergent patterns are more than the sum of their parts. As such the traditional reductionist methodology fail to describe how the macroscopic patterns emerge. Rather, holistic and totalistic investigatory approaches are applied that relate the simple rules and interactions of the simple adaptive agents to their emergent effects in a ‘bottom-up’ manner. Some relevant examples of CAS include: the development of embryos, ecologies, genetic evolution, thinking and learning in the brain, weather systems, social systems, insect swarms, *bacteria becoming resistant to an antibiotic*, and the *function of the adaptive immune system*. The field of Complex Adaptive Systems was founded at the Santa Fe Institute (SFI), in the late 1980’s by a group of physicists, economists, and others interested in the study of complex systems in which the agents of those systems change [21]. Perhaps one of the largest contributors to the inception of the field from the perspective of adaptation was Holland. He was interested in the question of how computers could be programmed so that problem-solving capabilities are built up by specifying: “*what is to be done*” (inductive information processing) rather than “*how to do it*” (deductive information processing). In the 1992 reprint of his book he provided a summary of CAS with a computational example called ECHO [216]. His work on CAS was expanded in a later book which provided an in depth study of the topic [218]. There is no clear definition of a Complex Adaptive System, rather sets of parsimonious principles and properties, many different researches in the field defining their own nomenclature. Popular definitions beyond Holland’s work include that of Gell-Mann [166] and Arthur [28].

Immune System as a Complex Adaptive System The acquired immune system of vertebrates is an often cited example of a natural complex adaptive system. Some examples of the acquired immune system phrased as a complex adaptive system include: A simulation of the immune system and the AIDS HIV using Cellular Automata [187]. A simulation of vaccination in the immune system using agents [10], a general summary of the immune system and simulation approaches in the context of CAS [11], and the simulation of an immune system and HIV using elements of classifier systems, genetic algorithms and cellular automata [383]. The connection between CAS and the field of Artificial Immune Systems has been observed. Hofmeyr and Forrest considers the relationship between the negative selection-based ARTIS, the acquired immune system, and CAS [214]. They proposed that representation is fine-grained where peptides are the basic unit of representation, and that coordination is emergent from the interaction of trillions of cells

and molecules. They also considered the importance of the systems tight-coupling with its environment, and specifically the immune systems ability to respond to perpetually novel stimuli. Vargas, et al. consider the immune system and Artificial Immune Systems as examples of complex adaptive systems [405], and phrase an AIS inspired Learning Classifier System [406, 404].

Beyond explicitly investigating immunology as a CAS, the formalisation of such systems can provide insights into the general properties of the immune system as a complex adaptive system. For example, Holland suggested that such systems respond instant-by-instant, and the importance of the systems ability to balance *exploration* (acquisition of new information or capabilities) with *exploitation* (efficient use of information or capabilities already available) [217]. Arthur proposed a definition of CAS in the context of economics with six properties [28], all of which resonate with the information processing properties of the acquired immune system. Generally these properties are: the *dispersed interaction* of possibly heterogeneous agents acting in parallel, the *lack of a global controller*, the *cross-cutting interaction* of units across hierarchical concerns, *continual adaptation* in the presence of *perpetual novelty*, and finally *dynamics* where the system operates far from global optimum and equilibrium. Finally, Jost suggests that the environment in which a CAS exists is more complex than the CAS itself and that CAS depends on regularities in its environment [237]. Jost provides a rigorous assessment of CAS in the context of *internal* and *external* complexity. External complexity is defined as the amount of input, information, energy the system obtains from the environment. Internal complexity is the complexity of the internal representation of the information it takes as input (model complexity). The goal of the system is to handle as much input as possible with as simple a model as possible, to attempt to increase the external complexity and reduce the internal complexity of the system.

Clonal Selection The immune system may be an instance of clonal selection, although the computational properties of clonal selection alone are not. Clonal selection describes the structure and function of an opportunistic improvement mechanism for specialising a repertoire of molecular pattern recognition cells. A broader perspective of clonal selection that takes into account more than receptors, lymphocytes, and antigen introduces complexity into the adaptive system that will make effective assessment difficult. A complexity perspective on adaptive systems may be exploited in the investigation of computational models of clonal selection by focusing effort on a top-down analysis emergent or systemic effects rather than a bottom-up analysis of the discrete components. Specifically, two top-down clonal selection centric concerns include: (1) the information composition of the system which may be explicitly aggregated and assessed in the case of distributed system (for example information in the holistic repertoire), and (2) the holistic system capability in a given environment potentially aggregated across distributed contact points (for example capability across the holistic repertoire).

3.5.4 Summary

Clonal selection is a theory that defines the adaptation of a general repertoire of lymphocytes with surface bound immunoglobulin, toward a repertoire with specialised capability for a hosts antigenic environment. Holland's adaptive system formalism provided a framework for defining the sub-elements of clonal selection as an adaptive plan, highlighting the generally underestimated role of the environment and the feedbacks and interactions between the components of the system. Specifically, the scope of reviewed clonal selection algorithms may be considered *first order* adaptive systems where the extent of the feedback to the system is presented in the utility function. Modern adaptive systems theory embodied in the study of Complex Adaptive Systems, highlighted the inherent complexity of the environment in which a system operates, and the difficulty in using bottom-up approaches to analyse such systems. The perspective suggested at the consideration of the immune system as an instance of a CAS, where clonal selection alone provides an adaptive quality to the system and is in and of itself not an CAS.

3.6 Cellular Strategy and Beyond

This section considers the cellular-focus criticism of clonal selection algorithms made in Section 3.3.3. Specifically, given that cellular clonal selection may be considered a motivating metaphor for adaptive information processing, elaborations of this foundational approach are considered including explicit investigation in the context of related immunological metaphors as well as related cellular interaction concerns. The problem of distributed clonal selection is considered in the context of the structure and function of the broader immune system. An agenda is outlined toward distributed clonal selection algorithms that suggests: (1) the exploitation of immune physiology with regard to the lymphatic system and lymphocyte migration, and (2) the exploitation of populations of interacting immune systems with regard to immunisation and evolutionary immunology.

3.6.1 Adaptive Cellular Strategy

This chapter has reviewed clonal selection from an immunological (Section 3.2), computational (Section 3.3), and adaptive systems perspectives (Section 3.5) providing a abundance of support to suggest at the information processing properties of the motivating metaphor. The theory is adaptive by definition as it describes how a general repertoire of lymphocyte receptors are specialised (adapted through iterative improvement) in the context of an environment. The adaptive theory relies on the replication properties of immune cells and iterative improvement depends upon the blind mutation-based affinity maturation to antigenic signals. Receptor interaction with antigen is a pattern recognition process involving the 'fitting' or 'matching' of pre-committed molecular structures, where each attempt at match with antigen requires a localised proximity and physical contact (trials within trials). The geometric paradigms from theoretical immunology phrase the scope and preparation of receptor configurations as a finite combinatorial search problem. Pair-wise relationships between configurations can be assessed in the context of a speciali-

sation of Wrights fitness landscape formalism that stresses the adaptation of the ‘goodness of fit’ of structures.

A first-pass exploitation of the underlying information processing qualities of the theory elicited a suite of adaptive algorithms that superficially show a strong resemblance to parallel mutation hill climbing algorithms for the purposes of global optimisation and vector quantisation. A deeper assessment of the underlying adaptive strategy in the context of adaptive systems theory provided insights into the behaviour of the approach, and motivations for elaboration. Regarding behaviours, the strategy uses adaptation toward information acquisition in a population memory structure using accumulated trial-and-error with pre-committed structures, each of which is individually redundant. More interestingly, this redundancy may promote degeneracy in contextual emergent specificity in addition to general fault tolerance. Two specific areas were highlighted for elaboration, as follows:

Cellular Interactions The interactions between cells conventionally belongs to the domain of the immune network theory rather than clonal selection, although cellular interactions represent a fundamental concern of clonal selection as an adaptive strategy and an immunological theory. Specifically, the cells interaction when competing for selection and ultimately for limited resources in the repertoire. In the immune system, lymphocytes compete in a spatial environment, and varied immune cell types interact with each other to coordinate a response such as in the case of the two-signal activation of B-lymphocytes by antigen and helper T-cells. Cell interactions add an additional level of complexity of acquired information and behaviour to the adaptive strategy beyond raw improvement of structures. This approach of elaborating clonal selection with so-called ‘more detailed’ immunological inspired models aligns with the general approach advocated in the field (Section 2.6). Specifically, the added detail provided by cellular interactions was advocated by Timmis, et al. in the context of CLONALG, where it was suggested that “*whilst [CLONALG is] intuitively appealing, [it] lacks any notion of interaction of B-cells with T-cells, MHC, or cytokines.*” [389].

Antigenic Environment The explicit phrasing of clonal selection as an adaptive system highlighted the distinct lack of a well defined environment in which the system is situated and interacts with. Antigen provide the catalyst for adaptation in clonal selection, and more generally the extent of endogenous and exogenous antigen the system may be exposed to over its lifetime of usefulness defines the scope of information with which the system must be concerned. The classical primary and secondary response to antigen highlighting the improvement capability of the clonal selection process highlights the exposure-centric nature of this environment (environmental trigger and system response). Without an antigenic environment, the information processing qualities of clonal selection have little meaning, highlighting the importance in elaborating this fundamental aspect of the adaptive strategy.

3.6.2 Toward Distributed Clonal Selection

The discrete population basis of the clonal selection strategy and trigger-based adaptation provide an inherently parallel process suitable for concurrent implementation. This was shown to have been highlighted and demonstrated explicitly with AIRS for classification, although parallisation alone fails to capitalise on the promise of distributed realisations of the underlying adaptive strategy. Toward this end, the structure and function of the broader immune system provide motivating metaphors for extending and constraining the clonal selection strategy for decentralised adaptation in a distributed environment. Specifically, two broader perspectives of clonal selection are proposed as paths towards such approaches, as follows:

Host of Tissues A clone of lymphocytes possess the same receptors and are therefore suited to the same antigen, where a repertoire (pool or population) of cells defines an aggregation of such groups. In an organism (such as a mammal), this repertoire is spatially distributed throughout the tissues of the host, where diverse groups of lymphocytes provide antigen surveillance by migrating around the body via the vascular system, whilst other groups position themselves in tissues awaiting the arrival of pathogen. As such, the structure and function of the immune organs, tissues, and their connectivity throughout the general mammal organism provide a template for organising and constraining a tissue class of clonal selection approaches. Lymphatic tissues and lymphocyte migration has been proposed as a model for distributed Negative selection specifically in Hofmeyr's ARTIS (specifically as read-only broadcast of negative detectors, see Section 2.5.2) and more generally as an inspiration for decentralised Artificial Immune Systems by Dasgupta [106], although to the authors knowledge such approaches have not been implemented or investigated in the context of distributed adaptive clonal selection.

Population of Hosts A natural extension of the tissue perspective, is to scale the adaptive concerns of clonal selection to that of a population of hosts with interacting acquired immune systems. Although conceptually, a host perspective is removed from the concerns of receptor interactions with antigenic determinants, the general principles of population immunology depend on such consistent low-level adaptation. Specifically, the a population of hosts is concerned with immunisation of all types such as vaccinations, epidemics, and maternal immunity. Unlike tissues of a body, individual hosts in a population are redundant in the same manner as individual cells are, with life cycles, and evolution. The consideration of genetic evolution combined with ontogenetic evolution of an immune gene library using genetic and clonal selection was considered by Hightower, et al. [208, 207], although there has been little follow-up work, specific with regard to the role and relationship of clonal selection in such systems.

3.7 Chapter Summary

The detailed review of the clonal selection theory in Section 3.2 provided the context to assess the motivations and expected behaviours embodied in the state of the field of clonal selection algorithms. This insight aided in outlining the three proposed criticisms of the state of clonal selection algorithms in Section 3.3.3. The first regarding the placement and assessment of the field in a broader context of related Computational Intelligence approaches was addressed in Section 3.4 strengthening the relationship with Evolutionary Computation, and outlining the important relationships with Hill Climbing algorithms, Instance-Based Learning, and Competitive Learning. The second limitation suggested that the lack of explicating investigating clonal selection as an adaptive strategy which was addressed in Section 3.5 highlighting the need for: (1) a clearly defined environment to complement the adaptive response of the strategy, (2) the importance of interactions between discrete components of the system, and (3) the focus on emergent behaviours. Section 3.6 addressed the third and final limitation of the cellular focus of Clonal Selection Algorithms, and in so doing outlined a research agenda for both (1) elaborating upon the existing cellular clonal selection algorithms, and (2) provided a basis for investigating distributed clonal selection algorithms. This agenda occupies the remainder of this dissertation. Specifically, Chapter 4 investigates the Cellular Clonal Selection adaptive strategy including both the definition of an Antigenic Exposure Paradigm, as well as algorithms that incorporate additional forms of intra-repertoire interactions. Chapter 5 defines and investigates the Tissue Clonal Selection Paradigm inspired by the structure of the lymphatic system and behaviour of migrating lymphocytes, and Chapter 6 defines and investigates the Host Clonal Selection Paradigm inspired by the immunisation and evolution of the acquired immune system.

Chapter 4

Cellular Clonal Selection

4.1 Chapter Overview

This chapter consolidates the findings of the previous chapter and proposes the first of three constrained perspectives of the clonal selection strategy. Specifically, this chapter is concerned with clonal selection as constrained by the antigenic, cellular, and molecular interactions of cells within a repertoire called the *Cellular Clonal Selection Paradigm*. Importantly the perspective of clonal selection as an adaptive strategy considered in this chapter both encapsulates the state of the art in clonal selection algorithms, and provides a bedrock of understanding for the integrated hierarchy of perspectives considered in the remainder of this dissertation. Section 4.2 considers the cellular perspective from an abstract perspective, dividing the concerns of the paradigm into that of *system* and *environment*, clearly delineating the responsibility of the adaptive strategy and the problem domain that provides the context for adaptation. The paradigm is realised in Section 4.3 with regard to specific algorithm and problem definitions and empirical measures for assessing the composition and capability of a given repertoire governed by clonal selection. The behavioural expectations of an archetype of the strategy are assessed and confirmed in a series of three empirical studies in Section 4.4. These findings provide a foundation from which three extensions of the strategy are proposed and investigated using empirical study to confirm behaviour expectations, including (1) a spatial context for the repertoire in Section 4.5, (2) mediation of response provided by the interaction different cell casts in Section 4.6, and (3) the promotion of higher-order cross-exposure structures in the repertoire via network inspired interactions in Section 4.7.

4.2 Abstract Cellular Paradigm

The Cellular Clonal Selection Paradigm is a rephrasing and restricting of the existing field of algorithms that realise the computational attributes of the clonal selection theory of acquired immunity. This section considers the abstract concerns of the paradigm focusing on (1) a quintessential cellular clonal selection adaptive strategy, (2) a system-environment phrasing of the strategies interaction with a given domain called the antigenic exposure paradigm, and (3) a generic problem domain for conceptualising and investigating cellular

clonal selection algorithms and their elaborations.

4.2.1 Quintessential Clonal Selection

This section considers an archetype of clonal selection adaptive strategy as the consolidation of the related immunological theory, computational principles and approaches, and adaptive systems theory from Chapter 3. Clonal selection is constrained to the concerns of a repertoire of cells and antigen called the *Cellular Clonal Selection Paradigm* that focuses on the separation and interaction of system-environment, and clonal selection as an adaptive strategy.

Constrained Scope

The clonal selection theory is defined at the cell, antibody, and antigen level of immunobiology and immunochemistry to describe what in complex system theory are described as *emergent effects* of host immunity. Therefore one may constrain the investigation of the information processing concerns of clonal selection at this level and consider the emergent immunity effects within the context of a repertoire of cells with regard to antigen. In this work, the study of clonal selection under this constraint is referred to as *Cellular Clonal Selection*, distinct from the consideration of clonal selection under multiple of such repertoires in *Tissue Clonal Selection* (Chapter 5), and the consideration of clonal selection across multiple whole immune systems in *Host Clonal Selection* (Chapter 6). Cellular Clonal Selection is primarily concerned with the management of a repertoire of discrete cells and its interaction with antigen. The scope of control for this management is the repertoire in its entirety. This suggests that the repertoire of cells in cellular clonal selection is centrally governed with access to complete information regarding the information content of the repertoire. Therefore the scope of cellular clonal selection algorithms considered in the reviewed taxonomy (Section 3.3) may be considered Cellular Clonal Selection Algorithms. As discussed in that section, this level of clonal selection algorithms are strongly related in mechanism and effect to the Genetic Algorithm (GA) and the Mutation Hill Climbing Algorithm (MHCA), as well as the related fields of Lazy and Competitive Learning.

System and Environment

A critical perspective on computational models of clonal selection highlighted by the adaptive systems theory in the previous chapter was the need for a clear delineation between system and the environment that the system interacts with. The system is defined as being comprised of a cellular repertoire managed by clonal selection and related information processing mechanisms. The environment is defined as the scope of endogenous antigen and exogenous pathogen to which the system may be exposed. This delineation is easily related back to the existing state of clonal selection algorithm research as a the separation between clonal selection algorithm and problem domain to which the algorithm is applied. The information content of the system and the environment may therefore be considered in the context of the shape-space and affinity landscape geometric paradigms from theoretical immunology as sets of discrete cells and antigen patterns that may be assessed

(specificity and antigenicity) against each other as a cost surface. The antigenic environment defines the scope of the information a system may acquire and apply (immunity), the full importance to clonal selection of which is considered in the Antigenic Exposure Paradigm in Section 4.2.2.

Clonal Selection as Strategy

The management of a cellular repertoire in the context of an antigenic environment is the responsibility of the clonal selection adaptive strategy. The focus of such a strategy is the acquisition and application of information (immunity) from an unknown information (antigenic) environment. The selectionist, adaptationist, and satisficing lens of the clonal selection strategy considered in the previous chapter highlighted the following computational principles of the theory:

1. Information is acquired through cumulative generate-and-test of pre-committed discrete structures.
2. Resource allocation is competitive, based on (in some way) the differential ‘goodness of fit’.
3. The continued acquisition and application of information is based on the premise that future antigenic exposures are much like past exposures.
4. Adaptation continues for as long as relative improvements can be made under all of the constraints of the strategy.
5. The emergent effects of the bottom-up exposure-wise interactions may be assessed in the holistic *composition* and *capabilities* of the information acquired in the cellular repertoire.

The implementation of such a strategy was shown in Section 3.3 to be commonly realised as management of a repertoire of binary strings under assessment of an oracle cost function resulting in duplication and mutation of low cost (high affinity) strings. The generality of such a realisation may be constrained by the principles or axioms governing cellular interactions with antigen from Section 3.2, as follows:

- *Assessment*: Affinity is an ordinal scoring of the ‘goodness of fit’ of a structure provided by an oracle.
- *Selection Competition*: Competition is promoted between cells specialised for the same antigen, and avoided between cells specialised for different antigen.
- *Selection*: A small and specific founding set of cells are selected to found a clonal response (oligoclonal).
- *Cloning*: The number of clones created from an exposure is larger than the founding set of cells selected for the clonal response.

- *Mutation*: The point-wise rate of mutation is higher than that of normal mitosis although is fixed.
- *Integration Competition*: The cellular repertoire is large although finite, promoting competition between cells for limited positions.
- *Integration*: Although many clones are created in a clonal response, few are retained in the repertoire.

The listed computational principles may be considered an extension and elaboration on the affinity maturation principle of de Castro and Timmis, and the clonal selection principle of Cutello and Nicosia in Section 3.3.1. These principles are the interpreted heuristics from reviewed immunological theory use in this work and are not intended to provide laws for the realisation of cellular clonal selection algorithms. Specifically, differential resource allocation can be realised via a variety of specialised mechanisms including: assessment, selection, cloning, mutation, and integration¹. These principles are used to assess and configure cellular clonal selection algorithms in this work, and are augmented with knowledge from the study of the related GA's and MHCA's. Specifically from Section 3.4.2: (1) the acceptance of structures with an improved or equal assessment scoring to permit the acceptance of varied structures with neutral effect on assessment to cross flat spots in the affinity landscape, and (2) the use of a $\frac{1}{L}$ fixed mutation rate as an approximation of the optimal mutation rate of bit strings on linear binary cost functions.

4.2.2 Antigenic Exposure Paradigm

This section proposes an artificial conceptual model of a generalised natural antigenic environment. This model is comprised of a set of external triggers that stimulate an internal activation and response by an acquired immune system. This conceptualisation is called an *antigenic exposure*, and provides a perspective on the adaptation-based learning and memory qualities of the acquired immune system, and a standardised framework for mapping problem domains onto the cellular clonal selection paradigm.

The immune system is embedded or situated in an *antigenic environment*, to which it (in general) passively responds. Therefore, a system has qualities of an intelligent agent (it is situated, intelligent, and acts autonomously), although does not act upon the environment, unlike a reinforcement learning system. A consideration of antigen other than superficial mapping of a given problem domain is largely ignored in the investigation of Artificial Immune Systems in general, as the investigatory concerns lie predominantly with the information processing qualities of the system. Antigen may be exogenous or endogenous in origin and may be benign or actively adversarial. The environment encompasses the scope of the information that may be acquired by the system, dictating what to learn and when to learn it. Three antigenic exposure regimes and resultant models are proposed, as follows: (1) single exposure, (2) multiple exposures, and (3) multiple antigen

¹As evidence, one may consider the variety of such mechanisms used for differential resource allocation in the related field of evolutionary computation.

with multiple exposures, that highlight the spatial-temporal consistency concerns of clonal selection at the cellular level.

Single Exposure Model

An antigenic exposure is an event that involves the arrival of antigens (information of external origin) to an immune system. The arriving antigen are identified (typically forced as a requirement of the domain) by the system stimulating a proportionate immune response. If correctly identified, the *external stimulation* results in an *internal activation* of some lymphocytes in the immune systems repertoire. The response involves raising a clone of lymphocytes to address the stimulation with regard to the magnitude and specificity of the response. The system is concerned with raising a good response quickly. The principle of an antigen is that it may cause harm to the host in the form of tissue damage promoting the immediacy of the response. The system must raise a clone that can address the magnitude of the exposure, therefore the effectiveness of the response should be proportional to the virulence of the antigen, and the response size (clonal expansion) should be proportional to the number of arriving antigen (amplitude or dose).

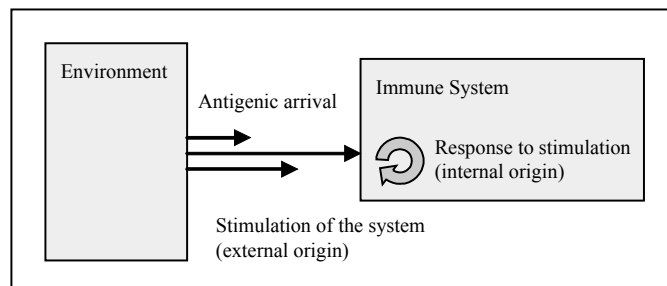


Figure 4.1: Depiction of external stimulation resulting in internal activation.

This *exposure virulence* is a property of the antigen thus is not likely to change for a given immune system unless the antigen itself changes. The virulence may define the penalty for not identifying a antigen, as well as the amount of resultant specificity desired or required in a clonal selection and expansion response. This concern (virulence-specificity) is difficult to conceptualise as it defines the relative amount of refinement-based learning required of the system for a given antigen. The concern of the size of an antigen dose may be conceptualised as the amount or *exposure amplitude* of information an antigen represents. The number of simultaneous antigen that arrive to the system in a given exposure defines the amount of work the system may have to support. The number of antigenic molecules that arrive in an exposure may be more than the number of lymphocytes in the repertoire. The result is that selection events may be more intense, resulting in the delay of some interaction until after the initial cells have proliferated or additional resources have been recruited. This *exposure amplitude mismatch* scenario has two speculated implications: (1) Antigen may not be neutralised as quickly if it arrives in unexpectedly large quantities, that would result in the antigen arrival being considered a resources which must be degraded (neutralised over time), and (2) The increased se-

lective pressure from the size mismatch may result in the rapid synthesis of a response large enough to address the exposure. This suggests dynamic and variable response size capabilities inherent in the clonal selection and expansion response.

One may consider the effects of varied exposure virulence and amplitude on an immune system. Both properties affect the amount of work required in an immune response, although in different dimensions. Virulence may determine the desired specificity of a response such that high damage-causing pathogens are neutralised effectively and low damage-causing pathogens are neutralised with less specificity. The amplitude of the exposure defines the number of cells and the size of the clonal response. Therefore, the immune system must trade-off response strategies to the various anticipated virulence and amplitude combinations. For a single exposure event, an immune system is only concerned with quelling the antigen with a response. The scope of the concerns of an immune system for a single exposure is the specificity of the lymphocytes to the antigen and synthesising the cells in sufficient numbers. Initial antigen-identification (to trigger a response) may be assumed. Further, all resources of the repertoire may be allocated to the response for the exposure, as there are no future exposures. Iterations within a system under these conditions are concerned with adaptation in the context of a fixed and known single exposure event.

Multiple Exposure Model

A single exposure-response event may be abstracted to multiple antigen exposure-response events. A principle property of a multiple exposure regime is the temporal dimension it adds to such events. *Exposure frequency* facilitates conceptualising multiple exposures as a series of atomic and quasi-independent single events, where the repetitiveness of the arrival of a given antigen is defined as its frequency. The exposure frequency may be measured as the pattern of *exposure* and *no-exposure* events in discrete time over an interval. The pattern may be defined as a probabilistic function, random, uniform, or may be a regular deterministic exposure function. The concept of a non-stimulation fostered by no-exposure events suggests a downtime when the immune system is not stimulated to respond. Such a non-activity event may also occur for an exposure event that is not identified by the exposed system. This inactive time is referred to as antigenic *exposure downtime*. The activity of a system may be considered trigger-based in which learning and maintenance processes (such as cell aging and removal) only occur when the system is exposed to pathogen. During non-triggered periods, the system may enter a period of stasis awaiting subsequent triggers (the idle repertoire awaiting exposure model was criticised in the Immune Network Theory). An alternative model is that the system persists (always on), continuing to execute such maintenance processes in the absence of stimulation. The two models highlight the potential for the decoupling the system processes of *repertoire maintenance* and *repertoire stimulation*. Such a decoupled system would require an effective memory system such that acquired immunity (information) was not lost by the high-turnover of lymphocytes in the repertoire. The lack of adequate memory for this system type in an extended period of non-stimulation may cause the

repertoire to devolve to a naïve (random) state.

Variations in the exposure frequency control when the system can and must respond. Variations in the exposure amplitude control the size (scope or how much) of the response that a system must synthesise. Variations in exposure virulence and amplitude imply variations in a systems response strategy in terms of response specificity and response quantity. A similar relationship exists with multiple exposures with the addition of exposure frequency and the systems memory of the response. There is an inversely proportional relationship between immunological memory and exposure frequency (with constant virulence) such that when the exposure frequency is low, the memory must be long lived. When the exposure frequency is high, memory is short lived, and in fact may be satisfied with the quantity and specificity effects achieved from the single exposure without memory. One may consider the effect of combinations of high and low frequency exposures with high and low amplitude pathogen arrival. From a system perspective, the frequency and amplitude of exposures may be compressed (multiplicative relationship) into the amount of work (response effort) required for the interval of time. This compression of exposure frequency and amplitude may be referred to as the *exposure regime* of an antigen to a system through time.

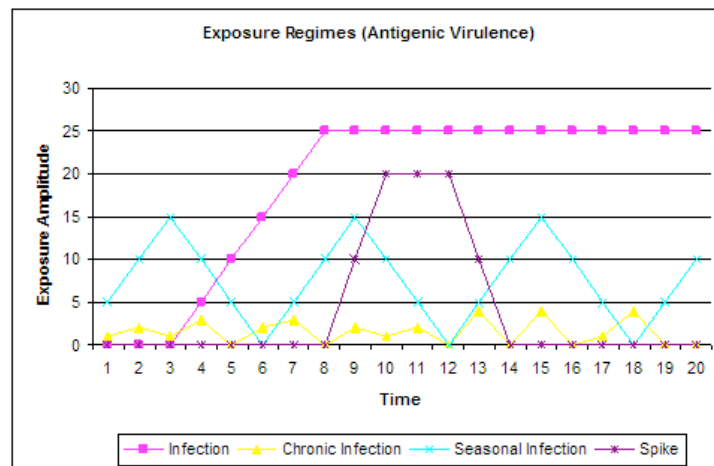
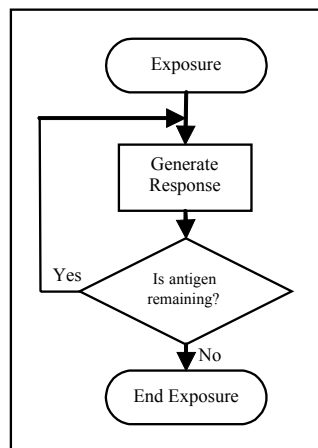


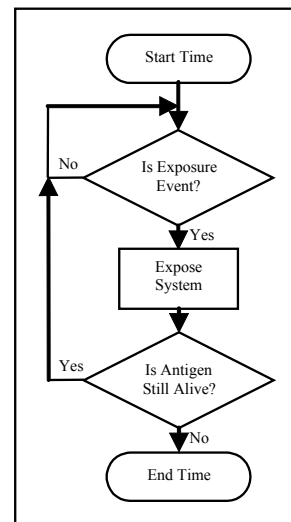
Figure 4.2: Example exposure regimes demonstrating some archetypical pathogen virulence behaviours.

The relationship between a pathogen exposure regime and an immune systems response effort is interesting. For a single exposure event, virulence defined the amount of tissue damage an antigen may cause on a host, therefore it may provide a conception of the amount of specificity refinement (effort) required of an immune system for the exposure. An exposure regime provides a way of crisply defining *antigenic virulence* (distinct from single exposure virulence) as a function of exposures through time. The amount of specificity refinement by the system for the antigen (response effort) is proportional to the exposure frequency and exposure amplitude (exposure regime). Antigenic virulence (and exposure regimes) may be expressed as an exposure amplitude graph against time. From this level of abstraction, one may give example exposure regimes as archetypical pathogen virulence strategies with which to subject an immune system (see Figure 4.2). In addition

to static antigenic virulence, it is possible for virulence to vary through time. The variance in virulence may be system-dependant such that virulence changes in response to the internal actions of the immune system such as in the case of an adversarial antigen (pathogen). For a multiple exposure event, the system is primarily concerned with devising a response to each exposure and ultimately using past exposures to anticipate the resources needed for future exposures. This exposure model supersedes the single exposure event, as in the case of a single exposure, all resources of the repertoire are allocated to the specific concerns of the antigen (as there are no other antigens). The learning that the system may achieve within the context of a single exposure may be limited, as a broader exposure regime may define how often the exposures arrive to the system. The exposure regime may impose a quantity and specificity capability constraints for a single exposure, meaning that a finite amount of specificity learning and/or quantity learning may be achieved for a single exposure event. The system must learn to anticipate (1) the required specificity of future exposures, and (2) the required quantity of future exposures, both in the face of non-exposure events.



(a) Single Exposure Model.



(b) Multiple Exposure Model.

Figure 4.3: Flow diagram of the single and multiple exposure models.

The exposure regime (frequency and amplitude) define the antigenic virulence and the systems required specificity. Repeated application of blind clonal selection facilitates improvement of a repertoire specificity for the antigen, and for a system that enters stasis during periods of non-stimulation (information acquired by the repertoire does not degrade), the density of information (clonal convergence and/or clonal densities) may define the quantity or scope of information required for the regime. Therefore, a repeated single selection on a relatively naïve repertoire fosters a model that matches the virulence of the antigenic exposure regime. For a repertoire with decoupled maintenance of cells and triggered adaptation, an explicit memory mechanism is required that retains acquired information proportional to the frequency of its exposure and complexity. For lower-

complexity information, a more efficient strategy may be to generalise or re-acquired as need.

Multiple Antigen Model

The natural extension of multiple antigen exposures is to subject an immune system to multiple concurrent exposure regimes. Each regime is expressed by a distinct *antigenic type* or variety. Antigen types differ in their composition (surface features) such that the immune system has to acquire different immunity characteristics for each. In addition to the varied immunity characteristics, each antigen type has its own exposure regime, thus in the context of an immune system, the regime elicits a distinct response effort. The aggregation of multiple antigen types is an aggregation of multiple antigen exposure regimes to a given immune system that may collectively be referred to as the exposure or *antigenic environment*. In responding to one given antigen, the repertoire may acquire a level of immunity (specificity) to another different and distinct pathogen type. This effect is called cross-reactivity of the immune response and may be conceptualised as reuse and generalisation of acquired information. One may consider an immune system to be exposed to the multiple different exposure regimes shown in Figure 4.2. In this example, the system is exposed to all four different antigen with varied amplitudes at the discrete time of $t = 11$. This may be referred to a *concurrent antigen exposure*. Exposure concurrency requires that the repertoire is (1) large enough (with regard to the quantity of lymphocytes), and (2) diverse enough (with regard to lymphocyte specificity) to address multiple pathogen types at the same time. It is expected that the efficiency of the system will be reduced in such a situation, as resources are allocated proportional to the relative virulence of each antigen type. This effect highlights the requirement of the system to be able to not only respond to multiple antigen types at the same time, but to integrate the results of the concurrent responses (expanded clones and shifted the repertoire specificities).

Unlike the previous two models that were able to allocate all the resources of the repertoire to a single exposure of antigen information, multiple antigens requires a proportional allocation of resources to each antigen and its exposure regime. The proportionality of the response and allocation of repertoire resources is defined by the relative antigenic virulence to other antigen in its environment. The previous concerns of optimising specificity, quantity, and the anticipation of these values may be reduced to the systems adaptation of information densities toward exposure regimes. The density is representative of the learned (aggregated from experience) cell quantity (clone size) and specificity in the repertoire which facilitates anticipatory responses. Each cell has an affinity response surface (affinity landscape) for each pathogen it has been (and may be) exposed to. The clonal densities are acquired from the antigenic virulence which is defined as the amplitude and severity of exposures (amount of learning opportunity offered) to the system through time. This model provides the pinnacle in environments for clonal selection at the cellular level and a complement for the quintessential clonal selection framework.

Exposure Models and Control

The environment and the system may be considered two interdependent and required components of which neither is meaningful in isolation. In modelling these two components, one may consider the variable (configurable) aspects of each of which may or may not be within the control of a given simulation. An intuitive example is to consider the environment (and therefore all aspects of antigen exposure) as outside of the scope of control, and all aspects of the system as inside of the scope of control. In this example the system must do all it can to cope with the specific properties of the exposures it is subjected to by the environment. The system must operate with finite resource, thus constraints are imposed on the control over aspects the system. This example may be extended further to consider that the evolution of antigen (pathogen) is influenced by the existence and evolution of the host system and that this relationship is reciprocal (co-evolution). Therefore, the system has influence (implicit influence via evolution) over aspects of antigen exposure over longer periods of time, and the antigen has influence (implicit influence via evolution) over aspects of the systems response over the same longer periods of time. This example may be generalised such that aspects of control are variable across both the environment and the system.

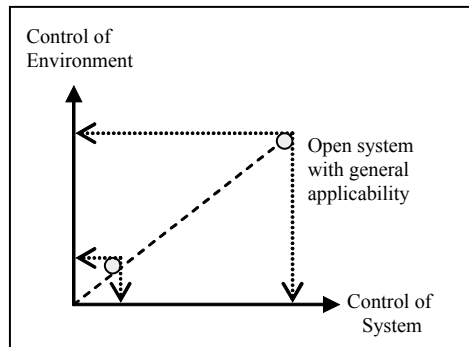


Figure 4.4: Depiction of the trade-off between control of the environment and control of the system.

In modelling *environment* and *system* and their *interactions*, control may be exhibited over all aspects of both sides of the interaction. It is important to consider the influence for control or lack there of (constraints on control) on a system and its environment as it provides insight into the mapping of domain properties. A constraint may be considered the prior selection of a configuration parameter of a system or environment. The fixing of a configuration parameter in the environment influences the suitability of other unconstrained (controlled) parameters in the interaction, and the influence is likely to be complementary (such that fixing an aspect of the environment strongly influences the suitability of related parameters of the system and vice versa). This point highlights the symmetry of constraint between *systems* and *environments*, and the need to specialise a given system for the constraints of a given environment in which it is to be deployed (mapped problem domain).

4.2.3 Colour-Space Domain

In his book detailing the Cognitive Theory of acquired immunity, Cohen used an example of cellular degeneracy in the retina as an analogy for the degeneracy of cell-bound receptors in the immune system [81]. Cohen described how individual cells in the retina have limited capability, responding to constrained wavelengths of light, although from these specific degenerate detectors higher-level information such as a coherent image can be perceived through aggregation. The Self-Organizing Map (SOM) is a competitive and unsupervised learning algorithm that is renowned for its feature extraction and dimensionality reduction capabilities (see Section 3.4.4). In his doctoral work, Honkela used a colour domain (red-green-blue or RGB), and a dataset of known colours in that domain [220] (page 14-18). The learning and ordering of the domain of RGB colours based on the dataset was used as an intuitive example of the topological preserving properties of the SOM algorithm. Although Honkela may not have been the first to employ this example domain to the SOM, it has become a canonical example domain.² Inspired by the simplicity and suitability of the colour domain to qualitatively (visually), and quantifiably (numerical measures) demonstrate the properties of the SOM, this section defines a *colour-space* domain from which a suite of test problem instances may be drawn to investigate clonal selection. Inspired by Cohen's analogy to the retina, colour-space is named for the relationship to the shape-space formalism, and is supported by proven use of the same principle for a related Competitive Learning paradigm.

Domain Definition

The domain does not provide a framework for investigating the so-called colour space theory, which refers to the study of colour models used in display and printer systems. Nor is the intention of the defined test domain for the study of colour theory, the investigation of colour segregation, colour recognition, palette compression, or colour clustering (or any other practical colour-based application domain). Given what the intentions are not, this section outlines a series of general goals for designing a test domain for adaptive systems, followed by the definition of a simple colour-space domain that meets these goals. The ambition is to contrive a domain from which trivial problem instances can be drawn to preliminary assess specific properties of adaptive models. These models may be engaged in processes that equate to pattern recognition and optimisation, and may involve the evaluation of broad characteristics such as learning, memory, and adaptation. See Table 4.1 for a list of the general design goals (requirements) for a problem domain suitable for this purpose.

The colour-space domain formalism meets the list of goals proposed in Table 4.1. The formalism is comprised of a series of terms to describe the characteristics of a colour space domain. The characteristics identified include the environment: the cardinality, the representation, and the mapping function (see Table 4.2 for their definition and summary).

The environment (E) is an n -dimensional hypercube (for visualisation purposes per-

²There are many applet and tutorial web sites on the Internet that use this example to demonstrate the SOM, the listing of which is not necessary.

G#	Goal
G1	The domain must provide a numerical and/or combinatorial basis that may be manipulated meaningfully by operators.
G2	The response surface must be correlated such that localised regions in the response surface have a gradient.
G3	The domain must be suitable for optimisation tasks.
G4	The domain must be suitable for pattern recognition tasks.
G5	The domain must be parameterised such that the relative complexity may be meaningfully adjusted.
G6	The domain must be easily and meaningfully visualised both online (dynamically) and off-line (end of run).
G7	The domain must be easy to understand, easy to implement, and easy to analyse.
G8	The properties of the domain (e.g. the response surface or visualisation) should facilitate rather than mask model behaviours being investigated.

Table 4.1: List of goals for a colour-space domain.

haps limited to 1-3 dimensions), where each dimension represents a different colour axis. The intention is that a distinct coordinate in a given environment (colour space) represent a distinct colour. For example, a monochrome (one-colour) environment would be implemented as a one-dimensional space (line) perhaps a gradient of white to black. An RGB colour space would be implemented as a three-dimensional (cube) environment. An environment is a volume of feasible discrete points, and the cardinality (C) defines the number of points in that volume. The granularity of each dimension in the environment also determines the number of colours, thus cardinality may be thought of as the palette of the environment. Further, the palette analogy may be exploited further in defining standard cardinalities, such as those common palettes that are used in computer display systems (for example, 8-bit colour (2^8) with a pallet of 256 colours). The coordinate-based representation (R) may be used directly by an adaptive system. Alternatively a symbolic or sub-symbolic representation may be used (such as bit-strings), which require a mapping function (M) to transform representation into a given environments coordinate system. The task of an adaptive system is called an intention (I), examples may include the optimisation of a randomly initialised system to a set of pre-selected points, or the adaptation of a system to a defined region within the environment.

One may measure the Euclidean distance between points in the colour space as a quantitative measure of difference. Distances measures may not be limited to Euclidean distance, for example Manhattan distance (and many others, see [261] pages 17-29) may be used in the colour space. Distance measures may be augmented with coordinate radii such that a point of interest in the space may represent a collective of points in its vicinity. Such regions may be defined as hyper-cubes or spheres, and may have a distance-based falloff such as linear, Gaussian or exponential. These distances and coordinate neighbourhoods are also qualitatively meaningful. Following a line (for example in a monochrome or RGB) in colour space shows a transition through intermediate colour coordinates (colours). These intermediate colours are quantifiably distinct (numerical coordinates), and qualita-

Term	Name	Summary
E	Environment	The search space (colour space) which defines the scope of feasible coordinates defined by its dimensionality ($1-n$), a boundary (limits of each dimension), and cardinality (C).
R	Representation	An encoding used and manipulated by adaptive models. It may be binary, integer, real, and may or may not directly represent a coordinate in the colour space (E).
M	Mapping Function	Converts a given representation (R) into a coordinate in the colour space (E). If the representation used matches the coordinate system of the environment, then the mapping function returns an unchanged coordinate.
C	Cardinality	The number of discrete points in the domain, and ideally (given display capabilities) the number of colours available for visualisation.
I	Intention	The goal of a system within the domain, a task or action it is to perform or achieve. This is a cover-all term for something to do in a defined colour-space.

Table 4.2: Summary of terms for the colour space domain formalism.

tively different (assuming the colour transition can be detected by the human eye). This correlation is also useful for visualisation: agents representing coordinates in the colour space may be implemented, which may permit qualitative measures and observations of system behaviour.

Illustrative Examples

This section proposes an *optimisation*, *pattern recognition*, and *classification* examples of Colour Space Problems that adaptive systems that may be employed or extended to investigate adaptive models.

- *Optimisation*: Involving the pre-selection of a coordinate in the space that is withheld from the adaptive system, and defining a cost function as a distance measure to the given coordinate. For example if a space was defined as an 8-bit one-dimensional domain (0-255), a goal coordinate of 127, and an Euclidean distance measure employed as a cost function, then the shape of the response surface would be a triangle with the apex at the selected coordinate.
- *Pattern Recognition*: The optimisation domain may be generalised with a set of goal coordinates (Colour Space Patterns) to which a given system is exposed. The intention of the system for this problem is to optimise (optimally match) the set of patterns, where the order of exposure of the patterns to the system is also withheld from the system.
- *Classification*: A classification task may be defined as an extension of the pattern recognition problem with the addition of categorical information. A geometry can be defined within the colour space and class labels can be assigned to the resulting concave regions. In addition to distance-based feedback, the system may be provided

with corrective categorical based feedback, relaxing the optimisation of the goal coordinates to that of accuracy of predicting pattern classes.

Criticisms and Limitations

This section addresses the important consideration's of the limitations of the formulated problem domain, as follows:

1. *Triviality*: The domain and resultant problem definitions are primitive. The problem instances are likely easily solved rapidly by standard deterministic techniques, and even by naïve approaches such as enumeration and random search.
2. *Transferability*: The results achieved on one or a set of derived problem instances are very likely to be not transferable to problem domains of interest (difficult, real-world). Conclusions regarding model performance are limited to the instances on which they are tested, and perhaps related trivial domains.
3. *Difficulty*: Although superficiality they appear trivial, there is no innate sense of the absolute or relative difficulty of the derived problem definitions.
4. *Visibility*: For high-cardinality domains (such as those at or above 24-bit or 32-bit RGB) the human eye cannot detect colour differences between coordinates within close proximity.
5. *Novelty*: The colour space domain and derived problem instances are not novel, they are generalisations and simplifications of existent test domains and benchmark problems.

It is prudent not to rebut these criticisms, but to acknowledge them as limitations (designed or otherwise) of the domain and framework that it provides. The domain and resultant problem instances are intended to be trivial (easy to solve) such that the attention remain on the model under study. The performance of a given model is not intended to be transferable, what is expected to be transferable are the general functional behaviours that the domain assists in isolating and assessing. The relative difficulty of problem instances may be determined by baseline strategies. Comparable relative and absolute difficulty may be defined using the tools of probability, statistical mechanics, and additional (suitable) mathematical apparatus. The visual distinctiveness of close-proximity coordinates in high-cardinality spaces may be exaggerated where appropriate. For example, a dynamically normalised colour scale may be allocated to coordinates of interest. Finally, novelty is not claimed, rather the domain is a *consistent reformulation* of classical optimisation and pattern recognition benchmark domains.

4.3 Realised Cellular Paradigm

This section provides a realisation of the concerns of the abstract *Cellular Clonal Selection Paradigm* presented in Section 4.2. This includes general definitions of a standardised

problem domain in the Antigenic Exposure Problem as a specialisation in colour space, and the Cellular Clonal Selection Algorithm that provides a basis for investigation into elaborated cell interaction schemes. A series of cell-based measures are presented for assessing cellular algorithms on instances of Antigen Exposure Problems that provide quantitative indicators regarding the concerns of information acquisition and application. Collectively, this section provides a basis for the implementation and exploratory empirical investigation into the Cellular Clonal Selection Paradigm, as well as a bridge between the biology (Section 3.2) existing approaches (Section 3.3), and the abstract clonal selection and antigenic exposure models (Section 4.2).

4.3.1 Antigenic Exposures

This section considers the realisation of the specialisations of the exposure paradigm outlined in Section 4.2.2 as a basis for empirical investigation and ultimate application of cellular models to problem domains. This realisation includes the definition of general antigenic exposure problem with exposure regimes, and a specialisation to the colour space domain.

Antigenic Exposure Problem

For the purposes of discussing exposure, a cellular clonal selection algorithm may be reduced to its information content, distinct from the specific clonal selection processes operating upon that information. Specifically, a cellular algorithm may be considered a population or repertoire (T , that stands for ‘tissue’ which will be made apparent in proceeding chapter) of discrete cells (C), as follows: $T = \{C_1, C_2, C_3 \dots, C_n\}$. Therefore an *Antigenic Exposure Problem (AEP)* is defined as the discrete interaction with a T , with a set (I , which stands for ‘infection’) of antigen (A) at the same level of abstraction, as follows $I = \{A_1, A_2, A_3 \dots, A_n\}$. An antigen is comprised of sub-antigenic information referred to as determinants (D), such that $A = \{D_1, D_2, D_3 \dots, D_n\}$.

Algorithm 4.1: Antigenic Exposure Problem (AEP).

```

Input:  $T$ ,  $N_{determinants}$   $N_{antigen}$ 
Output:  $T_{rs}$ 
1  $l \leftarrow 0$ ;
2 for  $i \leftarrow 0$  to  $N_{antigen}$  do
3    $A_i \leftarrow \text{CreateRandomAntigen}(N_{determinants})$ ;
4    $l \leftarrow A_i$ ;
5 end
6  $T_{rs} \leftarrow 0$ ;
7 while  $\neg \text{StopCondition}()$  do
8    $T_{rs} \leftarrow \text{Exposure}(l, T)$ ;
9 end
10 return  $T_{rs}$ ;

```

Algorithm 4.1 defines a simple antigenic exposure problem, with symmetrical information content for interaction with a given cellular system, where $N_{determinants}$ and $N_{antigen}$

define the number of determinants per antigen and total antigen within the scope of the AEP. The algorithm clearly shows the AEP's in control of the interactions with a given T , where discrete interactions are mediated via an $Exposure(I, T)$ operation. The algorithm definition also highlights that from an exposure perspective, the AEP is only concerned with a systems aggregate responses to exposures. The tissue result set denoted as T_{rs} may be considered to represent a cellular algorithm's (T) capability to address a given antigen exposure problem (I).

Cellular Exposure Regimes

The defined antigenic exposure problem delegates the specifics of algorithm-problem interaction to the $Exposure(I, T)$ operation, the behaviour of which may be referred to as a *Cellular Exposure Regime (CER)*. For example, the information content of the problem may be exposed to a system (1) consistently in the same order, (2) in a randomised order, or (3) in some biased manner resulting in an asymmetric perspective of such information content. Referring to such interactions as an exposure regimes also highlights the important point that the cellular algorithm is considered a passive information processing system that responds to the active exposure of antigenic stimuli. Therefore, depending on the specific problem domain, the concerns of exposure, re-exposure, and multiple exposures (from Section 4.2.2) may or may not be within the scope of control of the system. Algorithm 4.2 provides an example of a regular and consistent exposure regime $Exposure(I, T)$ operation where a system is exposed and must respond to the scope of the AEP each I exposure called an epoch. The example also delegates the responsibility of specific cell interaction (selection and response) to the cellular clonal selection algorithm in the $Exposure(A, T)$ operation. This operation encapsulates the scope of concerns of algorithms in the cellular paradigm.

Algorithm 4.2: Exposure Function (CER) for the Antigenic Exposure Problem.

Input: I, T
Output: T_{rs}

- 1 $T_{rs} \leftarrow 0;$
- 2 **foreach** $A_i \in I$ **do**
- 3 $C_i \leftarrow Exposure(A_i, T);$
- 4 $T_{rs} \leftarrow C_i;$
- 5 **end**
- 6 **return** $T_{rs};$

The cellular exposure regime represents the uncertainty in the information processing faced by a cellular algorithm, including but not limited to (1) the number of antigen ($N_{antigen}$), (2) the number of determinants for a given antigen ($N_{determinants}$), the diversity of the information in the antigenic environment (I), and the ordering of the discrete information exposures defined in the the exposure regime itself (CER). Finally, discrete cellular interactions with antigen and determinants may not yield specific information, rather a general indication. For example from an abstract point of view an $Exposure(A, C)$ or $Exposure(D, C)$ may result in a relative indication of fit or suitability (affinity or

avidity) rather than absolute cost. This is motivated by the pre-committed, iterative, and trial-based (selectionist) strategy reduced from the clonal selection theory.

Antigen Colour Space Problem

The antigenic exposure problem may be specialised in the colour space domain as a pattern recognition problem called the *Antigen Colour Space Problem (ACSP)*. The problem is defined as a set of Colour Space Patterns (CSP) each of which is defined as a colour comprised of three 64-bit colour components. The set of patterns are generated as a set of random binary strings ($A = \{0,1\}^L$, where $L = (3 \times 64)$) and decoded using the Gray Code method for each colour component (Equation 4.1, where \oplus denotes addition modulo two) to a three dimensional colour vector $A = [0,1]^3$. Cellular solutions (C) to each antigen colour space pattern (A) are 192-bit binary strings that decoded to 3-dimensional colour vectors using Gray Code. The affinity (avidity) function for a decoded cell, the $Exposure(A, C)$ is defined as the Euclidean distance between the two vectors (Equation 4.2).

$$GrayCode(A) = \frac{1}{2^{64} - 1} \left(\sum_{j=0}^{64-1} \left(\bigoplus_{k=1}^{64-j} A((i-1)64 + k) \right) 2^j \right) \quad (4.1)$$

$$EuclideanDistance(A, C) = \sqrt{\sum_{i=1}^n (A_i - C_i)^2} \quad (4.2)$$

This mapping of the colour space problem onto the antigenic exposure problem highlights the important point of the minimal cases of the problem, specifically: (1) minimum determinant, and (2) minimum antigen. From an abstract perspective, a cell matches onto one determinant, from an antigen of multiple determinants. Therefore, from a system perspective a one-to-one relationship exists between a given cell and its cognate antigen mediated via the specific sub-feature (antigenic determinant). A *Minimum Determinant Antigenic Exposure Problem (MDAEP)* may be defined where each antigen is comprised of a single determinant ($N_{determinants} = 1, \forall A \in I$). This case forces such a one-to-one relationship at the cost of the exclusion of the cross-reactivity of cells. In the same manner, the number of antigen in the set may be minimised ($N_{antigen} = 1$) to define a *Minimum Antigenic Exposure Problem (MAEP)* where a given system T is assessed based on its capability with a single antigen with a single determinant. This minimal case (single or multiple exposures of a single determinant), provides a minimal context in which to consider fundamental behaviours of cellular algorithms and their extensions.

4.3.2 Cellular Clonal Selection

This section considers the realisation of specialisations of quintessential clonal selection outlined in Section 4.2.1 as a basis for empirical investigation and ultimate application of cellular models to problem domains. This realisation includes the definition of a general cellular clonal selection algorithm as a homologue to CLONALG, and a replacement based extension.

Cellular Clonal Selection Algorithm

A *Cellular Clonal Selection Algorithm (CCSA)* is defined as the application of the computational properties of the clonal selection theory constrained by the concerns of the theory at the cellular-level. These constraints include but are not limited to (1) the interaction with antigen via their determinants, (2) the embodiment of information in a self-contained repertoire of discrete cells, and (3) the cellular response to active antigenic exposures. The extent of the taxonomy of clonal selection algorithms reviewed in Section 3.3 may be considered to reside in the cellular level (within the scope of the Cellular Clonal Selection Paradigm). This section considers an abstraction of cellular clonal selection algorithms in the context the concerns considered in quintessential Clonal Selection in Section 4.2.1. The Antigen Exposure Problem defined a cellular system in terms of its information content, where the problem is the active concern and the cellular system is passive until exposed with stimuli to which it reflexively responds. The scope of the concerns of cellular clonal selection algorithms are the management of information within the repertoire, and in particular the management of information under the discrete exposures of an AEP.

Algorithm 4.3: Initialisation Function for the Cellular Clonal Selection.

Input: N_{cells}

Output: T

```
1  $T \leftarrow 0$ ;  
2 for  $i \leftarrow 0$  to  $N_{cells}$  do  
3    $C_i \leftarrow \text{CreateCell}()$ ;  
4    $T \leftarrow C_i$ ;  
5 end  
6 return  $T$ ;
```

Algorithm 4.3 defines an initialisation operation for a the Cellular Clonal Selection Algorithm, where the *CreateCell()* may be specialised to create random cells with a representation appropriate for the specific problem domain (192-bit binary strings in the case of the ACSP). Algorithm 4.4 defines the *Exposure(A, T)* operation for the CCSA that provides a generalisation of the CLONALG, BCA, AIRS, and IA family of the reviewed clonal selection taxonomy. Specifically, the algorithm responds to antigen exposures by assessing the extent of the repertoire and selecting a subset ($N_{selected}$) to comprise a response. Each cell in the selected set creates a set of clones (N_{clones}) with mutations ($P_{mutation}$) that are assessed against the antigen and compete with the progenitor of the selected set for a position in the repertoire. This generalisation which strongly resembles CLONALG, implicitly manages the memory cell set and explicitly returns a best matching cell (*GetBestMatchingCell(T)*) for each exposure as the most suitable response within the scope of the repertoire. Importantly, the CCSA provides the flexibility to assume a variety of configurations, specialisations (such as the CSA taxonomy), as well as the extensions considered in this chapter.

The CCSA may be reduced to a minimal configuration to provide a complement to the Minimal Antigenic Exposure Problem called the *Minimal Cellular Clonal Selection Algorithm (MCCSA)*. This minimal configuration under the MAEP provides a realisation

Algorithm 4.4: Exposure Function for the Cellular Clonal Selection.

Input: $A, T, N_{selected}, N_{clones}, P_{mutation}$
Output: T_{rs}

```
1  $T_{rs} \leftarrow 0$ ;  
2 foreach  $C_i \in T$  do  
3   | Expose( $A, C_i$ );  
4 end  
5  $T_{selected} \leftarrow \text{Select}(T, N_{selected})$ ;  
6 foreach  $C_i \in T_{selected}$  do  
7   |  $T_{clones} \leftarrow 0$ ;  
8   | for  $i \leftarrow 0$  to  $N_{clones}$  do  
9     |  $C_i \leftarrow \text{Clone}(C_i)$ ;  
10    |  $C_i \leftarrow \text{Mutate}(C_i, P_{mutation})$ ;  
11    |  $T_{clones} \leftarrow C_i$ ;  
12  | end  
13  | Expose( $A, T_{clones}$ );  
14  |  $T'_{clones} \leftarrow \text{Select}(T_{clones}, N_{selected})$ ;  
15  | for  $i \leftarrow 0$  to  $N_{selected}$  do  
16    |  $C_{selected} \leftarrow T_{selected}$ ;  
17    |  $C_{clone} \leftarrow T'_{clones}$ ;  
18    | if  $C_{clone}.score \leq C_{selected}.score$  then  
19      |  $T.\text{Remove}(C_{selected})$ ;  
20      |  $T.\text{Add}(C_{clone})$ ;  
21    | end  
22  | end  
23 end  
24  $T_{rs} \leftarrow \text{GetBestMatchingCell}(T)$ ;  
25 return  $T_{rs}$ 
```

of of the ES(1 + 1) algorithm and connection with mutation based hill climbers made in Section 3.4. For example, ES(1 + 1) would be defined by the configuration: $N_{cells} = 1$, $N_{selected} = 1$, $N_{clones} = 1$, and $P_{mutation} = \frac{1}{L}$.

Replacement Cellular Clonal Selection Algorithm

An expected effect with the Cellular Clonal Selection Algorithm is that it results in the iterative adaptation of specialised and independent cells. This is expected (1) because it was observed in CLONALG with regard to a niching-like effect, and (2) because CCSA and CLONAL resemble a parallel mutation based hill climber. The *Replacement Cellular Clonal Selection Algorithm (RCCSA)* defined in Algorithm 4.5 is an extension of the CCSA where all clones created for an exposures are aggregated into a clonal set, which then competes with the cells in the repertoire on a per-cell basis for a limited position in the repertoire. This ‘clonal set-repertoire competition’ is intended to provide more opportunity for per-exposure resource allocation, by allowing clones to compete and potentially displace repertoire members that may or may not be the clonal progenitor (progenitor of clones), clonal siblings (same progenitor), or exposure clonal siblings (same exposure). This competition is promoted firstly through a cell-to-cell distance assessment that simu-

lates the competition the cells face during selection by antigen where similar cells compete for the same antigenic patterns, and in the affinity comparison that selects those cells that respond better to the specific antigen to which the system has been exposed. This competition mechanism decouples intra-repertoire competition for survival in the integration of clones from the antigenic-based selection that resulted in the creation of the clones. Such decoupling may be exploited in the integration of cells with varied origins.

Algorithm 4.5: Exposure Function for Replacement Cellular Clonal Selection.

Input: $A, T, N_{selected}, N_{clones}, P_{mutation}$
Output: T_{rs}

```

1  $T_{rs} \leftarrow 0;$ 
2 foreach  $C_i \in T$  do
3   |  $Expose(A, C_i);$ 
4 end
5  $T_{selected} \leftarrow Select(T, N_{selected});$ 
6  $T_{clones} \leftarrow 0;$ 
7 foreach  $C_i \in T_{selected}$  do
8   | for  $i \leftarrow 0$  to  $N_{clones}$  do
9     |    $Cl_i \leftarrow Clone(C_i);$ 
10    |    $Cl_i \leftarrow Mutate(Cl_i, P_{mutation});$ 
11    |    $T_{clones} \leftarrow Cl_i;$ 
12    | end
13 end
14  $Expose(A, T_{clones});$ 
15 for  $C_{clone}$  to  $N_{clones}$  do
16   |  $C_{selected} \leftarrow 0;$ 
17   | foreach  $C_i \in T$  do
18     |   if  $Distance(C_{clone}, C_i) < Distance(C_{clone}, C_{selected})$  then
19       |   |  $C_{selected} \leftarrow C_i;$ 
20     |   end
21   | end
22   | if  $C_{clone}.score \leq C_{selected}.score$  then
23     |   |  $T.Remove(C_{selected});$ 
24     |   |  $T.Add(C_{clone});$ 
25   | end
26 end
27  $T_{rs} \leftarrow GetBestMatchingCell(T);$ 
28 return  $T_{rs};$ 

```

4.3.3 Empirical Assessment

This section defines a series of general empirical measures that provide instantaneous information regarding a given CCSA on colour space specialisations of the Antigenic Exposure Problem. An important consideration of these measures in the context of investigation via exploratory experimentation is not their absolute value, but rather their relative change in value with changes to the systems being investigated. Two general classes of measures are considered, (1) error measures that assess the capability of a system in the context of

a given problem instance, and (2) diversity measures that consider the general state of the information content within a given system.

Average Cell Error

The error of a cell may be assessed as its affinity for a given antigen, and may be assessed using the problem domains specific $Exposure(A, C)$ operation, in this case Euclidean Distance (defined in Equation 4.2). Although not used directly, $Cell Error (CE)$ (defined in Equation 4.3) provides an error-centric building block for addition error assessments.

$$CellError(A, C) = EuclideanDistance(A, C) \quad (4.3)$$

An important application of Cell Error is its calculation for each cell in in the result set (T_{rs}) returned from a cellular algorithms $Exposure(I, T)$ operation. The Average Cell Error (ACE) defined in Equation 4.4 is a problem specific (ACSP) measure for assessing the general capability of a cellular algorithm on an antigenic exposure problem.

$$AverageCellError(I, T_{rs}) = \frac{1}{I_n} \sum_{i=1}^{T_{rsn}} CellError(A_i, C_i) \quad (4.4)$$

Average Cell Diversity

A diversity measure of the cellular repertoire provides some indication of the variation in the discrete information stored within the repertoire. Low diversity reflects a homogeneous repertoire, whereas high diversity reflects a heterogeneous set of cells. Given the selected bit string basis of the CCSA, diversity between two cells may be measured as the hamming distance between the cells (see Equation 4.6, where Hamming Distance is defined in Equation 4.5, and \oplus denotes exclusive-or (XOR)).

$$HammingDistance(X, Y) = \sum_{i=1}^L (X_i \oplus Y_i) \quad (4.5)$$

$$CellDiversity(C1, C2) = HammingDistance(C1, C2) \quad (4.6)$$

Cell Diversity (CD) can be calculated for each cell in the repertoire against all other cells in the repertoire to provide an indication the average number of bit differences of a given cell to the rest of the cells in the repertoires, called the Average Cell Diversity (ACD) defined in Equation 4.7.

$$AverageCellDiversity(T) = \frac{1}{T_n} \sum_{i=1}^{T_n} \left(\frac{1}{T_n} \sum_{j=1}^{T_n} CellDiversity(C_i, C_j) \right) \quad (4.7)$$

4.3.4 Trends and Behaviours

This section considers the expected emergent effects and behavioural trends of cellular-based clonal selection algorithms as defined in the specific realisations.

Primitive Behaviour Trends

The foundational expectation is that the realised Cellular and Replacement Clonal Selection Algorithm behave much like CLONALG and AIRS. Specifically, that the algorithm will acquire information via selection and adaptation. This section outlines the specific foundational behaviour trends of the algorithms in the context of the pressures and principles of the clonal selection strategy discussed in Section 4.2.1.

1. *Repertoire Size*: The repertoire size must be sufficiently large to ensure that selection of cells and integration of clones do not result in competition between cells specialised for different antigen. Such competition by an insufficiently sized repertoire is expected to result in unstable behaviour.
2. *Selection Size*: The number of clones selected per antigen exposure defines the amount of resources allocated to each antigen under fixed clonal integration. An increase in resource allocation via an increase in selection is expected to result in increased specialisation via the promotion of ‘concurrent redundant perspectives’ of a given antigen.
3. *Cloning Size*: The number of clones created per selected cell reflects the amount of computational effort expended on re-sampling modifications on a ‘known-to-be-good’ structure. Therefore an increase in cloning and acceptance of limited high-affinity clones is expected to result in an increased specialisation toward exposed antigen.
4. *Replacement Constraints*: Replacement constrained by a clonal progenitor is expected to result in a constrained and specialised hill climbing-like behaviour (MHCA). The relaxation of this constraint such that clonal progeny compete directly with the repertoire is expected to enhance the so-called ‘concurrent redundant perspectives’ to ensure that the maintained multiple perspectives are the most competitive, likely resulting in improved repertoire specialisation.

Contextual Adaptation Expectations

This section outlines the specific expectations regarding contextual adaptation both with regard to the localisation of activated and responding cells.

1. *Localised Activation*: Strong and non-overlapping (specialised for the same antigen) competition by selection is expected to sufficiently localise polyclonal activation and resultant adaptation of the clonal selection strategy (oligoclonal activation).
2. *Localised Response*: A top-down centralised aggregation of degenerate cells is expected to sufficiently localise a polyclonal response of selected cells under clonal selection.

4.3.5 Paradigm Agenda

This section outlines the research agenda regarding the cellular clonal selection paradigm in the remainder of this chapter. Specifically, the agenda is constrained to the investigation of the expectations made regarding primitive behavioural trends of the cellular clonal selection algorithms and the expectations regarding contextual adaptation of cells.

1. Investigate the expectations regarding the foundational clonal selection behaviour.
2. Investigate the expectations regarding contextual adaptation.
3. Investigate additional mechanism to influence the localisation of adaptation in the clonal selection strategy.

4.4 Cellular Clonal Selection

This section considers three empirical studies of the cellular clonal selection algorithm. Specifically, Section 4.4.1 considers the selection and cloning sizes in the Cellular Clonal Selection Algorithm (CCSA), Section 4.4.2 considers repertoire resource allocation in CCSA and the Replacement Cellular Clonal Selection Algorithm, and finally Section 4.4.3 considers contextual adaptation of symbolic and sub-symbolic degenerate cells under the clonal selection strategy.

4.4.1 Cellular Empirical Study

Aim

The aim of this empirical study was to investigate the CCSA as a viable realisation of cellular clonal selection and to assess the expected behaviour of this foundational algorithm with regard to the number of cells selected and the number of clones created under varied antigenic environments. Toward this end, the study had the following goals:

1. Assess the repertoire composition and capability of CCSA with small and large selection and clonal sets.
2. Assess such behaviour under small and large antigenic environments.

Method

Problems The colour space specialisation of the Antigenic Exposure Problem (AEP) was used called the Antigen Colour Space Problem (ACSP). Two variations of the problem were employed ACSP-1 with a single CSP ($N_{antigen} = 1$) also referred to as the Minimal Antigen Colour Space Problem, and ACSP-10 with 10 CSP ($N_{antigen} = 10$)³. Both prob-

³It is important to note that 10 patterns does not represent the expected extent of the approaches capability, rather relative increase over the minimal problem case of one pattern. The scalability of the pattern recognition capability of the investigated approaches is outside the scope of this thesis, although the approach is expected to be effected by the curse of dimensionality faced with similar pattern recognition systems (see Section 8.4 for further discussion).

lem instances used the minimal determinant configuration where the scope of the determinant was defined by the extent of each Colour Space Pattern ($N_{determinants} = 1$). Colour Space Patterns were randomly generated at the start of each algorithm-problem run. In the case of ACSP-10, CSP were randomly generated and accepted with the constraint that pattern binary strings must have a minimum of a 64-bit Hamming Distance from all other generated patterns for the run. The simple consistent Cellular Exposure Regime (CER) defined in Algorithm 4.2 was used to manage the interaction of the problem with the cellular algorithm.

Algorithms The study considered four different configurations of the Cellular Clonal Selection Algorithm (defined in Algorithm 4.4). The CCSA configurations for the ACSP-1 and ACSP-10 are defined in Table 4.3 and Table 4.4 respectively. The mutation rate ($P_{mutation}$) was fixed at $\frac{1}{L} = \frac{1}{192} \approx 0.005$ for all algorithm configurations. In addition to the CCSA parameters, the tables also list the derived values for the number of clones created per antigen exposure ($Clones (A)$), the number of positions in the repertoire under competition per antigen exposure ($Positions$), and the total number of clones created per infection exposure or epoch ($Clones (I)$). For all configurations, the number of cells was set to ensure that selection and cloning did not result in explicit competition for resources between cells specialised for different antigen.

<i>CCSA</i>	N_{cells}	$N_{selected}$	N_{clones}	$Clones (A)$	$Positions$	$Clones (I)$
<i>1+1</i>	1	1	1	1	1	10
<i>1+N</i>	1	1	10	10	1	100
<i>N+1</i>	10	10	1	10	10	100
<i>N+N</i>	10	2	5	10	2	100

Table 4.3: Summary of the assessed configuration of CCSA for ACSP-1.

<i>CCSA</i>	N_{cells}	$N_{selected}$	N_{clones}	$Clones (A)$	$Positions$	$Clones (I)$
<i>1+1</i>	1	1	1	1	1	10
<i>1+N</i>	1	1	10	10	1	100
<i>N+1</i>	100	10	1	10	10	100
<i>N+N</i>	100	2	5	10	2	100

Table 4.4: Summary of the assessed configuration of CCSA for ACSP-10.

Experiment Each algorithm used the *Maximum Epochs Stop Condition (MESOC)* defined in Equation 4.8 with $MaxEpochs = 1000$, where an epoch was defined as the complete exposure of an I to the T . The two cellular specific measures defined in Section 4.3.3 were collected from the state of the system after the triggering of the stop condition. These measures included the Average Cell Error (ACE) and the Average Cell Diversity (ACD). Each algorithm and problem received a new and different random number seed each run. Algorithm-Problem combinations were repeated 30 times.

$$StopCondition(Epoch_i) = (Epoch_i \geq MaxEpochs) \quad (4.8)$$

Results

Table 4.5 provides a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. The non-parametric Mann-Whitney U statistical test was calculated pair-wise for all algorithms in each problem group. Results are summarised as statistically significant if the null hypothesis that any two given populations in the group are drawn from the same distribution ($H_0 : \mu_0 = \mu_1$). Specifically, H_0 is rejected if the calculated p -value $< \alpha = 0.05$, meaning the results are statistically significant at the 5% level. Figures 4.5a and 4.5b show the ACD and ACE for all algorithms on ACSP-1 respectively, and figures 4.5c and 4.5d show the ACD and ACE for all algorithms on ACSP-10 respectively Figure 4.6 provides an example plot of the T_{rs} returned from the CCSAN+N algorithm on the ACSP-1 and ACSP-10 problems at after the triggering of the stop condition (same configurations as experiment, random seed of 5 and 1 for the problem and algorithm respectively).

Problem	System	ACD		ACE	
		\bar{x}	σ	\bar{x}	σ
<i>ACSP</i>	<i>CCSA</i>				
ACSP-1	CCSA1-1	0	0	0.165	0.074
ACSP-1	CCSA1-N	0	0	0.004	0.007
ACSP-1	CCSAN-1	85.267	1.104	0.041	0.027
ACSP-1	CCSAN-N	86.233	1.055	0.014	0.027
<i>Significant</i>		True ^a		True	
ACSP-10	CCSA1-1	86.363	0.808	0.161	0.037
ACSP-10	CCSA1-N	86.053	1.016	0.073	0.031
ACSP-10	CCSAN-1	95.006	0.103	0.053	0.013
ACSP-10	CCSAN-N	95.035	0.094	0.01	0.006
<i>Significant</i>		True		True	

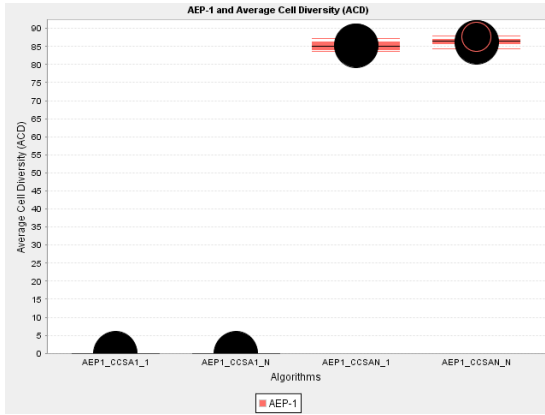
^aFalse for CCSA1-1 and CCSA1-N

Table 4.5: Summary of results for CCSA on ACSP-1 and ACSP-10.

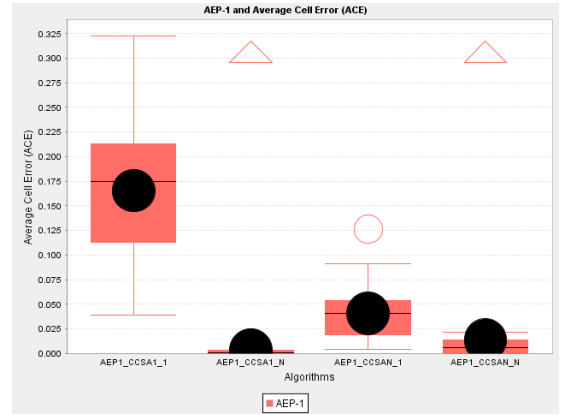
Analysis

This section provides an analysis of the results reported in the previous section in the context of the goals of the empirical study. Specifically, the analysis is broken down into the selection and clonal set size trends for small and large antigenic environments.

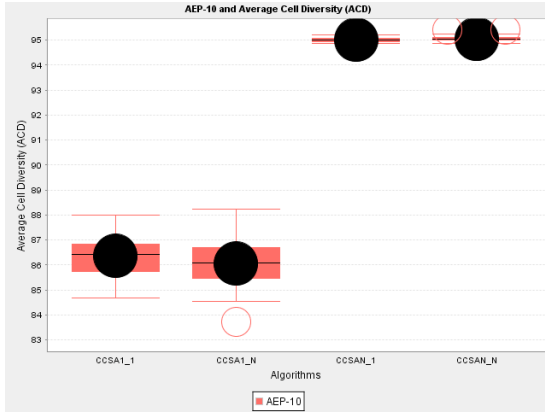
Selection and Clonal Trends (ACSP-1) This section considers the selection and clonal trends of the CCSA configurations on ACSP-1. The increase in the number of selected cells from 1 to N resulted in a large increase in diversity as expected given the required shift in increase in the number of cells in the repertoire to support the increased selection size. The increase in the number of cells selected resulted in a decrease in error with a single clone and an increase when multiple clones were created. The increase in the number of clones created for selected cells from 1 to N resulted in very minor differences in the repertoire diversity, although did result in a large decrease in the average cell error. The results confirmed the expectation that in the optimisation of a repertoire for a single



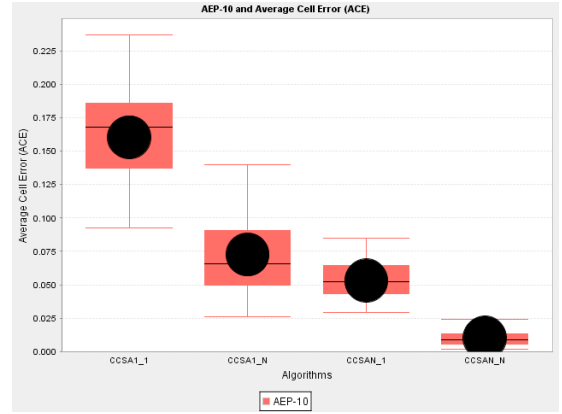
(a) Average Cell Diversity (ACD) on ACSP-1.



(b) Average Cell Error (ACE) on ACSP-1.

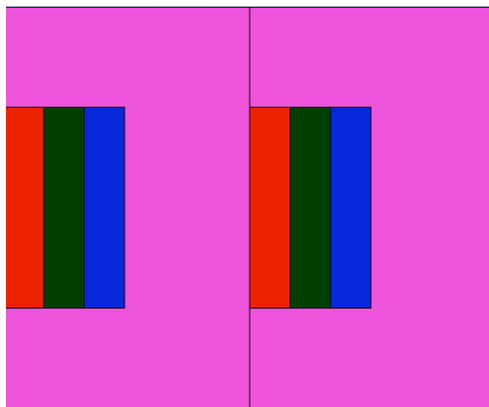


(c) Average Cell Diversity (ACD) on ACSP-10.

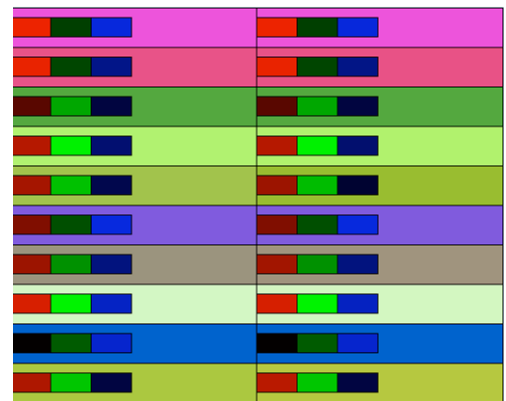


(d) Average Cell Error (ACE) on ACSP-10.

Figure 4.5: Box-and-whisker plot's from the Cellular Empirical Study.



(a) CCSA(N+N) on ACSP-1.



(b) CCSA(N+N) on ACSP-10.

Figure 4.6: Example plots of T_{rs} from the CCSA(N+N) at the end of the run on left section on both plots, for both ACSP-1 and ACSP-10, where the problem solution is represented on the right of each plot, and colour components within each CSP.

antigen, the creation of more clones results in an improved system capability given that each clone represents a new trial and potential improvement over an already relatively good structure.

Selection and Clonal Trends (ACSP-10) This section considers the selection and clonal trends of the CCSA configurations on ACSP-10. The increase in the number of selected cells resulted in a decrease in the system error with a small repertoire, although in this case the trend held for both small and large repertoire sizes. The increase in the number of clones resulted in only very minor changes to the average cell diversity, although resulted in a consistent decrease in the the average cell error. The increase in the size of the antigenic environment did not disrupt the general trend of increase number of clones resulting in increase system capability. Unlike the single antigen case, the increase in the number of selected cells also consistently resulted in a decrease in average cell error.

Conclusions

This section summarises the findings of the empirical study into the Cellular Clonal Selection Algorithm in terms of the primitives that were the focus of the study and the expectations that motivated the study.

1. The Cellular Clonal Selection Algorithm successfully addressed 1 and 10 antigen problems, demonstrating itself as a viable realisation of a foundational clonal selection adaptive strategy under the circumstances considered.
2. The increase in the number of selected cells (allocated resources) results in improved system capability with small cloning on the single antigen problem, although is consistently beneficial on the larger antigenic problem.
3. The increase in the number of clones (trials) results in an improved system capability across the tested antigenic environments.

4.4.2 Replacement Empirical Study

Aim

The aim of this empirical study was to investigate the properties of the CCSA and the RCCSA as a viable realisation of cellular clonal selection that promote improved utilisation of repertoire resources. Improved utilisation will effect the composition and may effect the capabilities of a repertoire, where the promotion and maintenance of redundant perspectives is expected to provide benefits in terms of fault tolerance. Toward this end, the study had the following goals:

1. Assess varied clonal integration mechanisms with CCSA that promote concurrent redundant perspectives of each antigen.
2. Assess varied clonal integration mechanisms with RCCSA that promote concurrent perspectives and compare to CCSA.

Method

Problems This study used the ACSP-10 problem used for the CCSA empirical study in Section 4.4.1.

Algorithms This section considers varied configurations of the CCSA and the RCCSA. The configuration parameters of each algorithm type are presented in Table 4.6. The chosen configurations of the repertoire, number of selected cells and number of clones facilitate the allocation of a maximum of six repertoire position per antigen. The CCSA configurations included the CCSA(N+N) from CCSA empirical study in Section 4.4.1 as well as two extensions. CCSA(N+N)-G aggregates all clones into a single group which collectively compete with the $N_{selected}$ selected cells from the repertoire for their position. CCSA(N+N)-GS also aggregates all clones into a single clonal group, each of which competes with the same number of the best matching cells as there are in the clonal set drawn from the repertoire. Two variations of the RCCSA algorithm (defined in Algorithm 4.5) were assessed, both of which used Hamming Distance (Equation 4.5) between cells in the replacement mechanism. The RCCSA-H configuration allowed clones to select the most similar cells in the repertoire to compete with for a position, whereas the RCCSA-H-ES configuration restricted selection to those cells in the repertoire that were not clonal siblings (hamming similarity excluding siblings).

<i>CSA</i>	N_{cells}	$N_{selected}$	N_{clones}	<i>Clones (A)</i>	<i>Positions</i>	<i>Clones (I)</i>
<i>N+N</i>	60	2	3	6	2	60
<i>N+N-G</i>	60	2	3	6	2	60
<i>N+N-GS</i>	60	2	3	6	6	60
<i>RCCSA-H</i>	60	2	3	6	6	60
<i>RCCSA-H-ES</i>	60	2	3	6	6	60

Table 4.6: Summary of the assessed configuration for the CCSA-RCCSA empirical study.

Experiment This study used the same experimental configuration including stop conditions, measures, as were used for the CCSA empirical study in Section 4.4.1. An additional measure was used to provide an assessment of the number of cells (C) in the repertoire (T) specialised to each antigen (A) in the scope of the problem (I). This measure is called the *Average Best Matching Cells Per Antigen (ABMCPA)* defined in Equation 4.9 that assessed and located the set of best matching cells for each antigen (Algorithm 4.6) and averaged the number of cells returned for each antigen.

$$AverageBMCPPerAntigen(I, T) = \frac{1}{I_n} \sum_{i=1}^{I_n} ABMCS(A_i, T)_n \quad (4.9)$$

Results

Table 4.7 provides a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. The non-parametric

Algorithm 4.6: Best Matching Cells (BMC's) for a given Antigen.

```

Input:  $\mathbb{T}$ ,  $A$ 
Output:  $T_{BMC}$ 
1  $T_{BMC} \leftarrow 0$ ;
2 for  $C_i \in \mathbb{T}$  do
3   Exposure( $A$ ,  $C_i$ );
4   if  $T_{BMC_n} > 0$  then
5      $C' \leftarrow T_{BMC_i}$ ;
6     if  $C_i.affinity < C'.affinity$  then
7        $T_{BMC} \leftarrow 0$ ;
8        $T_{BMC} \leftarrow C_i$ ;
9     else if  $C_i.affinity \equiv C'.affinity$  then
10       $T_{BMC} \leftarrow C_i$ ;
11    end
12  else
13     $T_{BMC} \leftarrow C_i$ ;
14  end
15 end
16 return  $T_{BMC}$ ;

```

Mann-Whitney U statistical test was calculated pair-wise for all algorithms. Figures 4.7a, 4.7b, and 4.7c show the ACD, ACE, and ABMCPA on all algorithms respectively. Figure 4.8 provides example plots of the state of the repertoire (left of each plot) compared to the problem domain (right of each plot) from each of the five algorithms (same configuration, random seeds of 1 and 5 for the algorithm and problem respectively).

Problem	System	ACD		ACE		ABMCPA	
		\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
ACSP	CSA						
ACSP-10	CCSA(N+N)	94.356	0.171	0.027	0.012	1	0
ACSP-10	CCSA(N+N)-G	93.995	0.24	0.023	0.014	1.82	0.124
ACSP-10	CCSA(N+N)-GS	88.456	0.937	0.062	0.015	4.907	0.528
ACSP-10	RCCSA-H	94.37	0.18	0.03	0.013	1	0
ACSP-10	RCCSA-H-ES	88.083	0.79	0.028	0.013	2.653	0.385
<i>Significant</i>		True ^a		True		True ^b	

^aFalse for CCSA(N+N) and RCCSA-H

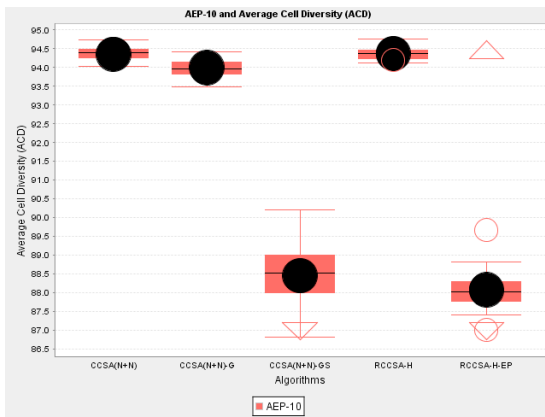
^bFalse for CCSA(N+N) and RCCSA-H

Table 4.7: Summary of results from the CCSA-RCCSA empirical study on ACSP-10.

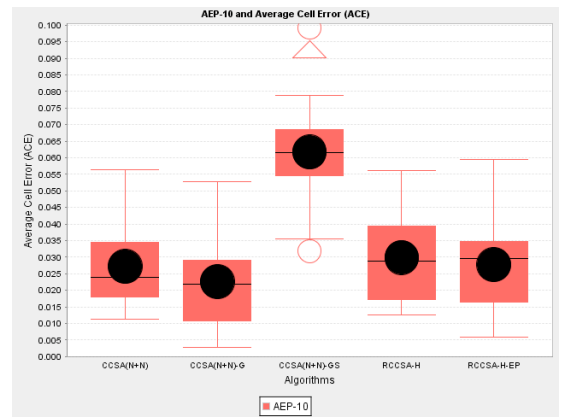
Analysis

This section provides an analysis of the results reported in the previous section in the context of the goals of the empirical study.

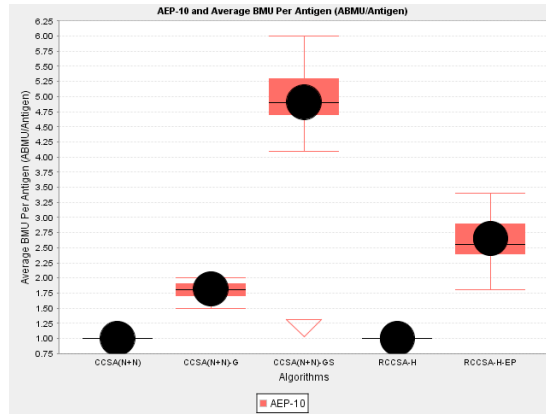
CCSA Trends With regard to average cell diversity, the aggregation and thus competition between the entire clonal set with the selected set resulted in a slight drop in diversity, although the relaxing of the competition from progenitor cells to the entire



(a) Average Cell Diversity (ACD) on ACSP-10.



(b) Average Cell Error (ACE) on ACSP-10.



(c) Average BMC Per Antigen (ABMCPA) on ACSP-10.

Figure 4.7: Box-and-whisker plot's from the Replacement Empirical Study.

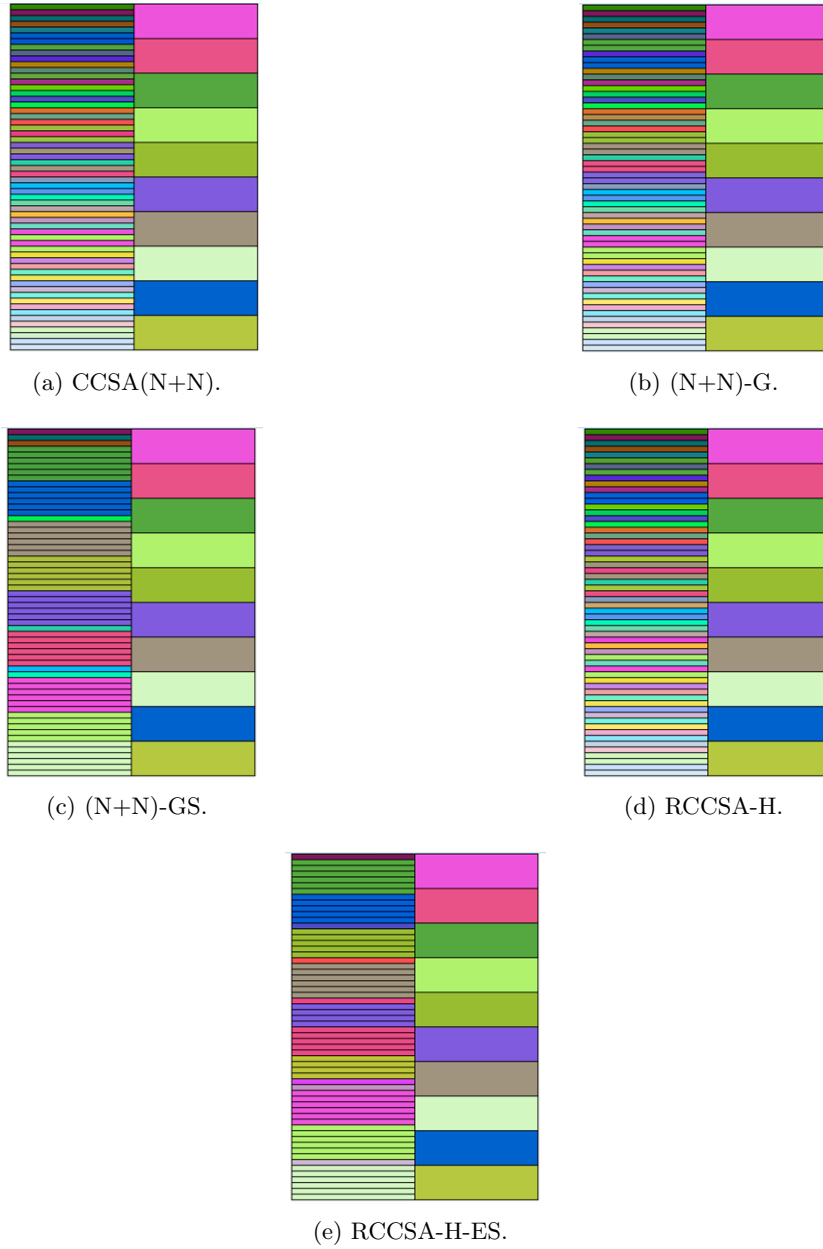


Figure 4.8: Plots of the ordered state of the repertoire after the triggering of the stop condition (left of each plot) for the assessed CCSA and RCCSA algorithms on the ACSP-10, where the right of each plot provides the problem optima.

repertoire resulted in a further larger decrease in diversity by ≈ 5.5 bits. Interestingly the CCSA(N+N)-G achieved a slightly lower error than CCSA(N+N) where the further relaxation of competition in CCSA(N+N)-GS resulted in a large relative increase in the ACE. Correlating with the decrease in ACD, the relaxation in the integration of the clones resulted in an increase in the number of best matching cells (BMC's) per antigen, rising to an average of ≈ 5 in the case of CCSA(N+N)-GS. The results demonstrate that multiple concurrent and redundant perspectives of antigen can be achieved by relaxing the replacement competition in CCSA which decreases repertoire diversity. The results also highlight that competition between the clonal set results in an improved result in terms of ACE and ABMCPA over CCSA(N+N), although the relaxation from progenitor set to whole repertoire dramatically increases ABMCPA at expense of increasing the systems general capability. The specialisation of each antigen's cellular imprint in the repertoire provided by CCSA(N+N)-GS is clearly demonstrated in Figure 4.8c compared to CCSA(N+N)-G in Figure 4.8b.

RCCSA Trends The exclusion of clonal siblings from replacement resulted in a significant decrease in diversity, the same seen between CCSA(N+N)-G and CCSA(N+N)-GS. The exclusion of progeny also resulted in a very slight decrease in ACE over RCCSA-H. Interestingly, the repertoire-wide competition between the aggregated clonal set in RCCSA-H resulted in a single average BMC per antigen, whereas, the exclusion of progeny during replacement promoted the competition between the concurrent clonal set such that ABMCPA increased to ≈ 2.5 . Also interesting was that the ACE and ABMCPA results for RCCSA-H were not significantly different from CCSA(N+N), demonstrating that unrestricted similarity-affinity based replacement into the repertoire behaves much like CCSA(N+N) confirming the expectation that motivated the exclusion of clonal progeny: that they are more similar. Also interesting is that the increase in ABMCPA with RCCSA-H-ES also resulted in a relative increase in ACE compared to CCSA(N+N) and CCSA(N+N)-G, suggesting that for the mechanisms used, increasing the specialisation of the cellular footprint for each antigen comes at the expense of general repertoire competence.

Conclusions

This section summarises the findings of the empirical study into the CCSA and RCCSA in terms of the primitives that were the focus of the study and the expectations regarding proportional resource allocation.

1. Relaxing the constraints for competition in the integration clones into the repertoire results in increased number of concurrent redundant perspectives of an antigen in the repertoire for CCSA and RCCSA.
2. Strong inter-clonal competition in CCSA results in improved performance with a slight increase in size in the number of redundant perspectives.

3. Repertoire-wide integration restrained by only similarity-and-affinity in RCCSA approximates CCSA(N+N) in behaviour.
4. Repertoire-wide integration restrained by affinity in CCSA or by similarity-and-affinity-exclusion in RCCSA results in a large increase in the size and quality of the number of redundant perspectives at the cost of general repertoire capability under both algorithms.

4.4.3 Degenerate Empirical Study

There is a lack of a one-to-one correspondence between antigen and receptors in the acquired immune system. A given antigen may have the capacity to trigger a very larger number of receptors, therefore resulting in a polyclonal activation, although a response is typically oligoclonal. This may be accounted for by the competition between clones for selection by the antigen. In addition, a given receptor may respond to a large number of antigens, thus resulting in a polyclonal response, something that cannot be accounted for by the clonal selection theory. In the Cognitive Theory of Immunology, Cohen refer to this polyclonal response of a given receptor as *cellular degeneracy*, where those receptors that are without a context are cross-specific, including auto-specific [82]. A polyclonal response is the antithesis of clonal dominance (a feature that underlies the clonal selection theory). The solution as proposed in cognitive theory is that specificity is an emergent phenomenon [80, 205]. Unlike the clonal selection theory, where specificity is a property of antigen-receptor interaction, emergent specificity is a down-stream effect that occurs after the initial interaction. A collection of varied and communicating cell types respond to the antigen in context. This is the so-called meta-response of cognitive theory, called the co-response or correspondence. The degeneracy of cell signalling is proposed as the basis of plasticity both in the brain and in the immune system, and is the feature exploited by antibiotics and pharmaceuticals.

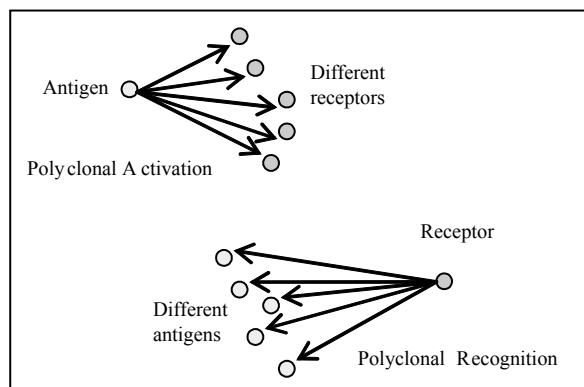


Figure 4.9: Depiction of a polyclonal activation and a polyclonal response.

Clonal selection accounts for the restriction of a potential polyclonal activation of the repertoire to an oligoclonal activation by dominant clones out competing the less well fitted clones (referred to as clonal dominance). The classical theory does not account

for lack of one-to-one correspondence between cells and antigen (the cross-reactivity of receptors referred to as polyclonal response). Cognitive immune theory accounts for polyclonal response between antigen and receptors by suggesting (1) that cells in isolation are degenerate and cross-specific, (2) context defines the use for degenerate cells resulting in emergent specificity.

Aim

Section 4.3.4 considered the expectations of clonal selection on a degenerate repertoire, specifically with regard to the pressures required to facilitate polyclonal activation and response. The aim of this empirical study is to investigate these expectations in the context of cellular clonal selection. Toward this end, the study had the following goals:

1. Investigate a clonal selection with a degenerate representation.
2. Assess selective pressures and explicit aggregation to constrain polyclonal activation and response.

Method

Problems The definition of the Colour Space Domain in Section 4.2.3 highlighted a ‘retina and colour perception’ analogy made by Cohen to describe cellular degeneracy and emergent specificity. The problems in this section are motivated by that analogy. This study used the ACSP-10 problem used for the CCSA empirical study in Section 4.4.1, although evaluated on a per colour component basis. To differentiate this mode of evaluation from the holistic ACSP, it is referred to as the Determinant Colour Space Problem (DCSP-10). A degenerate colour representation was used called component degeneracy where each colour was divided into the its Red, Green, and Blue components. Components are denoted D , where a $A = \{D_1, D_2, D_3\}$ for each colour space pattern. Degenerate component cells were assessed using Equation 4.10, where each degenerate component cell was defined by a 64-bit bitstring and decoded to a real value using Gray Code (Equation 4.1). Components were considered a symbolic degenerate representation as each component has direct meaning via their linear aggregation in the the context of a colour space pattern.

$$ComponentDistance(D, C) = |D - C| \quad (4.10)$$

Algorithms The study considers a *Degenerate Cellular Clonal Selection Algorithm (DCCSA)* as an extension of the RCCSA (defined in Algorithm 4.5) that uses a degenerate representation. RCCSA was selected as a base algorithm because (1) it achieves adequate results in terms of repertoire capability, and (2) because it achieves multiple high-affinity concurrent perspectives (proportional repertoire composition) on antigen to which it is exposed. The RCCSA was configured with the following configuration: $N_{cells} = 100$, with $N_{selected} = 3$, and $N_{clones} = 1$ per determinant exposure, resulting in a resource allocation of three cells per component, or nine cells per antigen. A Hamming-based similarity assessment was used in replacement, with clonal sibling exclusion. Two variations

of the degenerate cellular component selection were considered as follows: (1) explicit per-component selection called DCCSA-E, and (2) pre-committed component cells called DCCSA-P. Both approaches aggregated the best degenerate cell for each component into a holistic solution to address each given antigen in the ACSP. Explicit per-component selection maintained a repertoire of uncommitted components that were assessed and selected in the context of each component of each colour space pattern. This allowed degenerate cell reuse across the components of a single antigen and across the components of multiple antigen. Pre-committed component cells involved the management of one sub-repertoire for each component, where the size of each repertoire was a fraction of the total number of cells ($\lfloor \frac{1}{N_{components}} \times N_{cells} \rfloor$).

Experiment This study used the same experimental configuration including stop conditions, measures, as were used for the CCSA empirical study in Section 4.4.1. An additional measure was introduced for the degenerate component algorithms that assessed the extent of the cross-reactivity of the repertoire called the *Average Polyclonal Response Error (APRE)*, defined in Equation 4.11. APRE provides an indication of the average repertoire error per component.

$$APRE(I, T) = \frac{1}{I_n \times A_{in}} \sum_{i=1}^{I_n} \left(\sum_{j=1}^{A_{in}} \left(\frac{1}{T_n} \sum_{k=1}^{T_n} ComponentDistance(A_{ij}, C_k) \right) \right) \quad (4.11)$$

Results

Table 4.8 provides a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. The non-parametric Mann-Whitney U statistical test was calculated pair-wise for all algorithms.

Problem	System	ACD		ACE		APRE	
		\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
DCSP-10	DCCSA-E	31.638	0.047	0.003	0.001	0.333	0.018
DCSP-10	DCCSA-P	31.627	0.059	0.008	0.003	0.334	0.019
<i>Significant</i>		False		True		False	

Table 4.8: Summary of results from the DCCSA empirical study on DCSP-10.

Analysis

This section provides an analysis of the results reported in the previous section in the context of the goals of the empirical study, specifically (1) restrictive selection, and (2) explicit aggregation.

Restrictive Selection The results showed little significant difference on the selected measures between the explicit and pre-committed selection mechanisms, other than a slightly improved cell error for the explicit approach. The strong similarity in results, in

particular the Average Polyclonal Response Error demonstrate that the degenerate component clonal selection was generally unaffected when the repertoire was split into sub-repertoires per component compared to an integrated cross-component competitive repertoire. The minimum Hamming distance between generated colour space patterns likely resulted in little cross-colour space pattern reuse given the per-pattern/per-component specialisation of the repertoire, and the lack of difference between segregated and integrated component repertoires. An important consideration is that CCSA and RCCAS implicitly rely on restrictive selection to bound polyclonal activation of cells. This is clear if we consider a principle of cellular clonal selection that the scope of the repertoire available allowing N_{cells} interactions (assessments and activations). Without constrained selection, the scope of repertoire would respond in its entirety each antigen exposure.

Explicit Aggregation The clonal selection strategy by-design continues regardless of the integration of the product of exposure events. With respect to the problem domain, aggregation of the product of exposures is the critical concern of a degenerate representation. The chosen degenerate component representation was chosen because of its easy linear aggregation of the best matching components, always resulting in feasible aggregated responses at the per- D , per- A , and per- I levels. One may consider a sub-symbolic representation for the ACSF that does not easily aggregate into holistic colour space patterns solutions, such as sub-bitstrings. This is an important example because it highlights that such representation does not hinder the clonal selection strategy. Degenerate sub-strings may be assessed based on hamming distances from exposed A 's or D 's providing enough information for selection, cloning and integration. The explicit aggregation of sub-strings is less natural than the component case, requiring perhaps a per-bit position frequency assessment and a deterministic or probabilist generation of a viable holistic solution at the A scale. The example highlights that explicit aggregation may be suitable for the linear aggregation of components, although becomes more difficult with sub-symbolic representations one which clonal selection can still operate effectively. Explicit aggregation also highlights that this problem-centric concern is scale independent, where the components of a colour, and the colours of a colour set are the linear aggregation of selected information which may or may not be the case for more elaborate applications.

Conclusions

This section summarises the findings of the empirical study into the DCCSA in terms of the primitives that were the focus of the study and the expectations regarding proportional resource allocation.

1. *Degenerate Components*

- (a) Clonal Selection on components provides a viable realisation of cellular degeneracy under the circumstances considered where strong selection and aggregation can constrain polyclonal activation.

- (b) A RCCSA results in little difference in the partitioning or aggregation of the repertoire and thus competition of degenerate cells, suggesting clonal independence promoted via replacement is sufficient for non-overlapping clonal specialisation (assuming a large enough repertoire).
- (c) Explicit aggregation provides a context in which to bound polyclonal response to strongly activated (selected) structures.

2. Selection and Aggregation

- (a) Strong selection is required in all cellular clonal selection algorithms to manage the implicit polyclonal activation that occurs as a result of each antigens interaction with the entire repertoire.
- (b) Aggregation is independent of the principle clonal selection operations, requiring explicit integration into the process (for those applications that operate on sub-solution structures).

4.5 Spatial Repertoire

This section considers a spatial context as a method for forming relationships between units undergoing clonal selection and expansion. A spatial clonal selection paradigm is investigated which outlines a fixed lattice structured repertoire on which clonal selection operators are applied.

4.5.1 Spatial Clonal Selection

The processes of selection and maturation of lymphocytes occurs in the spatially distributed confines of the host organism lymphatic system [19]. Further, the spatial organisation may be required to provide context to guide emergent specificity [81]. The competition in the clonal selection principle is between cells for resources in the repertoire. Competition can be facilitated through differential selection of an activated set which results in differential allocation of resources for clones from the activated set. In envisaging the repertoire as a spatial structure, an additional level of competition is introduced called *spatial competition*. This competition puts pressure on cells in the same spatial neighbourhood to compete with each other. This pressure may be used for either the activation of cells during *selection* or *replacement* by clones of activated cells. A simple one or two-dimensional lattice structure may be used in which receptors occupy grid positions of the lattice, and the ends of the structure wrap around (ring or toroid), removing edge effects. The structure may be an equally arbitrary number of dimensions, although low dimensionality facilitates visualisation. Such a spatial repertoire structure provides a manifestation of the space complexity limits imposed on repertoire size, and imposes relationships between arbitrary neighbouring receptors for a given antigenic stimulus.

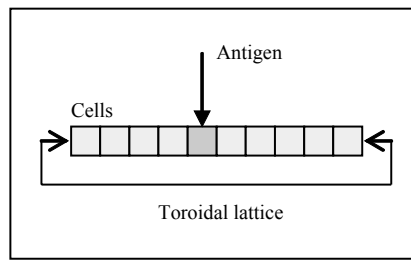


Figure 4.10: Depiction of a one-dimensional toroidal lattice.

Spatial Selection

The spatial structure may be exploited in the selection of the activated set for a given antigen. In the selection scheme the repertoire is evaluated, scanned for the highest affinity cell, which is returned to address the needs of the antigen, and an activated set is selected from the high-affinity cells. Therefore, the conventional activated set may be conceptualised as the selection of cells from across the spatial structure. One may partition antigenic signals and match them to regions (logical partitions) of the spatial structure, such that regions are directed in a top-down manner to be responsible for specific antigenic signals. The input space can be partitioned by a measure relevant to the representation of the antigen or the cells in the repertoire. This partitioning scheme would be implemented such that activated sets of receptors may only be drawn from the allocated region of the spatial structure, therefore putting pressure on the region to produce and develop receptors suitable to the allocated portion of the input space. A concern with this approach is that information effective for representational partitions may be developed outside of allocated regions and thus would not be exploited (for example the cross-reactive response for generalisation). Further, without inter-region competition, the partitions may be considered isolated repertoires such that the partitioning prevents the take-over or resizing of spatial regions with proportion to signal frequency and complexity (an effect maintained by differential resource allocation across the whole repertoire).

Spatial Replacement

The Replacement CCSA investigated in Section 4.4.2 may be elaborated to exploit a spatial structure such that replacement competition is *spatially localised* or constrained to the locations in the repertoire that responded strongest to an antigenic stimulus. Competition for resources may be constrained to the spatially local neighbourhood of high affinity cells. The principle parts of a spatial replacement mechanism are as follows: a structured repertoire or *lattice* data structure in which grid positions represent cells, a *neighbourhood function* that identifies the scope of competition (see Figure 4.11), and a *competition function* that defines how resources (lattice positions) are allocated to clones of selected cells. Spatial competition introduces spatial meaning between positions in the lattice, such that visualisation of the lattice can highlight relationships and facilitate the extraction of qualitative information from the repertoire.

The spatially restricted ‘back-end’ selection (neighbourhood function) is expected to

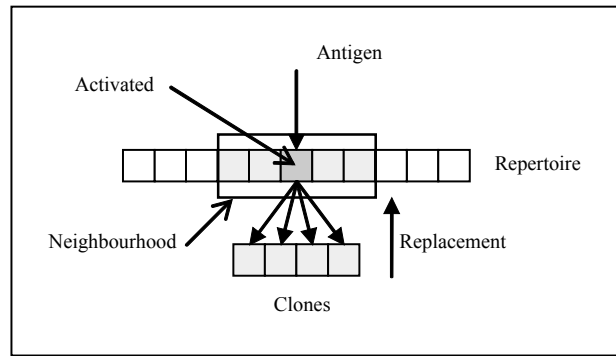


Figure 4.11: Depiction of clones competing in the neighbourhood of their progenitor.

result in the assignment of antigenic-responsibility to regions of the spatial repertoire structure⁴. Therefore, the neighbourhood function constrain's resource allocation resulting in a localised clonal dominance effect, where those cells that win upfront selection will produce clones that dominate sub-structures in the lattice. In effect, the spatial replacement strategy facilitates a mapping of the antigenic space onto the lower-order topology of the spatial repertoire. The clonal selection axiom of 'similar cells taking responsibility for similar antigen' is bonded to the lattice structure, facilitating the self-organisation of receptors to input signals. Ultimately, the expected organisation of similar cells into regions results in a broader competition between regions for activation, cloning, and the resulting ongoing retention of the region for spatial resources. The selection of the replacement function can play an important role in the maintenance behaviour of cell clusters. Clones are expected to have the same or similar affinity for the antigen as the activated cell given relatively low mutation rates. A clone that displaces a neighbouring cell is likely to be an activated cell in the following exposure. A similarity or affinity based replacement function is expected to maintain a cell cluster tied to a spatial locality and a random replacement function is expected to shift the centroid of the cluster around such that the fringe of the cell group may interact with neighbouring groups. In the case of the a stable spatial locality, upfront dominant cells are expected to put pressure on their neighbourhood to conform to the antigenic signals to which they dominant. Dominance leads to neighbourhood takeover which results in the emergent effect of self-organised responsibility.

Explicit and Implicit Organisation

Exploitation of the spatial structure as an upfront pressure results in the partitioning of input signals to regions of the spatial structure. This has the benefit of explicitly limiting the scope of interest of input signals space for regions of the lattice, with the problems associated with making assumptions regarding the partitioning the signal space. All resources are employed for a specific task, although not all the allocated resources may be required for their allocations. The exploitation of the spatial structure as a back-end pressure results in the implicit partitioning of the input space and self-organisation of receptor clones to take responsibility for the automatic partitions. This has the benefit of

⁴'Back-end' refers to competition at the end of an iteration of the algorithm cycle.

proportionate allocation (specialisation) of resources for the automatically identified complexities of the input space with the computational and delay costs for the self-organising process. Resource allocation is determined automatically, although not all resources may be utilised effectively. An interesting observation is that one method will lead to the other, such that the explicit exploitation of the spatial competition of both ends is not required. The application of upfront spatial partitioning will force regions to adapt to input signals, thus result in replacements occurring in the same region, reinforcing the partitioning. The application of back-end partitioning will self-organise the responsibility for antigenic signals via replacement, which will reinforce future input signals being directed to responsible areas of the spatial structure.



Figure 4.12: Depiction of the explicit versus implicit partitioning of input signals and the effects on the repertoire.

4.5.2 Spatial Repertoire Empirical Study

Aim

The aim of this investigation is to investigate cellular clonal selection constrained by a spatial repertoire structure. Toward this end, the study had the following goals:

1. Investigate the effects on the repertoire composition and capability in adopting a spatial repertoire structures.
2. Assess the effects on the repertoire in exploiting the spatial structure to localise clonal integration via replacement.

Method

Algorithms This study considers the Replacement Cellular Clonal Selection Algorithm (RCCSA), and two variations of the Spatial Cellular Clonal Selection Algorithm (SCCSA). The RCCSA defined in Algorithm 4.5 was configured with $N_{cells} = 100$, $N_{selected} = 2$, and $N_{clones} = 5$ to promote an allocation of ten cells per antigen in ACSP-10. Unlike RCCSA, SCCSA does not aggregate and integrate clones as a group into the repertoire, instead clones are replaced individually, with per-progenitor clonal sibling replacement exclusion. Two variations of the SCCSA algorithm were investigated: (1) a variation that organised the repertoire into a lattice although was constrained by the spatial organisation called holistic replacement (SCCSA-HR), and (2) a variation that exploited the spatial repertoire

and constrained the integration of clones to the spatial neighbourhood of the progenitor cell and with exclusion of clonal siblings (like RCCSA and SCCSA-HR) called neighbourhood replacement (SCCSA-NR). The neighbourhood was defined as the square of the nine cells surrounding and including a given position in the toroidal lattice. Both SCCSA used the same configuration as RCCSA, although the lattice dimensions used the square root of N_{cells} resulting in a 10-by-10 2-dimensional repertoire.

Problems This study used the ACSP-10 problem used for the CCSA empirical study in Section 4.4.1.

Experiment This study used the same experimental configuration including stop conditions, measures, as were used for the RCCSA empirical study in Section 4.4.2, including the Average Best Matching Cells Per Antigen (ABMCPA). An additional measure was used that calculated the average diversity of the 9-cell neighbourhood for each position in the lattice, called the Average Cell Neighbourhood Diversity (ACND). The diversity was calculated using the Average Cell Diversity (Equation 4.7) method for each neighbourhood, and averaged across all 100 neighbourhoods.

Results

Table 4.9 provides a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. The non-parametric Mann-Whitney U statistical test was calculated pair-wise for all algorithms. Figures 4.13a, 4.13b, 4.13c, and 4.13d show the ACD, ACE, ABMCPA, and ACND on all algorithms respectively. Figure 4.14 provides example plots of the state of the repertoire from the two spatial-based algorithms (same configuration, random seeds of 1 and 5 for the algorithm and problem respectively).

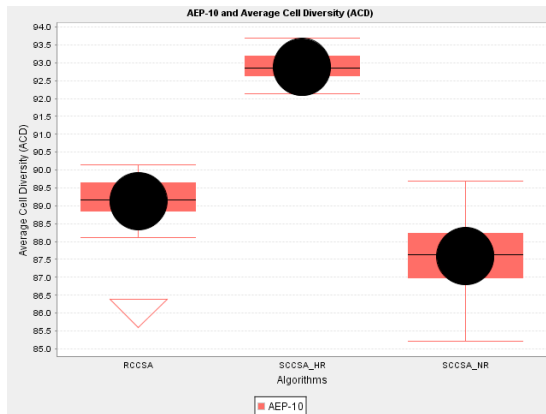
Problem	System	ACD		ACE		ABMCPA		ACND	
		\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
ACSP-10	RCCSA	89.128	0.732	0.002	0.002	3.94	0.803	N/A	N/A
ACSP-10	SCCSA-HR	92.876	0.405	0.012	0.012	2.83	0.352	83.257	0.508
ACSP-10	SCCSA-NR	87.592	0.968	0.002	0.001	3.417	0.424	65.119	1.514
<i>Significant</i>		True		True ^a		True		True	

^aFalse for RCCSA and SCCSA-NR

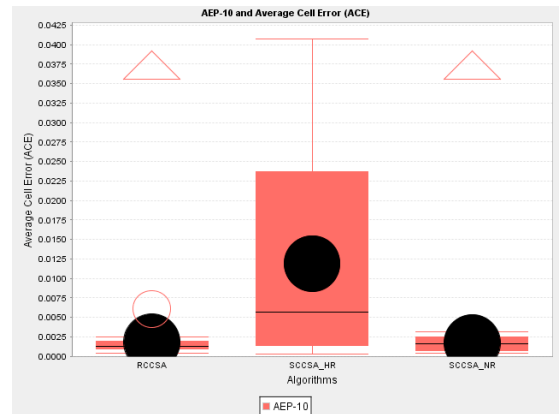
Table 4.9: Summary of results for SCCSA on AEP 10.

Analysis

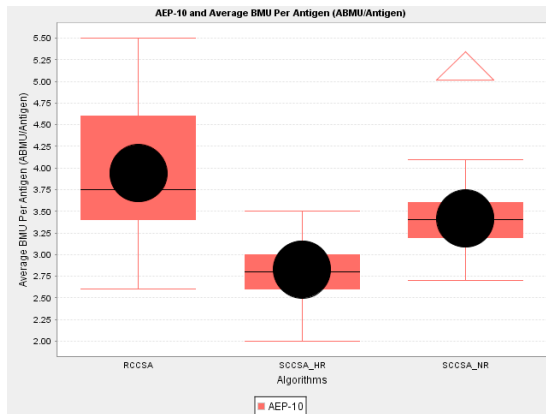
This section provides an analysis of the results reported in the previous section in the context of the goals of the empirical study.



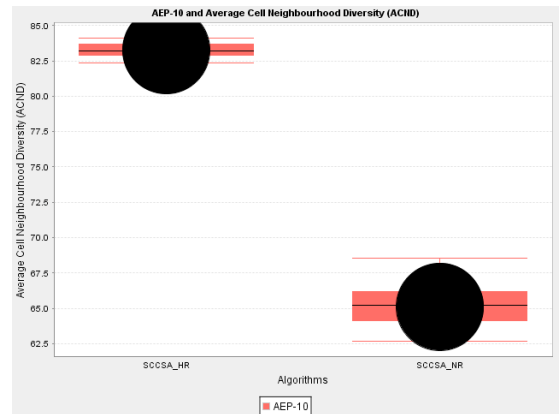
(a) Average Cell Diversity (ACD) on ACSP-10.



(b) Average Cell Error (ACE) on ACSP-10.

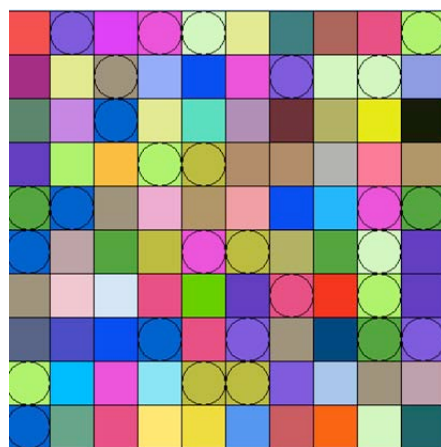


(c) Average BMC Per Antigen on ACSP-10.

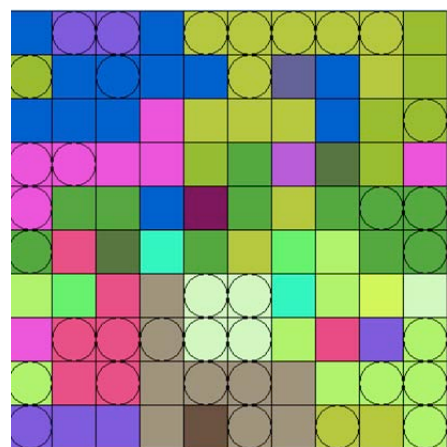


(d) Average Cell Neighbourhood Diversity on ACSP-10.

Figure 4.13: Box-and-whisker plot's from the Spatial Repertoire Empirical Study.



(a) SCCSA-HR on ACSP-10.



(b) SCCSA-NR on ACSP-10.

Figure 4.14: Example plots of the spatial repertoire from two variations of SCCSA at the end of the run on ACSP-10 with BMC represented with circles.

Spatial Structure Trends This section considers the effects of storing the repertoire in a spatial structure by comparing the results of RCCSA to SCCSA with holistic replacement. The use of the spatial structure in SCCSA-HR resulted in a small increase in repertoire diversity and in particular response error where ACE also showed a large increase in the variance of final ACE scores. The increase in diversity and error correlated with a decrease in the average number of BMC's. The spatial structure was used in SCCSA-HR to hold the repertoire although was not exploited in any way. The difference in the repertoire composition and capability between the approach and RCCSA demonstrate the effect of the per-progenitor clonal set replacement used in SCCSA compared to the aggregated clonal set replacement in RCCSA.

Localised Replacement Trends This section considers the exploitation of the spatial structure in the neighbourhood replacement used in SCCSA compared to holistic replacement. The SCCSA-NR resulted in a the increased organisation of the repertoire as expected by localising cells for the same antigen into areas on the spatial repertoire structures. This was demonstrated by the decrease in repertoire diversity compared to both SCCSA-HR and RCCSA, and a small increase in ABMCPA compared to SCCSA-NR. Interesting the exploitation of spatial neighbourhood competition resulted in a ACE not significantly different from RCCSA, suggesting that such competition results in the same level of per-antigen specialisation within the repertoire as RCCSA. Importantly the increased organisation was demonstrated by the significant decrease in average repertoire diversity compared to holistic replacement, which was depicted in the example plot in Figure 4.14 showing that this decrease was caused as a result of the clear grouping of similar cells (similar colours) in the spatial repertoire.

Conclusions

This section summarises the findings of the empirical study into the Spatial Cellular Clonal Selection Algorithm in terms of the primitives that were the focus of the study and the expectations that motivated the study.

1. Neighbourhood replacement provides intra-clone competition that results in similar general behaviours as consolidating the clone for replacement in RCCSA.
2. Neighbourhood replacement results in the spatial localisation of clones on the lattice, the so called spatial responsibility effect as expected.

4.6 Response Mediation

This section elaborates on the clonal selection by introducing a second repertoire of intermediate (mediator) cells to provide an adaptive context for localising the differential resource allocation of clonal selection, providing a model of decoupled *feature detection* and *system response*.

4.6.1 Mediated Clonal Selection

Two-signal theories have been proposed in immunology as verification signals permitting the proliferation and differentiation of lymphocytes. Historically developed for the activation of B-lymphocytes (in the context of self-nonself discrimination), the two-signal approach has also been extended to the various classes of T-cells. The dominant type of B-cell activation is dependant on Helper T-cells that identify MHC molecules on the surface of activated antigen-presenting B-cells. The detection of MHC provides a secondary verification signal to the B-cells, allowing them to proliferate and differentiate. This process is called the co-stimulation of B-cells by T-cells or T-cell dependant activation of B-cells. Likewise, T-cell activation requires an antigen specific signal via an antigen-presenting cell and a co-stimulating ‘antigen non-specific’ verification signal. The theory of the two-signal activation of T-cells is called associative recognition theory [57].

Mediated Clonal Selection or so-called inter-repertoire recognition decomposes the clonal selection process such that a mediation process is responsible for selecting those cells from the activated set that may proliferate. The mediation process is controlled by a sister repertoire of cells that also perform a clonal selection process using the activated cells of the first repertoire as input signals. The Helper T-cell metaphor is adopted such that the repertoire of cells (first repertoire) represents B-lymphocytes, which require a second verification signal from the helper T-lymphocytes (second repertoire) before proliferating and differentiating. After providing the secondary signal to the activated set of B-lymphocytes, the helper T-lymphocytes proceed with the cloning and maturation of the activated T-cells. Therefore, the T-cell repertoire is an application of the clonal selection algorithm that accepts the activated set of another clonal selection algorithm as a pseudo-‘antigenic set’. Two concerns that must be reconciled in integrating the two repertoires are as follows: (1) The T-cells must select those B-cells from the B-cells activated set that may proliferate (2) The B-cell activated set provides multiple antigens simultaneously to the T-cell repertoire, to which it is possible to generate a T-cell activated set for each member of the B-cell activated set. The remainder of this section considers the implications of the second clonal selection governed repertoire from a variety of different perspectives.

Mapping Function

The proliferation *verification signal* for the activated B-cell’s may be *antigen-dependent* or *antigen-independent*. If antigen-dependent, the information provided by the verification signal may be considered a generalisation of the input signals provided to the B-cell repertoire. If the verification signal is antigen-independent, then the information provided is a generalisation of the cells themselves in the B-cell repertoire. This is a subtle but potentially important difference that effects the representation of information by the system.

An antigen-dependent mapping from B-cell to T-cell requires that some regularities of the input signal are preserved. The transformation may be of the whole antigen itself, or just the part of the antigen recognised by the B-cell receptor. The T-cell repertoire adapts to a consistent (B-cell invariant) generalisation of the antigenic input patterns. An

antigen-dependent mapping between the repertoires provides a top-down organisation of the relationship between the repertoires, where simply matching onto the second repertoire provides the verification signal. A potential concern with antigen dependence is that the mapping of the input patterns may be achieved in one repertoire, making the second repertoire redundant. An antigen-independent mapping disregards the regularities of the input signal, such that the T-cell is performing pattern recognition of B-cell receptors, rather than a pattern recognition of antigen (whether symbolic or sub-symbolic). The result is a T-cell repertoire that is developed independent of antigen, and dependent on the B-cell repertoire (which in turn develops in response to antigen). A potential concern with an antigen-independent mapping is that it is arbitrary such that T-cells are completely dependent on specific B-cell lineages, and regularity between B-cell patterns has no meaning other than ancestry (likely to make the mapping harder).

Formation of High-Order Structures

Interestingly, the relationship between an antigen and cells, as well as cells between repertoires may be considered in the context of the formation and maintenance of higher-order structures. For example in the case of a single repertoire, one antigen may map onto one cell providing a specialised mapping. This may be elaborated to include the extremes cardinality of relationship. For example, a one-to-many relationship between antigen and cells maybe considered a decomposition of feature extraction, particularly if the mapping is partial or sub-symbolic. Additionally, many antigen may be mapped onto many cells either directly (generalisation of one-to-one) or partially (cross reactivity), or onto a single cell, compressing the signal.

The relationship between repertoires adds a second level of complexity to the hierarchical structure. The same relationships apply although in this case between cells, with the importance difference that the relationships can be manipulated with mapping schemes, activated B-cell set size, and activated T-cell set size. For example, it may be desirable to coerce the B-cell repertoire to decompose antigen into regular features (many), and the T-cell receptor to generalise those features toward a specific meaning (few). Alternatively, the B-cell repertoire may be coerced to generalise (compress) based on common antigenic features (few), and the T-cell repertoire to decompose the B-cell information content into a variety of different meanings (many).

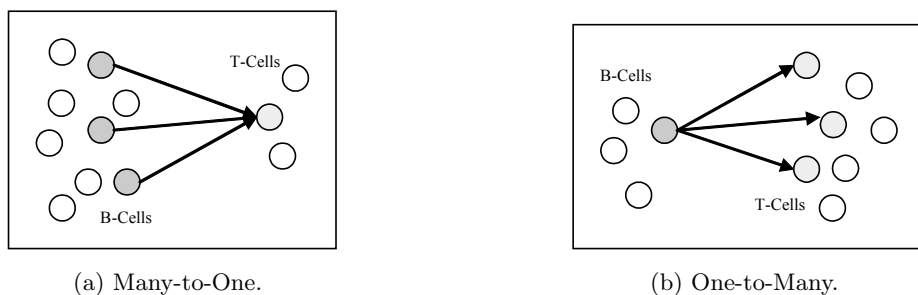


Figure 4.15: Depiction of inter-repertoire structures based on the mapping of cells.

Rather than treating each activated B-cell independently, the activated B-cells may interact with the T-cell population concurrently, such that multiple B-cells may activate a single T-cell. The T-cell repertoire may be configured such that a T-cell is required to match more than one B-cell before it may become activated. This provides a mechanism in which the T-cell repertoire facilitates the aggregation of multiple antigenic features into a single T-cell. The second repertoire facilitates the addition of a second layer of meaning. In the many-to-one case a T-cell may represent a high-order concept as an aggregate of multiple lower-level concepts (B-cells). The meaning is a single or small collection of concepts that are dependent on the presence of specific features (specific context). For example, *Many-to-One* provides a consensus forming operation in which many lower-order concepts (such as extracted features) are aggregated together to form a high-order concept. The *One-to-Many* relationship provides a descriptive operation in which a high-order concept is decomposed into a number of lower-order concepts (such as descriptive features). In both of these examples, the meaning implies a mapping that manages to preserve some regularities of the input signal, such that higher-order or lower-order concepts may be described.

Supervised Concept Formation

The incorporation of additional feedback into the algorithm allows an external process to assign meaning and supervise the mapping and formation of concepts. For the single repertoire algorithm, T-cell mediation provides an immunologically plausible basis for feedback (verification) to supervise the selection of activated receptor patterns in response to antigenic exposures. For the dual repertoire algorithm, tissue damage provides an immunologically plausible basis for feedback to supervise the selection of concepts formed in the T-cell repertoire, which in turn feeds back to the B-cell repertoire via the mediation process. The proposed integration of feedback results in two flows of information through the system. The top-down flow of antigenic signals results in B-cell receptors competing for activation, and T-cell receptors competing for activation of the B-cell receptor activated set. Those B-cells that receive the secondary signal, proliferate. The bottom-up flow of tissue-damage input results in T-cells competing for encouragement that the concepts that they represent are *useful*. Those T-cells that receive a positive (non-negative, and perhaps non-neutral) feedback, proliferate. Therefore, the application of feedback (if available) is to the activated set of T-cells. In the absence of feedback, there is simply the absence of corrective behaviour: the punishment, or reward of the activated T-cell concept.

Bottom-up supervision of concept formation provides a natural agenda for investigating mediated clonal selection beyond the basic principles of mappings and structure formation and maintenance (for example, see Figure 4.16). The integration of a feedback mechanism provide a natural relationship with the study of Reinforcement Learning (Section 3.4.5).

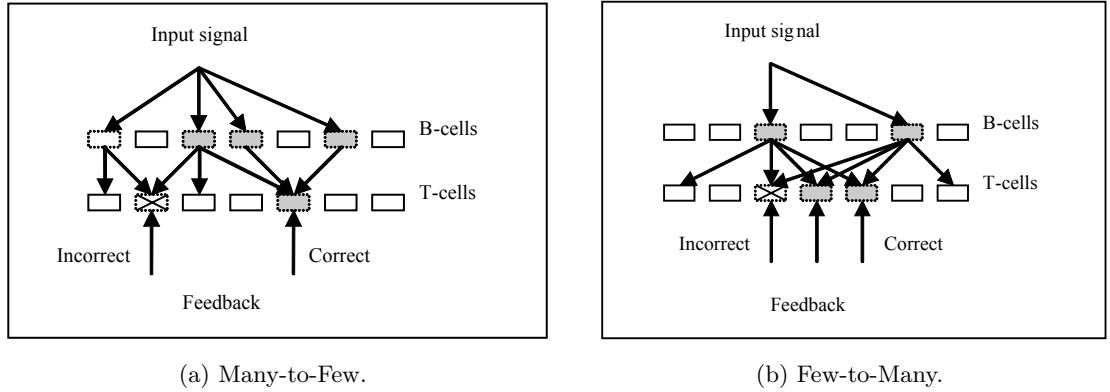


Figure 4.16: Depiction of the integration of feedback for supervised structure formation.

4.6.2 Mediated Empirical Study

Aim

The aim of this investigation was to assess the cellular clonal selection algorithm constrained by inter-cellular repertoire interactions. Toward this end, the study had the following goals:

1. Assess the behaviour of different inter-repertoire antigen-dependant mapping schemes.
2. Investigate the effects of varied cellular cardinality in the relationships exploited via inter-repertoire mapping.

Method

Problems This study used the ACSP-10 problem used for the CCSA empirical study in Section 4.4.1.

Algorithms Algorithm 4.7 defines the Mediated Cellular Clonal Selection Algorithm (ECCSA), where $Exposure(A, T)$ is a modified version of the $Exposure$ operation for RC-CAS (defined in Algorithm 4.5) that returns the selected set rather than the best matching cell. The RCCSA uses clonal sibling exclusion and a Hamming similarity function during replacement.

Algorithm 4.7: Exposure Function for Mediated Cellular Clonal Selection.

Input: $A, T_{bcells}, T_{tcells}, N_{bc-selection}, N_{bc-clones}, N_{tc-selection}, N_{tc-clones}, P_{mutation}$

Output: T_{rs}

- 1 $T_{rs} \leftarrow 0;$
 - 2 $T_{bc-rs} = Exposure(A, T_{bcells}, N_{bc-selection}, N_{bc-clones}, P_{mutation});$
 - 3 $T_{tc-rs} = Exposure(T_{bc-rs}, T_{tcells}, N_{tc-selection}, N_{tc-clones}, P_{mutation});$
 - 4 $T_{rs} \leftarrow SelectBestMatchingCell(T_{tc-rs});$
 - 5 **return** $T_{rs};$
-

The modified RCCSA $Exposure$ operation also permits the mapping of a selected set of cells (T_{rs}) to be treated as an antigen. This is achieved through use of an exposure

operation that provides a mapping function that aggregates affinity for a given cell against a stimulating set of cells. The first part of this study considers two different examples of the mapping function: (1) Euclidean Mapping, and (2) Hamming Mapping. The affinity scores are aggregated within each cell of the repertoire against the stimulating set to provide a general indication of a cells capability against a stimulating set. In this part of the study, both the B-cell and T-cell repertoires used identical configurations as follows: $N_{cells} = 100$, $N_{selected} = 1$, $N_{clones} = 10$ (10% of each repertoire per antigen). The second part of the study considered varied relationships between the two repertoires in terms of the size of the selected B-cell set which stimulates the T-cell repertoire, and the size of of the activated and responding T-cell set. A series of one-to-one and many-to-many relationships were assessed with the intermediates, where the number of clones and this repertoire assigned per antigen was kept constant at six (10% of each repertoire per antigen) with regard to the selection size. The configurations are summarised in Table 4.10. A Euclidean mapping was used between the repertoires in this second part of the study.

Relationship	B-Cells			T-Cells		
	$N_{b-cells}$	$N_{bc-selected}$	$N_{bc-clones}$	$N_{t-cells}$	$N_{tc-selected}$	$N_{tc-clones}$
<i>1-to-1</i>	60	1	6	60	1	6
<i>1-to-N</i>	60	1	6	60	6	1
<i>N-to-1</i>	60	6	1	60	1	6
<i>N-to-N</i>	60	6	1	60	6	1

Table 4.10: Summary of the configuration of the ECCSA with a series of inter-repertoire selection relationships.

Experiment This study used the same general experimental configuration including stop conditions as were used for the CCSA empirical study in Section 4.4.1. A series of new measures were defined for the experiment. The Average Cell Diversity measure (defined in Equation 4.7) was calculated for each of the B-cell and T-cell repertoires referred to as the Average B-Cell Diversity (ABCD), and the Average T-Cell Diversity (ATCD) respectively. This treatment was performed to the Average Cell Error measure (defined in Equation 4.4) referred to as the Average B-Cell Error (ABCE) and the Average T-Cell Error (ATCE) respectively, where the measure was taken as the best response from each repertoire against the antigen. The Average BMC per Antigen (defined in Equation 4.9) was adapted for the two re-repertoire configuration referred to as the Average Best Matching B-Cells Per Antigen (ABMBCPA) and the Average Best Matching T-Cells Per Antigen (ABMTCPA) respectively. The system’s response was taken as the best matching cell from the set of selected (activated) T-cells for each antigen exposure, called the Response Error (RE). A new measure was defined that calculated the T-cell error after mapping, averaged across each cell in the T-cell repertoire called the Average T-Cell Mapping Error (ATCME). This measure provided a general indication of the mapping error between the repertoire in the units of the specific mapping scheme used. The measure was also taken against just the set of T-cells activated and selected by the activated B-cells called the Average T-Cell Selected Set Mapping Error (ATCSSME). Figure 4.17 provides a depiction of the

repertoires, the antigen, the selected sub-sets and the system response as well as how all the collected measures relate to these principle attributes of the ECCSA.

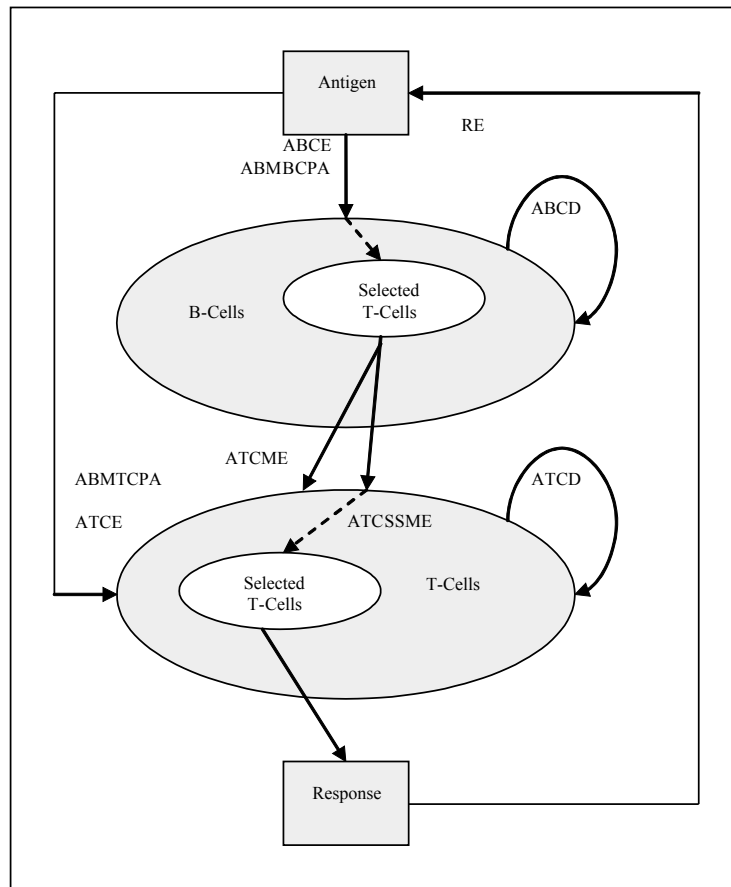


Figure 4.17: Depiction of the principle concerns of the MCCSA and the relationships in terms of the selected measures.

Results

Table 4.11 and Table 4.12 provide a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. The non-parametric Mann-Whitney U statistical test was calculated pair-wise for all algorithms. Figures 4.19a, 4.19b, 4.19c, 4.19d, and 4.19e show the ABCD, ABCE, ATCD, ATCE, and RE respectively. Figure 4.20 provide example plots for all four configurations on the second relationship part of the study. Results from all example plots are taken from the end of the run with algorithm and problem configurations matching those used during experimentation, and a random number generator seed of 1 and 5 for the algorithm and problem respectively.

Analysis

This section provides an analysis of the results reported in the previous section in the context of the goals of the empirical study.

Measure	Euclidean		Hamming		Significant
	\bar{x}	σ	\bar{x}	σ	
<i>ABCD</i>	89.399	0.822	89.399	0.822	False
<i>ATCD</i>	89.234	0.718	87.999	0.94	True
<i>ABCE</i>	0.003	0.002	0.003	0.002	False
<i>ATCE</i>	0.009	0.006	0.098	0.029	True
<i>RE</i>	0.009	0.006	0.617	0.123	True

Table 4.11: Summary of results for ECCSA with two different mapping schemes on ACSP-10.

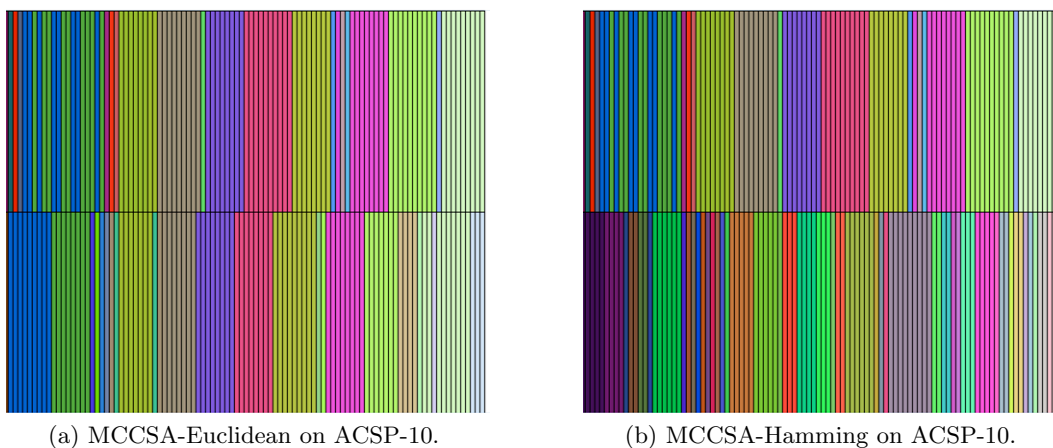


Figure 4.18: Example plots of the B- (top of each plot) and T-Cell (bottom of each plot) repertoires from the ECCSA at the end of the run on ACSP-10 for two different inter-repertoire mapping types.

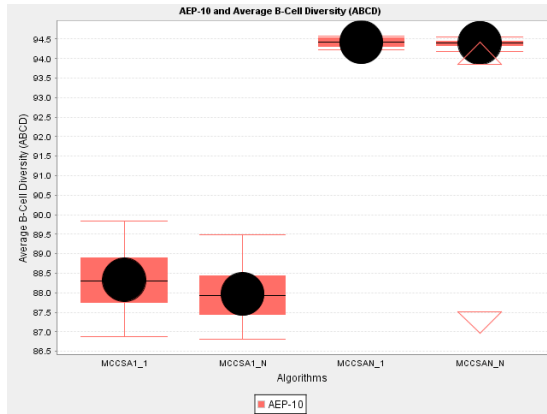
Measure	(1-1)		(1-N)		(N-1)		(N-N)		Sig.
	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	
<i>ABCD</i>	88.333	0.812	87.967	0.721	94.419	0.113	94.408	0.147	True
<i>ATCD</i>	88.005	0.755	94.298	0.146	88.103	0.924	94.309	0.105	True
<i>ABCE</i>	0.027	0.024	0.028	0.012	0.067	0.014	0.067	0.015	True
<i>ATCE</i>	0.052	0.021	0.076	0.015	0.102	0.019	0.093	0.013	True
<i>ATCME</i>	0.604	0.065	0.648	0.04	3.523	0.307	3.845	0.188	True
<i>ATCSSME</i>	0.042	0.02	0.158	0.023	0.844	0.129	1.161	0.121	True
<i>ABMBCPA</i>	2.713	0.33	3.13	0.426	1	0	1	0	True ^a
<i>ABMTCPA</i>	3.25	0.318	1	0	2.597	0.549	1	0	True ^b
<i>RE</i>	0.059	0.031	0.083	0.016	0.112	0.022	0.122	0.019	True ^c

^aFalse for ECCSA(N-1) and ECCSA(N-N)

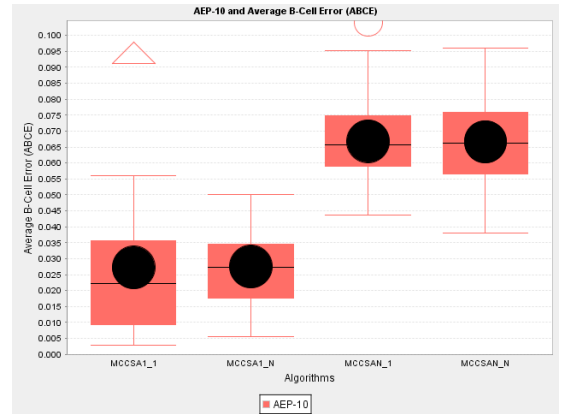
^bFalse for ECCSA(1-N) and ECCSA(N-N)

^cFalse for ECCSA(N-1) and ECCSA(N-N)

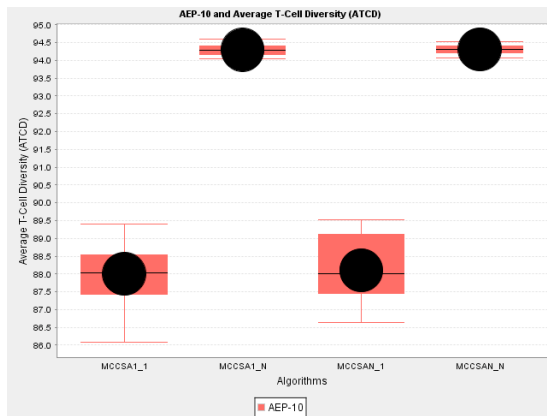
Table 4.12: Summary of results for ECCSA with four different inter-repertoire relationship schemes on ACSP-10.



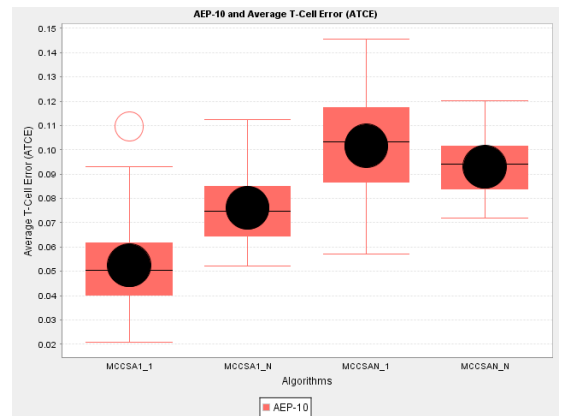
(a) Average B-Cell Diversity (ABCD) on ACSP-10.



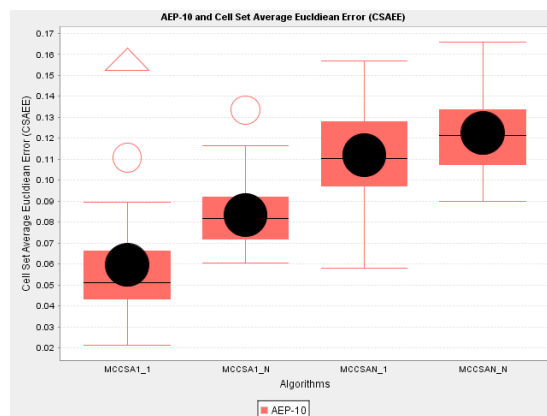
(b) Average B-Cell Error (ABCE) on ACSP-10.



(c) Average T-Cell Diversity (ATCD) on ACSP-10.

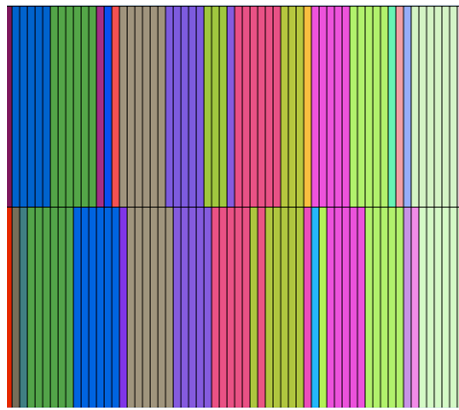


(d) Average T-Cell Error (ATCE) on ACSP-10.

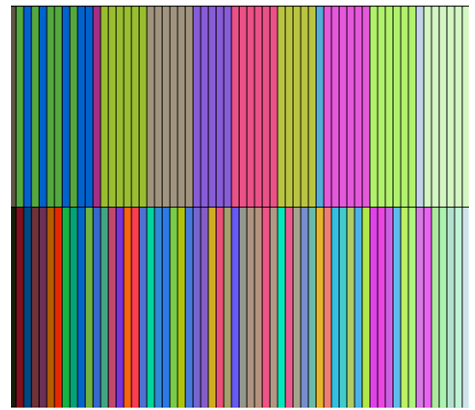


(e) Response Error on ACSP-10.

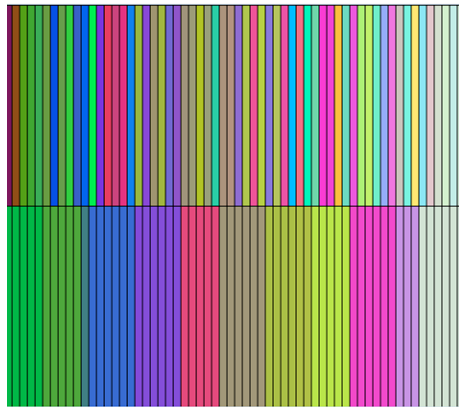
Figure 4.19: Box-and-whisker plot's for the relationship configuration with ECCSA.



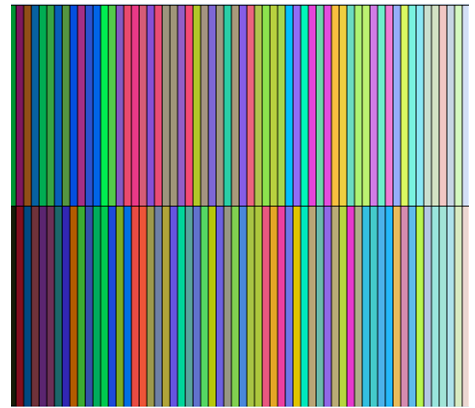
(a) MCCSA(1-1).



(b) MCCSA(1-N).



(c) MCCSA(N-1).



(d) MCCSA(N-N).

Figure 4.20: Example plots of the B- (top of each plot) and T-Cell (bottom of each plot) repertoires from the MCCSA at the end of the run on ACSP-10 for all four inter-relationship types.

Mapping Trends This section considers the effects of using the Euclidean and Hamming antigen-dependant inter-repertoire schemes. The B-cell diversity and error show no significant difference between the mapping schemes. This is an expected result, as the mapping scheme and the chosen ECCSA does not effect the B-cell repertoire. The T-cells showed a small decrease in diversity with the Hamming mapping, and more critically showed a factor of 10 increase in Average T-cell Error compared to the Euclidean mapping. This difference in error was reflected more strongly in the Response Error that showed a an increase in error of ≈ 70 . The example plots in Figure 4.18 clearly show well organised B-cell repertoires under both mapping schemes as expected, and the clear organisation and disorganisation in the T-cell repertoires between the Euclidean and Hamming schemes respectively. These results demonstrate that under the circumstances investigated, that Hamming distance is not a sufficient approximation for Euclidean distance, not providing enough information to promote the same structures in the T-cell repertoire as were selected in the B-cell repertoire. These results demonstrate the fragility of proxy-based response to the antigen-centric mapping function.

Inter-Repertoire Relationship Trends This section considers the mapping relationships between the number of selected (stimulating) cells in each repertoire. The diversity and error of the B-cell repertoire increased with the number of selected B-cells. This same general behaviour was observed with the T-cell diversity and the number of selected T-cells. The ATCE increased with the number of selected T-cells with 1 triggering B-cell, although showed a small decrease with the increase in selected T-cells with N selected T-cells. The average B-cell to T-cell mapping error showed an increase in the number of selected T-cells with both 1 and N selected B-cells, and effect also observed with the mapping error of the selected T-cell set. Together these observations suggest that the selection of a single cell and integration of its clones results in a more specialised (lower diversity and error) repertoire and mapping. Conversely, the result suggests that the selection of a large activated set and integration of a few clones from each results in a more diverse cellular perspective on the cellular and/or antigenic trigger. This observation lends support to the *oligoclonal axiom* of the clonal selection strategy outlined in Section 4.2.1. More specifically, in the context of the replacement-based resource allocation scheme used in the assessed algorithms, selection and integration of few cells with many clones results in a more specific and less diverse repertoire. This finding in the context of the bi-repertoire model with antigenic and cellular triggers lends support to the same finding in Section 4.4.2. This trend is observed poignantly in the Response Error from the relationship results where the 1-1 configuration achieved the lowest response-by-proxy error, increasing with the number of selected T-cells and selected B-cells. The plots in Figure 4.20 depict this in the organisation of the B-cell and T-cell repertoires in those configurations where a single cell is selected by antigen or cellular trigger resulting in the integration of a set of clonal siblings (1-1, 1- N , and N -1).

Conclusions

This section summarises the findings of the empirical study into the Mediated Cellular Clonal Selection Algorithm in terms of the primitives that were the focus of the study and the expectations that motivated the study.

1. *Mappings*

- (a) An antigen-centric mapping scheme results in the effective blind promotion (specialisation) of the T-cell repertoire and proxied response.
- (b) The quality of the mapping with regard to proxied responses in mediated clonal selection is as good as the amount of antigen-centric information mapped between the repertoires.
- (c) The remapping of the antigen-assessment Euclidean distance resulted in an effective proxy response, whereas the approximation of the assessment in the Hamming distance resulted in a markedly unsuccessful proxied response.

2. *Relationships*

- (a) The number of activated B-cells and thus stimulus for the T-cell repertoire was an important factor in response capability, resulting in a specialisation of feature detectors for remapping.
- (b) The one-to-one and one-to-many relationship configurations between the repertoires resulted in the better proxied responses, with more than a factor of two decrease in capability with an increase in the number of selected B-cells.
- (c) One-to-Many selection and clonal integration for a given repertoire results in improved repertoire capability (lower error), increase repertoires organisation (lower diversity), and an improved mapping given the specialisation it provides to the adaptive process.

4.7 Cell-Cell Recognition

This section elaborates on clonal selection by introducing a pattern recognition interpretation of the Immune Network Theory and its integration into the cellular clonal selection framework.

4.7.1 Network Clonal Selection

The Idiotypic Network Theory proposed by Jerne indicates that receptors (free or surface bound antibody) are selected by other receptors in addition to antigen (Section 2.4.2). As such, one may define two additional types of receptor-to-receptor interaction in addition to conventional interaction with antigen: (1) The activation of a receptor by the idio-type of another receptor that results in the creation of more anti-receptor receptors, (2) The activation of a receptor that is specialised toward an antigen by another receptor that results in the creation of more receptors for the antigen without the presence of the antigen

and anti-idiotypic receptors for the triggering receptor. The theory suggests that the aggregation of these low-level behaviours results in a network of receptors that may interact with antigen and each other providing both an antigen-recognition system and the self-regulation of immune response.

Recognition and Relationships

The two direct receptor relationships include: *Receptor-Antigen*: the traditional relationship where receptors for the antigen (target) are created with minor variations toward improving recognition, and *Receptor-Receptor*: the the idiotypic relationship where receptors for the triggering receptor (target) are created (anti-receptor receptors) with minor variations toward improving recognition. Both of these cases provide examples of *direct relationships*, that of a receptor and an antigen or a receptor and another receptor. In the first case, receptors have an *implicit relationship* with other receptors that also match for the same antigen in that they compete with each other for selection by the antigen. The second case is a lot more interesting, given the recurrent relationships that may result. These relationships are considered in the context of how a given receptor came to be, assuming a single matching source is responsible for receptor maturation. For example, a receptor may match to another receptor that has been matured for an antigen. Therefore, there is a direct relationship between the first receptor and the second, as well as a *proportional relationship* between the first receptor and the second receptors antigen (for example of both relationships see Figure 4.21).

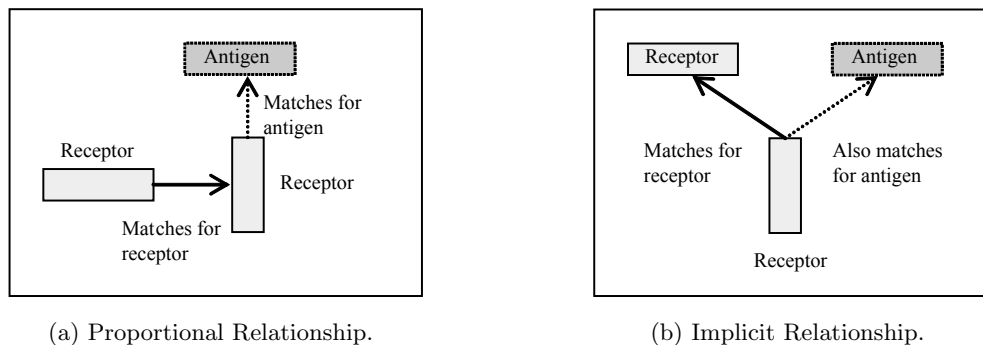


Figure 4.21: Examples of more complex relationships that may be formed given receptor-receptor recognition.

Matching Function

The clonal selection algorithm results in a repertoire that directly (relationship-wise) models features of the input space. When receptors themselves are treated as input signals a variety of representations may be used in the matching. The network theory proposes that receptors match onto a part of the receptor that is distinct for receptors of that lineage, although different from the receptors combining region (part of the other receptor that does the binding). A natural implementation scheme is to assign a random or remapping of the primary bit string to each naïve receptor that, like the primary string, is inherited

and matured by progeny cells. This results in the co-evolution (co-adaptation) of the primary string for activation and the secondary string for secondary effects (for example, see Figure 4.22).

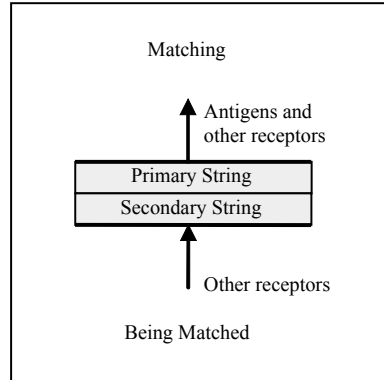


Figure 4.22: Depiction of the two-string representation scheme.

A natural representation (using the same string for both) will likely result in a positive feedback system, where the exposed antigen will be amplified by repeated simulated exposures. Any regularity of the antigenic pattern that is provided on the secondary strings (such as sub-strings or reordering) will likely have this amplification effect, proportional to the fidelity of the mapping of the regularity. This amplification of input signal may be useful to reinforce the acquisition of knowledge. A natural representation will continue to promote the same instigating signal for as long as activated members are promoted to antigen-like status.

Higher-Order Structures

A steady-state application of the approach emphasises some important properties regarding the formation of relationships as networks (graphs) of activation and exposure. In this interpretation, an antigen matches onto and activates a single receptor. The activated receptor is copied from the repertoire and is provided as an input signal in the next cycle. In this next cycle, both an antigen and the previously activated receptor are offered as input signals to the repertoire. A queue-based exposure scheme is used such that the more-recent activated receptors (instigation of relationship) are admitted, and older receptors are discarded. Therefore, the previously activated receptor is discarded, and the newly antigenic-activated cell is promoted to the next cycle as an input signal. The behaviour of this scheme may be described with an example (see Figure 4.23 and Table 4.13).

If a natural mapping is used for *string2*, then the scheme provides a simple signal-reinforcement scheme that provides a linear amplification ($t - 1$) reinforcement of signals past. If a remapping scheme is used for *string2* (such as a random string), then the scheme provides new information in subsequent exposures that may result in the formation of a connection between two different input signals. The simplest example is the selection of *R1a* by *A2*. In a natural mapping, if *A2* selected *R1a* then *A1* and *A2* would be the same input signal, and *R1* and *R1a* are likely the same receptor or clonal siblings (same

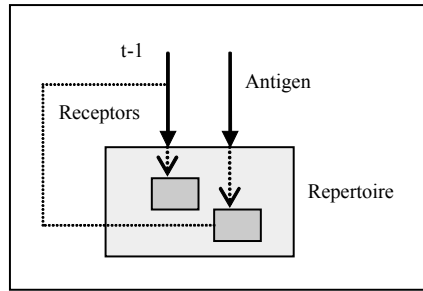


Figure 4.23: Example of a simple two-receptor promotion scheme.

Cycle	Exposed	Activated
1	A1	R1
2	A2, R1	R2, R1a
3	A3, R2	R3, R2a
4	A4, R3	R4, R3a
t	A_t, R_{t-1}	$R_t, R_{t-1}a$

Table 4.13: Exposure-activation scheme for a minimal implementation.

ancestor). In a remapping scheme (such as random), it is possible for the remapped string to match for a receptor that is also matched for by an antigen (see Figure 4.24).

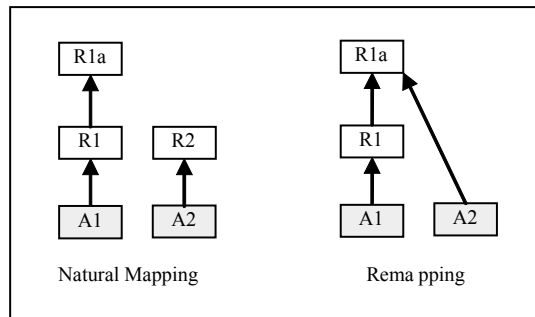


Figure 4.24: Depiction of the example difference between a natural and remapping scheme.

The remapping example provides an example of a relationship that may form between two different antigen ($A1$ and $A2$), facilitated by the remapping of receptors using the two-string scheme. If $A2$ is withdrawn, the relationship is fostered by $A1$, which activates $R1$, which in turn activates $R1a$. If $A2$ is withdrawn for all time, then the relationship will deteriorate due to genetic drift⁵, thus $A2$ provides a correcting influence to the relationship. If $A1$ is withdrawn then the other half of the relationship is reinforced ($R1a$), and again if $A1$ is not returned, then $R1a$ may drift such that $R1$ does not match to it any longer. The example relationship is unidirectional, in that the primary string of $R1a$ is a generalisation of $A2$ and $R1$'s secondary string. A natural extension is to consider the implications of

⁵The relationship is fostered by activation (usefulness), although here genetic drift refers to the adaptation of receptors (their progeny) in response to activation.

$R1a$'s secondary string. If $R1a$ matches for $R1$ then both $R1$ and $R1a$ provides surrogates for $A1$ and $A2$ reinforcing each other and the repertoires relationship between $A1$ and $A2$. It requires that $R1$'s primary string is a generalisation of $A1$ and $R1$'s secondary string. This network may be depicted as follows ($R1a$ is renamed $R2$).

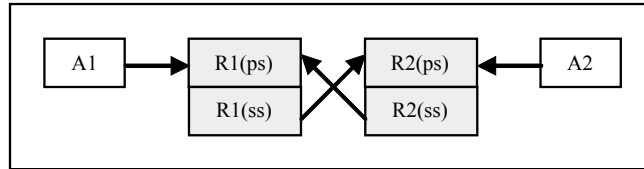


Figure 4.25: Depiction of the relationship between remapped antigen and receptors (ps is primary and ss is secondary string).

A larger queue allows more intra-antigen relationships (larger relationship structures in general), which in turn may prolong the activation of structures, but does not facilitate perpetual (antigen-independent) structure formation. Therefore, *Structure Durability* is defined when the exposure set size defines the sustainability of a structure in active memory (essentially defining short-term memory), in the absence of renewed promotion. Finally, the examples do not take into account the proliferation of selected receptors, therefore the density concerns were subsumed with activation counts that delineated receptor persistence. Additionally, in the structures proposed the secondary strings become surrogates for antigen strings, thus ‘take the form’ of antigen signals. This suggests that a mapping, such as substring’s or reordering of the primary string may be easier for the system to retrofit for such a purpose rather than a random-based mapping (assuming similarity rather than complementarity of the mapping between receptors and antigen).

4.7.2 Recurrent Empirical Study

Aim

The aim of this empirical study is to investigate some of the priming behaviours of the cellular clonal selection strategy with intra-repertoire (network) interactions, specifically the recurrent network model. Toward this end, the study had the following motivating goals:

1. Investigate the effects on the composition and capability of the repertoire under recurrent cellular exposures.
2. Assess a variety of cell promotion pressures in the recurrent network model.

Method

Problems This study used the ACSP-10 problem used for the CCSA empirical study in Section 4.4.1.

Algorithms Algorithm 4.8 defines the recurrent variation of the Network Cellular Clonal Selection Algorithm. Like ECCSA, two exposures occur each epoch, although unlike ECCSA both exposures occur to the same repertoire of cells. The second exposure involves the exposure of the Best Matching Cell (BMC) from the last *antigen* exposure (BMC_{t-1}) to the repertoire. A cellular exposure involves the assessment of the repertoire against the representation in the cell, specifically using the Euclidean affinity function between the representation as is used between an antigen and a cell (defined in Equation 4.2).

Algorithm 4.8: Recurrent Network Cellular Clonal Selection.

Input: $A, T, BMC_{t-1}, N_{selection}, N_{clones}, P_{mutation}$

Output: T_{rs}

- 1 $T_{rs} \leftarrow \text{Exposure}(A, T, N_{selection}, N_{clones}, P_{mutation});$
 - 2 $\text{Exposure}(BMC_{t-1}, T, N_{selection}, N_{clones}, P_{mutation});$
 - 3 **return** $T_{rs};$
-

A Natural Mapping variation of the approach was assessed to provide a baseline of performance referred to as RE-NCCSA-NM. The algorithm responded to each exposure using the replacement mechanisms of RCCSA, and used the following configuration: $N_{cells} = 100$, $N_{selected} = 1$, and $N_{clones} = 5$, assigning 5% of the repertoire per exposed stimulus. This approach was called natural mapping, because indeed no mapping was used, where BMC_{t-1} were exposed directly to the repertoire, likely causing their clonal siblings in the repertoire to respond. A remapping approach was used where each cell was provided with two data strings: a primary string which responded to exposed stimuli, and a secondary string which may be used as an antigenic stimulus (the two-string method described in Section 4.7.1). Table 4.14 defines a set of four different *replacement-based* promotion schemes used with the mapping variation of the algorithm. The schemes are replacement based in that the configurations define the specific pressures applied during the competition for limited position in the repertoire during the integration of clones. *Similarity* refers to the specific representation of cells (primary or secondary) used to match clones with cells in the repertoire, and *Assessment* refers to the affinity scoring (against the primary or secondary string) used during the tournament for the position in the repertoire between a clone and its most similar counterpart in the repertoire. The promotion of the antigen in the secondary representation is expected to promote similar cells in the recurrent exposure. The promotion of the cellular stimulus in the primary representation in the cellular exposure is expected to improve the mapping in the recurrent exposure.

R-NCCSA	Antigen Exposure		Cellular Exposure	
	<i>Assessment</i>	<i>Similarity</i>	<i>Assessment</i>	<i>Similarity</i>
<i>PP</i>	Primary	Primary	Primary	Primary
<i>PS</i>	Primary	Primary	Secondary	Secondary
<i>SS</i>	Secondary	Secondary	Secondary	Secondary
<i>SP</i>	Secondary	Secondary	Primary	Primary

Table 4.14: Summary of the various replacement configurations for the mapped Recurrent Exposure Network Cellular Clonal Selection Algorithm (RE-NCCSA).

Experiment This study used the same general experimental configuration including stop conditions as were used for the CCSA empirical study in Section 4.4.1. Two additional diversity measures were introduced to assess the composition of the repertoire with regard to secondary mappings. The Average Cell Mapped Diversity (ACMD) assesses the average Hamming distance of a given cell to all the other cells in the repertoire with regard to the secondary representation. The Average Network Cell Diversity (ANCD) provided a similar diversity measure although treats both representations of a given cell as one providing an indication of the composition of the repertoire independent of representations. Both diversity measures use the same mechanism as ACD in Equation 4.7 on their representations respectively. Finally, a measure was introduced to assess the state of the mappings between the repertoire and recurrent cellular exposures called the Average Mapped Best Matching Cell Error (AMBMCE), that averaged the Euclidean distance between recurrent cells secondary representation and the BMC in the repertoire for each epoch using the same mechanism as ACE in Equation 4.4.

Results

Table 4.15 provide a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. The non-parametric Mann-Whitney U statistical test was calculated pair-wise for all algorithms. Figures 4.26a, 4.26b, 4.26c show the ACE, ACD, and AMBMCE respectively. Figure 4.27 provide example plots of the response mapping and final repertoire. Results from all example plots are taken from the end of the run with algorithm and problem configurations matching those used during experimentation, and a random number generator seed of 1 and 5 for the algorithm and problem respectively.

System	ACD		ACE		ACMD		ANCD		AMBMCE	
<i>NCCSA</i>	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
NM	93.218	0.248	0.019	0.011	N/A	N/A	N/A	N/A	0	0
PP	91.974	0.413	0.09	0.025	91.777	0.46	183.751	0.669	0.184	0.026
PS	91.912	0.31	0.03	0.013	92.032	0.453	183.944	0.597	0.173	0.025
SS	91.805	0.429	0.211	0.04	91.783	0.419	183.588	0.609	0.147	0.034
SP	91.744	0.421	0.234	0.039	91.745	0.403	183.489	0.542	0.162	0.03
<i>Sign.</i>	True		True		True ^a		True ^b		True ^c	

^aFalse for RNCCSA-PP and RNCCSA-SS, RNCCSA-PP and RNCCSA-SP, RNCCSA-SS and RNCCSA-SP

^bFalse for RNCCSA-PP and RNCCSA-PS, RNCCSA-PP and RNCCSA-SS, RNCCSA-PP and RNCCSA-SP, RNCCSA-SS and RNCCSA-SP

^cFalse for RNCCSA-PP and RNCCSA-PS, RNCCSA-PS and RNCCSA-SP, RNCCSA-SS and RNCCSA-SP

Table 4.15: Summary of results for Recurrent Network Cellular Clonal Selection Algorithm (R-NCCAS) on ACSP-10

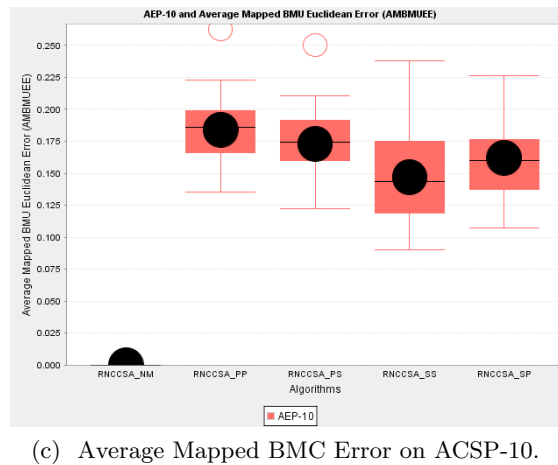
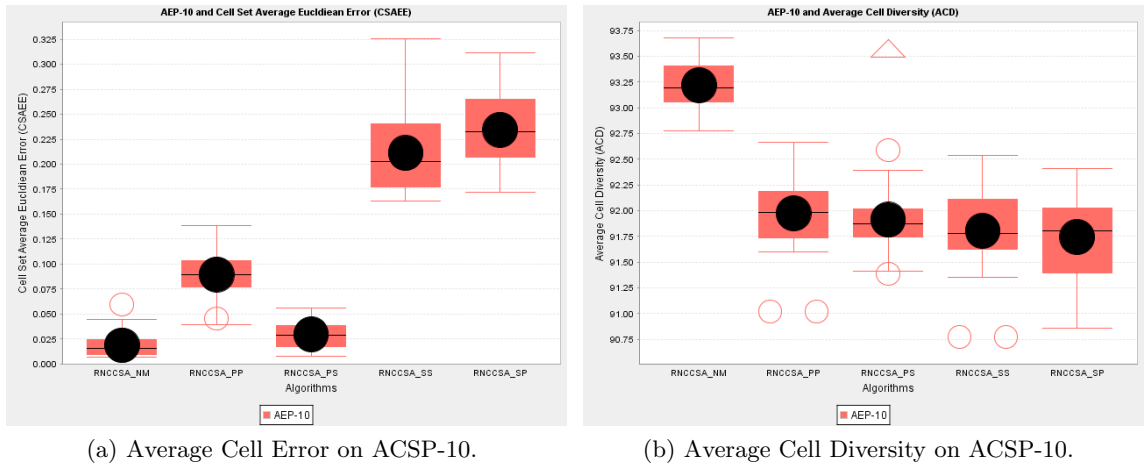


Figure 4.26: Box-and-whisker plot's from the Recurrent NCCSA Empirical Study.

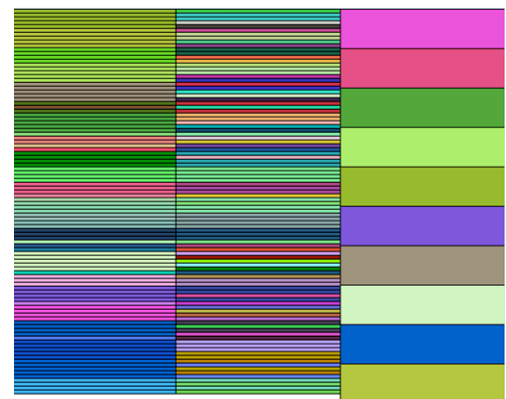
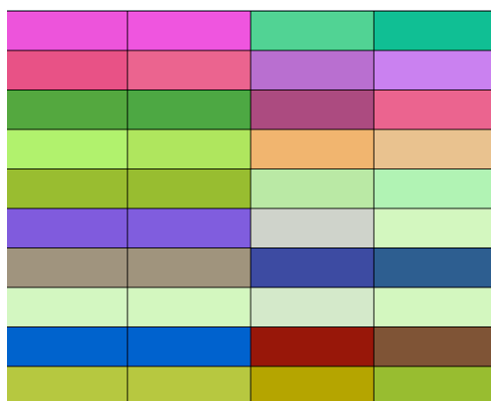


Figure 4.27: Example plots from of the response mappings and final repertoire from R-NCCSA-PS on ACSP-10.

Analysis

This section provides an analysis of the results reported in the previous section in the context of the goals of the empirical study.

Promotion Trends This section considers the results of the recurrent network algorithm in the context of the trends for the different promotion schemes. The natural mapping achieved the best results with regard to ACE as expected, showing a AMBMCE of zero suggesting that cells mapped onto themselves or clonal siblings with no error on average by the end of the run, providing an ideal (in the mapping sense) case for comparison. The various promotion schemes for the mapped recurrent algorithm showed little comparative difference between the three diversity scores: ACD, ACMD, and ANCD. The ACD scores showed a slight decrease with the remapping approaches compared to the natural mapping. The similarity between ACD and ACMD scores suggests that organisation of the information in the primary and secondary representations (pseudo-repertoires) was generally equivalent across the promotion schemes. The telling results were provided in the error measures. The two promotion schemes that promoted the primary representation achieved a much lower ACE as expected than those that promoted the secondary representation against the antigen. Interestingly the configuration that promoted the secondary representation against the cellular exposure (PS) performed better than the promotion of the primary string on both exposures (PP). The error results for the recurrent cellular exposures were poor compared to error scores for viable ACSP (< 0.1), where the promotion of the secondary string against antigen exposures resulted in relatively lower AMBMCE than the promotion of the primary string. A general observation of the state of the repertoire and exposure mappings during the execution of general runs revealed that consistent intra-repertoire mappings were unstable across the various promotion schemes (for example see Figure 4.27). This is reflected in the poor mapping error scores achieved.

Curtailing Instability The complexity of the approach makes analysis difficult although the observed behaviour may be explained by the instability of the general approach. This instability is likely the result of the competition between similar representations for antigen and for cells. For example in the case of PP, the promotion of primary strings in the cellular exposure may be considered the promotion of random secondary strings, with no explicit pressure of adding meaning to the secondary strings of antigen-selected BMC's. This is addressed in PS by the cyclic promotion of secondary strings which are expected to provide pattern-matching loops which counter some of the instability effective ACE. The results for secondary string promotion during antigen exposures suggest that the promotion of BMC's for antigen in the repertoire via proxy ($t+1$ cellular exposure), in particular R-NCCSA-SP, is insufficient for the repertoire to address the ACSP. This suggests that the de-coupled promotion of antigen-BMC within the repertoire may result in conflicting competition between primary and secondary representations.

Conclusions

This section summarises the findings of the empirical study into the Recurrent Network Cellular Clonal Selection Algorithm in terms of the primitives that were the focus of the study and the expectations that motivated the study.

1. *Primitives*

- (a) The Recurrent NCCSA provides a viable approach for investigating the subtle effects of intra-repertoire remapping of exposure signals and the subtleties of competition in such a mapping.

2. *Behaviours*

- (a) The recurrent exposure of cells via a natural mapping provides a reinforcement of antigenic signals
- (b) The recurrent exposure of remapped cells requires careful consideration of the competition between representations to avoid conflicts in such competition.

4.7.3 Dual Exposure Empirical Study

Aim

The aim of this study is to investigate the antigen-driven formation of intra-repertoire structures as outlined in Section 4.7.1, and Figures 4.25 and 4.25. Specifically, this empirical study is concerned with further investigating the subtle application of replacement pressure as was considered in the recurrent model, although toward the formation and promotion of antigen-dependant rather than cellular-dependant structures. Toward this end, this study has the following goals:

1. Assess the composition and capability of the repertoire under a dual-exposure model.
2. Investigate the formation of antigen-dependant high-order structures via the application of intra-repertoire promotion schemes.

Method

Problems This study used the ACSP-10 problem used for the CCSA empirical study in Section 4.4.1.

Algorithms Dual exposure refers to the way in which the repertoire interacts with the antigen, specifically two at a time (requiring the ACSP to be comprised of an even number of CSP). Dual exposure may be considered a variation on the linear Cellular Exposure Regime considered in Algorithm 4.2 where the first antigen in the pair is always exposed to the repertoire's primary string representation, and the second antigen to the secondary string representation (see Algorithm 4.9). Such exposure governs the selection of the BMC returned as a response from the repertoire to satisfy the requirement of the exposure represented.

Algorithm 4.9: Dual Exposure Network Cellular Clonal Selection.

Input: $A_1, A_2, \mathbb{T}, N_{selection}, N_{clones}, P_{mutation}$ **Output:** T_{rs}

- 1 $T_{rs} \leftarrow \text{Exposure}(A_1, \mathbb{T}, N_{selection}, N_{clones}, P_{mutation});$
 - 2 $T_{rs} \leftarrow \text{Exposure}(A_2, \mathbb{T}, N_{selection}, N_{clones}, P_{mutation});$
 - 3 **return** $T_{rs};$
-

As with the recurrent approach (defined in Algorithm 4.8) a series of configuration schemes are considered that provide subtle adjustments to the replacement competition of aggregated clonal sets against the repertoire. Unlike the recurrent approach, these configurations are concerned with the effects of varying or decoupling the context for similarity and assessment used during replacement (see Table 4.16). In this case, *Similarity* refers to the string representation (primary or secondary) used to locate the most similar repertoire member for a clone to compete with, and *Assessment* refers to which antigen the representation is assessed against (primary antigen or secondary antigen). As mentioned, the first exposure is always assessed against the primary representation, and the second against the secondary representation.

DE-NCCSA	First Exposure		Second Exposure	
<i>Name</i>	<i>Assessment</i>	<i>Similarity</i>	<i>Assessment</i>	<i>Similarity</i>
<i>PP-SS</i>	Primary	Primary	Secondary	Secondary
<i>PS-SP</i>	Primary	Secondary	Secondary	Primary
<i>SS-PP</i>	Secondary	Secondary	Primary	Primary
<i>SP-PS</i>	Secondary	Primary	Primary	Secondary

Table 4.16: Summary of the various replacement configurations for the Dual Exposure Network Cellular Clonal Selection Algorithm (DE-NCCSA).

Experiment This study used the same general experimental configuration including stop conditions as were used for the R-NCCSA empirical study in Section 4.7.2.

Results

Table 4.17 provide a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. The non-parametric Mann-Whitney U statistical test was calculated pair-wise for all algorithms. Figure 4.28 shows the Average Cell Error (ACE) for the four different dual exposure configurations. Figure 4.29 provides example plots of the four schemes at the end of a run on ACSP-2. The algorithms and problem use a random number generator seed of 1 and 4 respectively. The configuration of each algorithm was reduced to $N_{cells} = 20$ to ensure a proportional distribution of cells in the repertoire. The smaller ACSP-2 was chosen to clearly highlight the formation and/or lack of formation of bi-antigen structures in the primary and secondary representations.

System	ACD		ACE		ACMD		ANCD	
	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
PPSS	89.213	0.647	0.02	0.02	89.069	1.001	178.282	1.267
PSSP	89.231	0.678	0.018	0.016	89.382	0.766	178.613	1.09
SPPS	89.289	0.705	0.211	0.05	89.551	0.751	178.839	1.233
SSPP	89.301	0.876	0.214	0.041	89.581	0.835	178.882	1.382
<i>Sign.</i>	False		True		False ^a		False	

^aTrue for DE-NCCSA-PPSS and DE-NCCSA-SSPP

Table 4.17: Summary of results for the Dual Exposure Network Cellular Clonal Selection Algorithm (DE-NCCSA) on ACSP-10.

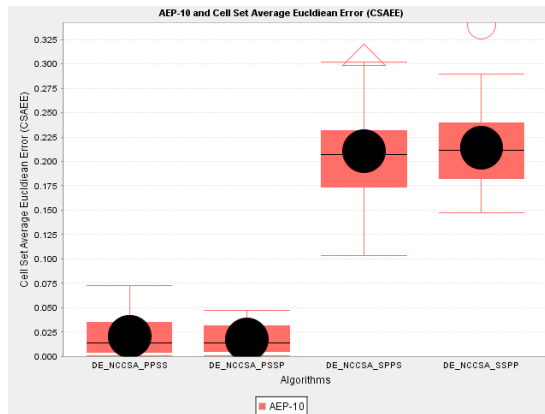
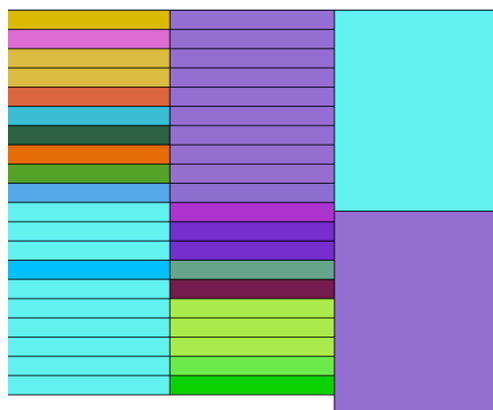
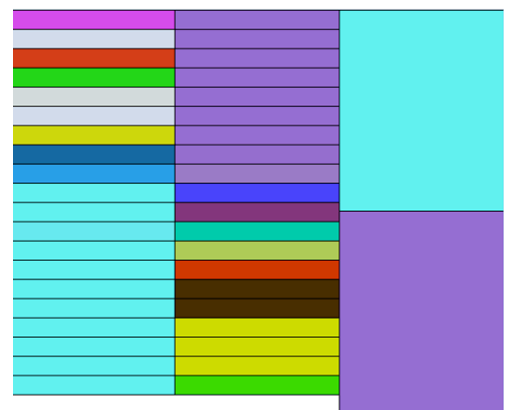


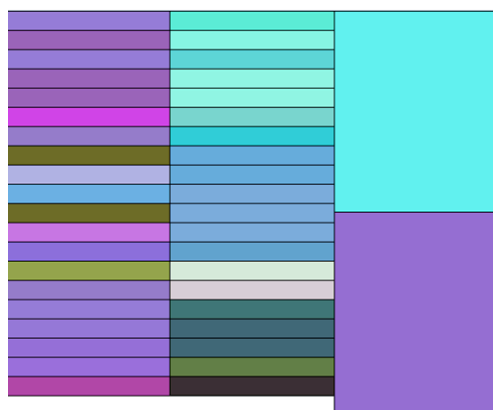
Figure 4.28: Box-and-whisker plot of the Average Cell Error for the DE-NCCSA schemes on ACSP-10.



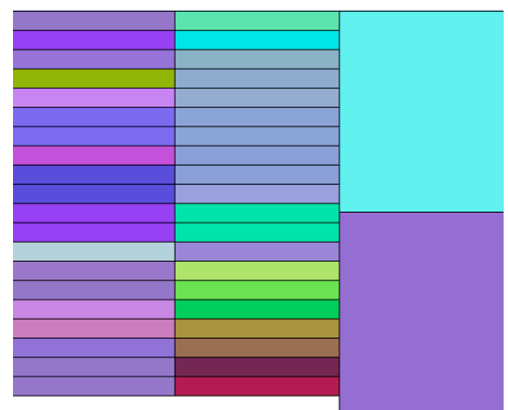
(a) DE-NCCSA-PPSS.



(b) DE-NCCSA-PSSP.



(c) DE-NCCSA-SPPS.



(d) DE-NCCSA-SSPP.

Figure 4.29: Example repertoire plots of the four Dual Exposure NCCSA configurations on ACSP-2 showing primary and secondary representations and antigen left to right.

Analysis

This section provides an analysis of the results reported in the previous section in the context of the goals of the empirical study.

Promotion Trends This section considers the promotion trends for the four different configuration schemes of the dual exposure algorithm. The results show little difference with regard to the collected diversity measures. Specifically there was no significant difference in the ACD and ANCD, and generally no difference in ACMD. As with the empirical study into recurrent network the effect of the promotion schemes was on the repertoire capability. For both schemes where the exposed antigen is promoted via assessment (PPSS and PSSP), the ACE demonstrated viable results against the ACSP-10. This suggests that regardless of whether the cells are competing with repertoire members for the same primary or secondary antigen, replacement tournaments must be decided based on the representation used to satisfy the antigen. Interestingly, this effect can be enhanced by competing with repertoire members with similar representations that are not the representation used for selection and assessment. Specifically, PSSP resulted in a slightly (although significantly lower) ACE than PPSS. The example repertoire plots provided in Figure 4.29 highlight the behaviour of this class of NCCSA. Specifically, the plots suggest that the promotion via assessment against selecting antigen results in the specialisation toward the respective antigen, although decoupled structures. In both cases where replacement competition occurred against the complementary antigen (SPPS and SSPP), the plots strongly suggest that such structures were formed in the repertoire (top of each plot for SPPS and SSPP). The ACE results recorded suggest that such structures were of a much lower capability than the promotion of selecting antigen, which can be explained by the reversal of the specialisation of the representations observed clearly in the example plots (primary string to second antigen and secondary string to first antigen as opposed to the reverse case expected by the exposures). Interesting this was not reflected in the ANCD. This suggests that the cross-promotion can form such structures, although the selective pressure of the antigen alone was insufficient to promote improved speciality for exposed antigen.

Conclusions

This section summarises the findings of the empirical study into the Dual Exposure Network Cellular Clonal Selection Algorithm in terms of the primitives that were the focus of the study and the expectations that motivated the study.

1. *Primitives*
 - (a) The dual exposure model provides a viable complementary approach to recurrent for investigating promotion pressures in NCCAS.
2. *Behaviours*

- (a) Antigen-dependant cross-exposure structures can be formed within the repertoire by biasing the integration of clones toward the complementary parts of the structure.
- (b) Such structures cannot be formed by only biasing the similarity pairing alone, requiring a biasing in the assessment used for competitive tournaments in similarity pairs.

4.8 Chapter Summary

4.8.1 Paradigm Review

The *Cellular Clonal Selection Paradigm* is a re-definition of the current state of clonal selection algorithms. This rephrasing is differentiated from the current state of the field in the following ways: (1) the separation of the concerns into a system (cellular algorithm) and an environment (antigenic environment), (2) the explicit definition of the computational properties of clonal selection as an adaptive knowledge acquisition strategy, and (3) the restriction of the concerns of the clonal selection systems and environments to the cellular-level. The cellular level is defined as the maintenance and operation of clonal selection on a repertoire of discrete cells, not limited to the antigenic, molecular, and cellular interactions that may occur in such a repertoire.

4.8.2 Paradigm Trends and Findings

The following summarises the important principles and findings from the definition and investigations into the Cellular Clonal Selection Paradigm:

Primitives

1. *Cellular*: Increasing the proportion of the repertoire dedicated to an antigen and/or the number of clonal trials, increases the specialisation and therefore the capability of the repertoire for an antigenic environment.
2. *Replacement*: Repertoire-wide integration of an aggregated clonal set using affinity tournaments between similar cells, with sibling exclusion provides a controlled specialisation of a footprint of concurrent redundant perspectives in the repertoire for each antigen.
3. *Degenerate*: The strong selection of activated cells is exploited by cellular clonal selection to constrain the inherent polyclonal activation of the repertoire to each antigen exposure. The explicit aggregation of repertoire responses is an artefact of the exposure paradigm constraining the polyclonal activation of the degenerate repertoire, an important principle not limited to DCCSA.

Extensions

1. *Spatial*: Constraining the integration of clones to the neighbourhood of the antigen selected cells provides a localising spatial pressure for areas of a spatial repertoire to take responsibility for specific antigenic signals.
2. *Mediated*: The use of cell casts to mediate responses to antigen relies heavily on an antigenic mapping between the casts and an oligoclonal (small founding set) relationship between the feature detector and concept formation tiers.
3. *Network*: Intra-repertoire cell interaction requires careful management of the pressures that govern the differential allocation of resources. The cross promotion can facilitate the formation of antigen-dependant structures across multiple exposures although at the expense of specificity.

4.8.3 Integration

The Cellular Clonal Selection Paradigm defined and investigated in this chapter provides a bedrock in understanding of clonal selection as an adaptive strategy in the context of an antigenic exposure paradigm specialised in the colour space domain. The features and follow-on information processing characteristics set the investigated approaches apart, the relative merits of which are considered in the context of application problem domains in Section 8.3 and Section 8.4. The chapters that follow build upon this bedrock providing a Tissue (Chapter 5) and Host (Chapter 6) constrained perspective on clonal selection as an adaptive knowledge acquisition strategy, that collectively with the Cellular perspective provide an integrated hierarchical framework for clonal selection algorithms (Chapter 7).

Chapter 5

Tissue Clonal Selection

5.1 Chapter Overview

This chapter exploits the principles and findings of the cellular paradigm by taking them for granted, instead investigating clonal selection where the cellular paradigm is a primitive component in a broader *Tissue Clonal Selection Paradigm*. Section 5.2 reviews the acquired immune system from the perspective of the tissues that provide the structure and manage the function of the lymphocytes central to the clonal selection theory. Section 5.3 considers an abstraction of the reviewed physiology and tissue-based immunology and the basis of the tissue paradigm. Most importantly in this abstraction are the spatial and temporal considerations of the discrete exposure of repertoires of cells to information, and the tissue architectures that provide base patterns for the design of inspired algorithms and implementations. The paradigm is realised as the investigation of Tissue Clonal Selection Algorithms in the context of an Infection Antigenic Exposure Problem, that provides the basis for empirical investigation in the remainder of the chapter. Tissue-based immunology strategies for organising information are defined and empirically investigated in an colour space specialisation of the infection problem, resulting in a series of algorithms: the Minimal, Recirculation, Homing, and Inflammation Tissue Clonal Selection Algorithms. The investigation of these algorithms confirms expectations regarding the natural organisation of information based on the regularity and consistency of exposure, as well as important findings as to how the defined strategies may be exploited when such regularities and consistencies are not known *a priori*.

5.2 Physiology of Lymphocyte Migration

The white blood cells and immunoglobulin within a host that comprise the recognition and response component of the acquired immune system are *mobile*. There are pools of lymphocytes that recirculate the blood stream and the lymphatic system. More interestingly there are lymphocytes that selectively home to tissues close to where the cells were created and there are pools lymphocytes that are recruited to sites of infection and inflammation. The lymphoid tissue of the lymphatic system provides the structural scaffold for the mobility of these cells and related molecules, although interestingly the migratory

behaviour is controlled at the finest level: through localised molecules and receptors on the cells themselves. The migration of cells involved in the immune system is highly complex and is still not completely understood. This section reviews some superficial features of lymphocyte migration in the immune system with the interest of extracting principles for use in clonal selection-based Artificial Immune Systems. Lymphocyte migration is reviewed in the context of the human and related mammalian immune systems, providing perspectives of cell movements from the lymphatic system, movement types, cell types, and cell classes.

5.2.1 Lymphatic System

The lymphatic system is a complex collection of lymphoid organs integrally related to the functioning of the immune system and responsible for the transport and filtering. Lymph is a clear bodily fluid that surrounds all tissue that forms when proteins and cells leak out of blood carrying venules and capillaries. The lymph carrying capillaries flow unidirectionally draining the lymph back to lymphoid tissues that make up the lymphatic system. Also carried in this lymph are antigens that may have entered the organism. The lymphatic system acts as a secondary circulation system (to the cardiovascular circulatory system) transporting lymphocytes between lymphoid organs, and carrying antigen to lymphoid organs for interaction with lymphocytes. There are two types of lymphoid organs (1) the primary or central tissues which are the sites of lymphocyte formation (the thymus and bone marrow), and (2) the secondary or periphery lymphoid organs where immune response take place (the spleen, lymph nodes and gut-associated lymphoid tissues). One may consider the rest of the bodies' tissues as tertiary lymphoid tissues which normally only contain few lymphoid cells, although on infection and inflammation, manage to recruit many immune cells.

Central Lymphoid Tissues

Primary lymphoid organs consist of the thymus and the bone marrow, and are responsible for differentiating stem cells into pre-immune cells (pre-B-cells and pre-T-cells) in a process that is believed to be independent of antigenic stimulation. The bone marrow provides a micro-environment for the production of blood cells. B-cells migrate to secondary lymphoid organs, whereas T-cells migrate to the thymus. The thymus provides an environment for the further antigen-free maturation of T-lymphocytes. Cells proliferate and differentiate in a process involving negative and positive selection where the majority of produced T-cells never leave the thymus. Those cells that do survive (less than 5% per day) may join the recirculating lymphocyte pool, or migrate to secondary lymphoid tissues. There are two types of primary lymphoid tissues, as follows:

- *Bone Marrow*: Tissue located inside large bones responsible for the production of many different blood cell types not limited to lymphocytes (white blood cells). Provides an environment for differentiation of naïve B-cells.

- *Thymus*: An organ located behind the sternum, providing an environment for the differentiation of naïve lymphocytes from bone marrow. These lymphocytes differentiate into T-lymphocytes, which involves a negative selection maturation process.

Peripheral Lymphoid Tissues

The secondary lymphoid tissue is responsible for collecting antigens at points in the organism where they may enter, and facilitate their exposure to recirculating lymphocytes. For example, the lymph nodes are primarily responsible for filtering the lymph for antigens, the spleen is responsible for filtering the blood for antigens, and the tonsils for the respiratory system. Goodnow puts this concisely suggesting that the “. . . *central function of secondary lymphoid tissues is filtering: collecting blood-borne, lymph-borne, or mucus membrane antigens and holding these so that they can be surveyed by immune cells before being destroyed*” [181] (page 6). The secondary lymphoid organs are as follows:

- *Tonsils*: A collection of lymphoid tissue on the side of the throat responsible for providing lymphocyte access to and protection of the respiratory system from antigen.
- *Lymph Nodes*: Small gland like structures found throughout the body that filter the lymph for foreign antigen material, which are then presented to lymphocytes and other immune system cells.
- *Peyer’s Patches*: A collection of lymphoid tissue found in the lowest section of the small intestine. There are numerous instances of these patches in the intestine and they act like lymph nodes, and provide centres for lymphocytes protecting the gastrointestinal tract.
- *Spleen*: An organ in the upper abdomen, responsible for the destruction of old red blood cells. It is the only lymphoid organ that crosses the blood stream, and thus provides a site where lymphocytes can interact with antigens carried in the blood.
- *Lymphatic vessels*: Vessels throughout the body that carry lymph. They are responsible for transporting lymph from tissues to the blood (vascular circulatory system) and to the lymph organs.

In addition, the secondary lymphoid tissue provides a suitable micro-environment for the development and maturation of an immune response, such as the formation of Germinal Centres (GC’s) for B-cell clonal expansion and affinity maturation. Germinal Centres are little understood dynamically generated regions in lymphoid tissue where activated B-lymphocytes clonally expand, undergo hypermutation, and ultimately differentiate into plasma and long-lived memory B-lymphocytes [37, 387]. GC’s are founded by a small number of activated B-cells (oligoclonal), which are in turn co-stimulated by helper T-cells [275]. It is believed that there are repeated rounds of selection, expansion, and mutation of B-cells within the GC. The majority of the produced B-cells may be plasma cells and have a lower affinity than the founding cells, although a few higher affinity clones are preferentially selected to survive as plasma or memory cells [277, 292]. Those cells

that are worse die due to the lack of positive selection. GC-like structures may also occur outside of lymph nodes, such as sites of infection and allergic inflammation [182].

The geography of the lymph nodes and other secondary lymphoid tissues plays an important structural role for lymphocyte selection and response, for example Zinkernagel, et al. suggests that the “...*secondary lymphoid organs present antigen optimally and enhance the chances of specific antigen encounter and specific cell interactions.*” and that “...*these requirements render chance encounters of antigen by lymphocytes and activation elsewhere extremely inefficient and of no biological relevance.*” [446] (page 201). It is likely that the organisation and function of the secondary lymphoid organs has been optimised for recirculating lymphocytes to efficiently detect and remove pathogens from the organism, for example Fu and Chaplin suggest that “...*secondary lymphoid organs are thought to be organised into structures that optimize cellular interactions that support the efficient removal of unwanted pathogens*” [156] (page 400). Finally, in addition to providing a scaffold for lymphocyte recirculation, the secondary lymphoid tissues possess a high regenerative capacity. When circulation routes are disrupted or severed, the tissue is able to adaptively regenerate a connection between the effected nodes [309].

Summary

There are no lymphocytes or lymphocyte recirculation without the lymphatic system. The tissues are responsible for the development and management of the cells throughout their life cycle. The following summarises the lymphatic system’s governing role over lymphocytes:

1. *Formation*: Provide an optimal micro-environment for the production and maturation of lymphocytes.
2. *Presentation*: Collect diverse populations of lymphocytes into organ systems that drain antigens from entry points into the organism.
3. *Regulation*: Regulate interactions of different classes of lymphocytes in drainage organ systems, such as the arrangement of B-cells and T-cells.
4. *Dissemination*: Disburse effector elements of the immune response throughout the organism, and disseminate and amplify the immune response systematically throughout the lymphatic system.

For more information regarding the anatomy and physiology of the human lymphoid system see any current anatomy and physiology text on the subject, such as [280]. Other references used include lymphoid organ summaries by de Castro and Timmis [109] (page 71), Anderson [19], and Swartz [377]. See Andrian and Mempel for a review of lymph node anatomy and physiology in the context of cell migration [411].

5.2.2 Lymphocyte Mobility

Cell movement is the basis of immunology. It is involved in inflammation, differentiation, adhesion, recruitment, and required for cellular and molecular interactions. The recir-

ulation of lymphocytes was confirmed experimentally almost 50 years ago by Gowans and colleagues [183, 184], although the study of the topic is often neglected in favour of cell interactions [442, 322]. Interestingly much of the modelling work with populations of lymphocytes does not take into account the spatial heterogeneity of the lymphatic system.

The adhesion properties of cells plays a critical controlling role in the movement of lymphocytes (see Anderson for a treatment [19]). Adhesion is used by cells to crawl through tissues, and is used by activated cells to stick to the molecule that activated them. Adhesive receptors (so-called ‘homing receptors’) and chemicals are little understood and are a topic for intense study. The adhesive characteristics of a cell are believed to be selected for and differentiated along with cell antigenic receptor characteristics. These adhesive interactions provide bottom up control at the finest level [324], believed to be the basis for homing, recruitment and may ultimately control the extent and scope of the immune response [416]. A procedure or series of adhesive-based decisions have to be made for a lymphocyte to be recruited into tissues, in particular from recirculation in the blood into lymphoid tissues. This process is called the multi-step extravasation or molecular regulation procedure, which has the following four steps: (1) primary cell adhesion, (2) rapid cell activation, (3) activation dependant arrest, and (4) diapedesis (the movement into surrounding tissue).

Movement Types

Physically, there are two ways lymphocyte cells can move: *crawling* on tissues and *flowing* between tissues. In crawling, cells use chemical receptors to adhere to their surroundings and use this method to slowly migrate through tissues, a place where they spend most of their time given the slow pace of movement. The second mode is movement in fluid space such as in blood or lymph where cells are capable of moving a lot faster, covering great distances within the organism. The process of lymphocyte migration was originally considered to be random, although it is now known that this is not necessarily the case. Lymphocytes may be preferentially recirculated, and are able to home in and target specific tissues. In addition, memory lymphocytes show different migration behaviour to naïve lymphocytes. For example, memory T-cells preferentially migrate to non-lymphoid tissues. If a memory cell was created in a lymph node near the skin, the cell will preferentially migrate into skin tissue near the lymph node, if created in a lymph node near the gut, it will migrate to neighbouring gut tissue. Naïve T-lymphocyte cells are activated by dendritic cells and educated as to the chemical properties of the site of infection. These prepared T-cells recirculate around the host organism in search of their cognate antigen, and upon detecting the chemical properties of the site of infection, home into the tissue. The directed trafficking behaviour is called T-lymphocyte homing [67, 66, 324, 416], and the information which controls the where T-cell traffic is expressed as receptors on the surface of the cell for the chemical properties of the site of infection [139, 343].

The immune system maintains a pool of recirculating lymphocytes that cycle around the blood and the lymphatic system. The pace of recirculation is high, the number of lymphocytes entering the blood from the lymph each day is 10 times the size of the

recirculating lymphocyte pool [19]. The number of lymphocytes recirculating at one time may be anywhere from 1%-2% of all lymphocytes in the body in young adult animals [442, 371]. Lymphocytes may only stay in circulation in the blood for about 30 minutes. Trepel provided a seminal, although outdated extrapolation of the number and distribution of lymphocytes in man [398]. Such numbers may be useful for a general guide, as follows: 2.2% in the circulating blood, 41.3% in the lymph node and tonsils, 15.2% in the spleen, 4.3% in the gut-associated lymphoid tissues, 10.9% in the thymus, 10.9% in the bone marrow, and about 15% in other tissues. One may summarise the primary types of lymphocyte migration, as follows:

- *Trafficking*: The directed (non-random) movement of cells from tissues, blood, or lymph. May refer to a cells trafficking route as it homes to a specific region in the body. For example: pre-B-cells and pre-T-cells migrate to secondary lymphoid tissues to further differentiate and mature.
- *Recirculation*: The movement of lymphocyte cells around the body from lymphoid tissue, to the blood, to the lymph, and back to lymphoid tissue to repeat the process (circulation or rolling lymphocytes). Naïve and memory cells are the main recirculating lymphocyte types.
- *Recruitment*: Accumulation (sequestration) of cells such as a site of infection or tissue damage, such accumulation may occur through chemotaxis (directed movement in response to a chemical gradient).
- *Homing*: The directed (preferential tendency) of lymphocytes activated in a particular region of the body, to return to that part of the body (localisation). May refer to the arrival of lymphocyte to lymphoid or non-lymphoid tissue from the blood stream. Also called tissue-selective trafficking. Those cells with a memory of where they were differentiated may localise back to these regions after a period of recirculation.
- *Stationary*: These are cells that selectively arrest their movement, or do not move from the location of their differentiation. For example in the recruitment of cells, the differentiation of cells such as plasma B-cells created in germinal centres that release large amounts of antibody, and sentinel effector cells.

Lymphocyte Types

There are both circulating and non-recirculating populations of B- and T-cells. Although the behaviours of both cell types are tightly interrelated, both have differing migration behaviours during their development and life cycle. The general migratory life cycle of T-lymphocytes is as follows: pre-T-cells migrate from bone marrow to the thymus for maturation, surviving T-cells may migrate to tissues and becomes sentinel cells. Other cells may recirculate and seek activation in secondary lymphoid tissue. Memory T-cells preferentially migrate to non-lymphoid tissues. For B-lymphocytes: pre-B cells migrate from the bone marrow to secondary lymphoid tissues. Some cells will be activated by

antigen in the secondary tissue and differentiate into plasma and memory cells. Other cells will be activated and migrate to the spleen before differentiating.

There is commonality in the development of both B- and T-cells, specifically in terms of the generalised classes of lymphocyte they may differentiate into. These include the classes: *naïve*, *effector*, and *memory*. Naïve cells are untested lymphocytes that seek a potential cognate antigen. Effector and memory cells result from the union of naïve cells and antigen. Effector cells, such as Helper T-cells and plasma B-cells remain in lymphoid tissue. Memory cells make up the majority of the recirculating pool and continue to circulate between the bloodstream and the lymphatic system, at a higher rate than naïve cells. One may define the migratory behaviours of the generalised lymphocyte classes or casts, as follows:

- *Naïve Lymphocytes*: Recirculate between blood and lymphoid tissue, primarily involved in responding to antigen presented in lymph nodes, differentiating into effector and memory cells. Relatively homogeneous in their recirculating behaviour. Naïve cells compete with each other for activation and contribution into the memory recirculating pool.
- *Effector Lymphocytes*: Typically do not recirculate, stay at the site of differentiation, such as plasma B-cells, which differentiate from naïve B-cells in lymph nodes or the spleen.
- *Memory Lymphocytes*: Recirculate around blood and lymphoid tissue, but also extravasate to other non-lymphoid tissues. They are heterogeneous in their recirculation behaviour, with restricted and selective recirculation circuits. They home to areas where they are most likely to encounter, or re-encounter their cognate antigen. The number of memory cells is maintained within a moderate range during adult life.

Cell Mobility as a Strategy

The anatomy and physiology of the immune system may be thought of a defence strategy for the organism, and cell mobility is an integral part of that strategy. Picker and Butcher propose such an evolved strategy where the solution to a complex antigenic environment is “... *to compartmentalize the principle functions of the lymphoid system into discrete organs and tissues in the body, and to connect and unify these organs through the operation of an elegant system of targeted lymphocyte trafficking and recirculation.*” [324] (page 562). Some resources of the immune system are fixed in position and distributed throughout the body such as lymphoid tissues, and so called sentinel T-cells. Draining lymph into lymphoid tissue provides a way to localise antigen. Keeping the majority of the immune responses in movement (the detectors and effectors) provides a patrolling approach to rapidly deploy resources to wherever they are needed. The immune system has to produce many cells to detect unknown antigen, and must collect antigen in such a way that rare detector cells are given a chance to detect them. It must provide a micro-environment for controlled proliferation and differentiation, and disperse effectors to where they are

needed. Further, Rosen, et al. consider a recirculating population as an efficient approach in the context of alternative designs where “*Recirculation contributes to the efficiency of peripheral immune responses by maximising stochastic probabilities of a productive meeting between antigen and its cognate T-cell receptor*” [339] (page 161). This tissue-recirculation adaptive strategy may be considered to have the following general properties:

- *Movement*: Different cells and cell types have differentiated migration potentials (so-called ‘differential migration’). The lymphatic system provides a pathway, which allows lymphocytes to recirculate between blood, peripheral tissues, and lymph nodes.
- *Balance*: The so-called stirring or mixing effect of the recirculating lymphocyte pool facilitates the survival of the fittest, that is the survival of the most appropriate clones in the repertoire. The segregation in the repertoire prevents competition between unrelated lymphocyte subsets. Tissues selectively facilitate the segregation of lymphocytes to where they are needed. Such a balance between recirculation and homing may be critical to lymphocyte homeostasis [66].
- *Coverage*: A repertoire of cells is distributed and rotated in an attempt to expose rare (specialised) cells to as many opportunities for activation as are available. This may be considered the maximisation of coverage, where the rolling population seeks to maximise the probability of specialised cells being exposed to their cognate antigen. Large-scale migration allows a wide repertoire of specific lymphocytes to exist at a very low frequency, yet still function efficiently. An example of this is T-cells. Perhaps one in ten thousand T-cells released from the thymus will have a receptor that can detect a given antigen. This means that statistically at least ten thousand inappropriate cells must encounter an antigen before the right one does [19].
- *Surveillance*: Lymphocytes may be thought of as patrolling the organism, seeking cognate antigen or recruitment. They may also be thought of as monitoring the organism for change to self for self-antigen such as tumours in what is known as *immunosurveillance*.
- *Amplification*: Recirculation permits recruitment, the dynamic reallocation (localisation) of specialised resources to locations where they are needed. Further, the system is capable of further specialising the recruited resource and disseminating that information.
- *Dissemination*: The pathway provided by the lymphatic system provides a highway to rapidly distribute information such as the immune effector cells, antibodies, and long lived lymphocyte cells that retain a refined memory of an antigenic exposure for a future rapid response.
- *Alternative*: It provides a distributed mobile defence strategy and is different to other subsystems of the body such as the nervous system, which is distributed but structurally fixed. Further, if the repertoire did not recirculate then each lymph node would be an isolated (island) repertoire. Thus, each such isolated repertoire

would either have to be as large as the entire recirculating pool, or a smaller pool resulting in less antigenic coverage.

Problems and Changes to Migration

The rate and nature of lymphocyte migration does not remain constant, and in fact, it may be desirable in certain circumstance to manipulate the recirculation behaviour of lymphocytes. Alternatively, the recirculating behaviour of lymphocytes that provides a dynamic defence may also facilitate significant problems for the host organism. During an immune response, the lymph nodes may rapidly recruit cells from blood, temporally blocking them from leaving the tissue. This in addition to clonal expansion causes the nodes to swell in size and become sore. This swelling of the glands is a common sign of illness. It highlights the point that when an antigenic stimulus is detected, it has a profound effect on the entire system, affecting production and circulation behaviour of lymphocytes. There are some diseases that disrupt and decrease the migration and homeostasis of lymphocytes [20], thus it may be desirable for the system to temporally increase the rate of cell migration. Some approaches may include increasing blood flow, increasing the expression of adhesive chemicals, and manipulating the tissues where immune responses take place. Alternatively, there are diseases where it is desirable to decrease the recirculation of lymphocytes. Examples include when tissues are transplanted and autoimmune diseases. In both cases the immune cells seek to destroy healthy tissue, thus strategies focus on affecting the recruitment of cells to these areas.

The lymphatic system may facilitate the spread of disease in the body. For example, cancers and tumour cells are able to spread quickly throughout the lymph nodes of a host. Bacteria can use the lymphatic system to disseminate throughout the body. An example mentioned by Andrian and Mempel is the bacterium *Yersinia pestis* one of the most devastating bacterial pathogen in human history (the cause of the bubonic plague) [411]. It manages to get in to tissue through flea bites and moves to the lymphatic system where it proliferates, quickly overwhelming lymph nodes and spreading throughout the body. Additionally widely cited examples include the human immunodeficiency virus (HIV) that can lead to the condition of acquired immunodeficiency syndrome (AIDS), *Mycobacterium* that causes diseases including tuberculosis and leprosy, and the bacterium *Bacillus anthracis* that causes the disease Anthrax.

5.3 Abstract Tissue Paradigm

Section 5.2 demonstrated that the tissues that house lymphocytes govern their life cycle, and that together they form a high-level strategy for partitioning and managing the systems interaction with antigen. This section abstracts and elaborates on this tissue-centric adaptive strategy and proposes a *Tissue Clonal Selection Paradigm* that takes the clonal selection of the Cellular Paradigm for granted, and focuses on the inter-tissue interaction and information sharing provided by bottom-up cell migration mechanisms.

5.3.1 Model Components

Architecture Components

Discretising the repertoire refers to transitioning from an abstracted clonal selection single and self-contained repertoire, to multiple of such repertoires. The direct approach proposed in this work is to focus on the lymphoid tissue as the nodes in a distributed repertoire and the vascular system (the veins and such that transport the lymphocytes between the lymphatic tissues) as the edges of the network. This graph-based network formalism provides an architecture that separates the transport of lymphocytes on the edges of a network structure from the cellular-based information processing (antigen-lymphocyte reactions) that occur in the nodes of that network. Refer to Figure 5.1 for a conceptualisation of this relationship. A ‘*tissue*’ is a generic node in the network topology model, and may be specialised to various types of tissue with differentiated functional roles. A tissue contains a repertoire of lymphocytes and facilitates antigen-lymphocyte interaction. Therefore, a tissue node is a reformation of what might be considered the repertoire in the Cellular Paradigm. Tissue nodes may be functionally specialised after the general tissues classes in the lymphatic system, such as primarily, secondary, tertiary lymphoid, and the germinal centre (Section 5.2.1).

Cellular Components

Lymphocyte cells are discrete information packets that collectively in a repertoire or repertoires are an internally generated model of the external domain. Antigen may also be considered information packets and/or direct interaction with the domain, themselves being of external origin with regard to the lymphocyte model. The cellular components are the interest of the Cellular Paradigm, although this section exploits the rise in the level of abstraction provided by tissue to re-define cells from an amorphous information packet into functional classes outlined in Section 5.2.2. The classes define different information process behaviour, particularly with regard to the mobility of the cells within the network topology architecture.

Movement Mechanisms

Movement operators define the fine-grained control of lymphocyte migration. Without movement, the discretised repertoire may be considered multiple independent instances of some variant of the Cellular Paradigm. Movement may be considered as a decision process at a per-lymphocyte level involving the interaction of the lymphocyte (lymphocyte class), and its locality (tissue type and location such as a germinal centre). More broadly, the aggregation of lymphocyte-wise movement between tissue may be considered an inter-tissue communication process of information sharing. A lymphocyte may possess a *movement intention* given its class, and so may a node, given its tissue type. Further movement operators may not be limited to lymphocytes, but encompass all so-called cellular-level components such as antigen. Movement influences both the (1) repertoire homeostasis (both discrete repertoire and the aggregate of discrete repertoires), and (2) the dissemina-

tion of localised acquired knowledge. Both of these points may be considered to influence a trade-off between *recirculation* and *localisation* of lymphocytes that movement operators must address. A node may have finite capacity for the number of cellular components. Further movement may be quantified in terms of the number of cellular components over time (a rate). This rate may be aggregate or fine-grained by component class. Seeking points of stability or equilibria (homoeostasis) may involve direct and dynamic manipulation of these movement rates¹. Homoeostasis is perturbed by movement operators themselves and can be controlled through the management of aggregate movement rates. Another influence is the dynamic creation of lymphocytes (potentially on a large scale) in germinal centres and the recruitment behaviour at sites of infection and inflammation. Such functional behaviours have a potential to overwhelm the homoeostasis capabilities of the model.

5.3.2 Antigenic Exposures

Discrete Repertoire Exposures

The *Antigenic Exposure Paradigm* (Section 4.2.2) was presented as a series of three exposure types which increase in complexity: (1) single exposure, (2) multiple exposure, and (3) multiple pathogens. The primary difference between the tissue model type and a single repertoire model, is that the system has multiple points of exposure which may be all or a subset of the repertoires of the a given tissue model. Figure 5.1 provides a conceptualisation of a simple chain of discrete repertoires situated in an antigenic environment with an asymmetric antigenic exposure pattern. A concern raised with regard to reconciling the exposure model with multiple points of exposure is referred to as the *Tissue Exposure Regime (TER)*. A given systems exposure pattern defines the selection of tissue repertoires of the model to expose with antigen in the event of an exposure event. For example, a natural inclination may be to use a uniform repertoire exposure pattern for a single exposure (see Figure 5.2a). In the case of multiple sequential exposures, a uniform exposure pattern may not be desirable. Antigen may be considered to arrive to the system in different locations, thus the antigen-to-repertoire patterns may be irregular (see Figure 5.2b). In the case of multiple antigens with multiple exposures, not only may the exposure patterns be irregular, but also the timing of antigen-type exposures. Exposures may occur concurrently to the system across different repertoires.

In order to provide symmetry in the problem domain with the raise in the level of abstraction in the system, a new abstraction of the antigenic exposure paradigm is required. The exposure paradigm at the cellular level was concerned with an antigen of determinants, thus a suitable metaphor for an exposure of a tissue is an *Infection of Antigen*. An infection is defined as the invasion and multiplication of a pathogenic micro-organism that may cause injury the the tissues. This definition provides a suitable metaphor for the point-wise exposure of tissue for two reasons (1) invasion and replication suggests a

¹An alternative strategy for direct influence of homoeostasis is that of cellular component deletion (cellular removal). Implemented models are expected to combine deletion and movement in an effort to seek homoeostasis, such as stable repertoire sizes.

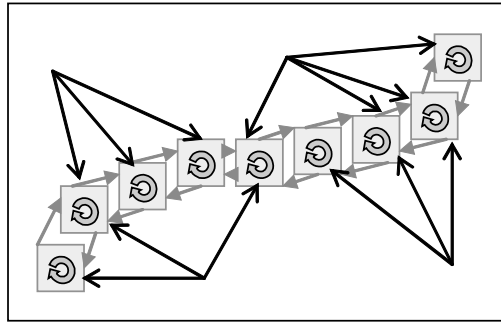
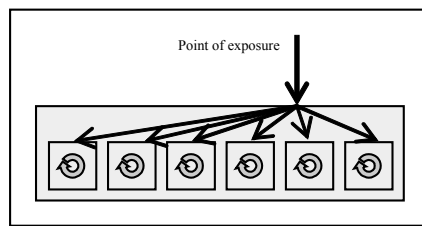
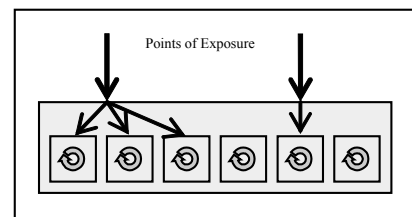


Figure 5.1: Multiple tissue model and the situated spatial properties that may affect antigenic exposure.



(a) Example Symmetric Tissue Exposure.



(b) Example Asymmetric Tissue Exposure.

Figure 5.2: Example of two different Tissue Exposure Regime with controlled selection of discrete repertoires on exposure events.

single antigenic type (pattern) perhaps with variation, and (2) tissue injury promotes urgency which may be interpreted as efficient and effective response, whether it is acquired in the tissue or imported from surrounding tissues. Although the abstraction is pathogen-centric given both the exogenous origin and selected metaphor, one may just as easily use an endogenous-centric metaphor such as tumour occurrence and growth. In both cases the abstraction properties of discrete (point-wise) information mediation hold.

Spatial-Temporal Exposures

The specifics of the Tissue Exposure Regime may or may not be in the control of tissue model, which is dependant on the application domain and model configuration. One may consider a possibility space of all exposure-to-repertoire interactions and consider questions of lymphocyte movement types in the context of desirable system behaviours under different regimes. Irrespective of the exposure properties, the concern of the system is to address *what repertoires* are exposed at *what times*. These concerns may be generalised to that of spatial and temporal consistency of exposure patterns. *Temporal consistency* refers to the systems anticipation of the amount of memory required. *Spatial consistency* refers to systems anticipation and the positioning of lymphocytes across the repertoires. A lack of consistency in either dimension removes the ability of a system to anticipate specifically, requiring general (all-repertoire, all-time) anticipation. It is important to highlight that the temporal memory applies system-wide without spatial memory (in the face of spatial inconsistency), and spatial memory applies for all time without temporal memory (in the

face of temporal inconsistency). Therefore, the default case is remember everything everywhere for all time in the face of spatial and temporal inconsistency. This is an important implication of the antigenic exposure paradigm as it defines the role of inter-repertoire cell movement as the exploitation of patterns in the combination of spatial and temporal consistency of system exposures such that the response memory may be generated and shifted (adapted) in response to changes in the spatial or temporal properties of the exposure.

The *Temporal Memory Effect* refers to the systems capability to retain an impression of an exposure event and recall that impression to provide an improved (compared to random) response under similar circumstances in the future. The effect is intuitive for a single repertoire model where it may be facilitated by memory cells or a change in cell compositional densities. The exploitation of the cells becomes a probabilistic function of the antigen encountering a memory in the repertoire. Figure 5.3 provides a depiction of the temporal memory effect in a single repertoire, where cell attrition via repertoire ‘turn over’ results in the fading of acquired information over time.

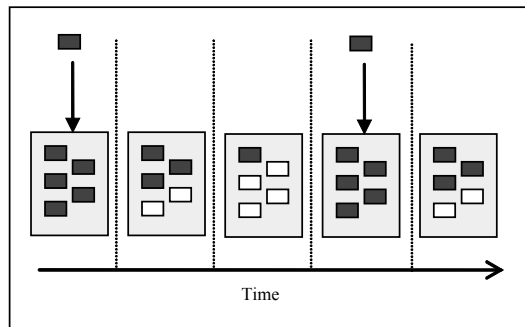


Figure 5.3: Depiction of the temporal memory effect in a single repertoire model over time, where the arrows at the top represent information flow into the system and the arrow along the bottom represents time.

The multiple repertoire model further requires the dissemination of memory across multiple repertoires. The memory represents a system-wide spatial anticipation of future exposures to the same or similar antigen. The following summarise the properties of temporal memory in the single repertoire model:

- *Short Term*: A response results in a large short-term memory in the form of effector cells (short lifespan), which are useful initially at the point of exposure, and disseminated throughout the system.
- *Long Term*: A response results in a small long-term memory in the form of memory cells (long lifespan), which also disseminate throughout the system.
- *Memory Size*: The larger the antigen dose, the larger the short-term and long-term memories.
- *Memory Stability*: With the increases in the frequency of exposures (over the systems lifetime), the larger and more stable the long-term memory. The memory size is expected to reach a point of stability where antigen-to-memory cell match is practically assured.

This *Spatial Self-Organisation Effect* is an emergent property of point-wise exposure of repertoires in the face of lymphocyte recirculation. The exposure and stimulation of a repertoire results in a response to the pathogen. The exposure is spatial, therefore the response is spatial. The following summarises the properties of spatial self-organisation in the discrete repertoire model:

- *Interaction:* (response maturation) The clonal selection principle raises a matured response to the antigen that results in effectors and memory cells. Some of the raised clone disseminates early, although the majority remains during the stimulation of the repertoire. This is the formation of a germinal centre and constraints are imposed on outbound lymphocytes.
- *Recruitment:* (response amplification) The repertoire tissue becomes inflamed resulting in the increased intake of recirculating lymphocytes, a response amplification effect via lymphocyte sequestration. The inbound cells may be naïve cells, effector cells, or memory cells and may or may not have improved specificity for the pathogen than the cells already in the repertoire.
- *Homing:* (preferential recirculation) Memory cells created in a specific repertoire, preferentially recirculate back to that repertoire, in an effort of spatial anticipation of the same antigen in the same place.

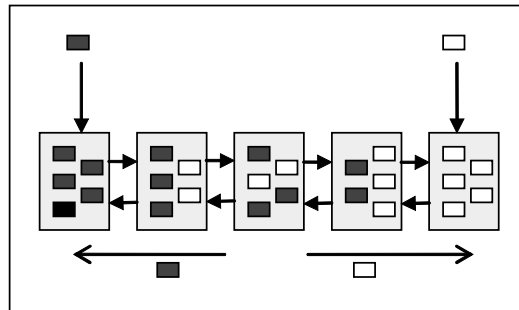


Figure 5.4: Depiction of the repertoire specialisation that results in the Spatial Self-Organisation effect, where the arrows from above represent information flow into the system and the arrows below represent dissipation of acquired information.

The result of the general behaviours of the tissues and the cells is a specific spatial response strategy that attempts to get the best from the system at the point of exposure in a bottom-up manner. The spatial organisation may extend through time, such that in addition to the patrolling effectors and memory cells that disseminate throughout the system, effectors may remain behind and provide a short-term memory and spatial anticipation of response. The following summarises the effects of spatial self-organisation in the discrete repertoire model:

- *Localisation:* Specific response at the time and the location of the exposure. Natural spatial-temporal response.

- *Surveillance*: Long-term and short-term memory in recirculation anticipating the exposure system wide. The premise that the system may be exposed to the same pathogen again the future.
- *Sentinels*: Short-term memory, stationary at the point of exposure, anticipating exposure at the same location. The premise that a location that has been exposed may be exposed again.

Spatial anticipation is facilitated in the short-term by sentinel effectors and in the long-term by memory cell homing patterns. In addition to the spatial self-organisation effect, the system mixes (perhaps disorganisation) the acquired immunity across the repertoires of the system. The principles of spatially mixing response effectors and memory provide a system-wide uniform (or consistent) response to a previously encountered pathogen. This ensures that immunity learned at one exposure point of the system, may be exploited elsewhere (anywhere) in the system.

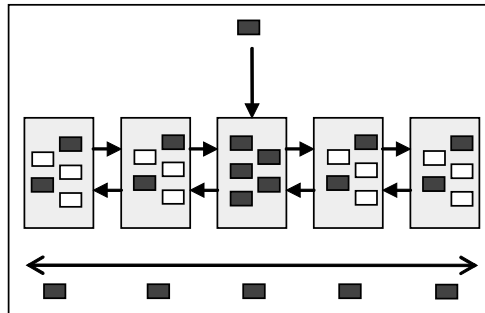


Figure 5.5: Depiction of the unbiased dissemination of information that results in the Spatial Consistency of Response, where the arrow above represents information flow into the system and the arrow below represents information availability throughout the system.

This effect called the *Spatial Consistency of Response* is achieved through the information dispersal properties of lymphocyte recirculation that ensure that a fraction of the larger lymphocyte pool is in motion between the repertoires (immunosurveillance). The effect promotes system-anticipation rather than the more specific spatial-anticipation, and provides a distributed knowledge or memory that may rapidly self-organise to varied repertoire exposure patterns. The consistency provided is not strictly uniform, rather it is quasi-uniform, driven by the bottom-up stochastic decision properties that govern lymphocyte recirculation. From a lymphocyte perspective, movement in the face of unknown exposure patterns provides the system-wide goal of *antigen-selection maximisation*. That is, lymphocytes exist to be selected by antigen, thus the behaviours of movement between, and residence within repertoires seeks to maximise the change of this event occurring. As already described, this is already facilitated by the homing behaviour of memory lymphocytes, and the patrolling recirculation behaviour of effectors and naïve lymphocytes. Some concerns of this behaviour are (1) the quantity of the immunity acquisition, and (2) the rate of dispersal of the memory. The following summarises the concerns with achieving consistency of response:

- *Acquisition Quantity*: The system response may be proportional to the exposure, although the memory (both short-term and long-term) must be sufficient to provide coverage to all repertoires of the system. This memory is additive, and is expected to expand in proportion to additional exposures by the same pathogen.
- *Acquisition Dispersal*: The memory of an exposure originates at the point of exposure, and the delay of dispersal is a function of the (1) quantity of the memory, (2) the size (number of repertoires) and connectivity of the system, and (3) the rate of recirculation. Dispersal of acquired immunity must be such that coverage (probabilistic) of all repertoires is obtained within a reasonable amount of time.

The dispersal or mixing effect in the consistency of response facilitates the blind (recirculation) and specific (recruitment) sharing of acquired immunity between repertoires. Sharing may refer to the exploitation of specific immunity at a exposure location different from that at which it was acquired. Further, sharing may also refer to the cross-reactive effects achieved through less specific acquired immunity. This effect includes the ability of the system to generalise from varied exposure types such that new variations and combinations may also be exploited.

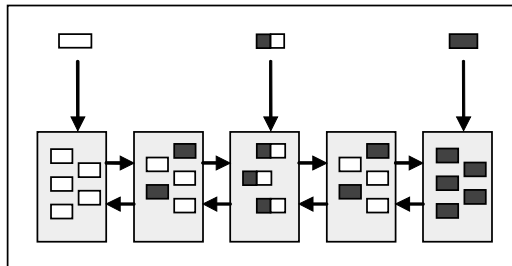


Figure 5.6: Depiction of information sharing and a cross-reactive response, where the arrows above represent the flow of information into the system as well as an implicit anticipation for the application of acquired information.

The integration of these features is a system that is expected to be plastic in its response capability, although is still capable of generating, maintaining, and self-organising a specific response to external stimulation. This behaviour may represent a trade-off in efficiency and efficacy of response between (1) system-wide consistency of response, and (2) segregation of response into specialised spatial compartments. The system may be configured to achieve one or the other of these effects, although achieving both effectively and concurrently represents the challenge for the system in unknown environments.

5.3.3 Tissue Architectures

Exploiting Tissue Responsibilities

This section describes a lymphoid tissue model in which the tissue types (nodes) constrain the cell-based information processing that may occur to the cells housed within. Tissues are considered the units of adaptation that may receive and transmit information. Cells

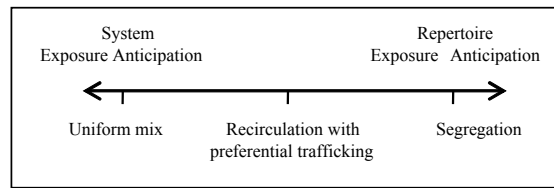


Figure 5.7: Depiction of the trade-off between the general anticipation of exposure and organised anticipation of exposure.

are the substrate for tissues knowledge representation, communication, and adaptation. This section describes a number of tissue-tissue interactions.

Cell Formation Naïve lymphocytes are created in primarily lymphoid tissue. A naïve lymphocyte is generated in a pseudo-random manner. It represents an untested guess of a unit of information that may be useful to the system. A naïve cell may be created in one primarily lymphoid tissue and migrate to another for specialisation, similar to the migration of pre-lymphocyte cells from the bone marrow to the thymus to become naïve T-cells. Those cells that remain in the bone marrow become B-cells. The primary tissue nodes may condition the formed naïve cells. For example, the cells may be subjected to a negative and positive selection process such as that subjected to T-cells in the thymus. The result of such a process is the assessment and ultimate judgement of the feasibility of the proposed information guesses that the cells represent. In addition to the migration of naïve cells between primary lymphoid tissue for maturation, naïve cells may uni-directionally migrate from primary to secondary lymphoid tissue. This represents an upgrade in status for the cell from unprepared to prepared naïve cells, ready for information processing in the secondary tissue. As an interaction between primary and secondary lymphoid tissue it represents a transmission of feasible quasi-random information that may be useful to the system.

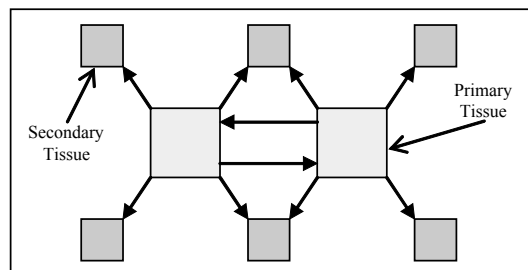


Figure 5.8: Depiction of the relationship between primary and secondary lymphoid tissues, where the arrows show the sharing of information between primary tissues and the outward push of information from primary to secondary tissues..

Cell Maturation Secondary tissues houses many naïve lymphocytes as well as some memory cells and effectors if available. The tissue accepts a steady stream of naïve lymphocytes from primary tissue and integrates them into the cell population. Secondary

tissue is responsible for exposing resident cells to antigen and facilitating any immune response that occurs as a result. Exposure is selection of cognate receptors for a given antigen, and response is the clonal selection process or similar as outlined in previous models. The result of the response are memory cells which have demonstrated their usefulness, and effector cells for addressing the immediate concerns of the antigen. Effector cells may or may not be given an impression of the tissue origin of the antigen (site of infection), as considered in the case of T-cell homing. Memory cells and non-homing effector cells recirculate between secondary tissues (lymph nodes for example), using the vascular system as the transport mechanism. The recirculation of memory and effector cells between secondary lymphoid tissue represents a transmission of acquired information regarding the environment. Portions of acquired information (individual cells) are shared between secondary tissues such that the information may be available for application or further maturation. Acquired information in the form of effector cells is transmitted from the location of maturation in the secondary tissue, to the location of application in the tertiary tissue.

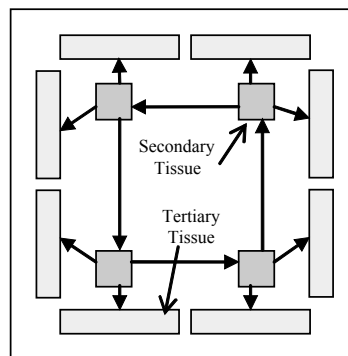


Figure 5.9: Depiction of the relationship between secondary tissues and secondary and tertiary tissues, where the arrows show the sharing of information between secondary tissues and the outward push of information from secondary to tertiary tissues.

Cell Application Tertiary tissues accept and house many effector cells. The tissues accept a steady stream of effectors from locally connected secondary tissues, which are integrated into the population. Tertiary tissues are responsible for being receptive to pathogenic infection, thus may be considered receptive fields for information from the environment. The effector cells are employed at the sites of infection as needed, addressing the major function of their life cycle: application. Samples of the antigen may be taken from the site of infection in the tertiary tissue and transported into secondary tissue by courier cells, taking in addition information as to where the antigen was collected (to facilitate homing). Samples of antigen and some effector cells may be drained back into locally connected secondary tissue. This draining effect is facilitated by the lymph fluid that permeates tertiary tissues. The secondary tissue collections and filters the lymph for antigen, which are presented to resident lymphocytes. Pathogen exposure to tertiary tissue represent samples of the environment collected by the system. Different tertiary tissues may collect samples in different ways (for example skin, food, respiratory). The couriering

and streaming of samples of pathogen back to secondary tissue provides a transmission of sensory information from the environment to areas of the system that can be adapted to address it.

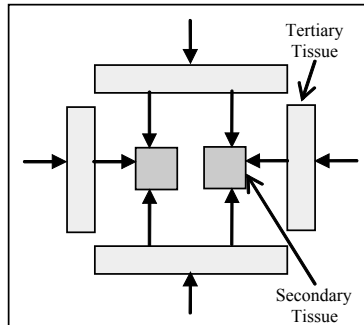


Figure 5.10: Depiction of the relationship between tertiary and secondary tissues and tertiary tissues and the environment, where the arrows depict the flow of information from the environment to the tertiary tissues and ultimately to the secondary tissues.

Integration

The aggregation of all three tissue types describes an integrated model with layered information processing. From the inside out, the information availability increases in maturation (usefulness). The tertiary tissue contains no other cells than those known to be immediately useful, the primary tissue contains no other cells than those whose usefulness is unknown. The secondary tissues act as the mediator between these two extremes, providing the testing and maturation processes for information acquisition. The outward flow of prior knowledge through the layers of the system, whereas information about the environment flows inward. Principle pathogen exposures occur in the tertiary tissue where it may be addressed by resident effectors. If sufficiently large or novel, samples of the pathogen find their way into secondary tissue (courier or stream) for presentation to the learning mechanisms of the system. Therefore, lymphocytes migrate from the centre of the system outward, increasing in proficiency as they approach the surface, whereas pathogen are housed in the outer layer and seep into the second tier.

Antigens that penetrate the system and manage to take-up residence in the tertiary tissue are drawn to the lymph nodes of the secondary tissue by fluid called lymph. Lymph drains into the lymph nodes, which in turn filter the fluid collecting antigens, which are presented to local and recirculating lymphocytes. From an information processing perspective, a lymph node is provided an information stream from the regionally located tertiary tissue. A proportionally large region of tertiary tissue feeds pathogen to a single lymph node, in effect proving multiple feeding information channels. The information is collected in specialised regions within the node. B-cells are exposed to the presented information, and if sufficiently activated may migrate to another specialised region on the node and form a germinal centre in which bouts of selection, proliferation, receptor maturation and ultimately cell differentiation occur. The product of this learning process are effector cells called plasma cells that produce antibodies for the triggering antigen

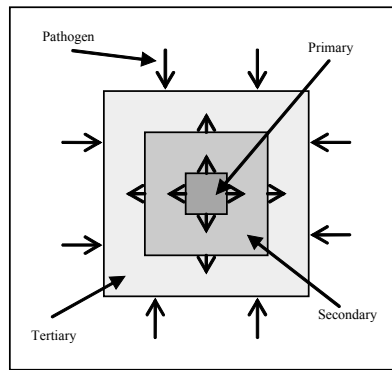


Figure 5.11: Depiction of the layered tissue architecture and information flows between the layers, the arrows show the flow of information from the environment into the system as well as the outward push of new detectors from the centre of the system.

in large numbers, and mature long-lived memory cells. The antibodies enter the blood stream and seek a serendipitous interaction with their cognate antigen. The memory cells recirculate between the secondary tissue and sometimes move into the tertiary lymphoid tissue, with the same mission. There is no homing effect for the B-cells. Their effectors flood the system, and their memory cells patrol. Importantly, the effectors and the memory cells are created and released in the vicinity of where the pathogen was encountered. Thus, they are more likely to encounter their cognate antigen sooner rather than later.

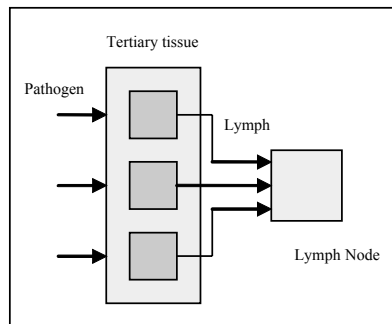


Figure 5.12: Depiction of the information streaming providing by lymph collecting in lymph nodes.

Another type of cell, called a dendritic cells, patrol tertiary tissue for pathogen. Once detected, these cells collect the material and migrate into the secondary lymphoid tissue (such as lymph nodes). In the secondary tissue, the dendritic cells present the pathogenic material to T-cells that respond, proliferate, and differentiate. The product of this learning process is mature T-cells in the form of effectors and likely long-lived memory cells. The effector T-cells are imprinted by the dendritic cell as to the location in the tertiary tissue of the pathogen exposure. The effectors recirculate the system and penetrate the tertiary tissue when the specific chemical signature is detected. Information is streamed from the tertiary tissue to the secondary tissue although the streaming is facilitated by courier cells. These cells are aware of where they collected their material and thus are capable of imprinting that information onto the produced effector cells. The result is an effect

that differentiates the lymphocyte movement behaviour from that of the B-cells (that blindly seek their cognate antigen), to the T-cells (that home into the chemical signature of infected tissue).

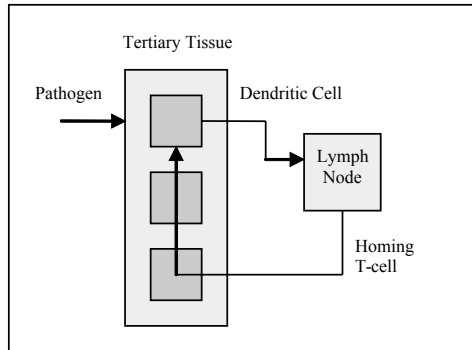


Figure 5.13: Depiction of the courier-like behaviour of dendritic cell's.

Tissue Modes

The information properties of the system have been described with regard to the transmission of the information substrate (immune cells) between the layers and within the layers. Table 5.1 lists the abstracted immunological functional behaviour required to meet the information processing needs outlined for the architecture.

Tier	Information Processing	Behaviour
1	Generation and feasibility testing	Negative Selection
2	Maturation	Clonal Selection
2	Homoeostasis	Elaborated Clonal Selection
2	Recirculation	Recirculation Algorithm
2-3	Courier/Homing	Homing Algorithm
2-3	Exposures/Streaming	Antigenic Exposure
2-3	Inflammation	Recruitment Algorithm

Table 5.1: Summary of information processing needs, and proposed algorithms that meet those needs.

The information processing constraints of each tissue type, outline a clear responsibility for each tier of the architecture. Primary tissue is concerned with information generation and the pre-processing of information before it is employed in a learning process. Secondary tissue is responsible for learning, for identifying which internal information is presently useful and attempting to improve it (maturation), whilst at the same time applying it (effectors). The second tier is also responsible for the generation and maintenance of longer term memory, and the selective application of acquired memory. The tertiary tissue is responsible for providing an interface to the environment, mediating the application of acquired knowledge to sensory signals, with the filtering of temporally novel signals to secondary tissues for response generation.

The majority of the clonal selection information processing occurs in the secondary lymphoid tissue. The other two layers may be easily incorporated into the secondary

Tissue Type	Responsibility
<i>Primary Tissue</i>	Information generation (preparation)
<i>Secondary Tissue</i>	Information maturation (learning and memory)
<i>Tertiary Tissue</i>	Information application (sensory)

Table 5.2: Summary of the principle responsibility of each tier of the architecture.

tissue, providing only conceptual attributes of the tissue model. This compression of responsibility has been implicit in previously proposed clonal selection, negative selection, and other cell-theory based algorithms. Given that the tissues may be compressed, it is important to consider the utility of separating the proposed responsibilities into a dependant hierarchy of layers.

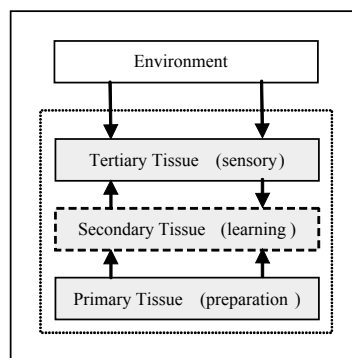


Figure 5.14: Depiction of the layered information processing properties of the lymphoid tissue architecture.

This section proposes three different models that employ the proposed lymphoid tissue architecture. Each model is inspired by a different aspect of the acquired immune system, from a lymphatic system (tissue) perspective. The principles for separated responsibility are (1) the centralised generation of naïve cells facilitating centralised prior knowledge, (2) consolidated secondary tissue for a networked learning architecture large enough to meet the needs of the system, (3) a large host consisting predominantly of tertiary tissue, which provides a distributed sensor platform for receiving and responding to environmental signals.

Filter Model

The filter model is inspired by the specialisation of the lymphatic system at the primary entry points for exogenous antigen (pathogen). Examples include the tonsils and related tissues for addressing pathogen that enter the host via the respiratory system, the Peyer's patch tissues in the intestines that addressing pathogen that enter the host via the digestive system, and the spleen that filters the blood for pathogen. The model contains a single point of entry for pathogen (tertiary tissue), which is monitored by one or a small collection of secondary lymphoid tissues. The secondary tissue is supported by a single source of naïve cells (primary tissue).

Although the model is called a *filter* (the basis for its inspiration) it is not restricted to

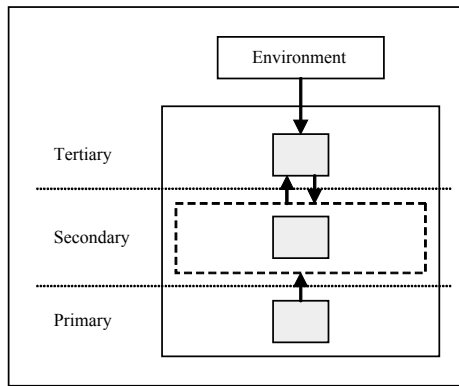


Figure 5.15: Depiction of the simple filter model.

filtering information processing tasks. The focus of this model is on the system possessing a single point of entry for pathogen (sensory signals) although at a high load (many signals), and one or a small number of specialised secondary tissues to address the needs of the tertiary tissue.

Tissue Type	Configuration
<i>Primary</i>	Single tissue feeding the secondary tissue.
<i>Secondary</i>	Single (or a small number) secondary tissues with high capacity of cells.
<i>Tertiary</i>	Single point of entry for the system to monitor.

Table 5.3: Configuration of the architecture for the filter model.

Lymph Node Model

The lymph node model is inspired by the network of lymph nodes, connected by the vascular system, that receive antigen as a stream in the lymph. The model is more complex than the filter model as it contains a network of moderately sized secondary lymphoid tissues that recirculate memory and effector cells. Further, unlike the filter model, it has a large receptive field of tertiary tissue, which provides a single broad sensory system.

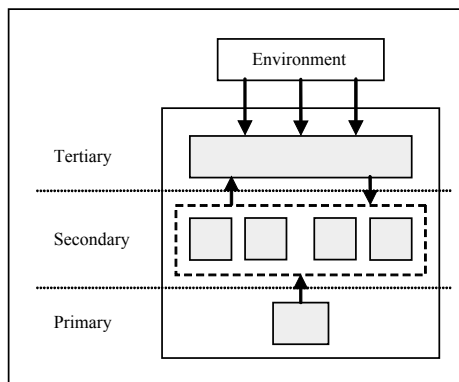


Figure 5.16: Depiction of the lymph node model.

The focus in this model is the large tertiary tissue receptive field, and the network

of secondary tissues (lymph nodes), in particular the capabilities of the lymph nodes to recirculate mature cells (share acquired information), and their capabilities in trafficking effectors into tertiary tissues (applying acquired information).

Tissue Type	Configuration
<i>Primary</i>	Single tissue feeding the secondary tissue.
<i>Secondary</i>	A consolidated network of moderate sized secondary tissues.
<i>Tertiary</i>	A large and uniform receptive field for input signals.

Table 5.4: Configuration of the architecture for the lymph node model.

Host Model

The host model is an integration of one or more filter models with the lymph node model. The resultant model is a larger-scale tissue architecture that may be equated to a lymphatic system of a host. The model is composed of a variety of different tertiary tissue types and thus different meanings for the secondary tissue to be provided with antigen to present to resident cells. In addition, the secondary tissue is made of a core network of lymph nodes, as well as collections of specialised lymphoid tissue to address principle pathogen penetration areas. Finally, the system has one or more primary lymphoid tissues for the creation and dissemination of naïve cells to secondary tissues.

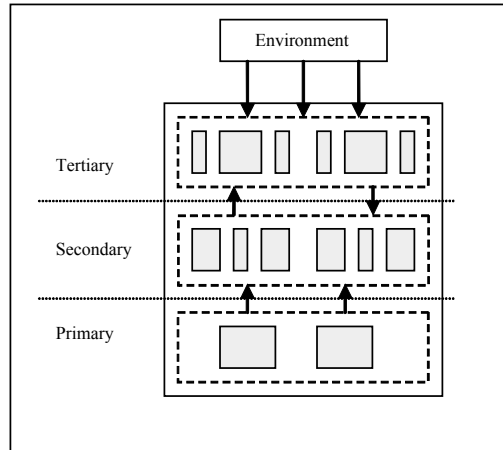


Figure 5.17: Depiction of the host model.

The focus of the model is the integration of the lymph node network (and related method for sampling antigen from tertiary tissue) with a number of filter-based models centred around primary pathogen access points. This host model represents a hybrid integration of the two sub-types of tissue architectures, and culminating (pinnacle in complexity) lymphatic tissue model.

Summary

The lymphoid tissue components, exposures, and architecture models provide a context for the previously defined adaptive models of the Cellular Paradigm concerning immune

Tissue Type	Configuration
<i>Primary</i>	Set of primary tissues feeding the secondary tissue.
<i>Secondary</i>	A backbone of a consolidated network as well as specialised regions for each major entry point (diverse).
<i>Tertiary</i>	Multiple tertiary tissues of varied size and scope (diverse)

Table 5.5: Configuration of the architecture for the host model

cells (*intra-tissue*), and a context for novel cell migration adaptive models to be investigated (*inter-tissue*). The applications of the architecture in the proposed lymphoid tissue models provides three specific settings for integrated lymphoid tissue-based information processing. This section suggested at the potential for Tissue Clonal Selection and more generally tissue-based Artificial Immune System algorithms, as a scale above the state of classical cell-based Artificial Immune System algorithms.

5.4 Realised Tissue Paradigm

A single immune system of tissues that may be exposed to infections provides an effective metaphor for considering the problem of the acquisition of information in a decentralised manner from a distributed environment, and strategies toward the anticipated application of acquired information. The presentation of the lymphatic system demonstrated that some infections can be localised to predictable lymphoid tissues (peripheral immune system) and others cannot (systemic immune system). Therefore a host immune system of lymphoid tissues must simultaneously manage predicted, unpredicted, and a variety of intermediate antigenic infections from its antigenic habitat. Toward this end, the host immune system considers the differential mobility of cells between tissues and differential cell types as a strategy for managing so-called anticipated and unanticipated antigenic infections across all tissues. This section provides a realisation of the concerns of the abstract Tissue Clonal Selection Paradigm presented in Section 5.3. This includes definitions of a tissue problem in the Infection Exposure Problem and specialisation in colour space, and the Tissue Clonal Selection Algorithm that provides a basis for investigation into cell trafficking schemes. A series of tissue-based measures are presented for assessing tissue algorithms on colour space realisations of infection problems that provide quantitative indicators regarding the concerns of spatial temporal exposures. Collectively, this section provides a basis for the implementation and exploratory empirical investigation into the Tissue Clonal Selection Paradigm, as well as a bridge between the the biology (Section 5.2) and abstract models (Section 5.3) of Tissue-centric Clonal Selection.

5.4.1 Antigenic Infections

This section considers the realisation of the exposure paradigm outlined in Section 5.3.2. In particular, this section abstracts the Antigen Exposure Problem considered for the cellular paradigm for the tissue paradigm called an Infection, and specialises an instance of the problem in Colour Space. In addition, the discrete properties of tissue exposures

are reconsidered and a number of specialised Tissue Exposure Regimes are defined.

Infection Antigenic Exposure Problem

The Antigenic Exposure Problem (AEP) defined in Algorithm 4.1 provides a general problem definition where a Tissue T is exposed to an Infection I . A given Cellular Clonal Selection Algorithm may be considered a single T in a Host of Tissues ($H = \{T_1, T_2, T_3 \dots, T_n\}$). Therefore the AEP must be generalised for interaction with an H . The *Infection Antigenic Exposure Problem (IAEP)* defined in Algorithm 5.1 provides a generalised exposure problem in which an H is exposed to a Habitat B of Infections I ($B = \{I_1, I_2, I_3, \dots, I_n\}$). The parameters $N_{determinants}$, $N_{antigen}$, and $N_{infections}$ define the properties of B with a symmetrical number of I which are passed to *CreateHabitat* to create a problem instance. *Exposure* refers to a given exposure mechanism, such as the variety of Tissue Exposure Regimes (TER) defined later in this section. An exposure regime returns a result set H_{rs} that represents the state of a given H with regard to its ability respond to the B under the given TER.

Algorithm 5.1: Infection Antigenic Exposure Problem (IAEP).

Input: H , $N_{determinants}$, $N_{antigen}$, $N_{infections}$

Output: H_{rs}

```

1  $B \leftarrow \text{CreateHabitat}(N_{determinants}, N_{antigen}, N_{infections});$ 
2  $H_{rs} \leftarrow 0;$ 
3 while  $\neg \text{StopCondition}()$  do
4   |  $H_{rs} \leftarrow \text{Exposure}(H, B);$ 
5 end
6 return  $H_{rs};$ 

```

Infection Colour Space Problem

The IAEP may be specialised as an infection extension of the Antigen Colour Space Problem (ACSP) defined in Section 4.3.1 called the *Infection Colour Space Problem (ICSP)*. An instance of the ACSP involved the exposure of a T to a set of Colour Space Patterns (CSP), where the goal of the problem is to minimise average error in the response T_{rs} . The ICSP specialisation of IAEP exposes a H to a set of sets of CSP (B of I 's), with the goal of minimising the average error in the response H_{rs} (defined in Equation 5.1). One may define a minimal configuration of the IAEP and thus the ICSP that focus the concerns of the problem on the highest level of abstraction, the Habitat B . This focus is provided by reducing the configuration at levels below the desired level to minimal levels of complexity. A *Minimal Infection Antigenic Exposure Problem (MIAEP)* is defined where $N_{infections}$ is variable, although $N_{antigen}$ and $N_{determinants}$ are fixed at 1. In the case of a minimal ICSP, the number of I defines the number of CSP (A), each of which has a single D comprised of all three Colour Space Components (Red, Green, and Blue).

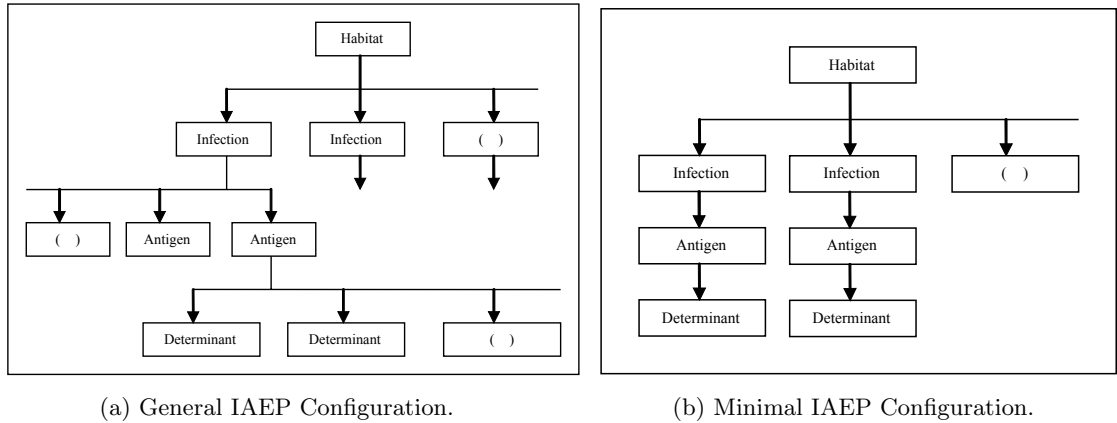


Figure 5.18: Depictions of general and minimal configuration examples of the Infection Antigenic Exposure Problem (IAEP).

Tissue Exposure Regimes

Section 5.3.2 defined the two important aspects of the problems that Tissue Clonal Selection Algorithms must address: (1) discrete tissue exposures, and (2) the spatial temporal properties of exposure anticipation defined by the regularity of a given tissue exposure regime. Toward this end, this section explicitly defines five different Tissue Exposure Regimes (TER) for specialising the $Exposure(H, B)$ of the IAEP providing examples from the range of possible regular and irregular exposures (spatial-temporal exposure of tissues) with some and all of the information in the problem domain. The schemes include: Symmetric Exposure, Asymmetric Exposure, Point Exposure, Random Exposure, and Probabilistic Exposure. The TER provides a mechanism by which to specialise an the infection exposures of a habitat to the tissues of a host, providing an generalisation of the limitations imposed by the structure, organisation, and properties of different tissues as described by the various tissue architectures in Section 5.3.3. Figure 5.19 provides a depiction of four of the five exposure regimes to a host of tissues.

Symmetric The *Symmetric Tissue Exposure Regime (STER)* defined in Algorithm 5.2 involves the exposure of each Infection I to each Tissue T (see Figure 5.19a). Thus, each tissue is exposed to and must address the scope of the antigenic habitat B .

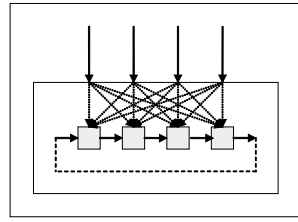
Algorithm 5.2: Symmetric Tissue Exposure Regime (STER).

```

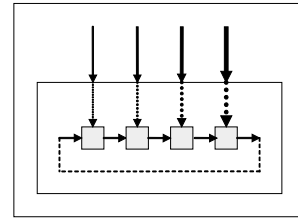
Input: H, B
1 foreach  $T_i \in H$  do
2   | foreach  $I_i \in B$  do
3   |   |  $Exposure(I_i, T_i)$ ;
4   | end
5 end

```

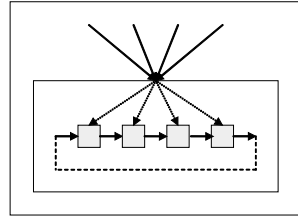
Asymmetric The *Asymmetric Tissue Exposure Regime (ATER)* defined in Algorithm 5.3 involves the exposure of each I Infection to a single Tissue T in a consistent manner, such



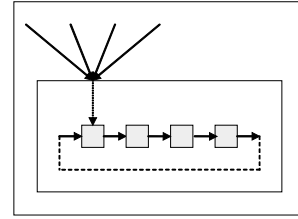
(a) Symmetric Tissue Exposure Regime (STER).



(b) Asymmetric Tissue Exposure Regime (ATER).



(c) Random Tissue Exposure Regime (RTER).



(d) Point Tissue Exposure Regime (PTER).

Figure 5.19: Depictions of the four different Tissue Exposure Regimes (TER).

that once established the same T -to- I mapping is used for each exposure (see Figure 5.19b). The number of T may match the number of I such that each T addresses $\frac{1}{N_{infections}}$ per host exposure.

Algorithm 5.3: Asymmetric Tissue Exposure Regime (ATER).

Input: H, B
1 **foreach** $T_i \in H \wedge I_i \in B$ **do**
2 | $Exposure(I_i, T_i);$
3 **end**

Point The *Point Tissue Exposure Regime (PTER)* defined in Algorithm 5.4 restricts the exposure of all I to a single and consistent T across all host exposures (see Figure 5.19d). The choice of T that becomes the point of exposure for the host may be randomly selected from H .

Algorithm 5.4: Point Tissue Exposure Regime (PTER).

Input: H, B, T_{point}
1 **foreach** $I_i \in B$ **do**
2 | $Exposure(I_i, T_{point});$
3 **end**

Random The *Random Tissue Exposure Regime (RTER)* defined in Algorithm 5.5 is similar to the asymmetric exposure regime, although the matching of T with I occurs randomly with re-selection each exposure (see Figure 5.19c). Thus, a given T may be exposed to a different subset of the antigenic habitat B each host exposure where it is possible for the set to be empty.

Algorithm 5.5: Random Tissue Exposure Regime (RTER).

Input: H, B
1 **foreach** $I_i \in B$ **do**
2 $T_{point} \leftarrow \text{RandomTissue}(H)$;
3 $\text{Exposure}(I_i, T_{point})$;
4 **end**

Probabilistic The *Probabilistic Tissue Exposure Regime (OTER)* defined in Algorithm 5.6 provides an intermediate specialisation of the RTER and ATER in which the selection of T is probabilistic based on past selections providing probabilistic spatial consistency of re-exposure. In addition, the I -to- T mapping is maintained for an extended number of host exposures before a probabilistic reselection is made, providing temporal regularity to the spatial exposures. The *duration* parameter defines the number of host exposures (the *HostExposureNumber* state variable) a selected T is exposed to a given I . The state variable *Histogram* keeps track of the frequency each T is exposed to each I , and *biasedRouletteWheelSelection(Histogram)* uses a biased roulette wheel selection routine on the histogram for a given I over all T to select a T to expose for *duration* host exposures. The $I_{tissues}$ state variable keeps track of which T each I has selected and should expose.

Algorithm 5.6: Probabilistic Tissue Exposure Regime (OTER).

Input: $H, B, duration, Histogram, HostExposureNumber, I_{tissues}$
1 **foreach** $I_i \in B$ **do**
2 **if** $(Epoch \bmod duration) \equiv 0$ **then**
3 $T_i \leftarrow \text{BiasedRouletteWheelSelection}(Histogram_{I_i})$;
4 $I_{tissues_i} \leftarrow T_i$;
5 $Histogram_{I_i T_i} \leftarrow Histogram_{I_i T_i} + 1$;
6 **end**
7 $\text{Exposure}(I_i, I_{tissues_i})$;
8 $HostExposureNumber \leftarrow HostExposureNumber + 1$;
9 **end**

5.4.2 Tissue Clonal Selection

The Cellular Clonal Selection Algorithm (CCSA) defined in Algorithm 4.4 describes the general clonal selection interaction between an Infection I and a Tissue T . The tissue paradigm requires a raise in the level of abstraction from a single T to a Host H of T ($H = \{T_1, T_2, T_3, \dots, T_n\}$) that responds to a Habitat B of Infections I . *Tissue Clonal Selection Algorithm (TCSA)* defined in Algorithm 5.7 as its names suggests describes a general tissue clonal selection algorithm that may be specialised based on the concerns and constraints of the tissue paradigm. In particular, the *TissueInteractions* operation provides a facility for specialising the interactions between T such as the trafficking of Cells (C). An important distinction between the rise in the level of abstraction from CCSA to TCSA's is the delegation of concerns. TCSA create and exploit multiple instances of the

CCSA without concern for the specifics of the initialisation and response to exposure. This provides flexibility in TCSA's to focus the tissue paradigm concerns such as cell trafficking and potentially making use of a heterogeneous set of CCSA, whilst at the same time forcing responsibility for tissue-level responses (T_{rs}) onto the CCSA. The construction of a H_{rs} is the linear aggregation of the specific T_{rs} from each T produced through the interaction of an I and a T as paired according to the TER. The TCSA is immune system centric unlike the IAEP that is environment centric, providing a flip in responsibility from a passive H responding to Exposures according to a TER in the case of IAEP, to an active H that invokes Exposures on a B , although still delegating the TER to the IAEP. This conceptual flip provides a computationally equivalent system-environment relationship that facilitates bi-modal realisations suitable for different implementation specific concerns. The use of a TER provides an important decoupling of the Host from the controlled governance with the information provided by the antigenic habitat, specifically with regard to the selection of the H_{rs} and the adaptation within exposed tissues. This decomposing in which control is a facet of the specific problem-system implementation provides a clear and simple platform for investigating the behaviour of decentralised clonal selection distributed across a set of potentially interacting independent repertoires of cells.

Algorithm 5.7: Tissue Clonal Selection.

Input: $B, N_{tissues}$
Output: H

```

1  $H \leftarrow 0$ ;
2 for  $i \leftarrow 0$  to  $N_{tissues}$  do
3    $T_i \leftarrow \text{CreateTissue}()$ ;
4    $H \leftarrow T_i$ ;
5 end
6 while  $\neg \text{StopCondition}()$  do
7    $B.\text{Exposure}(H)$ ;
8    $\text{TissueInteractions}(H)$ ;
9 end
10 return  $H$ ;
```

One may define a minimal implementation of the $TissueInteractions(H)$ that does not facilitate any interaction between tissues. This *Minimal Tissue Clonal Selection Algorithm* (MTCSA) of isolated tissues provides a baseline for comparison with extensions of the TCSA that do facilitate tissue interactions. The important behaviour that the MTCSA provides is an example of segregated adaptive tissues, highlighted in the trade-off between information organisation in a regular TER and information mix in an irregular TER (see Figure 5.7).

5.4.3 Empirical Assessment

This section defines a series of empirical measures derived from the measures used in the cellular paradigm (Section 4.3.3) that provide instantaneous information regarding a given TCSA on colour space specialisations of the Infection Antigenic Exposure Problem. The

measures are classified as *system* that provide general holistic information about the algorithm's information state and *component* that provide information about the algorithm's component information state. As with the cellular paradigm, an important consideration with the proposed measures in exploratory experimentation is not their absolute value, but rather their relative change in value with changes to the systems being investigated.

System Measures

Host Error The Average Cell Error (ACE) defined in Equation 4.4 provides a generalised measure of error for a single Tissue Result Set of cells T_{rs} against a given Infection of antigenic patterns I . For the purposes of investigating TCSA the ACE measure may be renamed *Tissue Error (TE)* that may be calculated for each T_{rs} in a Host Result Set H_{rs} and averaged to define a *Host Error (HE)* (see Equation 5.1). HE provides an indication of instantaneous average error for each T_{rs} for a given H_{rs} provided by a TCSA in response to an Tissue Exposure Regime. HE is in the units of the B problem space, for example Euclidean distance for a colour space specialisation. HE provides an indication of a systems capability of addressing a given Habitat problem B under the constraints of a given TER, in particular the spatial-temporal regularity or lack thereof in exposures.

$$HostError(B, H_{rs}) = \frac{1}{n} \sum_{i=1}^n TissueError(I_i, T_{rsi}) \quad (5.1)$$

Host Diversity *Host Diversity (HD)* defined in Equation 5.3 provides an instantaneous diversity a TCSA. Each Tissue T in a Host H is compressed to a *Bit Frequency Histogram (BFH)* defined in Algorithm 5.8 which records the occurrence of each bit for all bit vector based Cells C . The BFH provides a compression of the information content of a given T that may be directly compared with that a BFH from another T as the absolute *Bit Frequency Difference (BDF)* defined in Equation 5.2. The measure calculates the average BDF for a given T against all other T in H , which is then averaged across all T . HD provides a holistic indication of the diversity of the information content of an H . An increase in the HD suggests a more heterogeneous set of information, whereas a decreases in HD suggests a more homogeneous set of information that may indicate an improved system-wide specialisation of resources toward the H .

$$BitFrequencyDifference(BFH_i, BFH_j) = \sum_{k=1}^n |BFH_{ik} - BFH_{jk}| \quad (5.2)$$

$$HostDiversity(H) = \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{n} \sum_{j=1}^n BitFrequencyDifference(BFH(T_i), BFH(T_j)) \right) \quad (5.3)$$

Algorithm 5.8: Bit Frequency Histogram (BFH).

```
Input:  $\mathbb{T}$ 
Output: BFH
1 BFH  $\leftarrow$  0;
2 foreach  $C_i \in \mathbb{T}$  do
3   foreach  $c_j \in C_i$  do
4     if  $c_j \equiv 1$  then
5        $BFH_j \leftarrow BFH_j + 1$ ;
6     end
7   end
8 end
9 return BFH;
```

Component Measures

Average Tissue Diversity The Average Cell Diversity (ACD) defined in Equation 4.7 provides a measure of Hamming distance for the average cell from all other cells in a Tissue repertoire of cells T . As such, the ATD may be referred to as *Tissue Diversity (TD)* which may be calculated for each T in a Host H and averaged to provide the *Average Tissue Diversity (ATD)*. ATD provides an instantaneous measure of the diversity of the average T in a host, measured in bits. ATD provides a general indication about the distribution of information in the system. A relative increase in ATD suggests that the average T has a more diverse cell composition, whereas relative decrease in ATD suggests that the average T has a less diverse cell composition. Decreased homogeneity may suggest an increase in the organisation of information across the system and a decrease in homogeneity may suggest an increase in the mixing of the information content across the system respectively.

$$AverageTissueDiversity(H) = \frac{1}{n} \sum_{i=1}^n TissueDiversity(T_i) \quad (5.4)$$

Average Tissue Error The Host Error defined in 5.1 provides an indication of a Host system's error with regard to the H_{rs} prepared as a result of where a give TER interacted with the Host. Alternatively, a H_{rs} may be prepared systematically for each T in response to the Habitat B , and averaged across all T in H called the *Average Tissue Error (ATE)*. ATE provides an instantaneous measure of the error for the average Tissue component T in the system in the units of the problem space B . ATE provides a general indication of the capability of the average tissue in the system. A relative increase in ATE suggests the average T has a more generalise response capability, whereas a relative decrease may suggest a more specialised response capability. Importantly, an increased generalised response capability suggests an increased mixing of the information content across the system.

$$AverageTissueError(B, H) = \frac{1}{T_n} \sum_{i=1}^{T_n} \left(\frac{1}{B_n} \sum_{j=1}^{B_n} TissueError(I_j, T_{rsi}) \right) \quad (5.5)$$

5.4.4 Trends and Behaviours

Section 5.3.2 highlighted two points of concern regarding the behaviour of tissue-based systems under different exposure regimes: (1) the spatial organisation of information based on the spatial temporal regularity of tissue exposures, and (2) the system-wide consistency of response. These two concerns provide a trade-off in the organisation of cell-based information within a multiple tissue model between a uniform mix for a completely irregular tissue exposure pattern and isolated tissues for completely fixed regular tissue exposure patterns (see Figure 5.7). This provides a primarily research agenda for the Tissue Clonal Selection Paradigm for investigating Tissue Clonal Selection Algorithm behaviour under a variety of different Tissue Exposure Regimes. The results from such an investigation provide a basis for using such algorithms as tools in problems in which the exposure and/or use of information is controlled, uncontrolled or some combination (a consideration that is further investigated in the Host Paradigm, see Section 6.3.2).

Algorithm Mechanism to Behaviour

One may clearly distinguish between the measurable *emergent effects*, the *information management strategy* that such effects represent, and the grounded (*component-wise mechanisms*) of TCSA's that promote such strategies and effects (summarised in Table 5.6).

Effect	Strategy	Mechanism
<i>Consistency of Response</i>	Everything everywhere	Inter-tissue <i>mixing</i> of cells
<i>Spatial Organisation</i>	Only at point of use	Inter-tissue <i>isolation</i>

Table 5.6: Summary of the difference in effect, strategy and mechanism between the complementary concerns of tissue algorithms under discrete exposures.

The two identified emergent effects of interest are the systems holistic consistency of response and the systems spatial organisation. These represent two different information management strategies, specifically the dissemination and thus promoted availability of all acquired information across the entire system, and the specialisation and thus availability of acquired information at the point of use. As discussed in Section 5.3.2, these two strategies represent two different relationships between *acquisition* and *anticipated need*, specifically local acquisition with global anticipated need and local acquisition with local anticipated need. The first strategy may be promoted via the global mixing via continuous high-amplitude inter-component sharing of information, and the second strategy may be promoted via the segregation and thus isolation of acquired information. These strategies represent extremes in mechanism and behaviour with many intermediate states promoted via variation in mechanism, and thus variation in emergent effect.

Exposure Regimes and Trends

The five exposure regimes may be considered with regard to their exposure regularity and irregularity in terms of the scope of system information acquisition and anticipated need (summarised in Table 5.7). For example, the asymmetric regimes represents an

archetype of regular exposure with local exposure and anticipation whereas the random regime represents the archetype of irregular exposure with global exposure and anticipation. Both symmetric and point regimes expose the system in a regular manner and thus with expected local availability and anticipatory system expectation. This is the case for PTER, although given that STER is PTER applied to all components of the system, there is opportunity for the same information to be useful (anticipated) at any point of the system. The probabilistic regime provides an intermediate between the asymmetric regime and the random regime. For example, with a *duration* of 1 and frequency-biased re-exposure OTER provides a version of RTER with probabilistically increased regularity (local re-exposure and anticipation), whereas random is OTER with *duration* 1 and unbiased selection and ATER is OTER with an infinite duration. OTER with a moderate *duration* provides an explicit intermediate between the global-global behaviour of RTER with the local-local behaviour of ATER.

<i>Regime</i>	<i>Exposure/Acquisition</i>	<i>Re-Exposure/Anticipation</i>
ATER	Local	Local
STER	Local-Global	Local-Global
PTER	Local	Local
OTER	Local/Global	Local/Global
RTER	Global	Global

Table 5.7: Summary of the exposure and re-exposure behaviour of the five defined tissue exposure regimes and how they relate to system information availability and anticipated need.

Relative Measure Trends

The proposed measures provide information content both quantitatively as required, although more usefully qualitatively as relative differences between similar systems and the same systems under varied exposure regimes. Table 5.8 provides a summary of the system and component level measures and an interpretation of the relative movement trends. These general trends may be used as a guide for assessing results.

Measure	Movement	Description
<i>DH</i>	Up	System is more diverse (tissues are more heterogeneous).
	Down	System is less diverse (tissues are more homogeneous).
<i>HE</i>	Up	System performance has decreased.
	Down	System performance has increased.
<i>ATD</i>	Up	Tissues are more diverse (cells are more heterogeneous).
	Down	Tissues are less diverse (cells are more homogeneous).
<i>ATE</i>	Up	Tissues perform worse independently.
	Down	Tissues perform better independently.

Table 5.8: Summary of general trends and their interpretation for Tissue Clonal Selection Algorithms.

System Behaviour Expectations

Table 5.9 provides expected qualitative comparisons for a given tissue-based system with global and local anticipation strategies under local-local and global-global tissue exposure regimes. The qualitative expectations are given in terms of the system and component level diversity and error, where diversity reflects information content and error ultimately reflects correctness of anticipation. The ATER may be consider an archetype local-local (regular) TER given the piece-wise exposure of information to fractions of the information environment (antigenic habitat). RTER may be consider the archetype global-global (irregular) TER given the unbiased mapping between components and exposure. The remaining TER's provide intermediates for investigation. The error and diversity measures at host and tissue scope may be generalised to system and component, thus making the extended trends broadly applicable (such as in the case of the Host Paradigm in Chapter 6). As discussed, high-rate inter-tissue mixing provides an example of a global anticipation strategy, and isolation such as that used in the MTCSA provides an archetype of a local anticipation strategy. Remaining variations and elaborations of the TER provide intermediates between these strategies.

Measure		Local-Local TER		Global-Global TER	
<i>Scope</i>	<i>Type</i>	<i>Global Strategy</i>	<i>Local Strategy</i>	<i>Global Strategy</i>	<i>Local Strategy</i>
<i>System</i>	<i>Diversity</i>	Decreased	Increased	Decreased	Increased
<i>System</i>	<i>Error</i>	Increased	Decreased	Decreased	Increased
<i>Component</i>	<i>Diversity</i>	Increase	Decrease	Increase	Decrease
<i>Component</i>	<i>Error</i>	Decreased	Increased	Decreased	Increased

Table 5.9: Summary of the expected behaviours of the defined measures with regular and irregular exposure regimes on systems with mixing and isolated information dissemination strategies.

Expected system error predictably follows the matching between strategy and TER, and the inverse of the relationship. Thus, the expectation is that system error is a correlated match between TER and information management strategy, whereas component error is generally invariant to TER and is correlated with the information management strategy. Both system and component level diversity are also expected to be generally invariant to the TER and correlated with the strategy. System diversity is expected to decrease with a global strategy as the system becomes information-homogeneous and become information-heterogeneous with a local strategy. This relationship is expected to be inverted for component level diversity in which components generally increase in diversity with the global strategy and decrease with a local strategy given the correlated decrease and increase in intra-tissue specialisation respectively. A final axis to these expectation is responses, specifically the time allocated to address an environment given sufficient resources. It is expected that both system strategy types will arrive at the same or similar final system error after a prolonged period of time. Thus, the expected behaviours for this measure must be observed after a limited although reasonable amount of time.

5.4.5 Paradigm Agenda

This section outline a high-level research agenda for exploratory empirical experimentation for the Tissue Paradigm, considered in the remainder of this chapter.

1. *Specific Goals: Confirm Expectations*
 - (a) System-Error correlation between Strategy and TER, specifically Local strategy with Local TER and Global strategy with Global TER.
 - (b) System-Diversity and Component-Error and Diversity invariance on Strategy and TER, specifically Decrease in System-Diversity and Component-Error with Global strategy and Increase with Local strategy, as well as Increase in Component-Diversity with Global strategy and Decrease with Local strategy.
2. *General Goals: Consider Intermediates*
 - (a) Graceful falloff in measure with intermediate strategies and/or exposure regimes, generally a predictable continuum between the extremes in behaviour and thus measure.
 - (b) Intermediate information management strategies provide varied intermediate behavioural effects.

5.5 Lymphocyte Recirculation

5.5.1 Recirculation Metaphor and Strategy

Section 5.2 described the recirculation of lymphocytes around the vascular and lymphatic systems of a host such that a circulating pool of antibodies and cells are able to rapidly transit and disseminate between locations of need and potential need. Lymphocyte recirculation provides a metaphor for a computational mechanism investigating the ‘Consistency of Response’ emergent effect. Collectively, the lymphocytes in all tissues and thus the tissues themselves represent the acquisition of information (immunity) from antigenic habitat. The movement process facilitates the dissemination of information between the distributed tissue structure of the host. Ideally, all acquired information would be available in all tissues in all time for the system to be able to effectively respond to the unknown environment of spatial-temporal exposure patterns. This is not possible given (1) the limitations of carrying capacity of tissue locations and (2) the vast quantity of lymphocytes possessed by the entire system. The recirculation mechanism is a strategic exercise in anticipated information availability. An antigen must physically contact the receptors of cells to ‘test’ for a reaction, thus smaller lymphocyte pools may provide some efficiency benefits in terms of *local* trials. Further, not all information is required everywhere, rather specific information for a given antigen is needed at a specific location at an unknown time. Bounded lymphocyte recirculation provides a trade-off of these concerns. The *Lymphocyte Recirculation Strategy* is defined as a mobile distributed repertoire that promotes spatially distributed information availability of a limited resource to get the right *information* (cells) to the right *location* (tissue) at the right *time* (exposure).

5.5.2 Recirculation Tissue Algorithm

The tissue interaction step of the Tissue Clonal Selection Algorithm (TCSA) may be specialised for cell recirculation involving the *selection* of cells to depart a given tissue and the *integration* of cells arriving to the tissue. These may be separate decision processes although they are clearly related. For example, a simple model involves the removal of a single lymphocyte (posted to a neighbouring tissue) which frees a position in a fixed-size repertoire for one-arriving lymphocyte from a neighbouring tissue. The trafficking of lymphocytes may be elaborated as the configuration of a *migration* and *integration* operators, which may be specialised as the selection of a subset of a given repertoire for removal and integration into a neighbouring repertoire (a *remove-insert strategy*). A simplified tissue architecture may be defined as a ring network topology with uni-directional communication (*directed cyclic graph*) that superficially resembles a circulatory system between the tissue nodes. The migration configuration is defined by the number of cells per migration event (*amplitude*) and how often migration events occur (*frequency*). The number and consistency of the sample of lymphocytes posted to the neighbouring tissue ideally should be representative of the acquired information content of a given repertoire. The lymphocytes arriving from neighbouring tissue to be integrated into the repertoire should have the similar information content, thus meeting the goals of the recirculation strategy. In addition to providing a suitable sample of the receptors in the repertoire, the strategy should permit the further development of information in the repertoire. Arriving lymphocytes must be integrated into the repertoire. This may be more involved if there is a size mismatch between the number of cells selected and sent compared to the number arriving (for example if cells arrive from multiple tissue locations). Arriving cells may fill the vacant slots provided by the departed cells, and use a replacement strategy (as discussed for clone integration in the elaborated clonal selection algorithm) in the case of overflow.

Algorithm 5.9: TissueInteractions for Recirculation Tissue Clonal Selection.

Input: $H, N_{migration}$

```

1  $H' \leftarrow 0;$ 
2 foreach  $T_i \in H$  do
3    $T_i' \leftarrow \text{Sample}(T_i, N_{migration});$ 
4    $H' \leftarrow T_i';$ 
5 end
6 foreach  $T_i' \in H' \wedge T_i \in H$  do
7   if  $T_i \equiv T_N$  then
8      $\text{Integrate}(T_i, T_1');$ 
9   else
10     $\text{Integrate}(T_i, T_{i+1}');$ 
11  end
12 end

```

The *Recirculation Tissue Clonal Selection Algorithm (RTCSA)* is defined (Algorithm 5.9) as a specialisation of the *TissueInteractions* operation in TCSA in which a sample T_i' is taken $\text{sample}(T_i, N_{migration})$ of $N_{migration}$ cells from each Tissue T and trans-

mitted to a single and consistent neighbouring tissue T_{i+1} where they are integrated $integrate(T_{i+1}, T_i)$. The RTCAS is configured with a remove-insert migration strategy that selects a consistent amplitude ($N_{migration}$) random sample of cells for migration, which is applied each epoch of the algorithm (fixed frequency). Each tissue in the system participates in the trafficking which is applied in a lock-step manner of selection-removal-insertion after the exposures of the epoch have been satisfied. The default realisation of the $Sample(T_i, N_{migration})$ operation is to take a random sample (T_i), and the default realisation of the $Integrate(T_{i+1}, T_i)$ operation is concatenation to the repertoire filling the positions made by the sampling operation.

5.5.3 Recirculation Empirical Study

Aim

The aim of this empirical study was to investigate RTCSA as a viable mechanism for realising the ‘Consistency of Response’ emergent effect. Toward this end, the study had the following goals:

1. Confirm the expectations outlined in Section 5.4.5 using the RTCSA and the MTCSA.
2. Consider intermediates as outlined in Section 5.4.5 using variable migration sample size ($N_{migration}$) and the variety of Tissue Exposure Regimes.

Method

Algorithms The study considers two algorithms, the Minimal Tissue Clonal Selection Algorithm (MTCSA), and the Recirculation Tissue Clonal Selection Algorithm. MTCSA is a specialisation of the TCSA defined in Algorithm 5.7, that was configured with $N_{tissues} = 10$. Each T was an instance of the Replacement Cellular Clonal Selection Algorithm (RCCSA) defined in Algorithm 4.5, with the configuration $N_{cells} = 50$, $N_{selection} = 1$, and $N_{clones} = 5$. RTCSA is a specialisation of TCSA with a $TissueInteractions(H)$ that connects all T into a directed cyclic graph defined in Algorithm 5.9. The same $N_{tissues}$ and RCCSA configuration was adopted as was used in MTCSA. Three different migration sizes were used including: small (RTCSA-SML) with $N_{migration} = 5$ (10% of each T), medium (RTCSA-MED) with $N_{migration} = 20$ (40% of each T), and large (RTCSA-LRG) with $N_{migration} = 40$ (80% of each T).

Problems The colour space specialisation of the Infection Antigenic Exposure Paradigm (IAEP) defined in Algorithm 5.1 was used called the Infection Colour Space Problem (ICSP). The minimal variation of ICSP was used with $N_{infections} = 10$, and $N_{determinants} = N_{antigen} = 1$. Each I was a Colour Space Pattern (CSP) randomly generated at the beginning of each run. The five different Tissue Exposure Regimes (TER) defined in Section 5.4.1 were used for the ICSP. These included the ATER where the number of T matched the number of I (one-to-one), STER, RTER, PTER where T_{point} was fixed at tissue 1, and OTER with $duration = 15$.

Experiment Each algorithm used the Maximum Epochs Stop Condition (MESC) defined in Equation 4.8 with $MaxEpochs = 1000$. The four tissue specific measures defined in Section 5.4.3 were collected from the state of the system after the triggering of the stop condition. These included the system measures HE and HD, and the tissue measures ATE and ATD. Each algorithm and problem received a new and different random number seed each run. Algorithm-Problem combinations were repeated 30 times.

Results

Table A.1 in Appendix A.1.1 provides a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. Box-and-whisker plots are provided in which the results for each algorithm are aggregated across all TER for a each measure. Figure 5.20 shows HD, Figure 5.21 shows HE, Figure 5.22 shows ATD, and Figure 5.23 shows ATE.

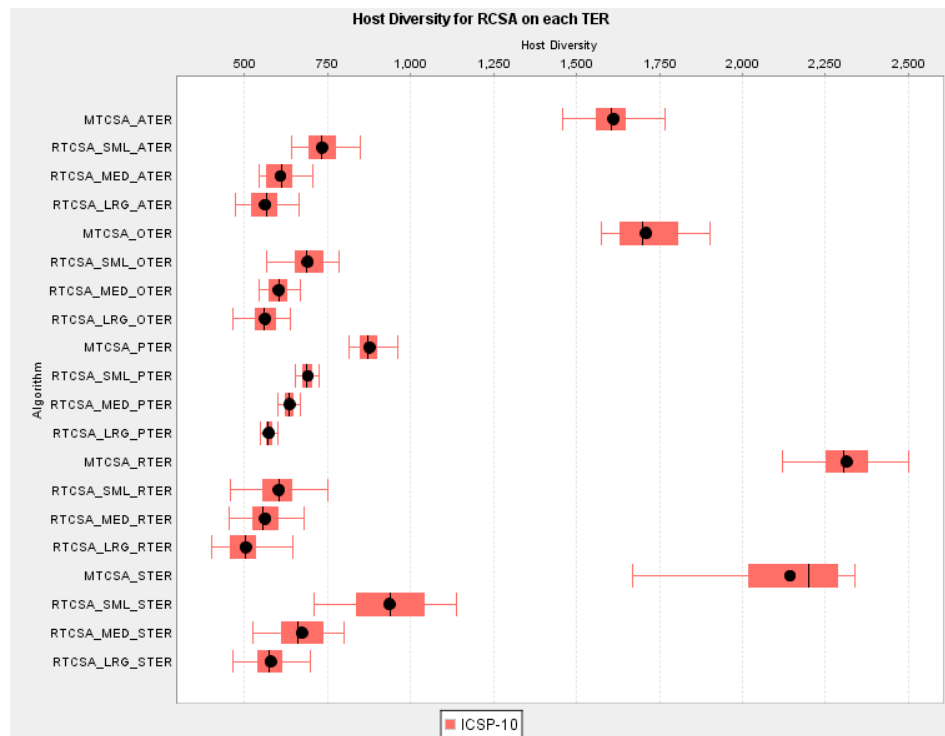


Figure 5.20: Box-and-whisker plot of Host Diversity (HD) across all TER for the RTCSA study.

Analysis

This section provides an analysis of the results summarised in the previous section. Table 5.7 was used to consider the general behaviour of each of the specific TER's. Table 5.6 was used to consider the general behaviour of the global and local information management strategies of RTCSA and MTCSA. Finally, Table 5.8 was used to qualitative interpret the summarised results.

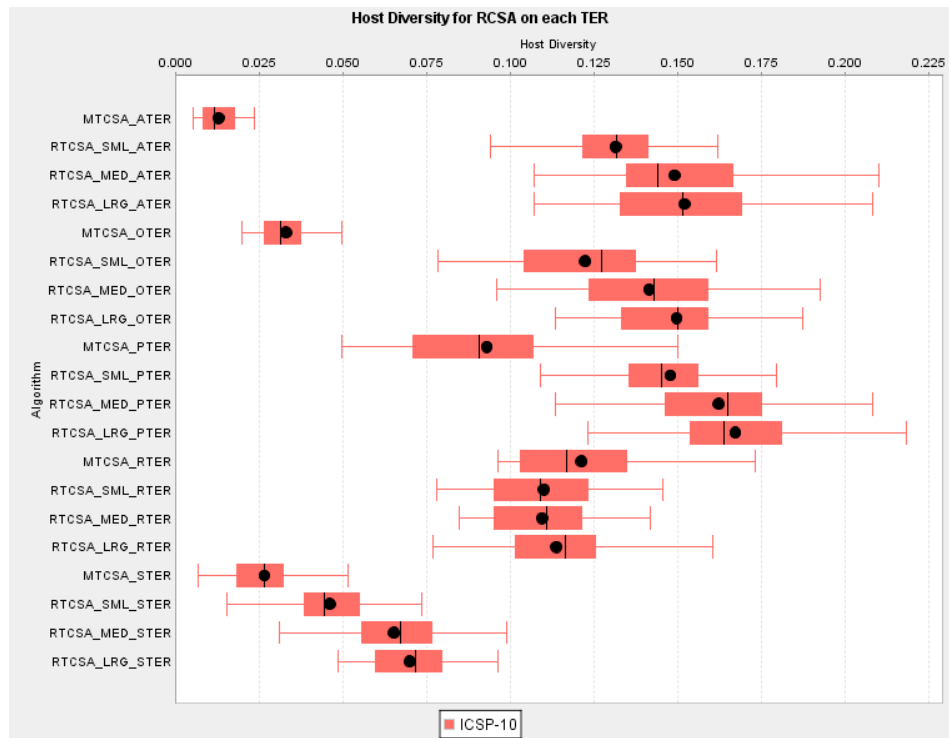


Figure 5.21: Box-and-whisker plot of Host Error (HE) across all TER for the RTCSA study.

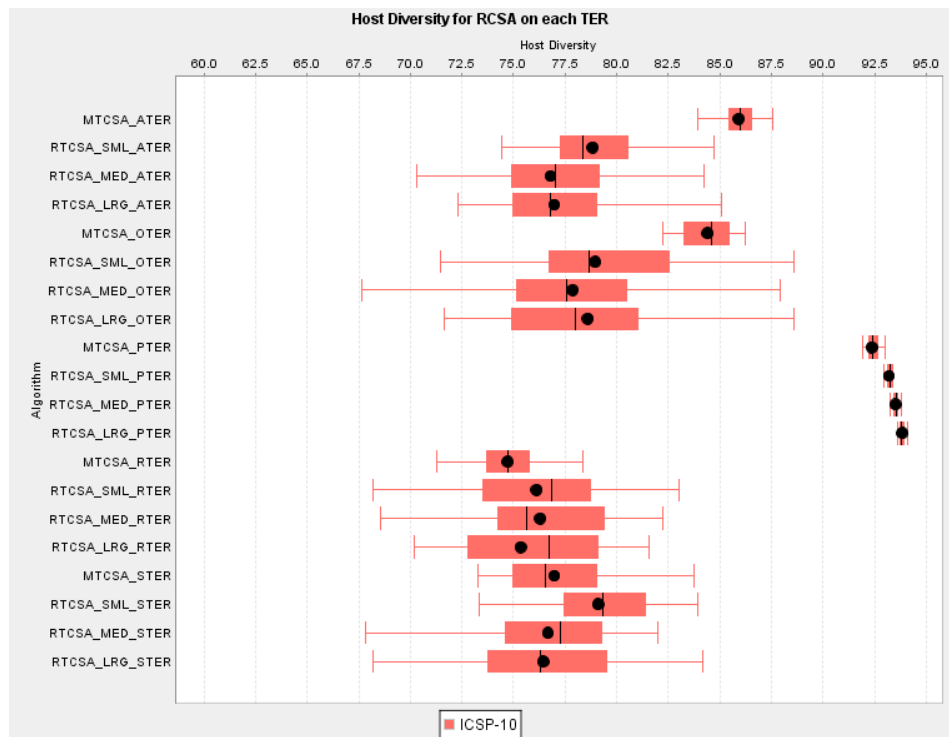


Figure 5.22: Box-and-whisker plot of Average Tissue Diversity (ATD) across all TER for the RTCSA study.

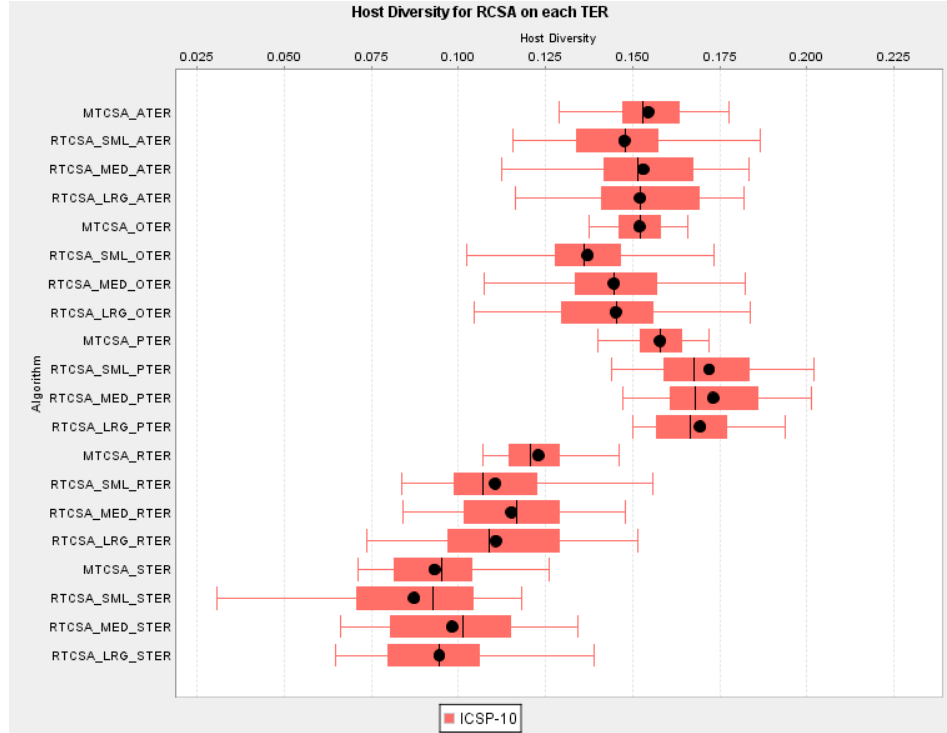


Figure 5.23: Box-and-whisker plot of Average Tissue Error (ATE) across all TER for the RTCSA study.

Information Management Trends This section considers the expected trends of the local and global information management trends under local and global information exposure regimes outlined in Table 5.9 from Section 5.4.4. ATER and RTER are considered local and global information exposure regimes (exposure and re-exposure) respectively, and MTCSA and RTCSA-LRG are considered local and global information management (anticipation) respectively. Table 5.10 provides a variation on the expected trends summary presented in Table 5.9 in which the specific measures of the selected algorithms are provided with qualitative comparison. Direct comparison between these tables addresses the primary aim of the experiment, to confirm or reject the behaviour expectations outlined in the tissue paradigm agenda in Section 5.4.5.

Measure	ATER			RTER		
	<i>RTCSA-LRG</i>	<i>MTCSA</i>	Sig.	<i>RTCSA-LRG</i>	<i>MTCSA</i>	Sig.
<i>HD</i>	562.136 (↓)	1610.68 (↑)	True	503.919 (↓)	2313.88 (↑)	True
<i>HE</i>	0.152 (↑)	0.013 (↓)	True	0.114 (↓)	0.121 (↑)	False
<i>ATD</i>	76.969 (↓)	85.903 (↑)	True	75.347 (↑)	74.7 (↓)	False
<i>ATE</i>	0.152 (↓)	0.155 (↑)	False	0.111 (↓)	0.123 (↑)	True

Table 5.10: Summary of the results for MTCSA and RTCSA-LRG on ATER and RTER and the relative quantitative difference is measures.

Considered first are the algorithm results for the local information exposure regime (ATER). From a system perspective both the HD and HE met the expectations in which the MTCSA resulted in decreased error and increased diversity, whereas the RTCSA-LRG

resulted in the inverse behaviour of increased error and decreased diversity. From a component perspective the comparison in diversity (ATD) did not meet expectations. Instead of an increase in average component diversity with a global strategy, this behaviour was observed by the local strategy. This means that the expectation that the global strategy would decrease component specialisation was interpreted incorrectly, demonstrating that a general increase in specialisation with a local strategy results in an increase in the diversity within the average component. Also unexpected was the lack of significant difference between the average component error between the two strategies. This was a surprising result as the mixing of acquired information provided by RTCSA-LRG provided no measured performance improvement over segregated repertoires on the local exposure regime.

The algorithm results from the global information exposure regime (RTER) are also considered. From a system perspective the diversity matched the expectation in which the segregation of MTCSA resulted in an increase in diversity compared to the decreased provided by recirculation in RTCSA-LRG. The expectation that system error would decrease with RTCSA-LRG was not confirmed as there was no significant difference between the two strategies. This is a surprising result as it generally suggests that consistency of response may not improve system performance under an irregular global information exposure regime². From a component level the expectation of component error was confirmed in which the global strategy resulted in a decrease in HE compared to an increase seen with the local strategy. The expectation of component diversity was not confirmed as there was no significant difference between the two strategies. This finding, together with the lack of significant difference in HE demonstrates that RTER alone was sufficient to engender a ‘consistency of response’ irrespective of the information management strategy, promoting a reconsideration of cause and effect between management strategies and exposure regimes. This is considered in the remainder of this section. Table 5.11 provides a reproduction of the expectations from Table 5.9 of global and local information management strategies and information exposure regimes, with those results that did not meet the expectations highlighted.

Measure		Local-Local TER		Global-Global TER	
<i>Scope</i>	<i>Type</i>	<i>Global Strategy</i>	<i>Local Strategy</i>	<i>Global Strategy</i>	<i>Local Strategy</i>
<i>System</i>	<i>Diversity</i>	Decreased	Increased	Decreased	Increased
<i>System</i>	<i>Error</i>	Increased	Decreased	Neutral	Neutral
<i>Component</i>	<i>Diversity</i>	Decrease	Increase	Neutral	Neutral
<i>Component</i>	<i>Error</i>	Neutral	Neutral	Decreased	Increased

Table 5.11: Summary of confirmed behaviours of global and local information strategies and exposure regimes, highlighting where expectations were incorrect.

Information Exposure Trends This section relaxes the notions of local and global information exposure regimes and considers general trends across the TER’s in the con-

²This suggestion is further supported by the lack of significance between MTCSA and all tested variations of RTCSA on RTER as shown in Table A.1.

text of both the local and global management strategies (MTCSA and RTCSA-LRG respectively). Table A.2 in Appendix A.1.1 provides a restricted summary of results that provides direct comparison and statistical significance between TER's for the local and global management strategies in which the effects of the exposure regimes may be teased from their information management strategies.

Observations may be drawn firstly from the behaviour of the five TER's with the local management strategy (MTCSA). From a system perspective, PTER resulted in the least diverse system whereas RTER resulted in the most diverse system. This suggests that the holistic adaptation the system was exposed to in RTER and other system-wide exposure regimes (ATER, STER and OTER) result in a system information content that is more diverse than the random repertoires under PTER (besides the single exposed tissue). ATER achieved the lowest error, which is unsurprising given the $T-I$ specialisation fostered by the TER, and RTER achieved the highest error which was also expected given it provides the least $T-I$ specialisation (the most irregular regime). From a component level, PTER achieved the highest ATD and ATE³ suggesting that from a component perspective point-wise exposure of information results in the expected behaviour of repertoires that individually perform poorly on average. This behaviour was not restricted to PTER as was also demonstrated to a lesser degree on the other restricted exposure regimes (ATER and OTER). RTER resulted in the lowest average component diversity, although with similar behaviour exhibited by STER, suggesting system-wide exposure of information results in repertoires that have consistent internal information content on average. STER achieved the lowest ATE followed by RTER demonstrating that the same effect on component diversity results in improved tissue capability.

Observations may be summarised from the behaviour of the five TER's with the global management strategy (RTCSEA-LRG). From a system perspective STER resulted in the lowest diversity and error. RTER resulted in the highest diversity although was not significantly different from many of the other TER's. Interestingly, PTER resulted in the highest error suggesting that global management presented a major disruptive influence to localised exposure regimes, demonstrated by the equally poor HE for ATER and OTER compared to the much improved STER and RTER. Further, this inverse in behaviour from the local management strategy results demonstrate the suitability of holistic exposure regimes to the global management strategy, and localised exposure regimes to the local management strategy. From a component level, PTER achieved the highest average component diversity and error demonstrating the same trend with the MTCSA management strategy. RTER resulted in the lowest ATD with no significant difference in result with the STER and ATER system-wide exposure regimes. Likewise STER resulted in the lowest ATE followed closely by RTER mirroring the pattern observed with MTCSA.

Toward teasing apart the influences of TCSA's and TER's the observations may be reduced into the following trends that describe the generalised relationship between information exposure strategy and information management strategy:

1. *Component Trends are Strategy Invariant*

³Difference in ATE was not significant between PTER and ATER for MTCSA.

- (a) Localising information exposure increases average component diversity, globalising information exposure decreases average component diversity.
- (b) Localising information exposure increases average component error, globalising information exposure decrease average component error.

2. *System Trends are Strategy Variant*

- (a) Globalising information management increases diversity and error under localised information exposure, and decreases diversity and error under globalised information exposure.
- (b) Localising information management decreases diversity and error with localised information exposure, and increases diversity and error with a globalised exposure strategy.

The component measure invariance trend with information management strategy demonstrates that the information exposure regime is the determining factor. This trend was observed independently for both the MTCSA and RTCSA-LRG strategies across the TER's, although specific and typically minor differences for strategies on specific TER's. The system measure dependence on information management strategy demonstrates the suitability of a localised strategy on localising information exposure and the suitability of a globalised strategy on globalising information exposure. This trend was observed independently and complementary (the inverse case) for both MTCSA and RTCSA-LRG strategies across the TER's, although with differences in the specific TER in each case. An important implication regarding TER's in addition to that of their spatial-temporal regularity is the scope of their exposure behaviour as either system-wide or restricted, providing a defining quality for localisation and dissemination management strategies. These trends suggest that TER's can predictably influence the component-wise organisation of information in terms of holistic problem error and information content diversity. The system-level suitability trends also suggest that MTCSA does demonstrate the spatial organisation (localisation) of information emergent effect and that RTCSA-LRG does demonstrate the consistency of response (dissemination) emergent effect, confirming the expectations regarding each approaches mechanism.

Recirculation Trends This section considers information management strategies beyond the two extremes already considered. Specifically, this section considers the range of recirculation sample sizes ($N_{migration}$) as intermediates between the two extremes, across all five TER's. These recirculation trends are drawn from the principle result summarises, specifically Table A.1, and Figures 5.20, 5.21, 5.22 and 5.23.

From the system perspective, migration size demonstrated the consistent trend of decreasing diversity (HD) where the larger the sample size the larger the decrease in HD compared to MTCSA. A consistent trend was also demonstrated with error, where the increase in recirculation size resulted in increased system error compared to MTCSA. This generalised error trend applied to all TER's with system-wide exposure scope (all tissues participated, not PTER). These system-level diversity and error trends provide

clear intermediates in system-level behaviour (localisation and dissemination) as defined by the captured measures HD and HE. From a component perspective the trends are less bold as a the system level. ATD demonstrate no clear TER invariant trends with recirculation size, although there was clear commonality in behaviour. For both the ATER and OTER exposure regimes, the increase in recirculation size resulted in a step decrease in average component diversity that seemingly levelled off after the drop from MTCSA to RTCSA-SML. This demonstrates that a sufficiently small sample size was required to effect a decrease in the average tissue diversity under those regimes, after which further increases in sample size had little effect. Interestingly, a different effect was observed for RTER and STER where a small recirculation size resulted in a small increase in ATD, after which further increases in size resulted in decreases. These effects were not observed with PTER, where the increase in recirculation size resulted in the progressive increase in ATD. The variation in recirculation size demonstrated a common trend across all TER's except PTER for ATE, where a small increase resulted in a decreased error, and subsequently larger recirculation sizes increased error. This trend demonstrates that a sufficiently small recirculation size is needed to effect a drop in AE under those exposure regimes. The inverse of this trend was observed with PTER where small recirculation resulted in an increase in ATE and larger sample sizes decreased this error toward the MTCSA score.

The observed measure-based may be reduced into the following generalised trends regarding RTCSA recirculation size:

1. *Graceful Intermediate Dissemination Effect*

- (a) The RTCSA recirculation size provides a graceful intermediate in emergent behaviour (system-level) between localisation demonstrated in MTCSA and dissemination.
- (b) Small recirculation size provides a decrease in average component error where large recirculation sizes do not on system-wide exposure regimes with local and global re-exposure.
- (c) Recirculation decreases the diversity of the average component on system-wide exposure regimes with local re-exposure.
- (d) Recirculation increases the diversity of the average component on system-wide exposure regimes with global re-exposure.

Conclusions

This section summarises the findings of the empirical study into the Recirculation Tissue Clonal Selection Algorithm, in terms of the primitives that were the focus of the study and the expectations that motivated the study.

1. *Study Primitives*

- (a) The MTCSA is a viable realisation of the spatial organisation effect (localisation) via an isolation mechanism.

- (b) The RTCSA is a viable realisation of the consistency of response effect (dissemination) via a recirculation mechanism.

2. *Confirmations and Findings*

- (a) The behavioural expectations for the two different management strategies under two archetype exposure regimes were considered with a mixture of system and component level expectations being confirmed and rejected.
- (b) The localisation or globalisation (exposure/re-exposure) of an exposure regimes defines the component-level error and diversity that is generally invariant to the information management strategy
- (c) The localisation or globalisation of an exposure regime is suited to a localising and disseminating information management strategy with regard to system-level error and diversity.
- (d) In addition to the regularity and holistic information exposure/re-exposure properties of tissue exposure regimes, a third important consideration is the scope of system exposure, specifically system-wide or constrained exposure/re-exposure.

3. *Intermediates*

- (a) Recirculation size of the RTCSA provides a viable behavioural intermediate between localisation and dissemination with predictable systemic effects and identified component trends on varied exposure regimes.
- (b) There is a need for a system to balance localisation and globalisation of acquired information when presented with an *unknown* exposure regime.

This final point about intermediates motivates the two remaining investigations in this chapter, and likely the whole Tissue Clonal Selection Paradigm.

5.6 Lymphocyte Homing

5.6.1 Homing Metaphor and Strategy

Section 5.2.2 considered the directed trafficking of cells, with the specific example of the imprinting of T-cells by dendritic cells as to the chemical properties of a site of infection. The educated cells then recirculate until they locate the imprinted chemical signature and seek their cognate antigen. The RTCSA demonstrated that continual lymphocyte migration provides a background disruptive behaviour to a tissues ability to specialise a response to regular exposures. The education and homing mechanism for cells may be generalised to a specialisation of recirculation that empower cells to express a preferential residence in tissues for which they may be useful. This bottom-up behaviour is called a *Preferential Residence Strategy* and is comprised of *Cell Imprinting* and *Differential Migration* mechanisms to realise the emergent effect.

In addition to specialising the resources of the system to the antigenic exposures, the specialised resources (lymphocyte receptors for antigen) must be positioned such that they are used in the right locations. The distribution (space and time) of antigenic exposures is unknown to the system (defined by a TER), thus the system recirculates specialised resources between the locations of exposure to maximise the chance application of the resource in an exposure event. Lymphocytes interact with each other and interact with antigen in this spatial structure. Each tissue type associated with a region of the host has its own unique chemical makeup, which may be used to differentiate tissues types and thus regions of the host. Lymphocytes may interact with tissues using specialised receptors that detect and differentiate between tissue types. This interaction occurs when the lymphocyte recirculates around the host, and may be used by the cell to travel specific pathways and to enter and remain in a specific tissue in search of its cognate antigen. These navigation receptors are not expressed in naïve lymphocytes, instead are expressed in matured cells that are produced after an interaction with antigen. A mature receptor (memory or effector) expresses navigation receptors for the tissue surroundings when it is conceived. Therefore, mature cells are specialised both for an antigenic stimulus, and for the spatial location (tissue) that is known to have been exposed to the antigenic stimulus. This *Cell Imprinting* process occurs irrespective of whether a selected (high-affinity) lymphocyte is mature or naïve. In effect, specialised resources may be re-educated in that their progeny may be imprinted with a different location than themselves (presuming the cell was selected in a tissue different from that it may home to). This imprinting mechanism facilitates both the spatial anticipation of antigen exposure based on past exposures, and in addition to recirculation, it facilitates the use and re-education of lymphocytes.

Imprinting empowers individual mature lymphocytes to assert preference for their locations allowing tissues to discriminate preferential residence during the selection of cells to traffic. Homing requires an imprinting apparatus, which includes both a unique identity or signature for tissues, and a way for lymphocytes to be impregnated with that signature. The match of information between a given lymphocyte and a tissue may be used to bias the selection of cells from the repertoire to migrate. One may consider a Boolean match/no-match for lymphocytes in their home tissue, and those not in their home tissue. Those progeny lymphocytes created as a result of the clonal selection process are already in their home tissue, thus they have a match to their tissue's signature. All those lymphocytes that migrate from another tissue will not match. Thus, a migration sample selection policy that biases toward no-matches and away from matches (1) exercises a preference for lymphocytes (2) facilitates an automatic recirculation population of homeless lymphocytes. This process is called *Differential Recirculation* where the probability that a mature lymphocyte is selected in an outbound migration sample is biased by the match between a given lymphocyte and the tissue's signature. This mechanism facilitates homing but introduces the problem of recruitment into the recirculation pool. Those cells in the recirculating pool will continue recirculating until they return their 'home', unless they get used or killed elsewhere. The initial recruitment probability will be uniform across the repertoire as there are no migrant cells to consider. If a recirculating cell is replaced

(killed) by the proliferation strategy at another (non-home) tissue, then that tissue will fill the gap in the recirculating population.

5.6.2 Homing Tissue Algorithm

The Homing Tissue Clonal Selection Algorithm (HTCSA) is defined as an extension of the RTCSA with the addition of a cell imprinting and differential migration mechanisms.

Cell Imprinting A pattern-recognition method may be used that is somewhat similar to the approach used for Mediated Cellular Clonal Selection (Section 4.6) or Network Cellular Clonal Selection (Section 4.7), although rather than a lymphocyte matching onto other lymphocytes with a mock ‘surface feature’, lymphocytes are provided with an additional receptor for detecting and matching the tissue they are in. In keeping with the bitstring basis of the minimal clonal selection algorithm (and extensions), a simple and scalable method is to assign each tissue (or tissue group) a random binary string. Each mature lymphocyte (that is a lymphocyte created from a cloning and maturation process) is then assigned (imprinted) with a copy of the tissue bitstring specific to the location of its creation. A less elaborate approach involves assigning a distinct identification number to each T in H and assigning a tissues identification number to activated cells at the time of exposure. Algorithm 5.10 defines a generic tissue exposure function in which the cell repertoire is ordered by its affinity for an Antigen A ($OrderByAffinity(T, C_i)$) and a subset of the $N_{imprint}$ highest affinity cells for A are imprinted with the current tissue identification number.

Algorithm 5.10: Cell Imprinting for Homing Tissue Clonal Selection.

```

Input:  $I, T, N_{imprint}$ 
1 foreach  $A_i \in I$  do
2   foreach  $C_j \in T$  do
3      $Exposure(C_j, A_i)$ ;
4   end
5    $OrderByAffinity(T, C_i)$ ;
6   for  $j \leftarrow 1$  to  $N_{imprint}$  do
7      $C_j.TissuePreference \leftarrow T.id$ ;
8   end
9 end

```

Differential Migration The $TissueInteractions(H)$ operation of the RTCSA that governed tissue interactions in the general TCSA may be specialised to exploit the imprinted cells by (1) migrating those cells that are not in their preferred tissue or do not have a preference, and not migrating those cells that are in their preferred tissue. Such preferences may be deterministic or probabilistically implemented. Algorithm 5.11 provides a re-definition of the $TissueInteractions$ defined for RTCSA in Algorithm 5.9 that simplifies differential migration by not migrating those cells in their preferred tissue. The $SelectSingleCell$ operation draws cells randomly from the T .

Algorithm 5.11: Differential Migration for Homing Tissue Clonal Selection.

```
Input:  $H, N_{migration}$ 
1  $H' \leftarrow 0;$ 
2 foreach  $T_i \in H$  do
3    $T_i' \leftarrow 0;$ 
4   while  $T_i' \neq N_{migration}$  do
5      $C' \leftarrow \text{SelectSingleCell}(T_i);$ 
6     if  $C'.TissuePreference \neq T_i.id$  then
7        $T_i' \leftarrow C';$ 
8     end
9   end
10   $H' \leftarrow T_i';$ 
11 end
12 foreach  $T_i' \in H' \wedge T_i \in H$  do
13   if  $T_i \equiv T_N$  then
14      $\text{Integrate}(T_i, T_1');$ 
15   else
16      $\text{Integrate}(T_i, T_{i+1}');$ 
17   end
18 end
```

5.6.3 Homing Empirical Study

Aim

The aim of this empirical study was to investigate the Homing Tissue Clonal Selection Algorithm (HTCSA) as an intermediate information strategy between high migration RTCSA and the segregated MTCSA. Toward this end, the study had the following goals:

1. Contrast HTCSA against MTCSA and RTCSA as a mechanism for simultaneously promoting dissemination and localisation of acquired information.
2. Investigate intermediates of the HTCSA mechanism and subsequent effect by varying the $N_{imprint}$ parameter.

Method

Algorithms The study considers three algorithms, the Minimal Tissue Clonal Selection Algorithm (MTCSA), and the Recirculation Tissue Clonal Selection Algorithm (RTCSA), and the Homing Tissue Clonal Selection Algorithm (HTCSA). The MTCSA and RTCSA algorithm are configured as was defined for the RTCSA empirical study in Section 5.5.3 with $N_{tissues} = 10$, $N_{cells} = 50$, $N_{selection} = 1$, $N_{clones} = 5$ for MTCSA and RTCSA, and $N_{migration} = 5$ (10% of each T) fixed for the RTCSA. HTCSA is an extension of RTCSA with the addition of an imprinting operation as defined in Algorithm 5.10 for each exposure and a differential migration as defined in Algorithm 5.11. The imprinting mechanism was specialised such that a variable number $N_{imprint}$ of the best-matching cells were imprinted for each exposure, as follows: small (HTCSA-S) $N_{imprint} = 1$ (2% of T), medium (HTCSA-M) $N_{imprint} = 5$ (10% of T), and large (HTCSA-L) $N_{imprint} = 10$

(20% of T). A probabilistic sample selection scheme was used in differential migration, changing the configuration from 100% rejection if a cell was housed within its ‘home’ tissue to a probabilistic 20% (80% chance of home cells not being migrated per migrating cell selection).

Problems The same Infection Colour Space Problem and Tissue Exposure Regimes were used as was defined for the RTCSA empirical study in Section 5.5.3.

Experiment The same experimental setup was used as was defined for the RTCSA empirical study in Section 5.5.3.

Results

Table A.3 in Appendix A.1.2 provides a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. Box-and-whisker plots are provided in which the results for each algorithm are aggregated across all TER for a each measure. Figure 5.24 shows HD, Figure 5.25 shows HE, Figure 5.26 shows ATD, and Figure 5.27 shows ATE.

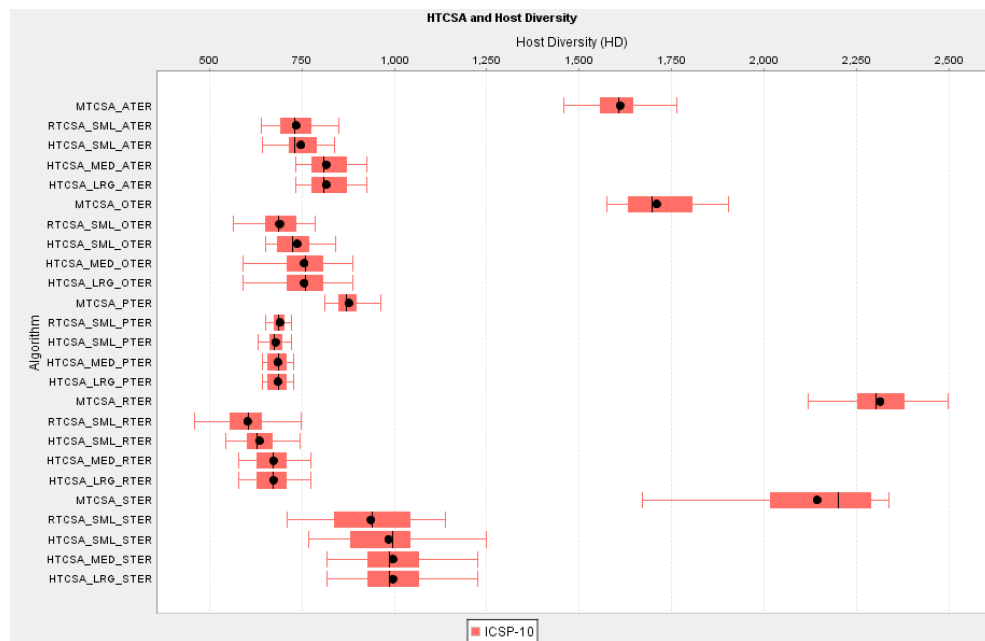


Figure 5.24: Box-and-whisker plot of Host Diversity (HD) across all TER for the HTCSA study.

Analysis

This section analyses the summarised results in Table A.3 in the context of the expectations from Section 5.4.4 and their confirmation and other findings regarding MTCSA and RTCSA from the empirical study in Section 5.5.3.

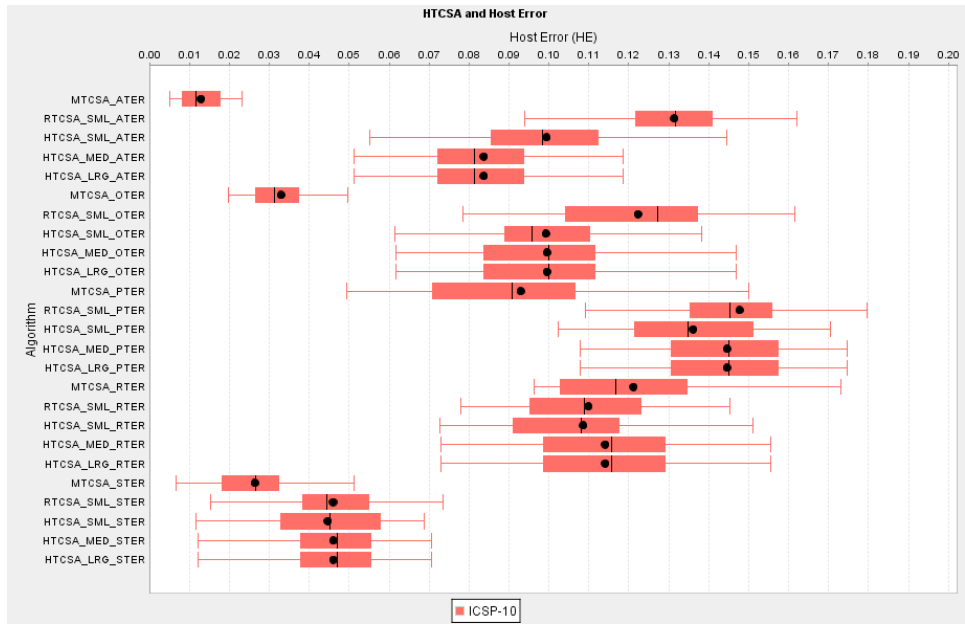


Figure 5.25: Box-and-whisker plot of Host Error (HE) across all TER for the HTCSA study.

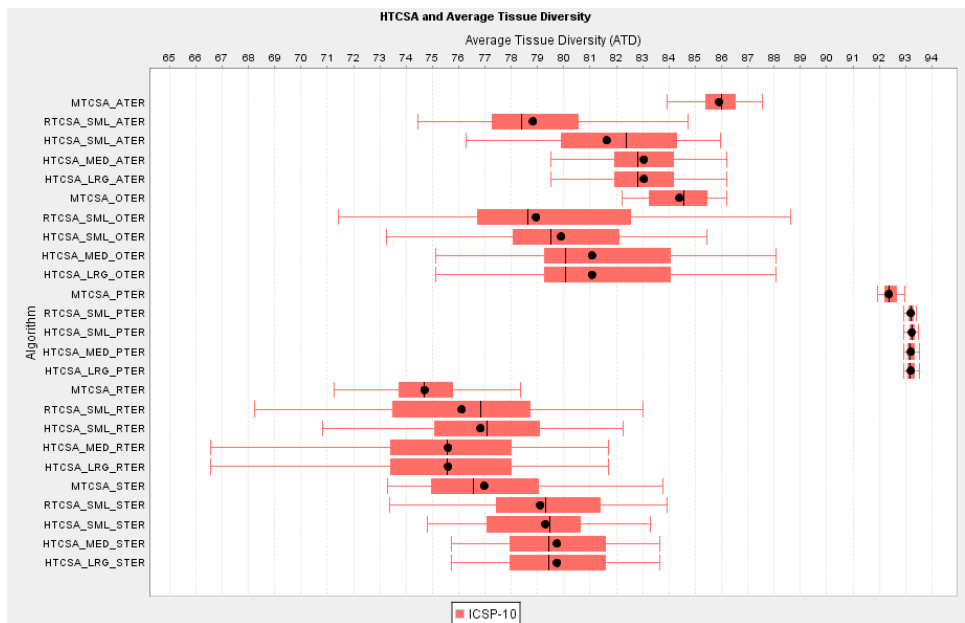


Figure 5.26: Box-and-whisker plot of Average Tissue Diversity (ATD) across all TER for the HTCSA study.

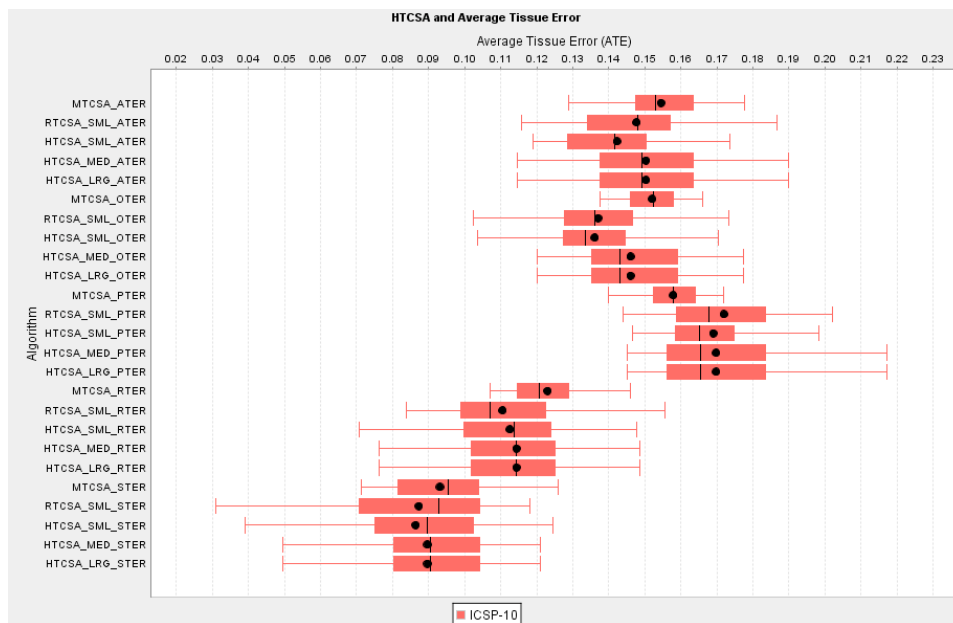


Figure 5.27: Box-and-whisker plot of Average Tissue Error (ATE) across all TER for the HTCSA study.

Localisation and Dissemination Trends This section considers a comparison in behaviour between the localising MTCSA, the disseminating RTCSA-S and preferential residence HTCSA-L. From a system perspective HTCSA-L performed in a similar way to that as RTCSA-S, both of which were contrastingly different from MTCSA. Both RTCSA-S and HTCSA-L were less diverse than MTCSA, although RTCSA-S more so than HTCSA-L. MTCSA achieved consistently lower HE, although HTCSA-L demonstrated a decreased error on constrained exposure regimes (ATER, OTER and PTER) compared to RTCSA-S, and no significant difference in error on the remaining (RTER and STER) exposure regimes. This was an important finding highlighting the increased spatial organisation capability for recirculation with preferential residence over recirculation without preferential residence on regular exposure regimes. From a component perspective the performance of HTCSA-L was distinct from MTCSA and similar to RTCSA-S, with generally little significant difference in the use or non-use of preferential residence. ATD was increased in HTCSA-L over RTCSA-S for the same constrained exposure regimes in which an improvement in system error was observed (ATER, OTER and PTER), and no significant difference for the remaining regimes. This result highlights that under regular and constrained exposures (different information in different locations in a regular manner), preferential residence changes the composition of components toward that of segregated components, whilst under the presence of recirculation, an effect that is not present for other exposure regimes. Specifically, preferential residence organises information toward the benefit of the system, when it is possible to do so as determined by the exposure regime. Homing had no consistent significant effect on recirculation regarding Average Tissue Error. The observations comparing and contrasting the three strategies may be reduced to the following general behavioural trends for the preferential residence strategy:

1. Homing reduces system error over recirculation on constrained (regular) exposure regimes whilst remaining consistent with recirculation on irregular and system-wide exposure regimes.
2. Homing increases system diversity (organisation) compared with recirculation.
3. Decreased system error is reflected in increased average component diversity on constrained exposure regimes over recirculation, whilst remaining consistent with recirculation on irregular and system-wide exposure regimes.
4. Improved organisation of homing does not effect average component error compared with recirculation.

Information Exposure Trends Section 5.5.3 identified the two trends that (1) average component measures of error and diversity are defined by the localisation (constrained) or globalisation (system-wide) of the exposure (generally invariant to the information management scheme), and (2) that globalising management is suited to system-wide exposure regimes (global-global) and localising management is suited to constrained exposure regimes (local-local). This section considers the HTCSEA in the context of these trends, specifically the degree to which the trends hold or are influenced by preferential residence. Table 5.12 summarises results for HTCSA-L across all the TER's, providing a complement to the same presentation for MTCSA and RTCSA-L in Table A.2.

System	Problem	HD		HE		ATD		ATE	
		\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
TCSA	TER								
HTCSA-L	ATER	816.023	53.374	0.084	0.019	83.046	2.228	0.15	0.018
HTCSA-L	OTER	756.301	64.802	0.1	0.019	81.077	3.568	0.146	0.015
HTCSA-L	PTER	685.395	27.472	0.145	0.017	93.206	0.161	0.17	0.018
HTCSA-L	RTER	673.063	58.672	0.114	0.022	75.595	3.745	0.114	0.018
HTCSA-L	STER	995.975	114.778	0.046	0.013	79.731	2.071	0.09	0.022
Significance		True ^a		True		True ^b		True ^c	

^aFalse for PTER and RTER

^bFalse for STER and OTER

^cFalse for ATER and OTER

Table 5.12: Summary of results from an empirical study into the HTCSA comparing TER's with the HTCSA-L information management strategy.

The results show the same trend with HTCSA-L as was shown with recirculation with a larger sample size in Table A.2 (RTCSA-LRG rather than RTCSA-SML that would be a fair comparison with any assessed HTCSA). This confirmation supports the comparisons of RTCSA, MTCSA, and HTCSA and outcomes in the previous section. Specifically demonstrating that the influence of the TER's on the system remained generally consistent with RTCSA, and that the differences between RTCSA and HTCSA are as a result of the HTCSA strategy.

Imprinting Size Trends This section considers the intermediate effects of varying the number of cells imprinted with each exposure ($N_{imprint}$), and thus the extent of each tissue that is empowered to express a preference during migration. From a system perspective some preferential residence was better than none, although the increases in $N_{imprint}$ generally did not have a measurable effect. This result was generally demonstrated for both Host Error and Diversity across all exposure regimes except PTER. The Point Exposure Regime demonstrated a clear and steady increase in the organisation effect of preferential residence with the increase in the number of cells imprinted per exposure. This suggests that homing can counter the disruptive effects of recirculation, a behaviour that gracefully increases with the number of imprinted cells per exposure. This effect is most pronounced with PTER because it is the regime that is the most susceptible to the disruptive effects of blind recirculation. From a component perspective a similar general result was observed for average component diversity where preferential residence increased diversity with little difference between the value of $N_{imprint}$, and only on the ATER and OTER regimes with no significant effect on the other regimes. No consistent significant effect was observed for average component error in varying $N_{imprint}$, including a value of zero (RTCSA-S). The observations regarding the variation of the number of imprinted cells per exposure may be generalised to the following trends:

1. Homing is generally better than no homing (recirculation) and the specific choice of $N_{imprint}$ is not critical.
2. Increase in effect with increase in $N_{imprint}$ is pronounced on constrained local exposure (PTER), although is not significant under other exposure regimes.

An interesting consideration for future investigation is the correlation of decreased system diversity and increased system error with migration size, and the general increase in organisation and decrease in error with preferential residence. In particular the knowledge that the homing effect does scale when isolated in PTER. One may investigate the relationship between the recirculation size ($N_{migration}$) and the imprint size ($N_{imprint}$) in order to find general configuration trends toward a trade-off and balance of the dissemination and preferential organisation effect.

Conclusions

This section summarises the findings of the empirical study into the Homing Tissue Clonal Selection Algorithm, in terms of the primitives that were the focus of the study and the expectations that motivated the study.

1. *Homing*
 - (a) The preferential residence (homing) effect realised in the imprinting and differential migration mechanisms provides a strategy that can exploit the dissemination of information by recirculation whilst countering the effect by localising used information.

- (b) Cell localisation (20% of a repertoire) is insufficient to counter all of the disruptive effects of continuous recirculation.
- (c) Localising used information allows progeny of such information to freely (blindly) recirculate.

This final point provides a suggestion as to the basis of the mechanism in the homing approach, and a basis for increasing the probability of migration for cells without a preference. The second important finding for the homing approach (besides it behaving as intended) is that even with a large number of cells imprinted each exposure, it is unable to completely counter the disruptive effects of continuous recirculation, rather it is limited to complementing the process. Countering the disruptive effects would likely require the imprinting of some or all of the progeny of exposures, which is expected to impact the dissemination properties of recirculation. This theme of allowing sufficient specialisation under regular exposure regimes whilst also disseminating acquired information is further explored via a Tissue Inflammation metaphor in the next section.

5.7 Tissue Inflammation

5.7.1 Inflammation Metaphor and Strategy

When a tissue is infected by a pathogen that damages the local area, the response is inflammation at the site. Inflammation has many symptoms, not limited to swelling and the increased carrying capacity of lymphocytes in the tissue. In addition, the blood vessels that lead to the area dilate, increasing the flow of lymphocytes (information) to the site, and typically involves the use of chemical signals or attract recirculating cells and/or arrest cell movement. An increase in carrying capacity allows more lymphocytes created at the site of infection to remain there, and more cells to migrate and home into the tissue and collect at the site of infection. Inflammation provides a mechanism to temporally decrease the competition between lymphocytes for limited resources within a tissue, providing a tissue-level *direct recruitment mechanism*. One may consider the tissue paradigm from the perspective of active distributed recruitment strategies as opposed to passive distributed information organisation strategies (already considered). Information is streamed from the tertiary tissue to the secondary tissue although the streaming is facilitated by courier cells. These cells are aware of where they collected their material and thus are capable of imprinting that information onto the produced effector cells. The result is an effect that differentiates the lymphocyte movement behaviour from that of *Recirculation* where cells blindly seek their cognate antigen, and from *Homing* where cells home into the chemical signature of infected tissue. The Homing model couriers information about the location of the pathogen invasion to the cells which are imprinted with that information. This is a ‘fire and forget’ response strategy. Like the recirculating of effectors, the homing effectors recirculate the system in search of their cognate antigen, although if the cells detect the imprinted signature of the site of infection, they take up residence and provide a local search strategy. Further, less effectors are created than in the recirculation model (for

example mature T-cells as opposed to antibodies), thus the response is both smaller and more directed. One may generalise two response strategies as follows:

- *Undirected Response*: Flood the system with effectors that will generally address the specific site of invasion and any other sites where the pathogen may have spread, or penetrated on multiple fronts (recirculation method).
- *Directed Response*: Produce a smaller set of effectors that home into the specific site of infection, and have a low chance of addressing additional sites of infection (homing method).

Both the undirected recirculation model and the directed homing model allocate (recruit) resources to the requirements of a specific antigenic exposure. Further, both strategies address the needs of the distributed information environment (resident receptors and inbound pathogen information). These two approaches may be considered to describe *Distributed Recruitment Strategies*. The undirected approach is a flooding strategy that does not use any structural information other than the generation of a response in the vicinity of the pathogens first detection. The directed approach is a homing strategy in which structural information is used directly to get the resources to where they are required. Recruitment may be phrased as resource allocation with the desire of maximising payoff. Thus, *Distributed Recruitment* may be defined as a strategy in which information dispersed across a spatial structure is sought (for example selected) and employed for a specific task (antigen neutralisation). The two proposed models both relate to the changing of lymphocyte trafficking behaviour, treating the lymphocyte (a space in the repertoire) as a limited resource.

5.7.2 Inflammation Tissue Algorithm

Inflammation is a strategy that promotes making more information available in the local area to address the infection, and there relaxing the competition between receptors for limited positions in the repertoire. This effect may be realised by temporarily increasing the capacity of a tissue for lymphocytes around the time of exposure. The increase in carrying capacity would be stepwise at the time of exposure such that the clones created at the time of exposure may be housed in the tissue with less competition. The decrease in carrying capacity would not be stepwise, rather it would decay over time (for example as a linear function of the increased size over time). Ideally, the capacity increase would facilitate the approximate optimal amount of local information needed to address the exposure. The decay rate would facilitate the approximate optimal amount of time that the local information may be required (for example antigen endurance or antigen exposure frequency). The amount of inflammation (increase in carrying capacity) may be a fixed amount (step size) and the decay function may also be fixed. In addition, an upper limit of lymphocytes for a site is required (upper bound on space complexity). For each exposure, the increase would be applied. The effect would be the allocation of space for lymphocytes temporally and proportional to exposures. Thus, inflammation allows space complexity to

be treated as a resource whose allocation is adapted by the system based on an exposure (needs) basis. Adaptation of lymphocyte capacity allows not only increases when required, but also decreases when it is not needed. The following lists some additional mechanisms that may complement a local increase of the carrying capacity of a tissue:

- *Lymphocyte Trapping*: Lymphocytes may be prevented from leaving a site of infection for a temporally time. This provides symmetry to the increase in carrying capacity, which puts less pressure on the cells in the present tissue. This pressure is reduced further by migrating (out-bound) fewer lymphocytes.
- *Upstream Dilation*: The tissues ‘upstream’ of an exposed tissue may be influenced to increase their down-flow migration of lymphocytes. Upstream and down-stream refers to the uni-directional flow of lymphocytes around the directed cyclic graph structure in which tissues are arranged. Increased flow of lymphocytes provides more localised information to address the exposure.
- *Effector Discrimination*: Tissue behaviour may be extended further such that it may actively discriminate as to the effectors to accept and integrate into the local repertoire from migration, and those that are decidedly not useful during an exposure may be migrated away from the tissue. Thus, tissues may be given the ability to discriminate between effector lymphocytes.

Algorithm 5.12: Increased Capacity for Inflammation Tissue Clonal Selection.

Input: $B, H, N_{maxcells}$

```

1 foreach  $I_i \in B$  do
2    $H' \leftarrow \text{SelectTissuesToExpose}(H)$ ;
3   foreach  $T_{i'} \in H'$  do
4      $\text{Exposure}(T_{i'}, I_i)$ ;
5      $T_{i'}.N_{cells} \leftarrow N_{maxcells}$ 
6   end
7 end

```

The *Inflammation Tissue Clonal Selection Algorithm (ITCSA)* is defined as an extension of the TCSA with the addition of a mechanism to increase the carrying capacity of exposed tissues and a mechanism to decay carrying capacity of expanded tissues back to default levels. Algorithm 5.12 defines a Habitat B exposure of a Host B in which the role of the TER is generalised to that of the selection of subset ($T_{i'}$) of H to expose to each I_i of B . This generalisation allows the explicit increase in carrying capacity of selected and thus exposed tissues from the default N_{cells} to a maximum $N_{maxcells}$. This additional carrying capacity is filled by progeny created from future exposures in a given expanded tissues, although alternatively may be used to integrate recirculated cells from a RTCAS specialisation of ITCSA’s *TissueInteractions* operation. Algorithm 5.13 provides a definition of a default *TissueInteractions* operation for ITCSA in which the carrying capacity of expanded tissues is progressively decreased by one cell each host exposure (epoch). This is realised by the selection and removal of a randomly selected cell from the repertoire.

Algorithm 5.13: TissueInteractions for Inflammation Tissue Clonal Selection.

Input: H, N_{cells}
1 **foreach** $T_i \in H$ **do**
2 **if** $T_i.N_{cells} > N_{cells}$ **then**
3 $T_i.N_{cells} \leftarrow T_i.N_{cells} - 1$;
4 **end**
5 **end**

5.7.3 Inflammation Empirical Study

Aim

The aim of this empirical study was to investigate the Inflammation Tissue Clonal Selection Algorithm (ITCSA) as an intermediate information strategy between disseminating RTCSA and localising MTCSA. Toward this end, the study had the following goals:

1. Investigate carrying capacity as a recirculation-independent mechanism for promoting the localisation of acquired information.
2. Investigate the capability of ITCSA to counter the disruptive effects of recirculation.

Method

Algorithms The study considers three algorithms, the Minimal Tissue Clonal Selection Algorithm, and the Recirculation Tissue Clonal Selection Algorithm, and the Inflammation Tissue Clonal Selection Algorithm. The MTCSA and RTCSA algorithm was configured as was defined for the RTCSA Empirical Study in Section 5.5.3 with $N_{tissues} = 10$, $N_{cells} = 50$, $N_{selection} = 1$, $N_{clones} = 5$ for MTCSA and RTCSA, and $N_{migration} = 5$ (10% of each T) fixed for the RTCSA, referred to as RTCSA-S (small migration). The ITCSA is a specialisation of the TCSA that increases the carrying capacity of tissues exposed to infections to $N_{maxcells}$ defined in Algorithm 5.12, and progressively decreases the tissue to N_{cells} each epoch defined in Algorithm 5.13. Two variants of ITCSA were investigated, one based on the MTCSA with no T - T interaction referred to as ITCA-N (no migration), and one based on RTCSA with recirculation ($N_{migration} = 5$) referred to as ITCSA-S (small migration). Both ITCSA used the same carrying capacity increase of $N_{maxcells} = 60$ (20% increase over $N_{cells} = 50$).

Problems The same Infection Colour Space Problem and Tissue Exposure Regimes were used as was defined for the RTCSA empirical study in Section 5.5.3.

Experiment The same experimental setup was used as was defined for the RTCSA empirical study in Section 5.5.3.

Results

Table A.4 in Appendix A.1.3 provides a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure val-

ues. Box-and-whisker plots are provided in which the results for each algorithm are aggregated across all TER for a each measure. Figure 5.28 shows HD, Figure 5.29 shows HE, Figure 5.30 shows ATD, and Figure 5.31 shows ATE.

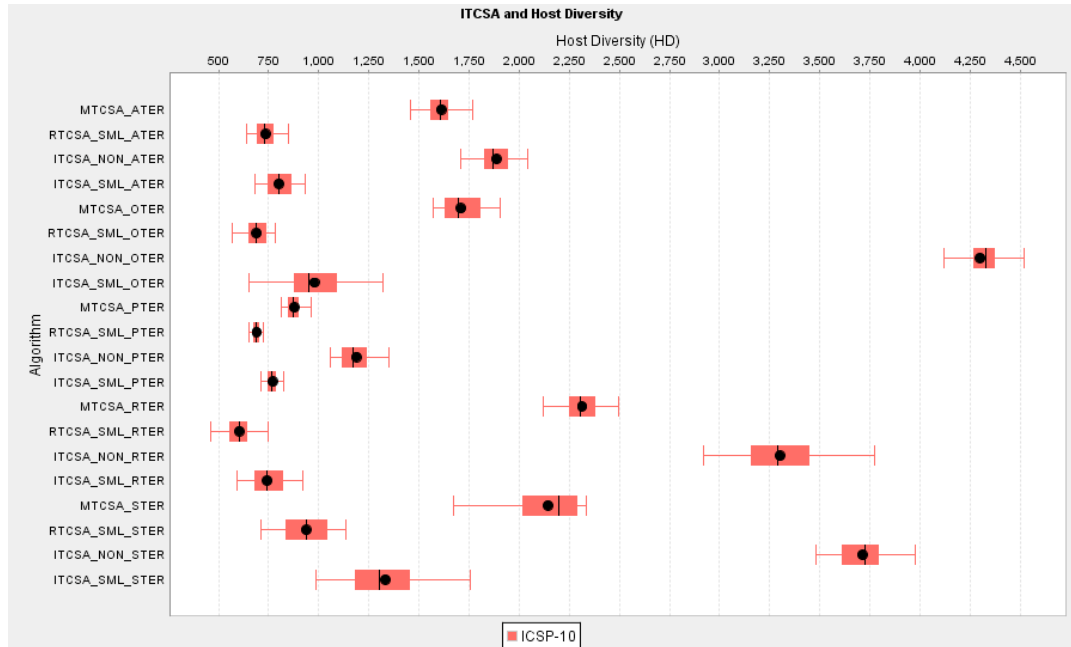


Figure 5.28: Box-and-whisker plot of Host Diversity (HD) across all TER for the ITCSA study.

Analysis

This section analyses the summarised results in Table A.4 in the context of the aims outlined for this empirical study.

Localisation Trends This section considers the localisation effects of increasing carrying capacity in ITCSA with out recirculation as compared to MTCSA. From a system perspective, ITCSA-N resulted in large increases in diversity compared to MTCSA, an effect expected by a strong localising method (Section 5.4.4). Unexpected was the concurrent large increases in host error across all exposure regimes. This suggests that the additional resources were exploited by specialising the system, although at the cost of resources that were required for improving the response. From a component level, ITCSA-N resulted in a general decrease in the average tissue diversity, and a general increase in the average tissue error. This increased tissue internal homogeneity is expected given that increased capacity is exploited by local clonal progeny, although the impaired average tissue capability across the TER's correlates with the systems increase in response error. This suggests that an artefact of the mechanism may be responsible for disrupting the acquisition of information. The observations may be generalised into the following trends:

1. Inflammation facilitates component and system level increases in diversity as expected by a strong localising effect.

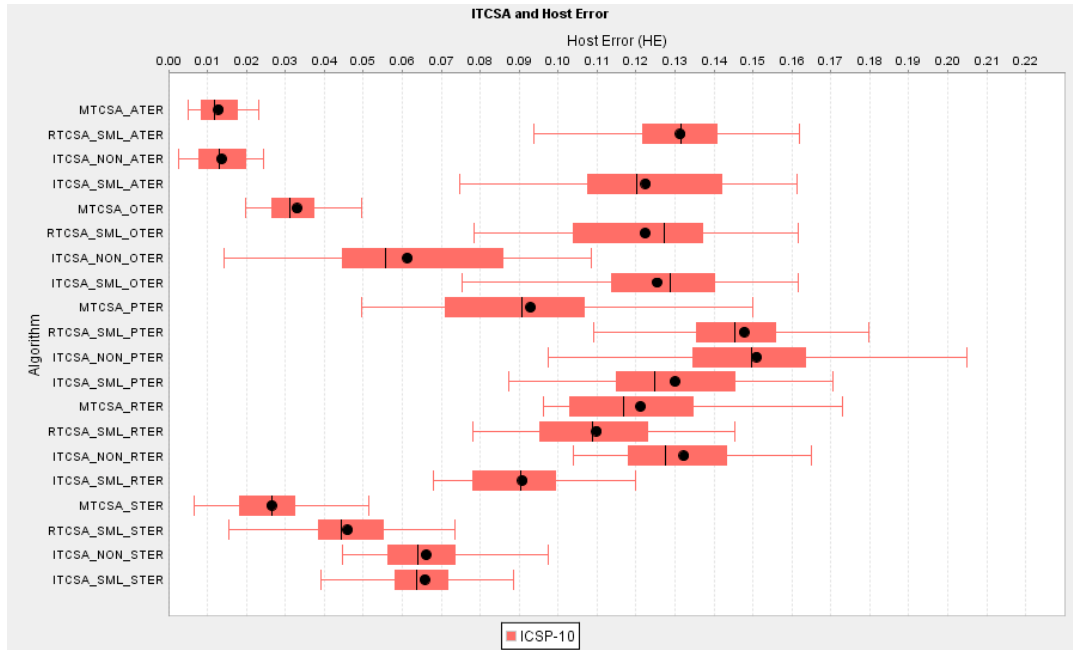


Figure 5.29: Box-and-whisker plot of Host Error (HE) across all TER for the ITCSA study.

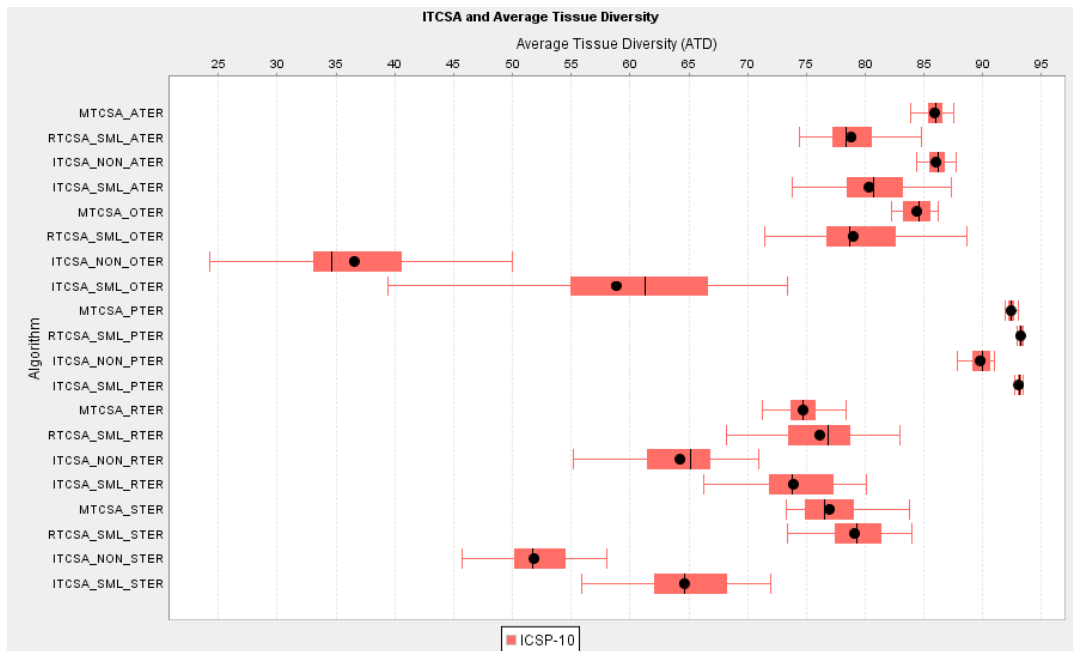


Figure 5.30: Box-and-whisker plot of Average Tissue Diversity (ATD) across all TER for the ITCSA study.

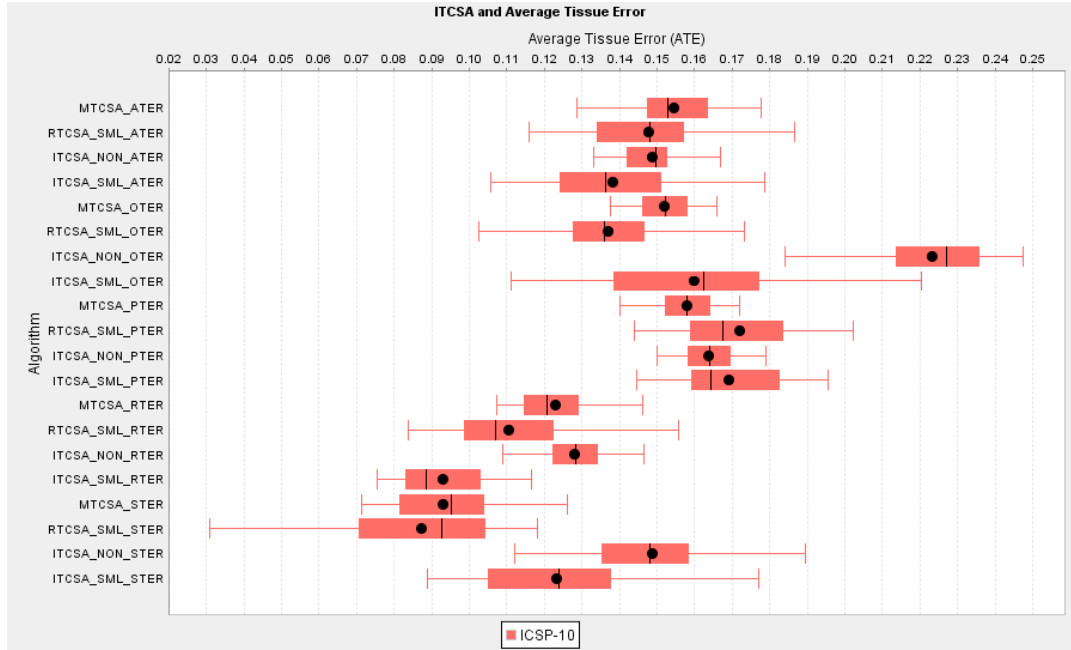


Figure 5.31: Box-and-whisker plot of Average Tissue Error (ATE) across all TER for the ITCSA study.

2. Inflammation's promise of localisation is broken by the increase in system and component level error, likely an artefact of the specific implementation.

Counter-Disruptive Trends This section considers the capability of ITCSA (specifically ITCSA with recirculation: ITCSA-S) to counter the disruptive effects of recirculation to the local exposure regimes such as ATER, OTER and PTER. From a system level ITCSA-S resulted in an increase in diversity compared to RTCSA-S, although less a dramatic increase than between ITCSA-N and MTCSA given RTCSA also decrease diversity compared with MTCSA. The diversity increase aligns with the expectation of increased localisation, and with the observation in the previous comparison. ITCSA-S resulted in an increase in error on ATER and STER and a decrease in system error for PTER and RTERT over RTCSA-S. From a component perspective inflammation resulted in a general decrease in the average tissue diversity, with an increase observed on ATER. Component error was increased for OTER and STER, although decreased with inflammation on ATER and RTERT.

The important finding and confirmation that inflammation can counter the disruptive effects of recirculation is the decrease in HE on PTER and RTERT. This finding becomes clear when one considers the global effects of varying carrying capacity with each exposure. In the case of regular system-wide exposure regimes (these are regimes under which each tissue is exposed to some information each host exposure) such as STER and ATER, the carrying capacity for each tissue is always increased to $N_{maxcells}$. For those regimes where regular system-wide exposure is not assured, such as OTER, PTER and RTERT, the irregularity promotes variability in the carrying capacity. Given that system error was not decreased for STER and ATER it suggests that the holistic increase in repertoire

size does decrease the disruption of recirculation, rather the inflammation provided an effect that increased error on ATER and STER. In the case of PTER and RTER system error was reduced over recirculation suggesting that the variable repertoire sizes across the host and the inflammation mechanisms method for exploiting the extra size (with progeny of selected cells) can only be exploited under these circumstances. Interestingly inflammation provided some increase and decrease in average per-component error that did not align with the regular and irregular system-wide exposure correlation, and in particular improved average component capability for both the archetype regular and irregular exposure regimes (ATER and RTER). These observations are reduced to the following general trends:

1. Inflammation generally counters the disruptive effects of recirculation with regard to system diversity, although system-level benefits are observed only under those exposure regimes that have irregular exposure regimes with regard to the number of tissues exposed to information per host exposure (epoch).
2. Inflammation under consistent system-wide exposure regimes results in a holistic increase in repertoire size that under the ITCSA inflammation mechanism decreases system capability with regard to host error.

Conclusions

This section summarises the findings of the empirical study into the Inflammation Tissue Clonal Selection Algorithm, in terms of the primitives that were the focus of the study and the expectations that motivated the study.

1. *Inflammation*

- (a) Inflammation promotes localisation that is apparent with and without recirculation, although it generally disrupts system capability (error).
- (b) The increase in carrying capacity mechanisms is only effectively exploited as a counter measure to the disruptive effects of recirculation under exposure regimes that promote disparity in the resulting effect (non-uniform tissue exposure).

The results suggest at the potential of the emergent effect although raise questions regarding the specifics of the chosen mechanism. Specifically, the method by which increased carrying capacity is exploited (what cells fill the space), and the method by which carrying capacity is decreased (what cells are selected and deleted from the repertoire). It is likely that viable cell lineages are stopped under random cell deletion resulting in the general increase in system error rather than an expected neutrality and potential decreases provided by increase specialisation. An alternative mechanism may be increased progeny tournaments for progeny during integration into the repertoire after creation, providing a natural mechanism for competition for increasingly limited resources. There is much room for elaboration on the primitive realisation of the inflammation effect, not limited to the discussed manipulation of the recirculation effect (increased tissue inflow and decreased

tissue outflow), as well as the clear integration of preferential residence retain specialised resources at the point of use (homing).

5.8 Chapter Summary

5.8.1 Paradigm Review

It is important to note that tissue models of the immune system have been considered before by Twycross and Aickelin [400] concerned with the innate immune system, and by Bentley, et al. [36] concerned with tissue growth models. The principle difference and contribution of the presented approach is the integrated and subsumed relationship with the existing clonal selection paradigm. The *Tissue Clonal Selection Paradigm (TCSP)* was defined as the investigation of the cellular clonal selection paradigm as constrained by the concerns of (1) multiple discrete repertoires of cells called *Tissues* and their interactions, and (2) the concerns of regimes of discrete repertoire exposures with information called *Tissue Exposure Regimes (TER)*. The paradigm exploits the metaphor of a holistic immune system called a *Host*, arranged into a lymphatic system of lymphoid tissues that may inter-communicate using the differential trafficking of lymphocyte cells called a *Tissue Clonal Selection Algorithm (TCSA)*. The principle information processing interest of the paradigm as gleaned from the metaphor are the information management strategies that may be employed to address the known and unknown regularities and irregularities in an antigenic *Habitat* of information called an *Infection Antigenic Exposure Problem (ITEP)* toward continuous acquisition of information and the anticipation of needs of acquired information. An abstract foundation for the paradigm was defined in terms of model components, discrete exposures with spatial-temporal regularities, and a series of architectures that may be exploited for arranging the components of tissue models.

5.8.2 Principles and Findings

The following summarise the important principles and findings from the definition and investigations into the Tissue Clonal Selection Paradigm:

General Principles

1. *Natural organisation based on exposure:* Given sufficient time and resources, information is naturally organised in a fixed structured system based on regularity and consistency of its exposure. For example, system-wide exposure naturally promotes system-wide anticipation and point-wise exposure promotes point-wise anticipation.
2. *Problem is in the exposure regime:* Resources are not always sufficient, therefore information management strategies promote more efficient and effective organisation of information in the face of unknown regularities and irregularities in information exposure and anticipation.
3. *Strategies based on tissue-level immunology:* The immune system inspires strategies for this problem as the survival of the host relies on the effective acquisition

of information and efficient intra-host application of acquired information to fight infections.

Specific Findings

1. *Recirculation*: Scope of system exposure (number of tissues) is best addressed with a match in scope of information management strategy at a system level, although the general trend at the component level is invariant to the strategy used. For example the specialisation of response is suitable for regular constrained exposure, whereas the unbiased dissemination is suitable for irregular and/or unconstrained exposure, although such strategies do not counter the general natural component-wise trends in organisation.
2. *Homing*: The dissemination of information under a constrained exposure regime when such dissemination is unneeded provides a disruptive effect that may be countered by empowering cells to exhibit a preferential residence, whilst not impairing dissemination under those regimes that benefit from the effect.
3. *Inflammation*: The increase in carrying capacity toward decreasing the competition for local resources at those points of the system that are exposed provides a second mechanism for decreasing the disruptive effects of recirculation under those exposure regimes where disparity of the effect is promoted, at the cost of providing a negative effect under those regimes that do not.

5.8.3 Integration

The *Cellular Clonal Selection Paradigm* defined and investigated in Chapter 4 was concerned with the adaptation as a method for information acquisition, and the how that method is influenced by degenerate information and inter-cell interactions. The Tissue Clonal Selection Paradigm defined and investigated in this chapter took information acquisition via adaptation for granted, using it as a primitive component in a Tissue Clonal Selection Algorithm raising the level of abstraction regarding information acquisition and adaptation from cells to tissues. The interest in this chapter was how different decentralised information management strategies effected the organisation acquisition and use of information under different information exposure regimes. The abstract tissue paradigm is grounded in Chapter 8 in the context of two general problem domains, highlighting the potential benefits the tissue algorithms provide over the cellular clonal selection algorithms. The subsumption theme of raising the level of abstraction in information processing is continued in Chapter 6 in which a Tissue Clonal Selection Algorithm is made a primitive component in a broader *Host Clonal Selection Paradigm* that is concerned with information management strategies in a population of hosts.

Chapter 6

Host Clonal Selection

6.1 Chapter Overview

This chapter extends the Tissue Paradigm, by raising the level of abstraction for clonal selection to that of a distributed population of self-governed immune systems interacting with each other in an antigenic environment. Section 6.2 considers a holistic perspective of the mammalian immune system and reviews the ways in which populations of immune systems may interact with each other both directly and indirectly. Specifically, host immune system interactions are considered in the context of the passive and active forms of immunisation in a population and across generations, as well as the evolutionary concerns for the basis of immunological function. Section 6.3 abstracts the background biology of host immunisation and immune system evolution and presents the *Host Clonal Selection Paradigm*. The paradigm is presented in terms of shared and evolved immunity interaction models, controlled and uncontrolled antigenic exposures, and population and generational host architectures. Collectively these concerns provide a foundation for the series of Host Clonal Selection Algorithms proposed and investigated in this chapter. Sections 6.5 and 6.6 investigate the inter-population sharing of the trigger and product of acquired immunity respectively in a transmission and shared immunity algorithms. Sections 6.7 and 6.8 investigate the the inter-generational sharing of the product and the evolution of the mechanisms of acquired immunity respectively. The information processing and management strategies of the host paradigm demonstrated similar localisation and dissemination qualities as were demonstrated in the tissue paradigm with the additional level of control over interactions and decoupling of cellular repertoires promoted by the perspective of clonal selection.

6.2 Populations of Interacting Immune Systems

Immune systems do not exist in isolation, in terms of the various organs and interrelated subsystems within the host organism, or in terms of the populations to which the host organisms belong. In raising the level of abstraction to that of populations of immune systems, one may be interested in the manners in which immune systems interact. This transition may be considered as moving from thinking of the Darwinian struggle of lym-

phocytes within one host holistically providing a defence to the organism, to populations of these systems holistically providing a defence to a population (or a species). This section considers populations of immune systems, specifically with regard to evolutionary concerns and natural and artificial immunisation, and ways in which these systems may interact. This general review is multi-disciplinary crossing such fields of study as population and evolutionary ecology, microbiology, epidemiology, and immunology. The intent of this section is to highlight some systems and processes that underlie the interactions between holistic immune systems to provide the basis for the inspiration of distributed computational Artificial Immune System models.

6.2.1 Immunisation

Generally, immunisation implies that a host has been exposed to an antigen (immunogen) inducing an immune response that has educated the system with respect to more effectively and efficiently detecting and neutralising the antigen in the future. When a pathogen enters the host and elicits a response (and the host survives), the host gains a level of immunity to the pathogen. This process can be induced artificially through inoculation (controlled exposure to the pathogen), and is the basis of vaccination. Interestingly, immunity is not limited to this reactive method, and evolution has designed a temporary and passive form of immunity that is employed by mothers to protect children from pathogen before and after birth. This section reviews the various natural and artificial forms of acquiring immunity in terms of those mechanisms that elicit an immune response, and those that do not. Table 6.1 provides a summary of examples.

	Natural	Artificial
<i>Active</i>	Infection	Vaccination, Inoculation
<i>Passive</i>	Maternal Immunity	Injection of antibodies

Table 6.1: Summary of the different types of immunity.

A developmental oddity of the acquired immune system is that the mass of the system is at its peak in early childhood, and decreases with time until early adulthood [289]. This raises questions as to the usefulness of decreasing the resources of the defence system with age. An intuitive consideration would suggest that the size of the system would increase with the increase of acquired knowledge over the lifetime of the host. An answer to this question provides insight into acquired immunity as a strategy. Children are more vulnerable to infection than adults (children under the age of 5 suffer 8 to 24 times more infection than other age groups). Each pathogen encountered by an infant is novel and must be responded to anew. As a strategy, acquired immunity learns the hosts' pathogenic environment early and quickly. It provides an attempt to increase the hosts' early and prolonged survival within its environment. During the early stages of life the system maximises diversity, maximises cell proliferation potential, and maximises the rate of memory cell creation. Cell numbers peak in the first six months of life then steadily decline. Further, the developing infant immune system is buffered to the antigenic environment by the

temporally inherited maternal immunity (natural passive) gained by the mother over her lifetime. This form of immunity is discussed in the following section.

Passive Immunity

Passive immunity is the transfer of active humoral immunity in the form of ready-made antibodies to a host organism. Antibodies are relatively short-lived molecules that have no reproductive capability, thus the defensive capability provided by them is fast, effective, and temporary. A common form of natural passive immunity is maternal immunity (for review see [42, 136, 196, 54]). Antibodies are provided by the mother to the foetus across the placenta and in breast-fed milk after birth (mucosal immunity). Vaccination of the mother during pregnancy may directly benefit the child as the antibodies produced by the mother will be passed to the foetus [172]. During breastfeeding, lymphocytes in the mother move to lymph nodes near the breast and release large amounts of antibodies into the breast milk. The antibodies enter the gut of the infant and provide protection against bacteria [363], and assists in oral tolerance (avoiding allergic reactions to food). The antibodies provide an initial and powerful boost in defence to the infant, which endures for a prolonged period after breastfeeding ceases. Breastfeeding stimulates the developing immune system, improving later response to vaccination, although initially suppresses the effector response of the infant immune system causing the failure of effective (artificial) vaccination during this period. It decreases infant mortality rate and provides benefits for the metabolism and disease resistance later in life [308].

Passive immunity may be provided artificially through the transfer of blood or serum, and is applied when there is insufficient time for the host to develop an immune response. Some examples include viral and bacterial infections and poisons such as snake bites [245]. Cell-mediated immunity may also be transferred in this manner (effector systems for detecting and neutralising infected cells rather than foreign molecules) and is called *adoptive immunisation*. This is the transfer of recirculating lymphocytes rather than antibodies, although is difficult as such cells may be identified as foreign and attacked by the host immune system, thus the process is rarely used in humans. Monoclonal antibodies (mAb) are an artificial passive immunity where large quantities of identical antibodies are synthesised from identical cloned lymphocytes [407]. Unlike polyclonal antibodies that vary in their specificity for an antigen, monoclonal antibodies all have the same specificity, and thus may provide a strong and directed defence. In theory it is possible to synthesise antibodies to recognise any substance. They are primarily used in biomedical research for detection and purification tasks, in the diagnosis of disease, and in the treatment of infections and cancer as delivery mechanisms [414]. The process for producing monoclonal antibodies was first described in 1975 [260], and later refined for human use [336].

Active Immunity

Active immunity involves the stimulation of the system by an antigen and the development of an immune response. Natural active immunity is the normal functioning of the acquired immune system for neutralising pathogens. Artificial active immunity involves

intentionally inducing the immune response by introducing an immunogen in a vaccine. A vaccine is typically administered before the patient contracts the disease, although it may be administered after, assuming the immune response to the immunogen is faster than the response to the infection. There are many different types of vaccines including inactivated (dead) pathogens, live attenuated (low virulence) pathogens, toxoids (inactive toxic compounds from pathogens), and subunit or pieces of pathogens. Vaccination does not prevent the host from contracting the disease, rather like natural active immunity it provides an improved response to the disease making it harder for it to spread throughout the population. Vaccination is one of the greatest achievements in medicine and public health [344]. The ethics of compulsory vaccination are still debated throughout the world, although the benefits of immunity to disease and disease eradication are generally considered to outweigh the costs of a minority of recipients contracting the disease the vaccine was designed to prevent. An aspect of population vaccination is an effect called *herd immunity* [140, 23, 22]. This is where the majority of the population (perhaps more than 90%) is vaccinated for a disease, which provides protective coverage for the entire population. An impact of this interesting effect is that vaccination schemes may not be concerned with immunising the remainder of the un-vaccinated population, rather the immunisation of enough of the population to reduce the chance (to a desired level) of a member of the population contracting the disease.

6.2.2 Evolution

The acquired immune system of mammals (such as humans, mice, rats, and cattle) is the most studied immune system. A consideration of the system in the context of the ‘tree of life’ reveals that this it belongs to a very small minority of taxa: the jawed vertebrates, raising the question as to why this system evolved when all other plants and animals survive without it. Evolution has designed a complex, specialised, and multi-layered defence in the immune system [105], although the system is not perfect: “... *vertebrates are exquisitely adapted [...] but not perfectly; autoimmunity and cancers present the costs and failures of immunity*” [337] (page 33). The very fact that mammals are vulnerable to disease highlights that the immune system is a sub-optimal adaptation [303]. Interestingly, some suggest the acquired immune system may be an evolutionary misstep given all the positive and negative feedback signals (checks and balances) needed to regulate an immune response [204]. This section provides a brief review of the innate and acquired immune systems from an evolutionary perspective.

The Innate and Acquired Immune Systems

The innate immune system is a collection of defence mechanisms, tissues, and specialised processes that defend the host from its pathogenic environment. The innate defence is inherited and defined in the genome, thus the specificity for the pathogens it may detect is ‘hard-wired’ throughout the hosts lifetime. The trade-off for this rigidity is that the defence it provides is fast, practically instantaneous. In the acquired immune system, evolution has designed a somatic learning system that may recognise any biochemical

molecule, although this tool of defence may also destroy the host which it is intended to protect. Acquired immunity is a personalised defence, learned within the context of the hosts' specific pathogenic environment. It is slow acting, requiring an aggregation of antigen encounters generally over about a week before it is ready to fight the pathogen (fast compared to the innate system which specialised over generational time). As a strategy, the system is anticipatory, assuming that future pathogenic encounters will be the same or similar to past pathogenic encounters. Table 6.2 contrasts some of the properties of the acquired and innate immune systems.

Receptor's Features	Immunity Systems	
	Innate	Adaptive
<i>Distribution in cell populations</i>	uniform	clonal
<i>Specificity</i>	low	high
<i>Affinity</i>	$10^3 - 10^4$	$10^6 - 10^{11}$
<i>Diversity (the number of variants)</i>	$10^2 - 10^3$	$10^9 - 10^{11}$
<i>Variability</i>	absent	high
<i>Ligand origination</i>	foreign (non-self)	foreign and self

Table 6.2: Summary of the differences between the innate and acquired immune systems, taken from [257] (page 565).

The acquired immune system is exceedingly complex, such that approximately 5% of the genome of humans is proposed to be for molecules used in the system. This antibody immunity is proposed to have evolved gradually over a long period of time from cells and tissues that previously performed other functions [256]. Klimovich considers some of the problems related to the evolution of the acquired immune system highlighting: (1) the system was formed in the presence not only of pathogen but also coexisting foreign organisms living in symbiosis, (2) many kinds of activity in the system are not protective in nature, instead serving alternate functions, and (3) the system performs the learning function *before* applying the protective mechanism (so called pre-committed) [257]. Hedrick suggests that perhaps long-lived complex jawed vertebrates need an acquired immune system, highlighting that the acquired immune system may be an approximate solution to the problem of parasitism, providing a small advantage in reproductive success that invertebrates could not afford the cost of resources to evolve, maintain, and exploit [204].

6.2.3 Summary

The evolution of immunity is the interaction of immune systems both with their environment, which may be adversarial in the case of parasites, and with each other in terms of reproduction and natural selection. Sexual recombination of the genetics of immunity provides interactions of immune systems in terms of new arrangements of existent traits. The pressure of natural selection results in the implicit competition between heterogeneous populations of immune systems in a spatial-temporal pathogenic environment. Immunisation provides a second example of the interaction between immune systems. Evolution designed the mammalian acquired immune system to be augmented in infancy by the lifetime immunological knowledge of the maternal immune system in a temporary

Lamarckian-style inheritance. This natural passive immunity provides a robust replacement defence for the developing system. Additionally this process may be reproduced artificially later in life by the removal of antibodies from one host and then injected into another to provide a short-lasting immediate defence to a pathogen.

6.3 Abstract Host Paradigm

This section abstracts and elaborates on a host-centric adaptive strategy and proposes a *Host Clonal Selection Paradigm* that takes the Tissue Paradigm and implicitly the Cellular Paradigm for granted, and focuses on the inter-host interaction and information sharing provided by bottom-up host-population and host-generational immunisation and evolutionary mechanisms. Conceptually, a *host* represents a rise in the level of abstraction from a *tissue* (conventional clonal selection algorithm or repertoire) to a collection of such tissues working together to provide holistic immunological function. Therefore a *Host Clonal Selection Algorithm* is concerned with a system comprised of a number of *Tissue Clonal Selection Algorithms* (Chapter 5) that in effect are the components or units of adaptation (hosts) within a host algorithm.

6.3.1 Interaction Models

The host paradigm is primarily concerned with the interaction of whole immune systems where the considerations of what happens within a single host is delegated to the Tissue Paradigm, and in turn the Cellular Paradigm for the concerns within a single tissue repertoire of cells. This section considered the two primary means in which immune systems can interact and more importantly effect each other: *explicitly* and *implicitly*. Explicit interaction between host immune systems is defined as *Effector Sharing* where the information acquired by an immune system (its effectors) are shared with other immune systems either by passive infusion or active response elicitation. Implicit interaction between hosts involves evolutionary processes over generational time that effect the basis and processes by which progeny immune systems acquire information called *Evolved Immunity*.

Effector Sharing

Immunity is acquired and refined over the lifetime of the host given antigen exposures and resultant clonal selection and affinity maturation. The effectors produced in this process (antibodies, memory and recirculating lymphocytes) may be shared with other immune systems. In this context *effector* is a generalised term that refers to some discrete information acquired by an immune system such as the product of a clonal selection and expansion event. This section describes two mechanisms for sharing effectors between systems in what is referred to as passive immunity: *Vertical Sharing* and *Horizontal Sharing*.

Vertical Sharing Vertical sharing is inspired by natural passive immunity such as the sharing of antibodies between mother and foetus across the placenta (maternal immunity),

and to the infant after birth through breastfeeding (mucosal immunity). Vertical sharing of immune effectors is a mechanism of inheriting acquired immunity between generations. A sample of the effectors are removed from a parent system and inserted into a child system. The shared effectors, such as all cellular components have a finite lifetime and may perish before being employed by the receiving system. A natural version of this mechanism is a parent-child relationship, but this is not a requirement. The mechanism does require (1) a generational population structure, (2) the selection of a donor from the previous generation, (3) the selection of a receiver from the next generation, and (4) a sampling method for the donor (effector selection mechanism).

Horizontal Sharing Horizontal sharing of immunity effectors is inspired by artificial passive immunity where effectors are synthesised and injected into hosts, or removed from one host and injected into another. This method of effector sharing provides a way of transferring learned immunity within a population of immune systems. The mechanism does not require a generational population structure, although, like vertical sharing it does require (1) a donor of effectors (within the population or synthesised externally), (2) a receiver system within the population, and (3) a sampling (selection) method for effector cellular components from the donor.

Evolved Immunity

Another way of sharing immunity is through genetic evolution involving (1) the assignment of immune system fitness, (2) selection of parent immune systems, (3) the recombination of parent genetic material to create children, and (4) the mutation of child genetic material. Those immune systems that demonstrate their relative utility (relative to other immune systems with the same generational population) may proliferate, and those that do not demonstrate utility, perish. The evolution of immune systems requires that the expressed phenotype (thing being evolved such as immune effectors or response strategy) be transcribed from a genome (representation) on which the evolutionary processes of selection, recombination, and mutation may operate. It requires a generational mechanism for implicitly sharing effectors (seed and process for creating effectors) rather than explicitly as in the case of vertical and horizontal effector sharing. Two evolved immunity mechanisms are presented as follows: (1) Evolution of an Innate Immunity, and (2) Evolution of an Acquired Immunity.

Innate Immunity Innate immunity may be thought of as hard-wired defence mechanisms that are learned in response to pathogens within a species over generational time. A simple example is a static recognition and response rule system for a host, which although remains fixed for the lifetime of a single host, evolves to protect the whole species. In this example, the rule system is fast acting on the lifetime scale, although slow to change on the same scale implying perhaps the increased expendability of individuals for the greater good of the species.

Acquired Immunity The acquired immune system, although is a somatic learning system rapidly adapting to the hosts pathogenic environment, may also acquire more generalised species level traits via evolution. Evolution may be used to evolve the acquired immune system by biasing the creation of naïve lymphocytes to specific sub-areas of the shape-space. This bias may accelerate the acquisition of immunity by individuals, requiring less work by the somatic learning system.

6.3.2 Antigenic Exposures

Discrete Host Exposures

Section 5.3.2 extended the *Antigenic Exposure Paradigm* of Section 4.2.2 to include the notion of multiple points of exposure on a Tissue Clonal Selection Algorithm (each repertoire). The attributes of discrete exposure were bundled into a concept called the *Tissue Exposure Pattern*. This concept was specialised to incorporate spatial and temporal regularities of exposure location and duration in order to assess the information acquisition and dissemination properties of tissue models called an *Antigenic Infection*. The aggregation of multiple antigenic infections on a tissue was referred to as situating the tissue model (host) in an *Antigenic Habitat*. Like Tissue Clonal Selection Algorithms, Host Algorithms also provide multiple points of exposure in an antigenic environment and therefore have similar concerns. Figure 6.1 provides a conceptualisation of a spatial population of discrete hosts, each situated in their own antigenic habitat.

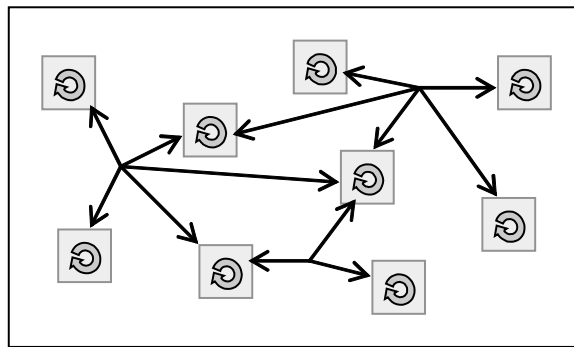


Figure 6.1: Host Clonal Selection Model and the situated spatial properties that affect antigenic exposure, where the arrows represent interaction between the discrete components (hosts) in the system.

There are two distinct and important differences between the Tissue and Host paradigms with regard to antigenic exposures: (1) exposures of antigen to hosts may be both controlled (elicited by other hosts) and uncontrolled (environmental) whereas tissue models have no control over such exposures, and (2) the level of abstraction regarding exposures is raised from that of an habitat of infections for a single host, to multiple and potentially different (asymmetric) habitats across multiple hosts called an *Antigenic Environment*.

Controlled and Uncontrolled Exposures

The acquired immune system learns through exposure, detection, and response to antigens, thus the antigenic environment defines what is learned and when. The immunity that a system receives through recognising and responding to antigens is called active immunity and may be induced naturally (uncontrolled) and artificially (controlled).

Uncontrolled Exposure Uncontrolled antigen exposure is a natural way for a system to achieve active response-based immunity. This form of antigen exposure is an elaboration of the previously defined standard model. An aspect of the distinctiveness of acquired immune systems in a population is the spatial and temporal variations of antigen exposures. These varied exposure patterns in conjunction with the varied naïve repertoire of each individual result in distinct initial and learned specificities to antigens. For a given antigen, exposure may be described as a function over space and time that may be skewed for many reasons such as antigen virulence, host density, and other pathogen-behaviour and spatial-population effects.

Controlled Exposure Control over antigen exposure provides control over immunity of individual systems and virtual immunity over the entire population through effects like ‘herd immunity’. Artificial exposure of systems to antigen provides control of (1) which systems are exposed to (2) what antigen (3) when. A vaccination exposure strategy may involve administering a vaccine to a large number of systems quickly. An inoculation strategy may involve exposing a select few systems to the antigen in a short amount of time. Control of exposures may also extend to preventative steps such as environment sterilisation or individual system isolation from the antigenic environment.

Antigenic Habitats and Environments

The Tissue Exposure Regime of the Tissue Paradigm may be extended to a *Host Exposure Regime (HER)* that defines which hosts are exposed to antigen and when they are exposed. This means that the system is less concerned with how well a given host can specialise a response and internally disseminate acquired information, and more concerned with how that information is acquired and disseminated between hosts in the population. Unlike tissues, hosts are not restricted to fixed network topologies like the ring network structure of the Lymphocyte Recirculation Algorithm, therefore asymmetries in the Host Exposure Regime and the host interaction patterns define the capability and speed of information dissemination throughout the population. Given the shift in focus from the internal process of information acquisition within a host to the internal information acquisition of a population of hosts, the antigenic environment may be conceptualised as a *Pathogenic Environment* consisting primarily of exogenous antigen. This does not mean hosts may not respond and communicate information regarding endogenous antigen, rather it suggests that the focus of the Host Paradigm is hosts and their relationships with external antigen that may come the environment and other hosts. In addition to the Host Exposure Regime, the rise in the level of abstraction in the Host Paradigm facilitates broader

concepts regarding antigen as self-contained entities, in particular (1) the co-evolution of pathogen with a population of hosts, and (2) the contraction of a pathogen by a host and transmission to neighbouring hosts.

Evolvable Pathogen A pathogen may explicitly evolve or co-evolve with the population of immune systems. Evolution may refine the strategy of the pathogen affecting configuration such as virulence and transmissibility. An individual acquired immune system may be open to multiple infections by the same pathogen during its lifetime if the rate of evolution is sufficiently rapid. Such a pathogen may require tracking and ultimately evolving countermeasures by the host population, perhaps facilitating a host-parasite interaction.

Transmittable Pathogen After exposure to a host, a pathogen may jump from host-to-host using a number of mechanisms (vectors). Further, transmissibility may be influenced by any number of epidemic-based host models such as density-dependence, spatial models, and genetic-similarity models.

6.3.3 Host Architectures

The abstraction of immune system interaction models of immunisation and evolution suggests two distinct Host architectures as follows: (1) *Populations of Hosts* that focus on the natural and artificial elicitation of immune response and transfer of acquired immunity, and (2) *Generations of Populations of Hosts* that focus on hosts with a finite life-cycle that include inter-generational elicitation and transfer of immunity and the evolutionary change of the properties of lifetime innate and acquired immunity. This section outlines these two distinct host architectures.

Population Host Model

An acquired immune system does not exist in isolation, rather it is a collection of organs and tissues that collectively represent an integrated part of a host organism, which in turn belongs to a population of such organisms. One may consider the hosts of a population interacting with each other, more specifically interacting with regard to their immune systems. The class of model in which a population of clonal selection-based immune systems interact with each other is referred to as a *Population Host Model*. The focus of this model is the host as the unit of adaptation, each with a distinct starting point, and likely distinct antigenic environment including interactions with the environment and other hosts. This diversity provides robustness in the population with regard to information acquisition and dissemination, abstracting concerns of population survivability in an unknown environment. A population of non-interacting hosts facilitate the development of multiple varied perspectives of an antigenic environment that remain isolated. Integration of this information without host interaction mechanisms would require augmentation from a higher level of abstraction (the population level). The *isolated perspectives* property of the population model is made more apparent in an asymmetrical pathogenic environment. In such an environment, not only does the heterogeneous population acquire information in different

ways and rates, but the information content of the environment to which they are exposed also varies. Therefore, although the population may collectively represent a map of the environment, there is no way for individual systems to share acquired information. This concern is further highlighted if we consider that the asymmetrical pathogenic exposure regimes for the systems in the population vary over time. In this case, the information acquired by one system is directly useful to other systems, although cannot be exploited. A population of non-interacting hosts provides a foundation on which to incorporate intra-population (inter-host) interactions. Intra-population re-transmission of captured antigen, and intra-population effector sharing provide two examples that motivate the investigation of decentralised intra-population interactions.

Generational Model

Host organisms in a population reproduce such that a generational structure may be defined between hosts where progenitor-hosts produce progeny-hosts. In this mechanism, the population of immune systems (the present generation) are responsible for creating a new set of immune systems (the next generation), which replace members of the present generation. This principle defines the *Generational Host Model* that provides an extension to the proposed Host Population Model. Hosts may use an asexual reproduction mechanism where one host contributes a variation of itself to the next generation. Alternatively, hosts may collaborate in their contributions to successive generations (sexual reproduction). Further, the contributions of various hosts to successive generations may not be homogeneous, such that more successful hosts (by some measure of success) may contribute more than less successful hosts. The principle components of the generational approach are (1) the reproductive scheme, and (2) the mechanism that triggers the generational change.

- *Reproductive Scheme*: The mechanism by which one generation contributes to and prepares the successive generation. Examples include sexual and asexual reproduction where hosts may or may not cooperate in their contributions to the next generation. Further, the contributions of any one host to the next generation may or may not be homogeneous with respect to the other hosts of the present generation.
- *Generational Trigger*: The condition, that once satisfied requires the application of the reproductive scheme to the present generation to create the next generation of immune systems. An example is a time trigger in which immune systems are given a representative exposure to the pathogenic environment.

To be clear, the host generational approach is the same in principle to classical *generational* clonal selection algorithms in that one generation represents one iteration of the systems execution. The important difference is that one generation of the host generation system involves multiple generations of the classical clonal selection algorithm (cellular) within each host (hosts' tissues).

The premise of the generational mechanism is that a given instance of a host immune system may acquire different information in the same antigenic environment over successive

trials (generations). A generational model that does not allow one generation of hosts to convey information to progeny, rather it provides a clean slate for each system to repeat a trial. Ultimately such a simplified generational model ‘resets’ each host to initial conditions at the beginning of each new generation. The restart principle gives a population of immune systems, each with their unique and ongoing perspective of the environment, successive chances of learning a perspective of the environment. The inability of the systems to share acquired information between generations, means that each generation must re-learn (start from scratch) those things learned in the previous generation. If a generational-accumulative model is desired, it must be implemented as a mechanism at a higher scale than the generational model. Such a mechanism is responsible for explicitly integrating information across generations. The hope of the generational model is that each fresh-start and subsequent run (trial with the environment) results in a varied perspective and thus varied acquired information perhaps at the individual system-level, and more importantly at the population-level. The concern of a non-interacting generational model is that operating with a fresh-start is likely too computationally intensive (inefficient) with regard to the information acquired and explicit integrated. More importantly, there is little guarantee that the ‘varied perspective’ principle will hold for a given environment. The model may be extended by proposing two different vertical sharing mechanisms by which populations can share information between generations of the a algorithm: *explicitly* by the inter-generational sharing of effectors, and *implicitly* by the evolution of characteristics that influence the response to antigenic exposures. The intent of sharing information between generations is that subsequent generations can learn from the success and failures of the previous generation, and the generations that came before it. Such mechanisms allow the generational-population (species) to implicitly integrate information about the environment over generational time (rather than explicitly by an external mechanism). Further, such mechanisms do not restrict subsequent generations to the limitations of the past, rather integrating generational knowledge into the ‘fresh-start’ systems facilitating the desired trait of the non-interacting generational model of acquiring a varied perspective (if such a variation is possible).

6.4 Realised Host Paradigm

A population of hosts each with a distinct immune system, exist within an antigenic environment where interactions are mediated by an unknown exposure regime. This provides an effective metaphor for considering the problem of the acquisition and application of information using a decentralised strategy with distributed information processors. As was the case with the Tissue Paradigm, information management strategies are concerned with the effective localisation and dissemination of information, taking the internal acquisition of such information for granted. From the perspective of a population of hosts, the specifics of information acquisition within each host are taken for granted, relegated to the investigatory concerns of the Tissue Paradigm. Host immunisation methods provide a rich metaphor for considering host-host interactions in a population toward eliciting immune response and sharing the product of the immune response in both a static population

and generational population structure. These interactions, in addition to implicit host interactions in an evolutionary process, motivate the realisation of distributed information management strategies in the host clonal selection paradigm. This section provides a realisation of the concerns of the abstract Host Clonal Selection Paradigm presented in Section 6.3. This includes a general problem definition that captures the concerns for investigating systems at the host-level of abstraction, as well as its specialisation in the colour space domain. A general Host Clonal Selection Algorithm is proposed providing a basis for population-based host architecture and as well as its extension, the generational host architecture. Empirical measures are proposed for the investigation of developed strategies, and general behaviour expectations are proposed motivating confirmatory research.

6.4.1 Antigenic Habitats

This section considers the realisation of a specialisation of the Antigenic Exposure Paradigm outlined in Section 6.3.2. In particular, this section abstracts the Infection Exposure Problem considered for the Tissue Paradigm, to a Habitat for the Host Paradigm, and a specialisation of the problem to Colour Space. In addition, the discrete properties of tissues exposures are considered in the context of hosts in a population and specialised Host Exposure Regimes are defined.

Habitat Antigenic Exposure Problem

The *Habitat Antigenic Exposure Problem (HAEP)* defined in Algorithm 6.1 provides a general problem definition where a Population P is exposed to an Habitat B . A given Tissue Clonal Selection Algorithm (TCSA) may be considered as a single H in a Population of Hosts ($P = \{H_1, H_2, H_3, \dots H_n\}$). Likewise the Infection Antigenic Exposure Problem (IAEP) defined in Algorithm 5.1 may be considered a single H in an Environment of Habitats ($E = \{B_1, B_2, B_3, \dots B_n\}$). The parameters of HAEP are the same as IAEP with the addition of $N_{habitats}$ defining the number of B in E . The *CreateEnvironment* operation creates a new environment that may be exposed to a P via the *Exposure* operation that returns a subset of the population P_{rs} representing the populations ability address the specific exposure requirements of the environment under the given Host Exposure Regime.

Algorithm 6.1: Habitat Antigenic Exposure Problem (HAEP).

Input: $P, N_{determinants}, N_{antigen}, N_{infections}, N_{habitats}$

Output: P_{rs}

```

1 E ← CreateEnvironment( $N_{determinants}, N_{antigen}, N_{infections}, N_{habitats}$ );
2  $P_{rs} \leftarrow 0$ ;
3 while ¬StopCondition() do
4   |  $P_{rs} \leftarrow$  Exposure( $P, E$ );
5 end
6 return  $P_{rs}$ ;
```

Habitat Colour Space Problem

The HAEP may be specialised as a habitat extension of the Infection Colour Space Problem (ICSP) defined in Section 5.4.1 called the Habitat Colour Space Problem (HCSP). An instance of the HCSP involves the exposure of a H to a set of Colour Space Pattern Sets (CSPS) where the goal of the problem is to minimise the average error in the response P_{rs} . As was the case with IAEP, one may define a minimum configuration of HAEP and thus HCSP that focuses the concerns on the highest level of abstraction in the problem (an E of B). The focus is provided by marginalising the configuration at the lower levels. A Minimal Habitat Antigenic Exposure Problem (MHAEP) is defined where $N_{habitats}$ remains variable, although $N_{determinants}$, $N_{antigen}$, and $N_{infections}$ are fixed at 1. Therefore, in the case of HCSP, the number of B defines the number of Sets of CSP (I), each of which is comprised of a single CSP (A) where all three components are treated a single determinant (D).

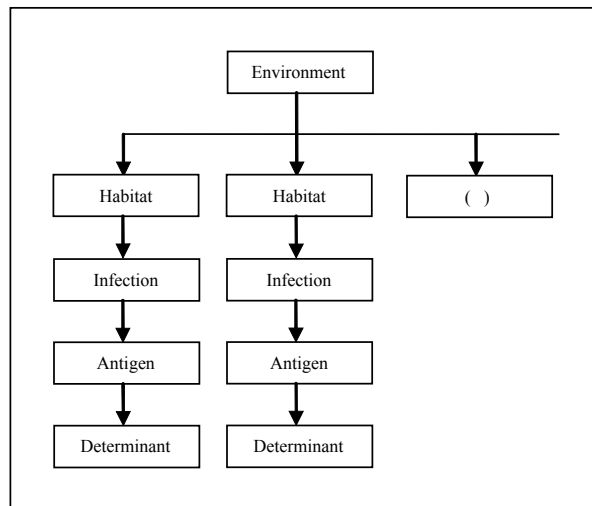


Figure 6.2: Elaboration on figure 5.18b to depict the Minimal Habitat Antigenic Exposure Problem (MAHEP)

Host Exposure Regimes

The HAEP may be specialised with a series of Host Exposure Regimes (HER) in the same manner as the IAEP was specialised with a series of Tissue Exposure Regimes (TER) in Section 5.4.1. This reuse is facilitated because both the Tissue and Host paradigms are concerned with strategies for organisation information in the context of unknown scopes and regularities of information exposure. As such, the series of five TER's may be modified for use with the Host Paradigm such that the algorithms remain the same and the parameters are changed from H and B to P and E . The following Host specific exposure regimes are defined as AHER, OHER, PHER, RHER, and SHER.

6.4.2 Host Clonal Selection

Population Host Clonal Selection Algorithm The Tissue Clonal Selection Algorithm (TCSA) defined in Algorithm 5.7 describes the general interaction between a Habitat B and a Host H . The host paradigm requires a rise in the level of abstraction from a single H to a Population P of H ($P = \{H_1, H_2, H_3, \dots, H_n\}$) that responds to an Environment E of Habitats B . The *Host Clonal Selection Algorithm (HCSA)* defined in Algorithm 6.2 describes a general host algorithm that may be specialised based on the concerns and constraints of the Tissue Paradigm, specifically the *HostInteractions* operation that provides facility for specialising the interactions between hosts. The focus of the HCSA is host interactions within a population, therefore to differentiate the approach from the generational extension (described next), the HCSA may be referred to as a *Population Host Clonal Selection Algorithm (PCSA)* to clarify its relationship with the generational extension. In the tradition of the CCSA and TCSA, one may define a minimal implementation of *HostInteractions* in which there are no interactions between hosts in the population. This *Minimal Population Host Clonal Selection Algorithm (MP-HCSA)* provides a baseline for comparison with extensions of the PHCSA that facilitates intra-population interactions. In addition to no host-interactions, the MP-HCSA codifies the focus of the Host Paradigm on host interactions, taking intra-host concerns considered in the Tissue Paradigm for granted. Specifically, each host in the MP-HCSA has a single tissue ($H = \{T_1\}$), allowing each host to be treated as a generic information processing component in the system.

Algorithm 6.2: Population Host Clonal Selection.

Input: E, N_{hosts}
Output: P

```

1  $P \leftarrow 0$ ;
2 for  $i \leftarrow 0$  to  $N_{hosts}$  do
3    $H_i \leftarrow \text{CreateHost}()$ ;
4    $P \leftarrow H_i$ ;
5 end
6 while  $\neg \text{StopCondition}()$  do
7    $E.\text{Exposure}(P)$ ;
8    $\text{HostInteractions}(P)$ ;
9 end
10 return  $P$ ;

```

Generational Host Clonal Selection Algorithm In addition to a population-based host organisation Section 6.3.3 proposed a generational-based host organisation as an extension to the population-based approach in which a new population of hosts is periodically created from an existing population. Algorithm 6.3 defines the *Generational Host Clonal Selection Algorithm (GHCSA)* as an extension of the PHCSA defined in Algorithm 6.2. The *GenerationalChange* operation defines the frequency of restarts that may for example be based on a consistent number of epochs (population exposures). The

CreatePopulation operation defines the mechanism for creating a new P from an existing P that has interacted with the antigenic environment E . A minimal implementation of the *CreatePopulation* can be defined in which the creation of the new population of hosts is independent of the existing population (no interaction). This *Minimal Generational Host Clonal Selection Algorithm (MG-HCSA)* provides a baseline for comparison with extensions of the GHCSA that promotes intra-generational interactions.

Algorithm 6.3: Generational Host Clonal Selection.

```

Input:  $E, N_{hosts}$ 
Output:  $P$ 
1  $P \leftarrow 0$ ;
2 for  $i \leftarrow 0$  to  $N_{hosts}$  do
3    $H_i \leftarrow \text{CreateHost}()$ ;
4    $P \leftarrow H_i$ ;
5 end
6 while  $\neg \text{StopCondition}()$  do
7   while  $\neg \text{GenerationalChange}()$  do
8      $E.\text{Exposure}(P)$ ;
9      $\text{HostInteractions}(P)$ ;
10  end
11  if  $\neg \text{StopCondition}()$  then
12     $P \leftarrow \text{CreatePopulation}(P)$ ;
13  end
14 end
15 return  $P$ ;

```

6.4.3 Empirical Assessment

This section defines a series of empirical measures derived from the measures used in the Tissue Paradigm (Section 5.4.3) that provide instantaneous information regarding a given HCSA on Colour Space specialisations of the Habitat Antigenic Exposure Problem. The measures are classified as *system* that provide general holistic information about the algorithm's information state, and *component* that provide generalised information about the algorithm's component information state. As with the Cellular and Tissues Paradigms, an important consideration with the proposed measures in exploratory experimentation is not their absolute value, but rather their relative change in value with changes to the systems being investigated.

System Measures

Population Error The Host Error (HE) defined in Equation 5.1 provides a generalised measure of the error for a single Host Result Set of tissues H_{rs} against a given Habitat of antigenic infection B . This measure may be calculate a H_{rs} for each H in a P to formulate an P_{rs} that may be averaged to define a Population error (PE) (see Equation 6.1). PE provides an indication of instantaneous average error for each H_{rs} for a given P_{rs} provided

by a HCSA in response to an Host Exposure Regime. PE is in the units of the E problem space, for example Euclidean distance for a colour space specialisation.

$$PopulationError(E, P_{rs}) = \frac{1}{n} \sum_{i=1}^n HostError(B_i, H_{rsi}) \quad (6.1)$$

Population Diversity The Host Diversity (HD) measure defined in Equation 5.3 provides an instantaneous diversity a TCSA. The same approach used in HD for calculating the diversity between repertoires of cells may be used at the population level if the tissues of each host are compressed to a single repertoire. The measure assumes such a reduction permitting the direct calculating of a Bit Frequency Histogram (defined in Algorithm 5.8) for each H , permitting the comparison of Hosts using Bit Frequency Difference (defined in Equation 5.2), and ultimately the calculation of diversity according to Equation 5.3. For consistency in the measures used in the host paradigm, this measure is re-branded *Population Diversity (PD)*.

Component Measures

Average Host Diversity The Average Tissue Diversity (ATD) defined in Equation 5.4 provides an instantaneous measure of the diversity of the average T of a host measured in bits. This measure may be averaged across all hosts in the population P to provide a measure of the average host diversity as the average of the diversity of tissues (defined in Equation 6.2). Alternatively, Average Host Diversity may be calculated as the Host Diversity (defined in Equation 5.3) averaged for all hosts in the population. This measure will provide more fidelity for population configuration in which hosts have more than a single tissue.

$$AverageHostDiversity(P) = \frac{1}{n} \sum_{i=1}^n AverageTissueDiversity(H_i) \quad (6.2)$$

Average Host Error In the same manner as the Average Tissue Error (ATE) in Equation 5.5, an average error may be defined that exploits the system-level error calculation for each component of the system. Equation 6.3 defines the *Average Host Error (AHE)* that calculates the Population Error (defined in Equation 6.1) for E against each host H in P , that is averaged by the number of hosts in the population.

$$AverageHostError(E, P) = \frac{1}{H_n} \sum_{i=1}^{H_n} \left(\frac{1}{E_n} \sum_{j=1}^{E_n} HostError(B_j, H_{rsi}) \right) \quad (6.3)$$

6.4.4 Trends and Behaviours

This section considers the expected emergent effects and behavioural trends of host-based systems in population and generational architectures as defined in the specific realisations. Although specific to the Host Paradigm, the general behavioural trends are expected to be qualitatively similar to those expected and observed in the Tissue Paradigm. The

primarily reason for this is the reuse of varied exposure regimes as a decoupling mechanism between systems and environment for the spatial-temporal regularity or lack there of in the interaction between the two.

Algorithm Mechanism and Strategy

Section 6.3.3 proposed a population and a generational host architectures providing two different foundational organisations of hosts. This section considers the emergent effects of these two organisations contrasting their mechanism and information management strategies. Table 6.3 provides summary of the two models in these terms.

Model	Mechanism	Strategy
<i>Population</i>	Population of whole systems.	Independent perspectives of problem.
<i>Generation</i>	Systems with life-cycles.	Independent perspectives with multiple trials.

Table 6.3: Summary of the difference in mechanism and strategy between the complementary concerns of host algorithms under discrete exposures.

From a population perspective, each host provides a different independent perspective of the antigenic environment. Inter-host interaction promotes sharing and improvement of the independent toward improving the competence of individuals and ultimately the entire population. This provides a motivation for information management strategies for the population model, highlighting the concerns of information dissemination and localisation in a population (similar to that in the Tissue Paradigm), toward population capability robustness. From a generational perspective, the population turnover provides an opportunity for individual perspectives to be refreshed, and for individual host configurations (starting conditions) to be re-trialled in the same or similar antigenic environment. From the abstract perspective, the generational model builds upon the population model, taking the concerns of inter-population interaction for granted, focusing on the mechanism and effects of inter-generational host interactions. The population in each generation is independent from the generation before and after it, thus the motivating concerns of information dissemination and sharing from the population model are directly applicable for motivating intra-generational information management strategies. Therefore the general motivations and expectations of the minimal and extended strategies for both host architectures may be summarised as follows:

1. *Population Strategies*

- (a) Minimal population promotes multiple and potentially variable perspectives on a problem.
- (b) Intra-population interaction promotes dissemination between isolated perspectives.

2. *Generation Strategies*

- (a) Minimal generational promotes multiple trials from multiple and potentially variable perspectives on a problem.

- (b) Intra-generational interaction promotes dissemination about acquired information between trials.

Host Exposure Trends

In the tissue paradigm the urgency of acquisition and effective application were promoted by the metaphor of injury by unattended infection, such urgency is not promoted in the host paradigm. Instead, hosts unlike tissues are more expendable as long as some of the hosts population acquire and employ an effective immunity, the ‘system’ (population) will survive. This *disposability of components* is fostered more so in the generational architecture as the mechanism for population assessment and turnover (new component creation) are explicitly defined. An important consideration in the Host Paradigm is the independence of perspective, in particular the capability of different hosts to respond in varied ways to the same stimulus. This may be investigated by varying host configuration and starting conditions within a population and exposing them to a symmetric antigenic environment to assess emergent effects. Alternatively, the host exposure regimes provide an opportunity for investigating the same effects using the opposite case of the same host configuration exposed to regimes varying in regularity and scope of population exposure. The integration of these two complementary cases represents the pinnacle host-model complexity, where the variation of perspective promoted by the variability of exposure regime provides an idealised and controlled model with strong similarity to model used in the tissue paradigm. This reuse of exposure regimes in Tissue and Host Paradigms promote similar expectations as to their general effects in terms of *localisation* and *dissemination* across the components of the system although under different constraints. Specifically, a population of host systems interact in different ways from that of the fixed intra-tissue interaction.

Table 6.4 classifies the five Host Exposure Regimes in terms of their general attributes, including information *distribution*, *nature*, and *consistency*. One may consider the *distribution* of the exposure of distinct information to the components of the system with the different host exposure regimes. Symmetrical exposure suggests that each component has the opportunity of being exposed to any piece of information in the environment, whereas asymmetrical constraints this expectation. The opportunity may be deterministic, such as a certainty of being exposed to all or some of the information environment or probabilistic in those regimes with a stochastic element. *Scope* provides a general notion as to how much of the system has the opportunity for interacting with the antigenic environment irrespective of the specific information it may be exposed to. The *nature* of the regime defines the basis of the exposure mechanism which directly relates to the consistency of the component-information match ups each exposure event.

Localisation and Dissemination Trends

Information dissemination promotes *generalised* capability whereas information localisation promotes *specialised* capability. In the tissue paradigm these two contrasting management types were described with regard to their emergent effects in the context of a tissue

Regime	Distribution	Scope	Nature	Consistency
<i>AHER</i>	Symmetrical	System-Wide	Deterministic	Regular
<i>OHER</i>	Symmetrical	System-Wide	Probabilistic	Semi-Regular
<i>PHER</i>	Symmetrical	Constrained	Deterministic	Regular
<i>RHER</i>	Asymmetrical	System-Wide	Probabilistic	Irregular
<i>SHER</i>	Asymmetrical	System-Wide	Deterministic	Regular

Table 6.4: Classification of Host Exposure Regimes in terms of general attributes.

model, specifically: the ‘consistency of response’ and ‘spatial organisation of information’ respectively. These two effects may be rephrased in the context of the host models as the following:

1. *Dissemination*: Consistency in host perspective of the environment across the population and generations.
2. *Localisation*: Specialisation in host perspective of the environment across the population and generations.

The general observational trends for both management types are expected to hold for hosts as they did for investigations into Tissue Paradigm. Specifically, an increase in the population generalisation results in Average Host Error that approaches the Population Error with decrease in system diversity as components are more similar, contrasted with a relative decreased in system error with population specialisation resulting in an increase in system diversity as component become more heterogeneous.

6.4.5 Paradigm Agenda

The research agenda for this chapter and perhaps the Host Clonal Selection Paradigm is to investigate the population generational models in the context of the expected emergent effects and behavioural trends. Specifically the relative qualitative comparison of the localisation and dissemination of information of different decentralised information management strategies under a variety of information exposure regimes. This may be summarised as follows:

1. Investigate strategies to promote intra-population (population model) localisation and dissemination of acquired information.
2. Investigate strategies to promote intra-generational (generational model) localisation and dissemination of acquired information.

6.5 Population Elicited Immunity

6.5.1 Antigen Transmission

Eliciting an immune response is the way in which the immune system acquires information about the antigenic environment. As reviewed in Section 6.2.1 this active form of immunisation may occur naturally through normal interactions with pathogen and artificially

by exposing a system to harmless molecules with similar features to the antigen of interest that elicits the same response. The artificial elicitation of immunity is the principle behind vaccination's against disease. In both the artificial and natural cases, eliciting an immune response provides control over what information a host should acquire and when it should be acquired, although this control may or may not reside with the host. This control may be abstracted to the general *sampling* of antigen from the environment by a host and future *transmission* to other hosts in the population. This sampling and transmission provides a mechanism by which a given immune system may share information about the antigenic environment with other immune systems in the population. Individual immune systems must collect and store samples of antigen to which they are exposed for transmission to other systems in the population. This mechanism is referred to as the *Antigen Sampling Scheme*. The sampled antigen must then be passed on to other immune systems with which a given immune system interacts. This passing on is referred to as the *Antigen Transmission Scheme*. The antigen sampling mechanism is dependent on the antigenic environment, more explicitly the host antigenic exposure regime the system is subjected to by the environment and other hosts (Antigenic Habitat). The antigen transmission mechanism is dependent on the intra-population (inter-host) interactions facilitated by the population structure, host environment, and any interaction intentions of the individual systems themselves.

The specialisation of the Population Host Clonal Selection Algorithm (defined in Algorithm 6.2) with an antigen sampling and antigen transmission scheme is called the *Transmission Host Clonal Selection Algorithm (THCSA)*. Algorithm 6.4 defines a generic antigen sampling scheme where all habitats exposed to a population are recorded by the hosts for future transmission. The principle of the algorithm is for hosts in a population to inform other hosts in the population in either a controlled or uncontrolled manner regarding antigen of interest, and to allow those hosts to elicit their own internal response. The algorithm is reasonably general, such that it may be specialised in a number of different ways inspired by population-immune system interactions. Two specialisation of the THCSA are defined: (1) uncontrolled transmission scheme inspired by pathogen dynamics, and (2) a host controlled transmission scheme inspired by vaccination dynamics.

Algorithm 6.4: Stored Exposures for Transmission Host Clonal Selection.

```

Input: E, P
1 foreach  $B_i \in E$  do
2    $P' \leftarrow \text{SelectHostsToExpose}(P)$ ;
3   foreach  $H_{i'} \in P'$  do
4      $\text{Exposure}(H_{i'}, B_i)$ ;
5      $H_{i'}.exposures \leftarrow B_i$ ;
6   end
7 end

```

Pathogen Transmission Dynamics

A first simple specialisation of the THCSA is to that of an infectious pathogen that may be transmitted from host-to-host. Transmission is achieved via host contact, the number of hosts, and the time of infectiousness (storage time for transmission), all of which are a function of the pathogen. The approach facilitates pathogen behaviour not considered at either the cellular or tissue scales. So-called ‘host-mediated’ pathogenic exposure regimes facilitate emergent information processing that reformulate a given Host Exposure Regimes such as *epidemics* (rapidly spreading), and *pandemics* (complete or close to complete population exposure). The specialisation of the THCSA is defined as follows:

- *Sampling Scheme*: Samples are collected when a host is exposed to pathogen from the environment or another host. The longevity of the sample in storage is a property of the infectiousness of the pathogen. A host may be a carrier for more than one pathogen at a given time.
- *Transmission Scheme*: A host may transmit a carried pathogen to any systems to which it comes into contact with (for as long as the host is a carrier). Interactions with other immune systems may be random or mediated by proximity, such as in a spatial organisation.

The pathogen dynamics inspired approach to the transmission algorithm is called the *Pathogen Transmission Host Clonal Selection Algorithm (PT-HCSA)*. The concerns of pathogen infectiousness and host-carriers are simplified to the collection of all exposed antigenic information (as per Algorithm 6.4), the random selection and transmission of sampled antigen, and the clearing of sampled antigen for each population exposure (algorithm epoch). Two variations of the PT-HCSA are proposed: (1) in which transmissions occur between random pairings of infected hosts, and (2) where transmissions occur between random neighbouring hosts in a one-dimensional fixed-spatial environment. Algorithm 6.5 defines a specialisation of the *HostInteractions* for random pairing based transmission.

Algorithm 6.5: Random Pairings for Pathogen Transmission Host Clonal Selection.

```
Input: P
1 foreach  $H_i \in P$  do
2   if  $H_i.exposures \neq 0$  then
3      $H' \leftarrow 0$ ;
4     while  $H' \equiv H_i$  do
5        $H' \leftarrow \text{SelectRandomHost}(P)$ ;
6     end
7      $B' \leftarrow \text{SelectRandomExposure}(H_i.exposures)$ ;
8      $\text{Exposure}(H', B')$ ;
9   end
10 end
```

Algorithm 6.6 defines a specialisation of the *HostInteractions* for the random spatial pairings approach to pathogen transmission where hosts are arranged into a one-dimensional ring structure that limits the random pairing of hosts. Those hosts with sampled exposures transmit a single exposure to a randomly selected spatial neighbour.

Algorithm 6.6: Spatial Pairings for Pathogen Transmission Host Clonal Selection.

```

Input: P
1 foreach  $H_i \in P$  do
2   if  $H_i.exposures \neq 0$  then
3      $H' \leftarrow 0$ ;
4     if  $\text{RandomDouble}() < 0.5$  then
5       if  $H_i \equiv H_n$  then
6          $H' \leftarrow H_1$ ;
7       else
8          $H' \leftarrow H_{i+1}$ ;
9       end
10      else
11        if  $H_i \equiv H_0$  then
12           $H' \leftarrow H_n$ ;
13        else
14           $H' \leftarrow H_{i-1}$ ;
15        end
16      end
17       $B' \leftarrow \text{SelectRandomExposure}(H_i.exposures)$ ;
18       $\text{Exposure}(H', B')$ ;
19    end
20 end

```

This specialisation of THCSA highlights an important property of the transmission principle. Specifically that the antigenic environment (exposure regime to which a population is exposed) may be augmented by intra-population interactions between hosts. Further, the causal influences to inter-population interactions may also be considered indirect influences of host-transmitted pathogenic exposures, such as a hosts proximity to other hosts, host density, or interaction with other hosts from ‘high pathogenic exposure’ areas of the environment. A feature of host-based exposures is that a given host may be a carrier for one or more pathogens to which an interacting system may be exposed. In effect, a given host may represent a micro-environment of antigenic exposures, each potentially providing an exposure regime in and of themselves.

Vaccination Transmission Dynamics

The pathogenic dynamics configuration provides a specialisation of the transmission algorithm, in which the sampling scheme is defined by the pathogen, and is thus outside the control of the host system. In this configuration, the host is provided with control over both the sampling and transmission mechanism. A host may select both the antigen to collect and store, and a small sample of host systems from the population to which to transmit the antigen (exposure). This specialisation of the algorithm is inspired by vacci-

nation and inoculation principle from immunology in which a small sample of pathogen is given to a healthy system to illicit immunity to the pathogen. The scheme is simpler than that of the pathogen dynamics scheme in that (1) the host has control over selecting which pathogen to transmit, and (2) the host has control over selecting a ‘healthy’ system to which to transmit a sampled pathogen. The operations for the vaccination specialisation of the THCSA are defined as follows:

- *Sampling Scheme*: Samples are collected when a host is exposed to pathogen, although the host has discriminatory control over which pathogen are collected and stored and which are not. Further, the host may decide to discard stored pathogen in the acquisition of additional information from the environment (such as exposure frequency).
- *Transmission Scheme*: A host interacts with systems according to the properties of the environment and the population structure, although the host has control over selection of which encountered system to transmit a sample pathogen to. An example is that a host may select to infrequently inoculate those systems that have a lower probability of being exposed to a given pathogen.

As with PT-HCSA, a simplified exposure sampling scheme is used where a randomly selected sampled exposure is transmitted, where sampling is refreshed each epoch. This *Vaccination Transmission Host Clonal Selection Algorithm (VT-HCSA)* (defined in Algorithm 6.7) has a parameter $N_{vaccinate}$ that defines the number of hosts in the population a single randomly selected host may vaccinate with a single sampled exposure.

Algorithm 6.7: HostInteractions for Vaccination Host Clonal Selection.

```

Input:  $P, N_{vaccinate}$ 
1 foreach  $H_i \in P$  do
2    $P_I \leftarrow \text{SelectAllExposedHosts}(P)$ ;
3    $H_I \leftarrow \text{SelectRandomHosts}(1, P_I)$ ;
4    $B_I \leftarrow \text{SelectRandomExposure}(H_I.exposures)$ ;
5    $P_{subset} \leftarrow \text{SelectRandomHosts}(N_{vaccinate}, P)$ ;
6   foreach  $H_i \in P_{subset}$  do
7      $\text{Exposure}(H_i, B_I)$ ;
8   end
9 end

```

The host selection mechanism provides a powerful tool for the population of immune systems to artificially manipulate the pathogenic exposure regime. This manipulation will likely have the intention of improving ‘coverage’ by the population for the inoculated pathogen. A host may decide to push the limits of the transmission scheme (mass vaccinations) either under special circumstances (adaptive transmission), or all the time. Vaccination of a large sample of the population may provide benefits to the entire population given a ‘herd immunity’ effect.

6.5.2 Transmission Empirical Study

Aim

The aim of this empirical study was to investigate the Transmission Host Clonal Selection Algorithm, specifically the pathogen and vaccination configurations in the context of their information dissemination and localisation capabilities as compared to the Minimal Population Host Clonal Selection Algorithm under a variety of host exposure regimes. Toward this end, the study had the following goals:

1. Investigate the relative effects of the two defined host interaction methods for the uncontrolled pathogen transmission scheme.
2. Investigate the relative effects of the controlled elicitation of immunity of the vaccination transmission scheme under different vaccination sample sizes.

Method

Algorithms The study considered the Minimal Population Host Clonal Selection Algorithm (MP-HCSA) and the Transmission Host Clonal Selection Algorithm (THCSA), with two specialised configurations: the Pathogen Transmission Host Clonal Selection Algorithm (PT-HCSA) and the Vaccination Transmission Host Clonal Selection Algorithm (VT-HCSA). MP-HCSA is a specialisation of the HCSA defined in Algorithm 6.2, that was configured with $N_{hosts} = 10$. Each H was configured with a single tissue ($N_{tissues} = 1$) that was an instance of the Replacement Cellular Clonal Selection Algorithm (RCCSA) defined in Algorithm 4.5, with the configuration $N_{cells} = 50$, $N_{selection} = 1$, and $N_{clones} = 5$. The method for sampling all exposures in THCSA defined in Algorithm 6.4 was used for both specialised transmission regimes. Two different pathogen transmission schemes were investigated, the random pairings scheme defined in Algorithm 6.5 referred to as (PT-HCSA-RP), and the random spatial pairings defined in Algorithm 6.6 referred to as (PT-HCSA-SP). The vaccination regime defined in Algorithm 6.7 was used with a small vaccinated set (VT-HCSA-S) $N_{vaccinate} = 1$ (10% of the population) and a large vaccinated set (VT-HCSA-L) $N_{vaccinate} = 5$ (50% of the population).

Problems The colour space specialisation of the Habitat Antigenic Exposure Paradigm (HAEP) defined in Algorithm 6.1 was used called the Habitat Colour Space Problem (HCSP). The minimal variation of HCSP was used with $N_{habitats} = 10$, and $N_{determinants} = N_{antigen} = N_{infections} = 1$. Each B was a Colour Space Pattern (CSP) was randomly generated at the beginning of each run. The five different Host Exposure Regimes (HER) defined in Section 6.4.1 were used for the HCSP. These included the AHER where the number of H matched the number of B (one-to-one), SHER, RHER, PHER where H_{point} was fixed at tissue 1, and OHER with $duration = 15$.

Experiment Each algorithm used the Maximum Epochs Stop Condition (MESCC) defined in Equation 4.8 with $MaxEpochs = 1000$. The four host specific measures defined

in Section 6.4.3 were collected from the state of the system after the triggering of the stop condition. These included the system measure PE and PD, and the host measures AHE and AHD. Each algorithm and problem received a new and different random number seed each run. Algorithm-Problem combinations were repeated 30 times.

Results

Table A.5 in Appendix A.2.1 provides a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. Box-and-whisker plots are provided in which the results for each algorithm are aggregated across all HER for a each measure. Figure 6.3 shows PD, Figure 6.4 shows PE, Figure 6.5 shows AHD, and Figure 6.6 shows AHE.

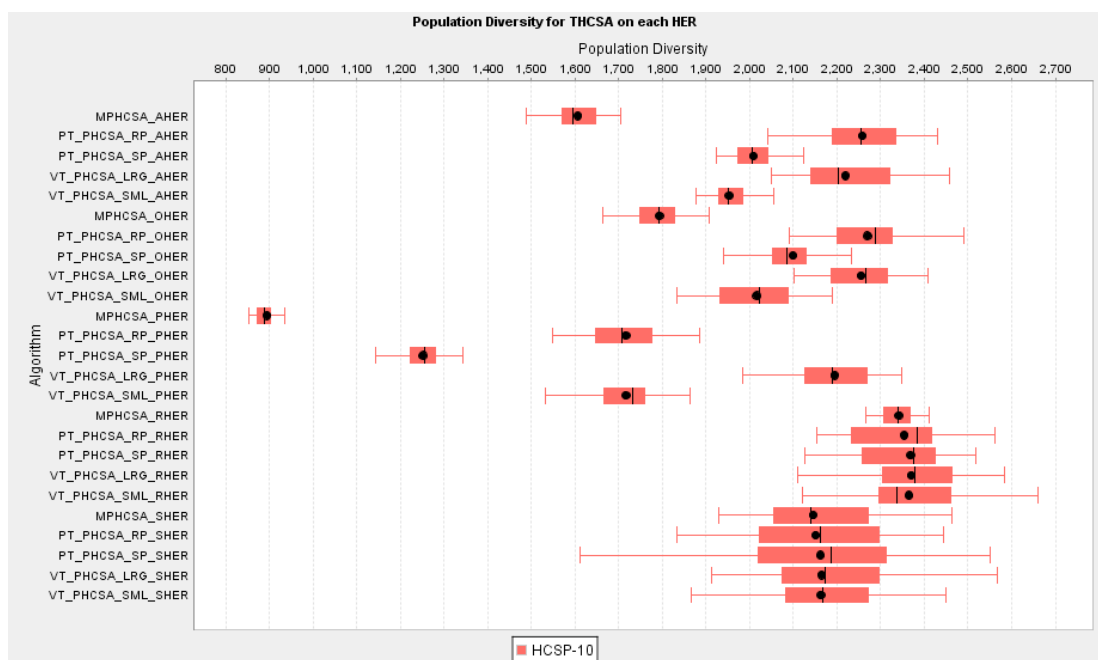


Figure 6.3: Box-and-whisker plot of Population Diversity (PD) across all HER for the THCSA study.

Analysis

This section provides an analysis of the results from the empirical study into the THCSA summarised in Table A.5. These analyses exploit the trends and expectations outlined in Section 6.4.4, and in particular the attributes of host exposure regimes in Table 6.4.

Pathogen Transmission Trends This section considers the constrained (spatial) and unconstrained random host pairing methods for the pathogen transmission scheme compared to the MP-HCSA that does not facilitate inter-host communication. From a system perspective the effect of pathogen transmission was restricted to those exposure regimes with an asymmetric distribution of information to the hosts in the population. Pathogen transmission resulted in an increase in population diversity and population error compared

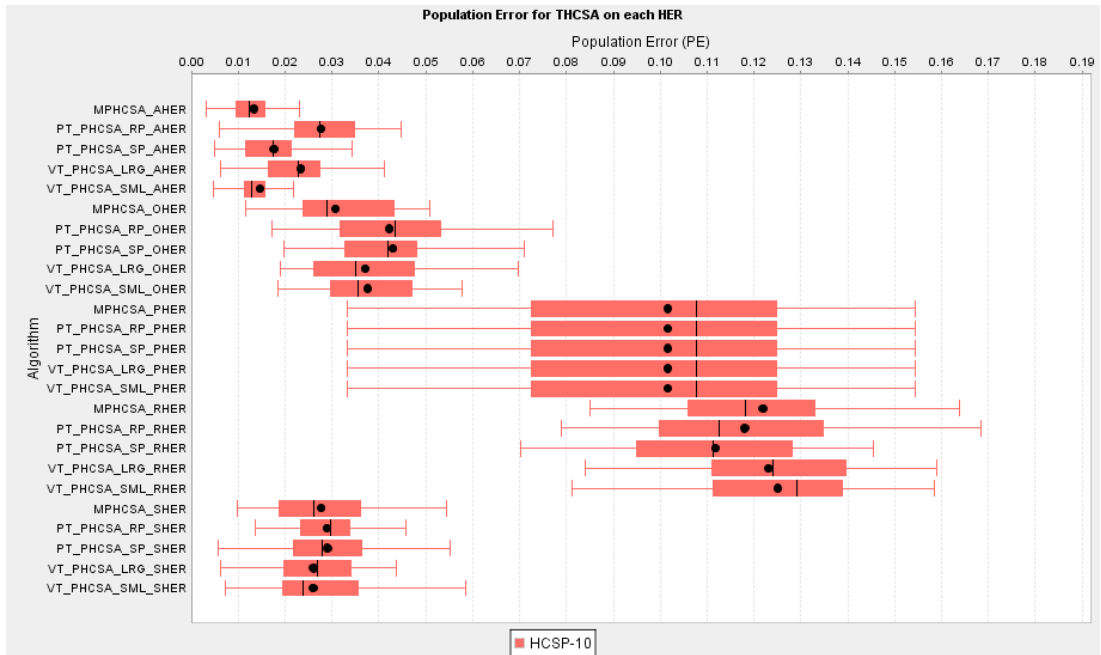


Figure 6.4: Box-and-whisker plot of Population Error (PE) across all HER for the THCSA study.

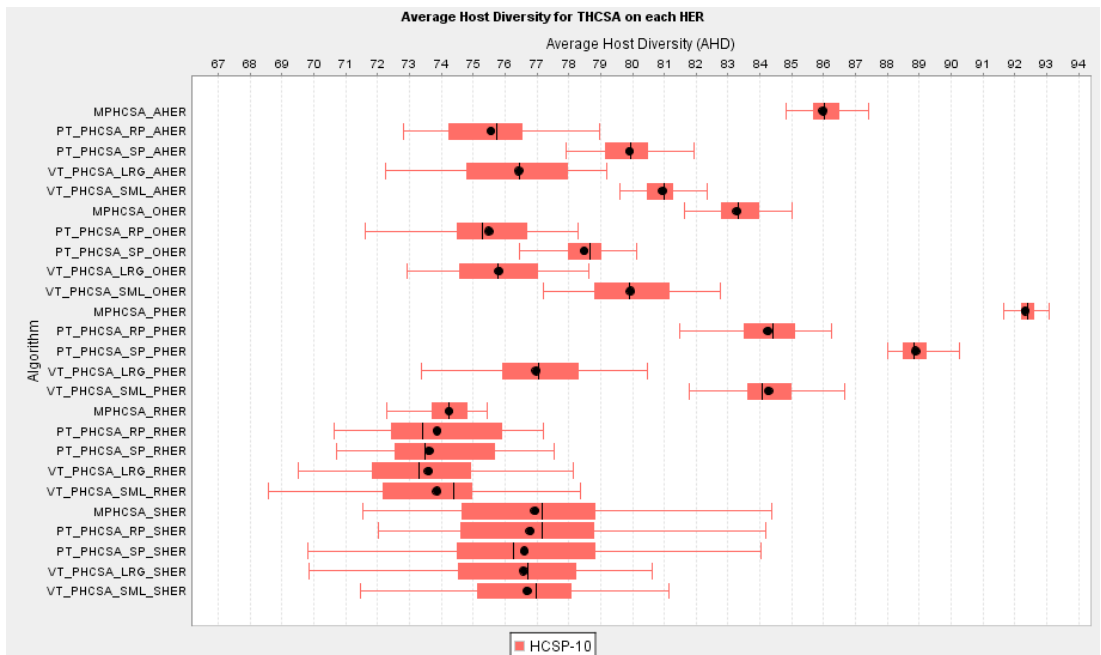


Figure 6.5: Box-and-whisker plot of Average Host Diversity (AHD) across all HER for the THCSA study.

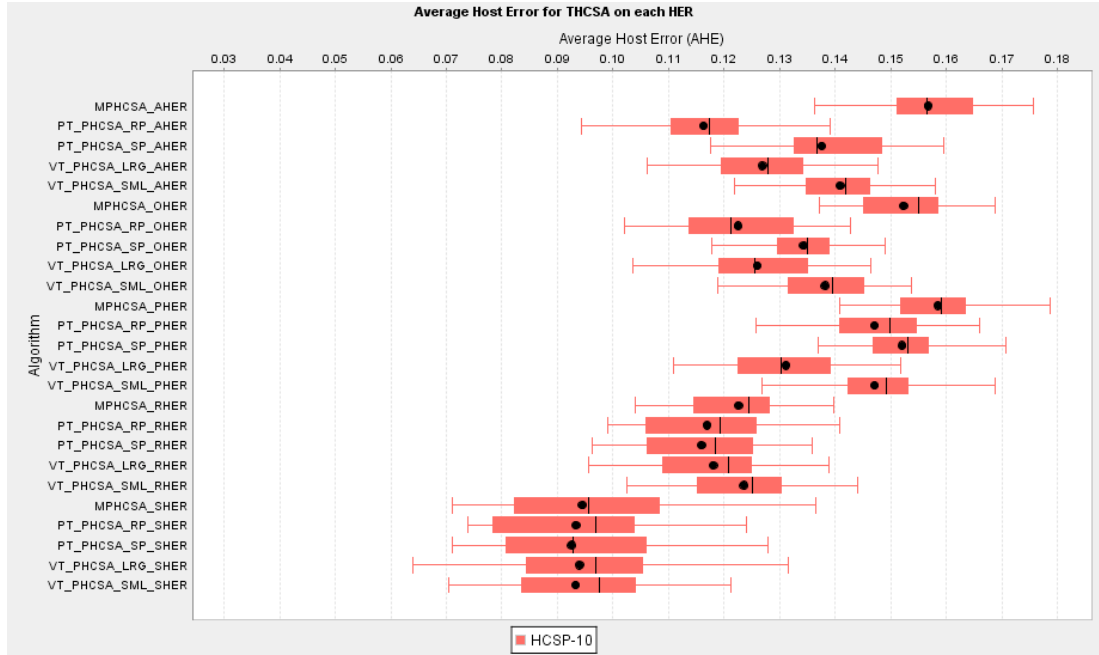


Figure 6.6: Box-and-whisker plot of Average Host Error (AHE) across all HER for the THCSA study.

to MP-HCSA on AHER, OHER, and PHER, with no significant difference between the three approaches at the system level on RHER and SHER. The random pairing method resulted in a higher system diversity and system error (increased effect) compared to the constrained spatial random pairing method. The same trend was observed at the component level with regard to the symmetry in information distribution by the exposure regimes. No significant difference was observed between the three approaches on RHER and SHER with regard to AHD and on SHER with regard to AHE. A relatively large decrease in component diversity and component error was observed with both pathogen transmission schemes on AHER, OHER and PHER. As was the case with the system measures, the component measures observed an increase in effect with the unconstrained random pairing method than the constrained random spatial pairing method. The observations regarding the two pathogen transmission schemes may be generalised to the following trends:

1. Independent perspectives of the information environment are promoted by asymmetrical and not symmetrical information exposure with the chosen host configuration.
2. Pathogen transmission increases system diversity and error, and decreases component diversity and error on asymmetrical information exposure regimes.
3. The pathogen dissemination effect was more pronounced with random pairing compared with spatial pairing.

Vaccination Size Trends This section considers the vaccination transmission scheme compared to the MP-HCSA, and in particular the effect of varying the number of hosts selected for vaccination per population exposure ($N_{vaccinate}$). From a system perspective

the vaccination scheme demonstrated the same general trend as the pathogen scheme with regard to the restriction of the dissemination effect to those exposure regimes with an asymmetrical information distribution. Vaccination resulted in an increase in system diversity compared to MP-HCSA, and generally no significant change to system error other than a small increase in OHER. The large vaccination size demonstrated a larger increase in system diversity over MP-HCSA than the smaller vaccination size. From a component perspective vaccination demonstrated a general decrease in the Average Host Diversity and Average Host Error on the asymmetric information distribution exposures. This observed dissemination effect increased with the increase in the vaccination sample size. This result relates to the herd immunity effect, as effective information dissemination is required to facilitate majority coverage in the population under a probabilistic exposure regime. A stronger connection to the herd immunity effect may be investigated in the future by considering a probabilistic variation of PHER with spatial-temporal consistency like the OHER where the dissemination effect for a one or a number of antigenic habitat's could be assessed under a variety of $N_{vaccinate}$ values. The expectation is that the herd immunity effect would be correlated with the improvements to the already measurable dissemination effect (relative decreases in average component diversity and error). The observations regarding the vaccination transmission scheme with varied vaccination sizes may be generalised to the following trends:

1. Vaccination promotes dissemination of information that increases system diversity, whilst at the same time decreases average component diversity and error only on exposure regimes with asymmetric information distribution across the population.
2. The effects of dissemination on the average component is more pronounced with the larger vaccination size (lower relative Average Host Diversity and Error).

Dissemination Trends Table A.6 in Appendix A.2 contains the results from the random pairing pathogen transmission configuration and the vaccination transmission with large $N_{vaccinations}$. These two approaches demonstrated better information dissemination for their respective approaches. This section compares these two different approaches across the five exposure regimes in terms of information dissemination. From a system perspective there was little significant difference in system diversity and error between the pathogen and vaccination transmission schemes. The only observed significant difference at the system level was on the PHER with an asymmetric information distribution. The average component perspective provided more insight into the observable differences between the two transmission schemes. The pathogen scheme demonstrated a slightly lower average host diversity on AHER, and a relatively slight increase on the PHER compared the vaccination. The vaccination scheme demonstrated a lower average error on the PHER, whereas the pathogen approach demonstrated a slight decrease in error over vaccination on AHER. This differentiation between the two methods suggests that pathogen transmission provides some benefit on exposure regimes with asymmetric distribution and a system-wide scope, where as vaccination is suited to the constrained scope exposure with asymmetric distribution. This is an intuitive result as one-to-many dissemination is

preferred for constrained exposures, and one-to-one dissemination for system-wide exposures. The observations made regarding pathogen and vaccination transmission may be generalised to the following trends:

1. There was no significant difference between the pathogen transmission and the vaccination transmission approaches on exposure regimes with symmetric information distribution.
2. Pathogen transmission schemes provide improved dissemination on system-wide exposure with asymmetric information distribution, whereas the vaccination transmission scheme provides improved dissemination on constrained exposure with asymmetric information distribution.

Conclusions

This section summarises the findings of the empirical study into the Transmission Host Clonal Selection Algorithm, in terms of the primitives that were the focus of the study and the expectations that motivated the study.

1. *Primitives*

- (a) Pathogen transmission with random pairing provides a viable metaphor for information dissemination via elicitation of response.
- (b) Vaccination transmission provides a viable metaphor for information dissemination via elicitation of response that improves with vaccination size.

2. *Transmission*

- (a) There was generally little difference in the dissemination effect between PT-HCSA-RP and VT-HCSA-L, except the pathogen approach demonstrated a relatively slightly better effect on system-wide exposure as opposed to the relatively slightly better effect observed with the vaccination approach on constrained exposure with asymmetrical information distribution.
- (b) Exposure regimes with an asymmetric information distribution are required to engender a varied perspective of the antigenic environment in hosts permitting the observation of the dissemination effect.

This last point highlights the important finding from the results that confirmed a base assumption of the empirical investigations into host-based system, which is that the chosen minimal configuration does not result in significantly independent perspectives on the same antigenic environment. This result was demonstrated in the results for the symmetrical RHER and SHER regimes. Alternatively, this result may be interpreted as the confirmation that an exposure regime with an asymmetric information distribution is required to observe varied perspectives of the antigenic environment with the chosen host configuration, allowing the dissemination effect to be observed and measured. The configuration of transmission schemes were simplistic, providing much room for elaboration

and improvement. An important consideration to motivate further investigation is the use of both pathogen and vaccination based transmission schemes in parallel providing the benefits of varied cardinality-based dissemination (one-to-one, one-to-many) suitable for an unknown exposure regime.

6.6 Population Shared Immunity

6.6.1 Effector Transmission

An alternative to actively eliciting an immune response is to allow hosts to sample and share the product of immunisation. Section 6.2.1 defined this as passive immunisation highlighting that the immunity provided by inter-host transfer of antibodies is fast, effective, and temporary. This section considers the artificial transplantation of passive immunity between the hosts as a strategy for disseminating specialised acquired information about the antigenic environment between hosts in the population. The *Shared Immunity* principle is defined as the sharing of acquired immunity between hosts in a population by explicit sampling their own cellular repertoire and transmitting sampled cells to other selected hosts in the population. The sharing requires a sampling scheme of a hosts internal cellular repertoire, and the explicit selection of one or more other sibling host systems in the population to which sampled cells are transmitted. This sampling and transmission facilitates *horizontal sharing* of acquired immunity within the population. The *sampling scheme* is responsible for selecting those mature cells most likely to be useful to those host systems to which the sampled cells are transmitted. There is expected to be a tight coupling between the selection (sampling) of a hosts cellular repertoire and the selection of the recipient host or hosts. Sampled cells will be mature in that they have been produced as a result of an interaction with the antigenic environment. More useful, are those cells that are created from a series of exposures, and thus represent more refined and presently useful information about the environment. The selection of recipient host systems via a *transmission scheme* is the selection of those systems in the population that would most benefit from the sampled acquired immunity. Examples include the transmission of cells from a healthy system (good performance with regard to some system-environment measure), to a system or set of systems that are less healthy (as defined by the same measure). In the absence of such a measure, a good heuristic is the transmission of a representative sample of acquired information to a host that is geographically distant (with regard to the antigenic environment). This heuristic provides a general sharing scheme in which the hosts of the system seek to provide general coverage to the population by sharing acquired immunity. After hosts are selected and the cells are transmitted, the recipient hosts must integrate the received cells into their cellular repertoire.

- *Sampling Scheme*: The identification (selection) and collection of acquired immune cells from a given hosts cellular repertoires for transmission to one or more other hosts (as defined by the hosts cell transmission scheme). Those cells sampled for transmission should be mature with regard to the antigenic environment, and representative of information acquired during the hosts recent exposures.

- *Transmission Scheme*: Requires the selection of sibling hosts within the systems population of immune systems, and the injection (transmission) of sampled cells to selected hosts. The principle concern in the selection of recipient hosts is that the transmitted cells are likely to be useful. This likelihood may be increased by selecting sibling immune systems in the population that are in turn likely to have been subjected to different pathogen and/or at different frequencies (assuming an asymmetric antigenic environment). Once selected, the transmitted cells must be integrated into the recipient systems cellular repertoire.

Algorithm 6.8: HostInteractions for Shared Immunity Clonal Selection.

Input: $P, N_{sharers}, N_{recipients}, N_{sharedcells}$

```

1 for  $s \leftarrow 0$  to  $N_{sharers}$  do
2    $H_s \leftarrow \text{RandomHost}(P)$ ;
3    $T' \leftarrow \text{RandomTissue}(H_s, N_{sharedcells})$ ;
4    $P' \leftarrow P$ ;
5   while  $P'_{t_n} \neq N_{recipients}$  do
6      $H' \leftarrow \text{RandomHost}(P')$ ;
7      $P'.\text{Remove}(H')$ ;
8   end
9   foreach  $H_{i'} \in P'$  do
10     $H_{i'}.\text{Integrate}(T')$ ;
11  end
12 end
```

The Population Host Clonal Selection Algorithm defined in Algorithm 6.2 may be specialised with the the cell sampling and transmission of shared immunity, referred to as the *Shared Immunity Host Clonal Selection Algorithm (SI-HCSA)*. Algorithm 6.8 provides a definition of the *HostInteractions* operation in which a number of hosts ($N_{sharers}$) select a given number of other hosts in the population ($N_{recipients}$) and share a fixed number ($N_{sharedcells}$) of randomly sampled cells. Transmitted cells are copied from the sharing host and are integrated into the receiving hosts using an integration operation (*Integrate*), and host selection both with regard to sharers and receivers is random (*RandomHost*). This provides a simple realisation of the shared immunity algorithm without cell or host selection bias.

6.6.2 Shared Immunity Empirical Study

Aim

The aim of this empirical study is to investigate the Shared Immunity Host Clonal Selection Algorithm in the context of the information dissemination and localisation capabilities as compared to the Minimal Population Host Clonal Selection Algorithm under a variety of host exposure regimes. Toward this end, the study had the following goals:

1. Investigate the relative effects of of small and large sharer and recipient sizes under shared immunity

2. Investigate the relationships between sharer and recipient sizes and information dissemination.
3. Assess whether the trend of information dissemination restriction of exposure regimes with asymmetric information dissemination holds.

Method

Algorithms The study considered the Minimal Population Host Clonal Selection Algorithm (MP-HCSA) and the Shared Immunity Host Clonal Selection Algorithm (SI-HCSA). MP-HCSA is a specialisation of the HCSA defined in Algorithm 6.2, that was configured with $N_{hosts} = 10$. Each H was configured with a single tissue ($N_{tissues} = 1$) that was an instance of the RCCSA defined in Algorithm 4.5, with the configuration $N_{cells} = 50$, $N_{selection} = 1$, and $N_{clones} = 5$. SI-HCSA was defined in Algorithm 6.8, where the integration operation (*Integrate*) was specialised to the Euclidean-based replacement as is used in the RCCSA. The number of shared cells was fixed $N_{sharedcells} = 5$ and the number of sharers and recipients were varied. A small number of sharers $N_{sharers} = 1$ (10% of the population) and a large number of sharers $N_{sharers} = 5$ (50% of the population) were assessed with a small number of recipients $N_{recipients} = 1$ (10% of the population) and a large number of recipients $N_{recipients} = 5$ (50% of the population).

Problems The same Habitat Colour Space Problem and Host Exposure Regimes were used as was defined for the THCSA empirical study in Section 6.5.2.

Experiment The same experimental setup was used as was defined for the THCSA empirical study in Section 6.5.2.

Results

Table A.7 in Appendix A.2.2 provides a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. Box-and-whisker plots are provided in which the results for each algorithm are aggregated across all HER for each measure. Figure 6.7 shows PD, Figure 6.8 shows PE, Figure 6.9 shows AHD, and Figure 6.10 shows AHE.

Analysis

This section provides an analysis of the results from the empirical study into the SI-HCSA summarised in Table A.7. These analyses exploit the trends and expectations outlined in Section 6.4.4 regarding information dissemination between the hosts in the population.

Sharer Size Trends This section is concerned with the effect of the number of hosts sharing acquired immunity in the population ($N_{sharers}$), irrespective of the number of recipients ($N_{recipients}$). As such, the observations are taken as the general trends in the small and large number of sharer hosts. From a system perspective the Population Diversity

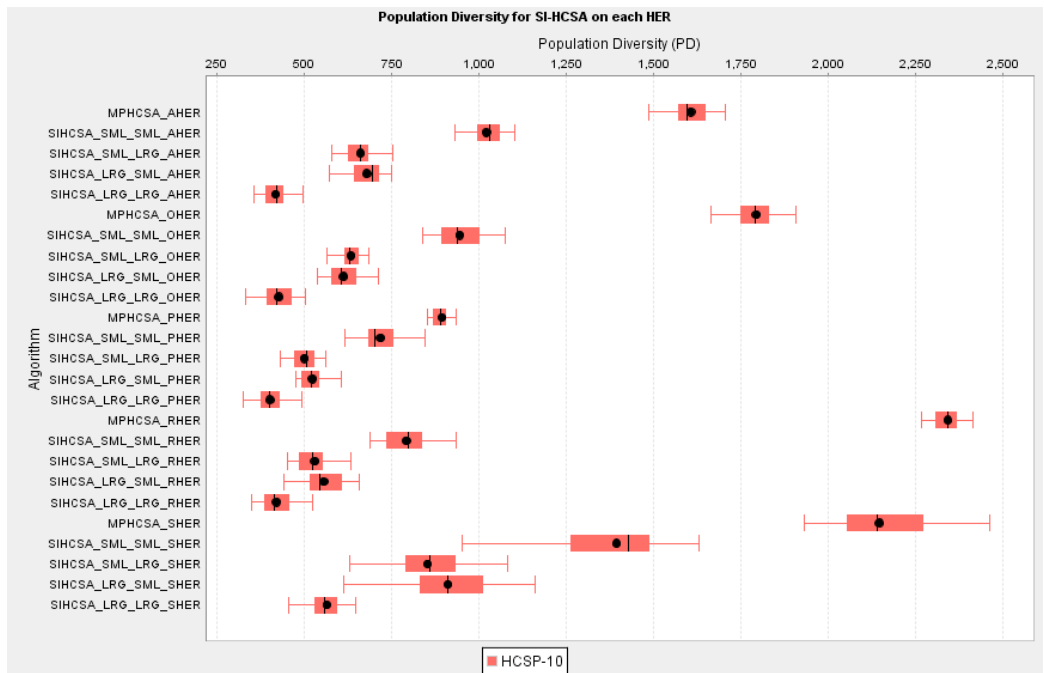


Figure 6.7: Box-and-whisker plot of Population Diversity (PD) across all HER for the SI-HCSA study.

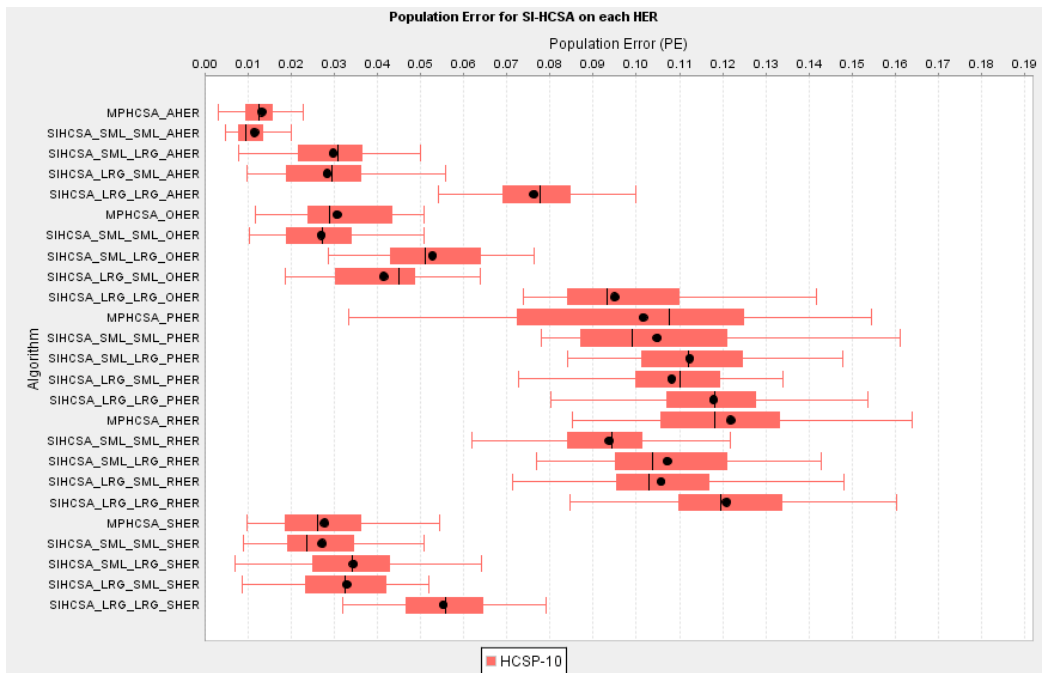


Figure 6.8: Box-and-whisker plot of Population Error (PE) across all HER for the SI-HCSA study.

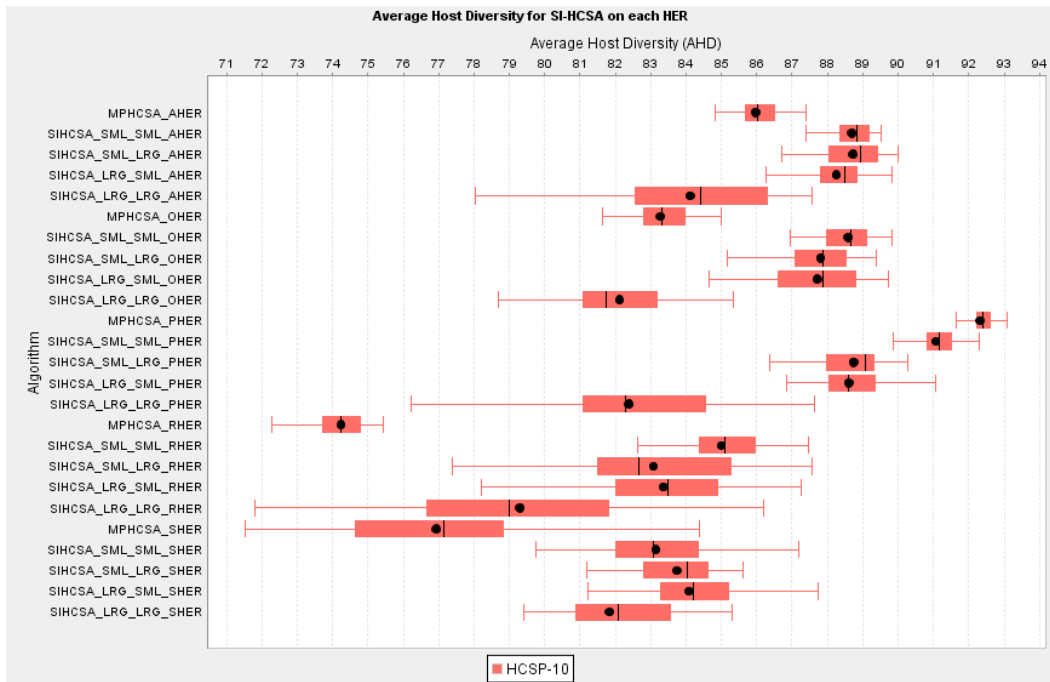


Figure 6.9: Box-and-whisker plot of Average Host Diversity (AHD) across all HER for the SI-HCSA study.

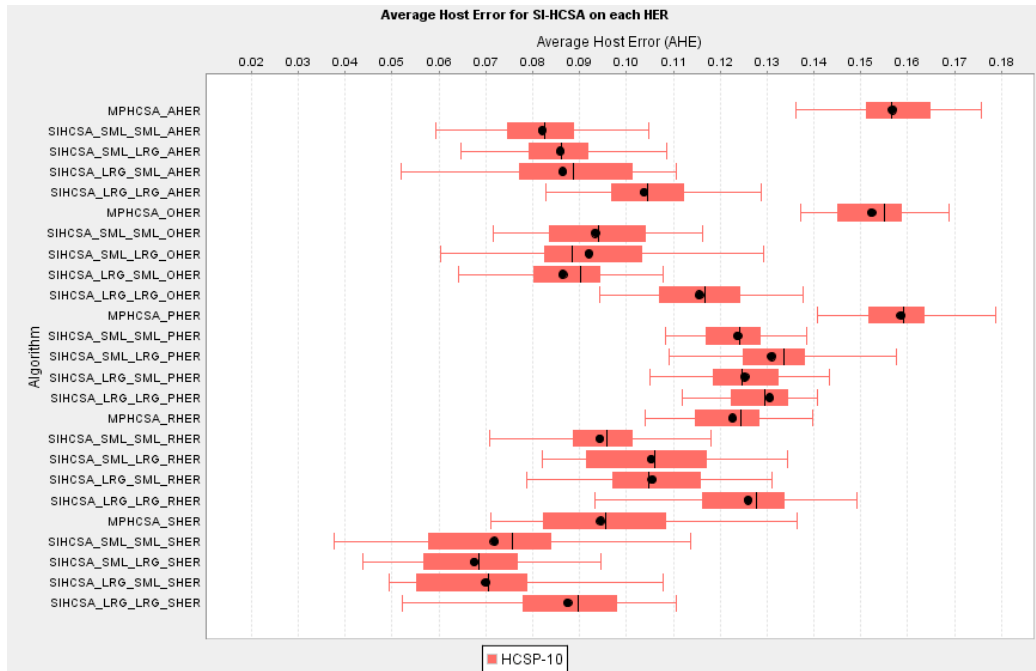


Figure 6.10: Box-and-whisker plot of Average Host Error (AHE) across all HER for the SI-HCSA study.

was generally lower with sharing compared to MP-HCSA across all exposure regimes. The increase in sharer size generally resulted in a decrease in diversity. A similar relationship was observed with system error, where sharing generally resulted in an increase in system error compared to MP-HCSA, an effect that increased with the number of sharers (increase in share size increased system error). From a component perspective shared immunity resulted in a general decrease in the Average Host Diversity than MP-HCSA. This effect decreased with the increase in number of sharing hosts, such that a AHD approached MP-HCSA on the AHER, OHER, and RHER domains. Interestingly, a similar effect was observed with the Average Host Error that was generally lower than MP-HCSA, although increased with the increase in the sharer size. This correlation in relationships at the system and component level with share size demonstrates a disruptive and information dissemination effect at the system and component level respectively that generally results in increased disruption at both scales with the increase in the number of hosts sharing immunity. The observations may be generalised to the following trends:

1. *System*

- (a) Shared immunity results in a generally lower population diversity and generally higher population error than MP-HCSA.
- (b) Increase in sharing size results in increase in system level effects, specifically decreases in system diversity and increases in system error.

2. *Component*

- (a) Shared immunity results in a general increase in Average Host Diversity, and a general decrease in Average Host Error compared to MP-HCSA.
- (b) Increase in a decrease in component level effects with an increase a decrease in Average Host Diversity, and a general increase in Average Host Error.

Recipient Size Trends This section considers the effects of varying the number of recipient hosts ($N_{recipients}$) irrespective of the number of hosts sharing. As such, the results are generalised across the configuration values for the number of sharers ($N_{sharers}$). From a system perspective Population Diversity and Error exhibited the same trends both with regard to the comparison to MP-HCSA and with regard to the increase in the number of recipients increasing the disruptive effect of sharing at the system level. The same relationship was demonstrated at the component level with a general increase in diversity and decrease in error compared to MP-HCSA, and the increased disruption (decreased average diversity and increased average error) with the increased in the number of recipients. These observation demonstrate an important symmetry in general behaviour between the number of hosts sharing acquired immunity and the number of hosts receiving acquired immunity. In particular increasing either results in a increase in the disruption at the system level and disruption to the dissemination of information at the component level. These observations may be summarised in the following generalised trends:

1. The system and component-level effects of recipient size mirror the effect trends of number of sharing hosts.
 - (a) System diversity decrease and system error which increases with recipient size.
 - (b) Component diversity increases and component error decreases, an effect that decreases with recipient size.

Sharer and Recipient Relationship Trends This section considers the trends in the relationships between the number of hosts sharing and the number of hosts receiving shared immunity, not considering the comparison the MP-HCSA. From a system perspective the highest population diversity was observed with a one-to-one relationship, whereas the lowest population diversity was observed with a many-to-many configuration. Generally there was no significant difference between one-to-many and many-to-one configurations across the exposure regimes, with PHER as the exception. The same general trend was observed with regard to population error, where the lowest error was achieved by one-to-one and the largest system error was achieved by many-to-many. Interestingly, no significant difference in system error was observed between SI-HCSA-SS and MP-HCSA on all exposure regimes except RHER where one-to-one sharing resulted in a lower system error. The component perspective demonstrated the same general trend where the highest average component diversity and least average component error were observed with the one-to-one configuration, and the most disruption in diversity and error caused by the large scale sharing in the many-to-many configuration. Also consistent was the lack of significant difference in effect between the one-to-many and many-to-one configurations. The consistency of the dissemination effect in the relationship between both the number of sharer and recipient hosts was confirmed both in the extreme cases with the least and most disruption caused by the least and the most amount of sharing, and in the intermediate cases confirming the symmetry in behaviour. These observations may be summarised in the following generalised trends:

1. Generally no significant difference in assessed measures between one-to-many or many-to-one with regard to the number of sharers and recipients across all HER.
2. The least disruptive system effect was observed with one-to-one, whereas the largest disruptive system effect was observed with many-to-many sharing.
3. The best dissemination effect at the component level was observed with one-to-one sharing, whereas the worst disruptive effect at the component level was observed with many-to-many sharing.

Conclusions

This section summarises the findings of the empirical study into the Shared Immunity Host Clonal Selection Algorithm, in terms of the primitives that were the focus of the study and the expectations that motivated the study.

1. Large scale sharing provides a disruptive effect at the system and component level.

2. Small scale sharing minimises the disruptive effect at the system level and maximise the dissemination effect at the component level.

An interesting aspect of artificial passive immunity is the concern that such shared information is initially useful although such utility is temporary. This property is expected to vary with exposure regime and provides an important extension to the investigation into the SI-HCSA.

6.7 Generational Maternal Immunity

6.7.1 Generational Transmission

Maternal immunity involves progeny receiving antibodies and immune cells from the mother, both across the placenta and in breast milk (mucosal immunity). This type of immunity is referred to as natural-passive immunity because the conferred acquired immune information is passed between the hosts without eliciting an immune response. Maternal immunity provides a strategy for one generation of immune systems with lifetime acquired immunity to transfer such acquired knowledge to the following generation. As with intra-population sharing of acquired immunity via passive means in Shared Immunity (Section 6.6), a cell sampling and transmission scheme are required. The concerns of cell sampling are predominantly similar to the concerns highlighted in the Transmission and Shared Immunity Host Clonal Selection Algorithms. It is important to highlight the trade-off of the sampled cells and their effect on freshly instantiated host systems. A selection of a sample that is too large or contains many dominant (with respect to clonal selection) pieces of acquired information (memory cells) will cause the recipient systems to (in effect) represent continuances of the transmitting host. The sample should be diverse and representative, and likely contain many effector cells. The selection of effector cells for transmission is useful for a number of reasons. Firstly, the biological inspiration (maternal, and more importantly mucosal immunity) confers mostly this type of acquired immunity in the form of antibodies.

- *Sampling Scheme:* The selection of mature acquired immune cells from host systems to be removed and transmitted to host systems of the subsequent generation. The selection scheme should draw a representative sample of the information acquired by the system over the course of its lifetime (trial period).
- *Transmission Scheme:* The selection of instantiated host systems of the subsequent generation by hosts of the present generation to which sample acquired immune systems cells will be transmitted. A simple host selection scheme is the selection of progeny host systems.

The host transmission scheme requires the selection of hosts in the subsequent generation to receive the sampled cells. A natural implementation of this scheme is the selection of progeny hosts. Specifically, the selection of hosts in the subsequent generation by hosts in the present generation, to which they are responsible for instantiating. The parent-child

transmission scheme may result in the formation of independent generational host-lines given the asexual basis of the approach. The concern of this effect is that if a given host system is lost, then the information acquired by that hosts generational line is also lost. This concern may be addressed by relaxing the parent-child transmission constraint, and allowing parent-hosts to potentially transmit to any host in the child generation. A biological basis for this configuration may be the use of manufactured formulae, or the use of a wet-nurse. An example of a decoupled transmission scheme is random host selection, with reselection. The reselection allows a given host in the child generation to potentially receive acquired immune information from more than one parent host system. The transmission of cells to non-child hosts provides redundancy between host lines, if there is a relationship between parent-child that may affect host-loss (such as reproduction in a hazardous spatial environment). A one-to-many transmission scheme between the generations will also foster redundancy of acquired information between the generations. The concern in a host system receiving acquired immune information from more than one host, is that the child system may become overly biased by the previous generation. A parent-child (one-to-one) transmission scheme limits the scope of received acquired immune information to a single host of the previous generation’s perspective.

The Minimal Generational Host Clonal Selection Algorithm (MG-HCSA) of the generational algorithm defined in Algorithm 6.3 provides a basis for comparison with a maternal immunity approach as it promotes independent generations. Algorithm 6.9 provides a definition of the *CreatePopulation* operation for the MG-HCSA. Equation 6.4 provides a generational change trigger based on a number of population exposures or epochs, where a user specified parameter $N_{genepochs}$ defines the maximum number of epochs before a generational change is triggered.

Algorithm 6.9: CreatePopulation for the Minimal Generational Clonal Selection.

Input: P, N_{hosts}
Output: P'

- 1 P' \leftarrow 0;
- 2 **for** $i \leftarrow 0$ **to** N_{hosts} **do**
- 3 $H_i \leftarrow$ CreateHost();
- 4 P' \leftarrow H_i ;
- 5 **end**
- 6 **return** P';

$$GenerationalChange(Epoch_i) = (Epoch_i \geq N_{genepochs}) \quad (6.4)$$

A *Maternal Immunity Host Clonal Selection Algorithm (MI-HCSA)* may be defined as the realisation of intra-generational sharing of acquired immunity that may be specialised to parent-child based sharing. Algorithm 6.10 defines the *CreatePopulation* operation for parent-child based maternal immunity, where the amount of inter-generational sharing is defined by a user defined parameter $N_{maternalcells}$.

Algorithm 6.10: CreatePopulation for Maternal Immunity Clonal Selection.

Input: $P, N_{maternalcells}$
Output: P'

- 1 $P' \leftarrow 0;$
- 2 **foreach** $H_i \in P$ **do**
- 3 $T'_{i} \leftarrow \text{SelectRandomTissue}(H_i, N_{maternalcells});$
- 4 $H'_{i} \leftarrow \text{CreateHost}();$
- 5 $H'_{i}.\text{Integrate}(T'_{i});$
- 6 $P' \leftarrow H'_{i};$
- 7 **end**
- 8 **return** $P';$

6.7.2 Maternal Immunity Empirical Study

Aim

The aim of this empirical study is to investigate the Maternal Immunity Host Clonal Selection Algorithm in the context of the information dissemination and localisation capabilities as compared to the Minimal Generational Host Clonal Selection Algorithm under a variety of host exposure regimes. Toward this end, the study had the following goals:

1. Compare and contrast generational clonal selection with and without inter-generational sharing.
2. Investigate the dissemination of information effect and its disruption by varying the amount of inter-generational sharing.

Method

Algorithms The study considered the Minimal Generational Host Clonal Selection Algorithm (MG-HCSA) and the Maternal Immunity Host Clonal Selection Algorithm (MI-HCSA). MG-HCSA is a specialisation of the HCSA defined in Algorithm 6.3, that was configured with $N_{hosts} = 10$. Each H was configured with a single tissue ($N_{tissues} = 1$) that was an instance of the RCCSA defined in Algorithm 4.5, with the configuration $N_{cells} = 50$, $N_{selection} = 1$, and $N_{clones} = 5$. The *CreatePopulation* operation for the MG-HCSA was realised as the creation of a new population of hosts each generational change as defined in Algorithm 6.9, and the epoch generational change condition defined in Equation 6.4 with $N_{genepochs} = 100$. The MI-HCSA was assessed with the parent-child specialisation of the *CreatePopulation* operation defined in Algorithm 6.10. The number of cells selected for inter-generational sharing was assessed with three different values as follows: small $N_{maternalcells} = 5$ (10% of the hosts repertoire), medium $N_{maternalcells} = 10$ (20% of the hosts repertoire), and large $N_{maternalcells} = 25$ (50% of the hosts repertoire).

Problems The same Habitat Colour Space Problem and Host Exposure Regimes were used as was defined for the THCSA empirical study in Section 6.5.2.

Experiment The same experimental setup was used as was defined for the THCSA empirical study in Section 6.5.2, except the maximum number of epochs of the EOSC stop condition was decreased from 1000 to 999. This configuration change was made such that the final measures recorded reflected the state of the system after 10 generations of 100 epochs before the last generational change (at the tenth generation).

Results

Table A.8 in Appendix A.2.3 provides a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. Box-and-whisker plots are provided in which the results for each algorithm are aggregated across all HER for a each measure. Figure 6.11 shows PD, Figure 6.12 shows PE, Figure 6.13 shows AHD, and Figure 6.14 shows AHE.

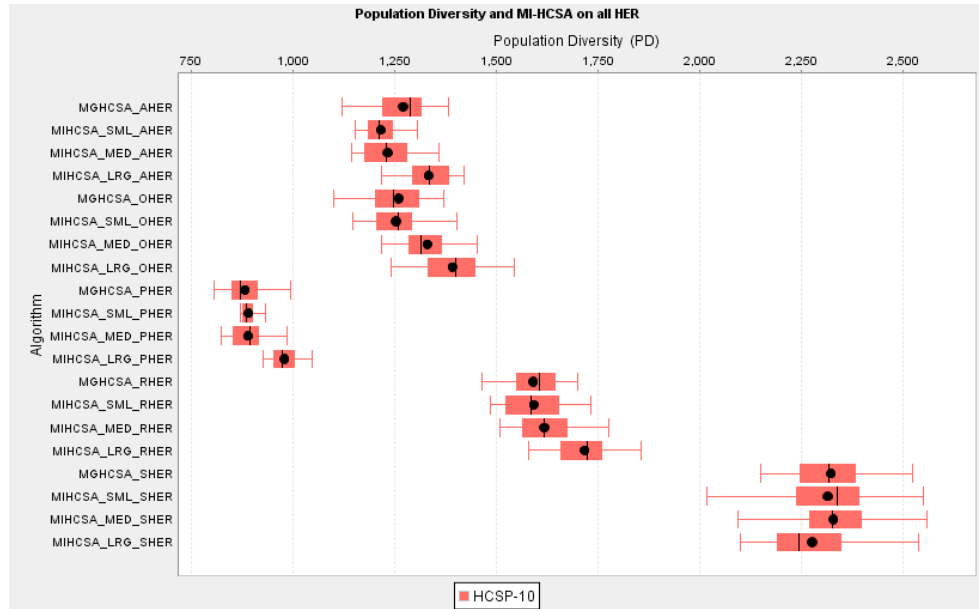


Figure 6.11: Box-and-whisker plot of Population Diversity (PD) across all HER for the MI-HCSA study.

Analysis

This section provides an analysis of the results from the empirical study into the MI-HCSA summarised in Table A.8. These analyses exploit the trends and expectations outlined in Section 6.4.4 regarding information dissemination between the hosts in the population.

Generational Trends This section compares and contrasts the minimal generational approach without sharing with the maternal immunity approach with inter-generational sharing. MG-HCSA is compared against MI-HCSA-L as it is expected to exhibit the largest effect of sharing and therefore difference to the minimal approach. From a system perspective MI-HCSA resulted in a generally population diversity and lower system error than MG-HCSA across all five of the exposure regimes. Interestingly this trend was not

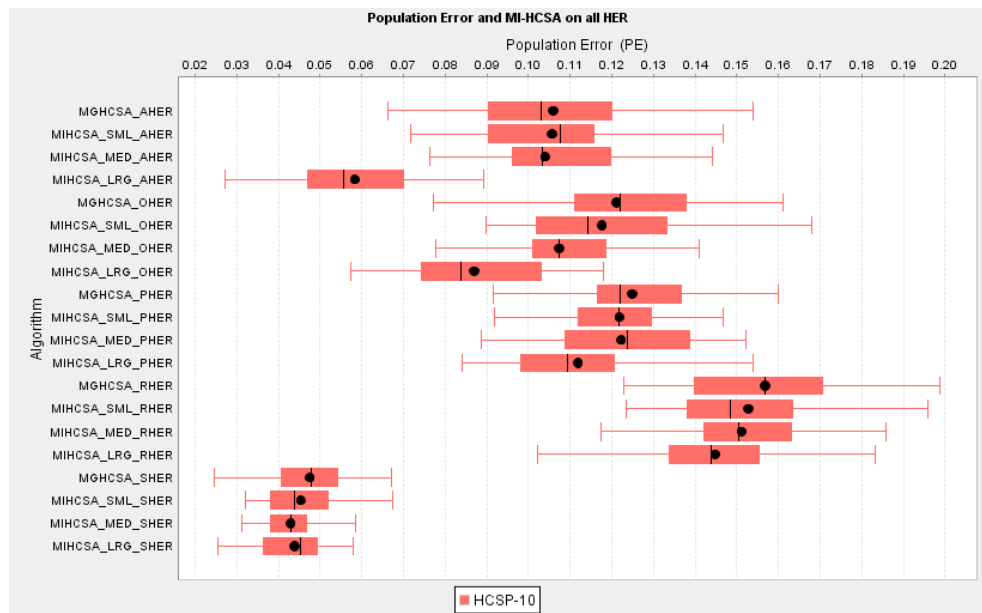


Figure 6.12: Box-and-whisker plot of Population Error (PE) across all HER for the MI-HCSA study.

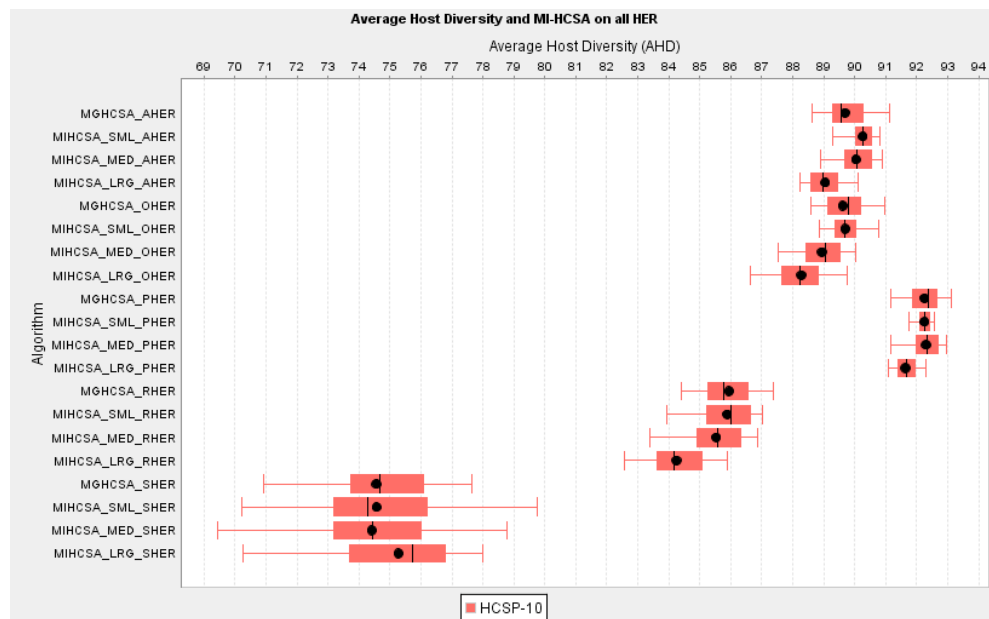


Figure 6.13: Box-and-whisker plot of Average Host Diversity (AHD) across all HER for the MI-HCSA study.

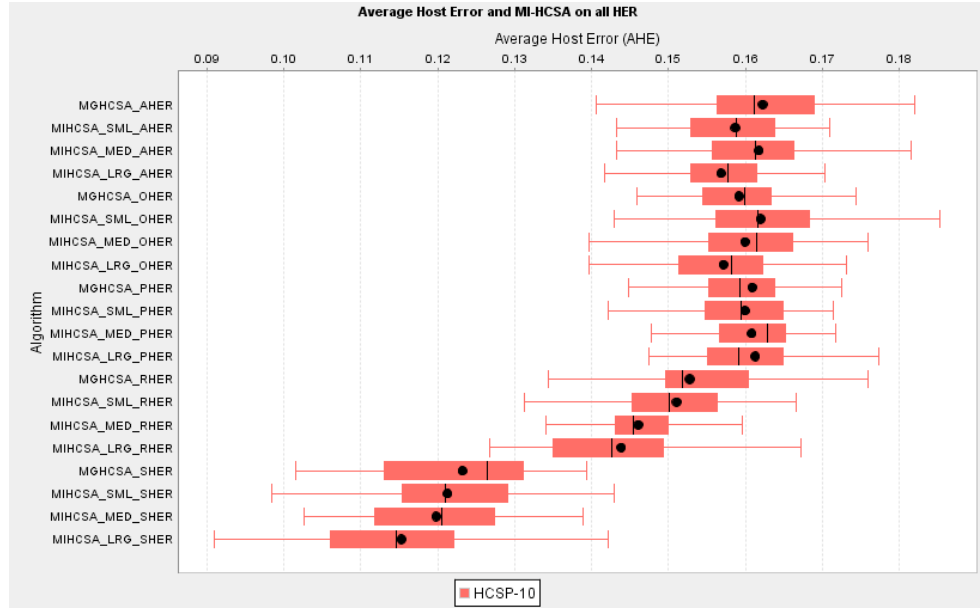


Figure 6.14: Box-and-whisker plot of Average Host Error (AHE) across all HER for the MI-HCSA study.

observed on the symmetric exposure regime, where no significant difference was observed. From a component perspective, maternal immunity generally achieved a lower Average Host Diversity and lower Average Host Error than the minimal approach. This was generally observed across the five HER's, except on SHER and PHER with regard to AHD and AHE respectively. The results show a clear localisation effect with decreased error and increase in diversity at the system level, with the interesting side effect of decreased average error at the component level. The general lack of effect on the SHER suggests that the inter-generational sharing does not provide any benefit under such a strongly specialising exposure regime. The observations may be generalised to the following trends:

1. Maternal immunity results in improved localisation compared to the minimal generational approach.
2. Parent-Child inter-generational sharing provides an inter-generational localisation method resulting in highly-specialised hosts.

Maternal Immunity Size Trends This section considers the effect of varying the number of cells shared between the generations defined by different values for the $N_{maternalcells}$ user parameter. From a system perspective the results demonstrated an increase in effect with the number of transmitted maternal cells. Specifically, increase in $N_{maternalcells}$ resulted in an increase in system diversity and decrease in system error. The effect was less pronounced with small sample sizes, typically with little significant difference between MG-HCSA and MI-HCSA-S as well as between MI-HCSA-S and MI-HCSA-M. From a component perspective the increase in the number of transmitted sampled cells resulted in relative increase in the effect as measured by decreases in Average Host Diversity and decreases in Average Host Error. The same general lack of significance was observed with

small, medium and empty sample sizes. The maternal cell size trend demonstrates that parent-child generational sharing promotes a strong localisation effect that increases with the amount of information communicated between the generations across the variety of exposure regimes. Importantly the effect is not apparent (significant) until a sufficiently large number of randomly selected cells are transmitted. The observations may be generalised to the following trends:

1. Localisation effect increases with the number of transmitted maternal cells.
2. The number of cells must be of a sufficient size (50%) for the effect to be consistently significant across the exposure regimes.

Conclusions

This section summarises the findings of the empirical study into the Maternal Immunity Host Clonal Selection Algorithm, in terms of the primitives that were the focus of the study and the expectations that motivated the study.

1. *Generational*
 - (a) The assessed specialisation of maternal immunity resulted in improved inter-generational specialisation than no inter-generational sharing.
2. *Parent-Child Sharing*
 - (a) Maternal Immunity with parent-child sharing is an inter-generational localisation method.
 - (b) The concern of maternal immunity that sufficiently large sample sizes would result in homogeneous generations was confirmed to the benefit of the localisation effect resulting in reduced system and component-wise error.

An important consideration not addressed is the differentiation from generational learning and lifetime learning. As mentioned in the method, the results were taken from the end of the run before a generational change, thus reflect generational learning over 9 generations of 100 epochs, and lifetime learning over 99 epochs in the final generation. Elaborative investigations may consider such a distinction and assess generational learning as the performance of the population before lifetime learning (at the epoch of creation), and lifetime learning before the generational change. When such a distinction is not made as in the case of this and the next empirical studies, the effects are still present, although aggregated (confounded) together. The parent-child approach assessed is perhaps the simplest of such maternal immunity approaches. One may consider shifts in the cardinality between sharers in the previous generation and receivers in the next generation, with a clear relationship to the investigation into SI-HCSA and whether the results hold across generations. Two additional important questions that may motivate future work include (1) the consideration as to how much of the new generation must be seeded with maternal cells for the effect to be observed, and (2) the effect of positioning children in different parts

of the antigenic environment (effecting the specialisation of children under the exposure regimes).

6.8 Generational Evolved Immunity

6.8.1 Generational Inheritance

The evolution of the immune system provides a inter-generational method for propagating functional aspects of the immune system based on lifetime performance. Unlike maternal immunity, evolution does not share acquired immune information directly in the form of cells, rather it shares it indirectly through selection and inheritance of the mechanisms that create and utilise those cells. An inherited generational algorithm requires a selection and reproduction mechanisms for hosts. The important differences in this specialisation of the generational algorithm is that a given host may contribute more or less than the other hosts in its generation. This is a realisation of natural selection, where a measure of the hosts in the population discriminates the reproductive fitness and genetic contribution to the next generation. The principle difference of this specialisation is the use of a genetic-basis for traits that effect hosts acquired immune system. The principle components of the inherited generational algorithm are defined as follows:

- *Selective Scheme*: A specialisation of the host selection scheme of the minimal generational algorithm that realises the principles of natural selection. A given hosts contribution to the subsequent generation is differentiated based on assessed fitness against an inherited trait that effects a hosts acquired immune system.
- *Genetic Basis*: A genetic code (genome) is used to define a trait of a hosts acquired immune system. This genetic code provides the basis of natural selection (differentiated reproductive success), and the medium for reproduction (genetic inheritance).
- *Reproductive Scheme*: The reproductive scheme of the minimal generational population-based algorithm, that may take into account the broader considerations of asexual and sexual reproduction. Reproduction provides the basis for inheritance with regard to the genetic basis of the trait or traits that effect a hosts acquired immune system, and manipulations to that genetic representation in the form of genetic mutation and/or genetic recombination.

The reproductive scheme provides the mechanism for inheritance of the genetic basis of an acquired immune system trait. Reproduction provides a duplication of the parents genetic material which may be modified in minor ways by genetic mutation (likely a lower rate of mutation than that of hypermutation in the immune response). This introduces variations of the trait for natural selection to differentiate reproductive success. A property of the host populations acquired immune system is defined by a parameter or set of parameters, which is encoded in a genome. The genome is the basis for inheritance, and the expressed trait is the basis for host selection. For example, an inheritance mechanism may be defined that encodes information that defines the generation of naïve immune cells in a

hosts acquired immune system. Unlike the maternal sharing scheme that directly reinforces the receptor configurations that are useful, an evolutionary algorithm indirectly reinforces receptor configurations in the generation of a hosts base repertoire, and ongoing naïve cells. This is a broader form of sharing that provides the flexibility for the selection mechanism and scope of the genetic basis for the trait to define the regions of ‘receptor configuration space’ that are beneficial for a hosts untested acquired immune system to sample. The mechanism provides an interesting example that combines both genetic-based learning over generational time and somatic-based learning over a hosts lifetime. Therefore, this class of host algorithm facilitates *learning at two scales*, specifically learning at the host-lifetime scale in somatic adaptations via clonal selection, and learning at the population-generation scale in genetic adaptations via natural selection. Examples of other traits that may be subjected to the pressures of evolution by natural selection include (1) the sensitivity of matching in the clonal selection algorithm (2) and the organisation and connectivity of tissue types in the lymphoid tissue algorithm. In addition to opening up parameterised models in the framework to the process of evolution via natural selection (generational learning), the minimal natural selection algorithm provides a connection of the host-level of the acquired immune system framework with the field of genetic algorithms.

Algorithm 6.11: CreatePopulation for the Evolved Immunity Clonal Selection.

Input: $P, P_{mutation}, P_{crossover}$
Output: P'

- 1 $P' \leftarrow 0;$
- 2 **foreach** $H_i \in P$ **do**
- 3 $H_iFitness \leftarrow HostError(H_i);$
- 4 **end**
- 5 $P_{parents} \leftarrow 0;$
- 6 **for** $i \leftarrow 0$ **to** N_{hosts} **do**
- 7 $H_i \leftarrow BiasedRouletteWheelSelection(P);$
- 8 $P_{parents} \leftarrow H_i;$
- 9 **end**
- 10 **foreach** $H_i, H_{i+1} \in P_{parents}$ **do**
- 11 $H_ichild_1 \leftarrow Crossover(H_i, H_{i+1}, P_{crossover});$
- 12 $P' \leftarrow H_ichild_1;$
- 13 $H_ichild_2 \leftarrow Crossover(H_i, H_{i+1}, P_{crossover});$
- 14 $P' \leftarrow H_ichild_2;$
- 15 **end**
- 16 **foreach** $H'_i \in P'$ **do**
- 17 $Mutate(H'_i, P_{mutation});$
- 18 **end**
- 19 **return** $P';$

The *Evolved Immunity Host Clonal Selection Algorithm (EI-HCSA)* is defined as a specialisation of the Generational Host Clonal Selection Algorithm (defined in Algorithm 6.3) that uses a genetic basis, host selection and host reproduction schemes as defined in the *CreatePopulation* operation in Algorithm 6.11. The $P_{mutation}$ parameter defines the probability of mutating a component in the hosts genetic representation during reproduction,

and $P_{crossover}$ defines the probability of creating two new host from a cross of two parental hosts genetic basis. The reproduction scheme is a rudimentary realisation of the classical genetic algorithm (discussed in Section 3.4.1) for a hosts initial repertoire.

6.8.2 Evolved Immunity Empirical Study

Aim

The aim of this empirical study is to investigate the Evolved Immunity Host Clonal Selection Algorithm in the context of the information dissemination and localisation capabilities as compared to the Minimal Generational Host Clonal Selection Algorithm under a variety of host exposure regimes. Toward this end, the study had the following goals:

1. Compare and contrasted generational clonal selection with and without an inherited initial repertoire.
2. Investigate the localisation and/or dissemination properties of inter-generational inheritance.

Method

Algorithms The study considered the Minimal Generational Host Clonal Selection Algorithm (MG-HCSA) and the Evolved Immunity Host Clonal Selection Algorithm (EI-HCSA). MG-HCSA is a specialisation of the HCSA defined in Algorithm 6.3, that was configured with $N_{hosts} = 10$. Each H was configured with a single tissue ($N_{tissues} = 1$) that was an instance of the RCCSA defined in Algorithm 4.5, with the configuration $N_{cells} = 50$, $N_{selection} = 1$, and $N_{clones} = 5$. The *CreatePopulation* operation for the MG-HCSA was realised as the creation of a new population of hosts each generational change as defined in Algorithm 6.9, and the epoch generational change condition defined in Equation 6.4 with $N_{genepochs} = 100$. The MI-HCSA with assessed with the evolutionary specialisation of the *CreatePopulation* operation defined in Algorithm 6.11. Host fitness was assigned using the AHE measure. The probability of crossover was fixed at 0.90 (90%) per host-pair creation. The mutation duration host creation was $\frac{1}{192}$ bits which is one mutation cell or 50 mutations per population creation ($\frac{50}{9600}$ bits).

Problems The same Habitat Colour Space Problem and Host Exposure Regimes were used as was defined for the THCSA Empirical Study in Section 6.5.2.

Experiment The same experimental setup was used as was defined for the MI-HCSA empirical study in Section 6.7.2, expect the maximum number of epochs on the EOSC stop condition was increased from 999 to 9999. This configuration change was made such that the final measures recorded reflected the state of the system after 100 generations of 100 epochs before the last generational change at one hundredth generation.

Results

Table A.9 in Appendix A.2.4 provides a summary of results for each algorithm-problem combination including the mean (\bar{x}) and standard deviation (σ) of collected measure values. Box-and-whisker plots are provided in which the results for each algorithm are aggregated across all HER for a each measure. Figure 6.15 shows PD, Figure 6.16 shows PE, Figure 6.17 shows AHD, and Figure 6.18 shows AHE.

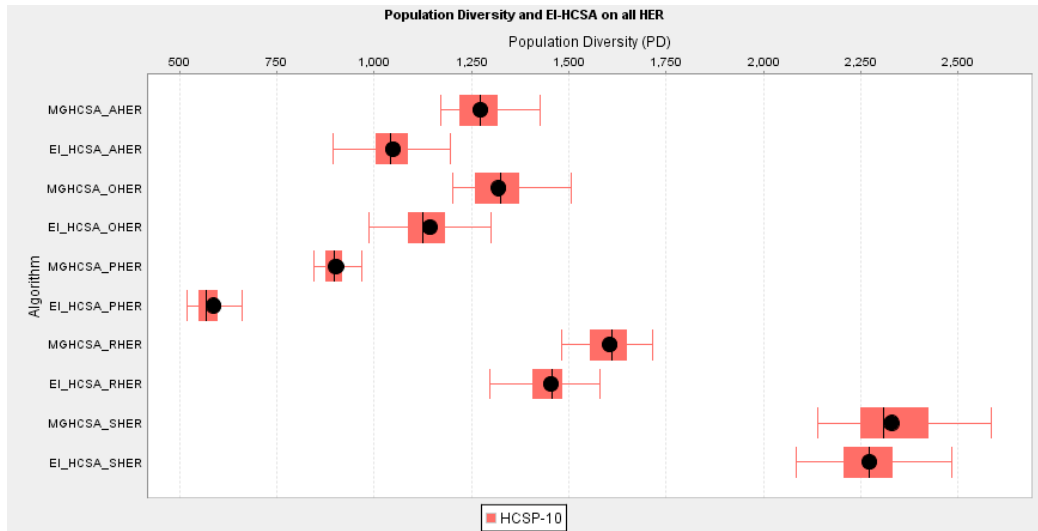


Figure 6.15: Box-and-whisker plot of Population Diversity (PD) across all HER for the EI-HCSA study.

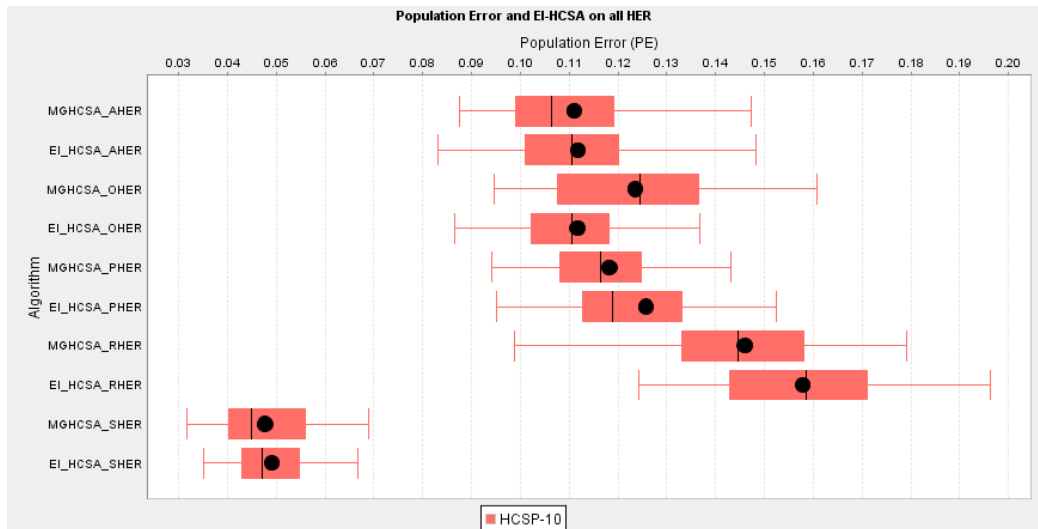


Figure 6.16: Box-and-whisker plot of Population Error (PE) across all HER for the EI-HCSA study.

Analysis

This section provides an analysis of the results from the empirical study into the MI-HCSA summarised in Table A.9. These analyses exploit the trends and expectations outlined in Section 6.4.4 regarding information dissemination between the hosts in the population.

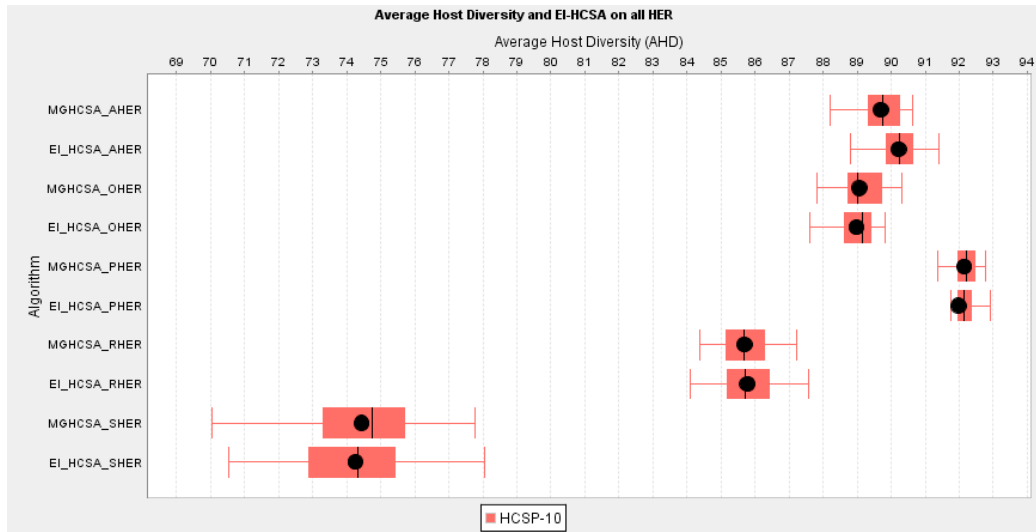


Figure 6.17: Box-and-whisker plot of Average Host Diversity (AHD) across all HER for the EI-HCSA study.

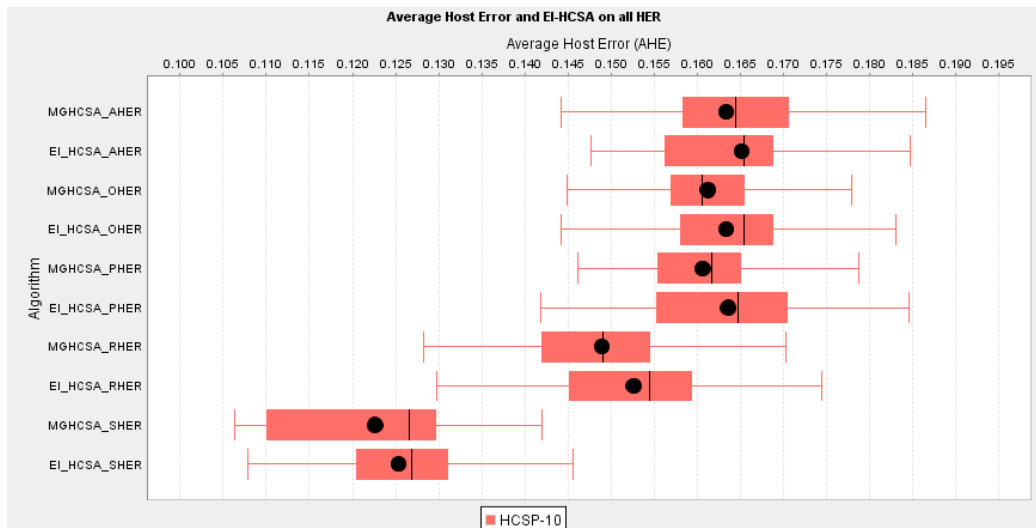


Figure 6.18: Box-and-whisker plot of Average Host Error (AHE) across all HER for the EI-HCSA study.

Generational Trends This section compares and contrasts the evolved immunity algorithm against the minimal generational approach. From a system perspective the evolved immunity algorithm generally exhibited a decrease in population diversity compared to the minimal generational algorithm, although the trend was not significant on the symmetric exposure regime. A difference in population error between the approaches was only significant on OHER and RHER, where a decrease and increase were observed respectively. From a component perspective there was practically no significant difference between EI-HCSA and MG-HCSA with regard to Average Host Diversity and Error, other than a small increase in diversity in the evolved approach on AHER. The results demonstrate little significant difference between the two approaches other than a general decrease in population diversity and some changes to system error. The change in population diversity provides confirmation that the evolutionary mechanisms were providing a convergence effect, resulting in increased homogeneity between the hosts in the population.

1. Generally there is little difference between inherited immunity and independent generations for the chosen configuration.
2. Decreased population diversity may suggest at an evolutionary effect that promotes a population of homogeneous genetic material (evolutionary convergence).

Conclusions

Two considerations that may have resulted in the lack of significance between the two approaches include (1) the amount of evolution permitted given the complexity of the problem, and (2) the noisiness of the fitness function. In the first case the configuration permitted 100 evolutionary generations each with 100 epochs providing a testing ground for the host genetic basis. Given that this genetic basis consisted of 9600 bits, it is more than likely that an increase in the number of generations (such as a factor of 10 or 100) would result in a significantly increased evolutionary effect toward a homogeneous host genetic footprint. In the second case, the objective function used (genetic system competence via host error) may have proven too noisy given the reduced amount of interaction with the problem (100 environment exposures). In addition, such an objective, although desirable may have been convoluted by the asymmetry of information exposure by some of the exposure regimes providing a potentially second level of noise. This may be tested by again increasing the number of generations and/or the number of epochs per generation, as well as testing on simpler symmetrical and deterministic exposure regimes.

The proposed inherited immunity approach aligned with the general progression in complexity with the proposed host clonal selection algorithms, although in the context of the potential of combining abstractions of evolution and the acquired immune system was very simplistic. The principle of combining the approaches begs for elaborated investigation.

6.9 Chapter Summary

6.9.1 Paradigm Review

The *Host Clonal Selection Paradigm (HCSP)* was defined as the investigation of the Tissue and Cellular Clonal Selection Paradigms as constrained by (1) multiple holistic immune systems called *Hosts* and their interactions, and (2) the concerns of regimes of discrete host exposures of information from an antigenic environment called *Host Exposure Regimes (HER)*. The paradigm exploits the metaphor of a collective of holistic acquired immune systems called a *Population* and the turn-over of populations called *Generations*. Intra-population communication is mediated by the hosts themselves as well as by pathogen that may spread through host-to-host contact. Intra-generational communication is mediated by parent-progeny relationships both explicitly and implicitly via an evolutionary process. The principle information processing interest of the paradigm as gleaned from the metaphor are the information management strategies that may be employed to address the known and unknown regularities and irregularities in an antigenic *Environment* of information called an *Habitat Antigenic Exposure Problem (HAEP)*. The concern of such strategies is the effective intra-population and intra-generational dissemination of information between spatially isolated hosts and temporally isolated populations toward improved host-wise and population-wise capabilities whilst maintaining the intrinsic localisation of the adaptive units independence.

6.9.2 Principles and Findings

The following summarises the important principles and findings from the definition and investigations into the Host Clonal Selection Paradigm:

General Principles

1. *Host-mediated dissemination*: Unlike the Tissue Paradigm that provides a natural intra-component communication channel, the focus of the Host Paradigm is the development of information strategies to promote the controlled dissemination of acquired information toward improved population capability.
2. *Strategies based on host-level immunology*: Population immunology such as immunisation and evolutionary immunology inspire information strategies for acquiring, localisation, and disseminating information in an unknown environment as the survival of a population of organisms with immune systems rely on such strategies for survival.

Specific Findings

1. *Transmission*: Elicited immunity via pathogen transmission exhibits an information dissemination effect on exposure regimes with an asymmetric information distribution with little difference in effect whether immunity is elicited by random-pairing based pathogen transmission or one-to-many based vaccination.

2. *Shared Immunity*: Shared immunity via effector transmission results in a disruptive effect at the system and component level with large numbers of shared effectors, which is minimised with small scale intra-population sharing maximises the dissemination effect at the component level.
3. *Maternal Immunity*: Generational transmission via maternal immunity in parent-child relationships results in an inter-generational localisation method exploiting the generational homogeneity promoted by the mechanism.
4. *Evolved Immunity*: Evolved immunity via generational inheritance demonstrated comparatively little difference over generational restarts other than a suggestion at evolutionary convergence toward a consistent genetic footprint, suggesting at the potential of the under investigated combination of generational learning via evolution with lifetime learning via the acquired immune system.

6.9.3 Integration

The Tissue Clonal Selection Paradigm in Chapter 5 took the Cellular Paradigm from Chapter 4 for granted, using it as a principle component in the Tissue Clonal Selection Algorithm, focusing on how different decentralised information management strategies effected the organisation acquisition and use of information under different information exposure regimes. In the same manner the Host Paradigm in this chapter took the Tissue Paradigm for granted, subsuming the scope of inter-tissue interaction concerns and using the system in its entirety as a principle component (Host) in a the Host Clonal Selection Algorithm. The population provided the first abstraction of multiple hosts, concerned with decentralised information dissemination strategies, whereas the generational-population abstraction subsumed the concerns of a given population, highlighting the related but different concerns of inter-generational decentralised information dissemination strategies. These two abstractions represent the pinnacle in the exploration of the clonal selection paradigm presented in this work. The following chapter considers the aggregation of all three paradigms (cellular, tissue, and host) in a unifying framework that provides both explanatory power for presented algorithms, as well as predictive power in highlighting unrealised algorithms and other facets of the considered abstractions. The abstract host paradigm is grounded in Chapter 8 in the context of two general problem domains, highlighting the potential benefits the tissue algorithms provide over the cellular clonal selection algorithms.

Chapter 7

Hierarchical Frameworks

7.1 Chapter Overview

The previous chapters outlined three distinct perspectives on the clonal selection adaptive strategy phrased as a *Cellular* (Chapter 4), *Tissue* (Chapter 5), and *Host* (Chapter 6) Clonal Selection Paradigms. This chapter considers all three paradigms together and aggregates the models, algorithms, and problems into a series of hierarchical frameworks. Specifically, Section 7.2 integrates the clonal selection models and algorithms, Section 7.3 integrates the antigenic exposure paradigm, and Section 7.4 combines the clonal selection and antigenic exposure frameworks into an integrated hierarchical clonal selection framework. The frameworks provide both explanatory tools for the design decisions and limitations of the investigated algorithms, as well as predictive power as to organisations and configurations that may reveal new and interesting findings regarding the underlying adaptive strategy. Finally, Section 7.5 considers the methodology behind the framework, the application of the framework to the broader field of AIS, and the general applicability of the methodology to related computational intelligence sub-fields.

7.2 Clonal Selection Framework

7.2.1 Paradigms

The three paradigms considered in this work provide tangible contributions to the field of Artificial Immune Systems, including: (1) varied perspectives of clonal selection as an adaptive strategy, and (2) algorithms within those perspectives inspired by the computational and information processing attributes of the clonal selection theory. The perspectives of clonal selection were selected to promote a natural hierarchical relationship where the tier that followed encompassed the concerns of the tier that proceeded. This section considers the integration of the clonal selection models and algorithms from the *Cellular*, *Tissue*, and *Host* Clonal Selection Paradigms into a *Hierarchical Clonal Selection Framework (HCSF)*. The focus of this framework is on the explicit integration of the information processing concerns of the each of the paradigms as abstractions of the clonal selection theory at difference scales and scopes of the acquired immune system. This integration

facilitates (1) inter-paradigm contrast and comparison, and (2) analysis and reasoning on the framework providing both explanatory and predictive capabilities regarding abstractions of the clonal selection theory. The focus of the framework is on the systems side of the system-environment relationship, and as such the concerns of the antigenic exposure paradigm may be compressed to a generic cross-paradigm antigen exposure problem (see Figure 7.1). This simplification of the environment directs focus on the elaboration, and inter-relationships on the systems.

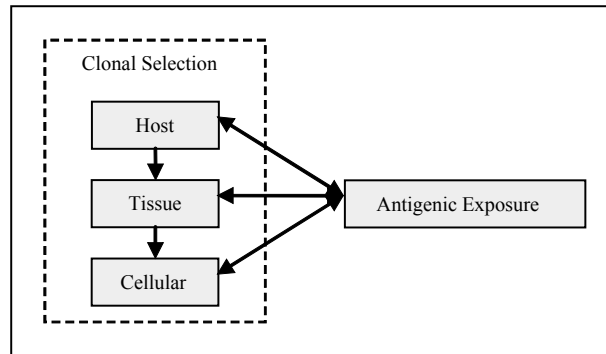


Figure 7.1: Depiction of the focus on systems and the compression of the environment in the Hierarchical Clonal Selection Framework (HCSF).

7.2.2 Comparison

This section directly compares the three paradigms beyond the obvious and superficial concerns of the perspectives and the algorithms within. Toward this end, the adaptive systems formalism from Section 3.5.1 is used as a basis to motivate the comparison. This formalism was selected because it explicitly provides a tool for considering the similarities and differences between adaptive systems (paradigms), and importantly between strategies (algorithms within a paradigm). The primary objects of the adaptive systems formalism provide a conventional perspective on the components of an adaptive system (see Table 3.5). From this high-level perspective the subtleties of the three paradigms are lost and one may consider the generalities of the clonal selection adaptive strategy, specifically: the antigenic environment (e), the clonal selection strategy in cells (s), and the affinity of cells to antigen (U). This consideration compresses the hierarchy to the essential concerns of the adaptive strategy (for example see Figure 7.2). This is an important perspective, as (1) it highlights the important primitives required in all so-called ‘Clonal Selection Algorithms’, (2) provides a context for comparing the general approach to other adaptive strategies such as genetic adaptive plans (the basis of genetic algorithms), and (3) provides a context for relating and integrating the general approach to other Artificial Immune System approaches (for example clonal selection as a sub-strategy in Negative Selection and Immune Network algorithms).

The secondary objects promote the consideration of the subtler details of an adaptive system, providing the detail of the formalism (see Table 3.6). Table 7.1 provides an interpretation of the three paradigms in the context of the secondary objects of the formalism

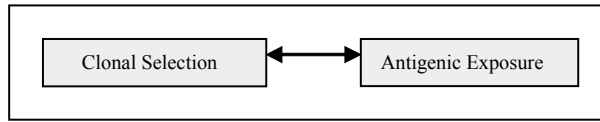


Figure 7.2: Depiction of the compression of both the system and the environment when the HCSF is compressed to the primary objects of the adaptive system formalism.

specifically motivated by Holland’s fundamental questions listed in Table 3.7.

	Cellular	Tissue	Host
<i>A</i>	Cell	Tissue	Host
<i>E</i>	Antigen	Infections	Habitats
<i>O</i>	Interaction of cells	Movement and capacity of cells	Movement of antigen and cells
<i>S</i>	Single repertoire	Discrete tissue and constrained connectivity	Discrete Hosts
<i>X</i>	Composition and capability	Acquisition and anticipation of need	Acquisition and anticipation of need
<i>I</i>	Selection	Spatial-temporal selection	Spatial-temporal selection
<i>M</i>	Proportional resource allocation	Spatial organisation	Spatial organisation

Table 7.1: Mapping of the cellular, tissue, and host clonal selection perspectives onto the secondary objects of the adaptive systems formalism.

The naming convention adopted for the three tiers highlights the structures or units under adaptation within each tier. This naming convention is mirrored in the areas of the environment to which each structure is adapting. The consideration of the adaptive structures highlights the already apparent subsumption of cellular concerns by the tissue tier, and the tissue concerns by the host concerns. Although discussed throughout the investigation of the paradigms, the implications of this point are less intuitive. The consideration of a tissue as a structure under adaptation requires a consideration of a repertoire of cells in its entirety as a unified structure. Adaptation of a tissue structure ($T \rightarrow T'$) may be considered to occur in response each infection exposure (aggregation of antigenic exposures). This transition is less concise when one considers continuous movement of cells via recirculation as a second transformation operator. In this context, the ‘structures’ of a single Host (tissue system) are considered in a constant state of constrained change that is distinctly different from the adaptation of tissues provided by infection exposure. Likewise, a similar state exists at the Host-level with the adaptive transition of hosts ($H \rightarrow H'$), the sharing of acquired information between adaptive units. One may generalise this transition to say that adaptation of structures occurs with each change at the finest level. This suggests new tissue and host structures are created with each shared piece of information and each clonal response to antigenic exposures. The scope (level of abstraction) of each perspective provides a frame of reference in which to choose when to aggregate such fine-grained changes and delineate or not (continuous) structure

transitions. An elaboration of the concerns of each tier may consider the subsumption of degenerate cells subsumed by the cellular tier, and the population subsumed by the generational variation of the host tier. Although the investigation of these concerns was tied to the cellular and host paradigms respectively, further elaboration of the sub-cellular and generational clonal selection algorithms may result in new perspectives with as much potential as the canonical three presented in this work (this consideration is elaborated in Section 7.4.2).

The comparison of strategies within each perspective (X) is interesting as both the tissue and host tiers are concerned with the effective acquisition of information in the context of spatial and temporal anticipation of need. This assessment is based on historic spatial-temporal selection ultimately of cells in infection and habitat exposures across the distributed components of each class of system (I). Importantly, this information and assessment motivate the strategies in each of these tier to (1) localise acquired knowledge for specialised spatial-temporal anticipation, (2) dissemination acquired knowledge for generalised system-level anticipation. Such organisation of information is directly reflected across the aggregation of the units of adaptation at each tier and thus may be considered a secondary memory (M). The spatial concerns of these two tiers is compressed in the cellular tier to that of the historic and thus anticipated temporal-properties of exposures. As with the other tiers, the aggregation of the structures under adaptation facilitate an additional super-structure (repertoire) in which the organisation, in this case proportional resource allocation under varied constraints, differentiate the strategies within the tier.

7.2.3 Bracketing Analysis

An important principle in defining the three perspectives of clonal selection was highlighting the specific constraints imposed at each level. This principle is referred to in this work as the *bracketing of concerns* after the suggestion by Goldberg in his integrated methodology for investigating adaptive systems (see section 2.7.3). Specifically, the bracketing of investigated systems occurred in two ways which are discussed in this section: (1) the bracketing of systems by the cardinality of structures under adaptation, and (2) the bracketing of systems by the interactions of managed structures under adaptation.

Component Cardinality

The hierarchical relationship between the tiers results in the adaptive units at one tier being comprised of an aggregation of the adaptive units from the previous tier. This section considers the cardinality in terms of 1 and many (N), and the bracketing of these components within specific algorithms. Table 7.2 summarises the principle algorithms from across the three tiers and their component cardinality in the full context of the previously defined Cells (C), Tissues (T), Hosts (H), and Populations (P). An additional component is defined below the cellular level called a sub-cellular component (c) that is used to differentiate a degenerate component in the context of degenerate cellular algorithms.

The table reveals much about (1) the specific component cardinality selected for the investigated classes of algorithms, and more importantly (2) suggests at the possibility

Paradigm	Algorithm	c	C	T	H	P
<i>Cellular</i>	DCCSA	N	1	1	1	1
	CCSA	N	N	1	1	1
	MCCSA	1	1	1	1	1
	*CCSA	1	N	1	1	1
<i>Tissue</i>	TCSA	N	N	N	1	1
	*TCSA	1	N	N	1	1
<i>Host</i>	PHCSA	N	N	N	N	1
	*PHCSA	1	N	1	N	1
	GHCSA	N	N	N	N	N
	*GHCSA	1	N	1	N	N

Table 7.2: Summary of the component cardinality across the principle abstract and realised clonal selection algorithms, where * represents a wild-card for descendant algorithms of a given base algorithm.

space of viable permutations of component cardinality across the hierarchical clonal selection framework. From a cellular perspective the Table clearly shows the trade-off in cardinality of the simplified minimal algorithm as a realisation of a (1+1) MHCA compared to degenerate and cellular algorithms. The explicit aggregation of degenerate components c into holistic cells for assessment is distinct compared to the less restricted CCSA, and the bracketed range of assessed cellular algorithms in Chapter 4. Strictly, the c was bracketed at $c = 3$ for the three colour components for investigated cellular algorithms and cellular primitives in later tiers. The table records this as $c = 1$ as there was no choice for selecting among the three c for each cell, the ordering (selection or aggregation) was specifically defined and fixed, therefore treating all three components as a single component. The investigation of tissue and host algorithms clearly show this same choice in bracketing, enforcing the aggregation and thus holistic adaptation of colour space patterns. The algorithms of the host tier also show an important bracketing of the tissue-level components. This choice of bracketing was made explicit to remove the spatial acquisition and anticipation of information effects that are the concern of the tissue paradigm, and compress them to the simpler terms of the cellular paradigm. This is reflected in the cardinality table, clearly showing each host in the population of investigated host algorithms as a single tissue repertoire of cells. An artefact of this choice in bracketing is that the investigation of host algorithms did not reveal any information on the compounded concurrent effects of spatial acquisition and anticipation of information in a population and within individual hosts.

This last point highlights the reason for bracketing as a tool for investigating adaptive systems, specifically to bound the complexity of an adaptive model to the areas of interest for an investigation. From this perspective, one may consider the cardinality of all algorithms irrespective of their categorising scope and consider the investigation clonal selection over the last three chapters as the relaxation of cardinality from the sub-cellular to the population level. This is evident if one considers the shift in the N (the shift in focus) from the left to the right of adaptive structures as the list is traversed top-to-bottom in Table 7.2. The power of bracketing the cardinality of the system is apparent when

one considers the alternative of investigating clonal selection at the three selected levels of abstraction from the perspective of a population or generational host clonal selection algorithm (see Figure 7.3). Such unbounded cardinality would likely make assessing principle behaviours difficult given the confounded effects of the delegation of responsibility through the hierarchy. This final point suggests at the focus of the Hierarchical Clonal Selection Framework, which is: *the investigation and predictable application of an integrated hierarchical system motivated by the emergent effects of a bottom-up clonal selection adaptive strategy.*

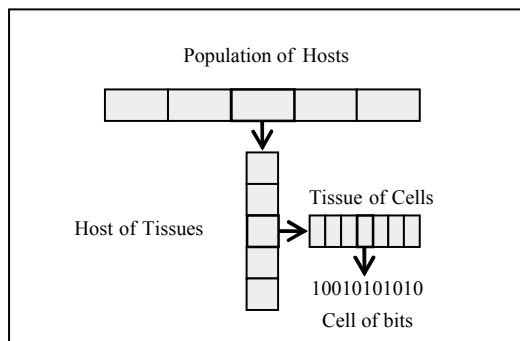


Figure 7.3: Depiction of an expanded realisation of a hierarchical clonal selection system with all three tiers.

Component Interaction

This section provides a complement to the analysis of cardinality bracketing by considering the bracketing of the interactions of the structures under adaptation across the investigated paradigms and algorithms. Interaction refers to direct competition or pattern recognition between structures (such as in the cellular paradigm), or the communication of information between structures (such as in the tissue and host paradigms). Bracketing of interaction refers to the specific constraints imposed on the interaction between structures at different scales within clonal selection algorithms. Table 7.3 provides a summary of which components interact across the algorithms of the three paradigms. The general algorithms for each paradigm, specifically CCSA, DCCSA, TCSA, PHCSA, and GHCSA explicitly list interactions for all structures, whereas specialised algorithms list interactions only for relevant structures.

Naturally, the suggested and adopted bracketing of interactions mirrors the bracketing of cardinality. Restricting cardinality was designed to directly restrict the effects of interaction, and the elaboration of the cardinality was designed to directly provide a context for investigation. An important point implicit in the bracketing of cardinality and explicit in the interactions is that cells and their interactions are required (almost) universally by all clonal selection algorithms, with the DCCSA as the single exception. This is because the structures under adaptation at higher levels of abstraction ultimately all subsume cells, specifically receptors on and excreted by cells. Beyond the verification of the primary objects of the adaptive systems formalism discussed in Section 7.2.2, this reinforces the

Paradigm	Algorithm	c	C	T	H	P
<i>Cellular</i>	DCCSA	yes	no	no	no	no
	CCSA	yes	yes	no	no	no
	MCCSA	no	no			
	*CCSA	no	yes			
<i>Tissue</i>	TCSA	yes	yes	yes	no	no
	MTCSA		yes	no		
	*TCSA		yes	yes		
<i>Host</i>	PHCSA	yes	yes	yes	yes	no
	MP-HCSA		yes	no	no	
	*PHCSA		yes	no	yes	
	GHCSA	yes	yes	yes	yes	yes
	MG-HCSA		yes	no	no	no
	*GHCSA		yes	no	no	yes

Table 7.3: Summary of the component interaction across the abstract and realised clonal selection algorithms, where * represents a wild-card for descendant algorithms of a given base algorithm.

fact that both the clonal selection theory and the inspired computational principles are defined in a cellular context. The implications of this fact reinforces the ‘perspectives on the clonal selection adaptive strategy’ consideration of the three investigated paradigms in this work. The interactions between structures was defined in Section 7.2.2 as the operators providing the defining characteristic between strategies within each paradigm. This is highlighted in the naming convention adopted where the *Minimum* of each algorithm class is defined with no interaction between the structures under adaptation, providing a baseline for comparison of the effects of such interactions.

7.2.4 Summary

The integration and analysis of the three clonal selection paradigms with a compressed antigenic exposure paradigm in the HCSF provided a context for both (1) retrospectively explaining generalised design and methodology principles for clonal selection algorithms, and (2) promoting consideration of permutations of cardinality and interaction that although are predicted by the framework were not investigated in this work. This section reviews the explanatory and predictive power of the ‘system biased’ Hierarchical Clonal Selection Framework, as follows:

1. *Explanatory Power*

- (a) *Cellular Basis*: Irrespective of the specific perspective taken (level of abstraction), the clonal selection theory and computational principles is a bottom-up pattern recognition-based adaptive strategy involving cells (receptors or detectors) and antigen (complementary structures).
- (b) *Bracketing*: The forward and backward bracketing of the cardinality in the context of a hierarchical complex adaptive system has been demonstrated to provide a powerful tool for investigating such a system.

- (c) *Naming Conventions*: A perspective invariant naming convention promotes the naming of the perspective for the structures under adaptation, and the design of a minimal strategy of bracketed interactions between such structures as a baseline for measure the effects of such interactions.

2. *Predictive Power*

- (a) *Varied Bracketing*: An analysis of the specific bracketing considered in the investigation highlighted the bracketing that was not considered, as well as the removal of bracketing all together to consider a broader hierarchical clonal selection system, providing a rich basis for future investigation.
- (b) *Compound Effects*: The investigation and application of an unbracketed clonal hierarchical selection based system provides an opportunity for identifying and assessing the compound and expected non-linear aggregate effects of cross-component and cross-tier interactions.
- (c) *Alternative Perspectives*: The consideration of the three paradigms in aggregation provided a context to consider (1) the potential splitting of the Cellular and Host paradigms into new and/or sub-paradigms, and (2) suggested at the potential for perspectives of clonal selection not yet considered.

7.3 Antigenic Exposure Framework

7.3.1 Paradigms

The antigenic exposure paradigm was defined as a complement to the clonal selection strategy, and elaborated in scale along with the adaptive systems across the three paradigms. Although the clonal selection abstractions had a basis in reviewed immunophysiology, the antigenic environments were contrived from the general properties of endogenous and exogenous antigen. The reason for this was that although an antigenic environment is required to stimulate a clonal selection system, it was not the focus of the investigation. This section integrates all of the scales of the antigenic exposure paradigm investigated in this work from across the three paradigms, specifically: antigen (and determinants), antigenic infections, and antigenic habitats. The integration of these concerns is referred to as the *Hierarchical Antigenic Exposure Framework (HAEF)*. As with the hierarchical clonal selection framework, the focus of this integrated framework is on the explanatory and predictive power provided via an analysis of the constraints imposed (referred to as bracketing) at each level in the hierarchy. Unlike the clonal selection framework, this section is not concerned with comparing and contrasting the tiers in the context of an adaptive systems formalism. The reason for this is that the antigenic environment is not an adaptive system, rather a required component of the clonal selection adaptive system. As such, the perspective on the antigenic environments is elaborated, compressing the concerns of clonal selection to provide an inverse perspective as to that considered in the integration of clonal selection (Figure 7.4)

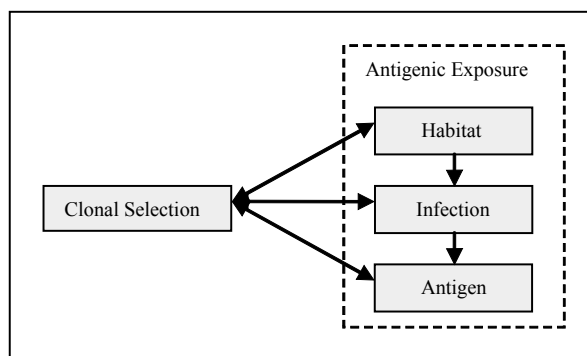


Figure 7.4: Depiction of the focus on environment and the compression of the systems in the Hierarchical Antigenic Exposure Framework (HAEF).

7.3.2 Bracketing Analysis

Bracketing in the antigenic exposure paradigm involved constraints imposed on the cardinality of the *triggers for adaptation* and the regimes that governed the piece-wise exposure of environmental information to a given system. This section considers the bracketing of configurations imposed on these two antigenic environment concerns.

Component Cardinality

The bracketing of the cardinality of the information in the environment was configured to match the scale of the units of adaptation in the adaptive systems. Unlike clonal selection, the antigenic environment was not motivated in its interaction with the systems (for example adversarial, dynamic, and/or adaptive). The neutrality of the environment was intentional to focus attention on the emergent system behaviour under low-complexity environment concerns. Table 7.4 summarises the cardinality of the abstract antigenic environments and the realised colour space antigenic environments in terms of the previously defined: determinants (D), antigen (A), infections (I), habitats (B) and environments (E).

Paradigm	Problem	D	A	I	B	E
<i>Cellular</i>	AEP	N	N	1	1	1
	ACSP	1	N	1	1	1
	DCSP	N	N	1	1	1
<i>Tissue</i>	IAEP	N	N	N	1	1
	ICSP	1	1	N	1	1
<i>Host</i>	HAEP	N	N	N	N	1
	HCSP	1	1	1	N	1

Table 7.4: Summary of the component cardinality across the principle abstract and realised antigenic exposure problems.

The table shows the potential of each abstract exposure problem and the strong bracketing of the colour space realisations for the structures under adaptation in the systems in each tier of the hierarchy. The ACSP, ICSP, and HCSP all exploited the same bracketing

method, specifically the treating of each Colour Space Pattern (colour) as a trigger for adaptation at the respective tier. For example, a set of antigen (ACSP) is equivalent to a set of infections (ICSP) and a set of habitats (HCSP), in that they can all be reduced to a set of Colour Space Patterns. This equivalence in problem was designed to minimise the complexity of the information to be acquired across the tiers to a level that was easy to interpret, whilst introducing complexity the way in which the information was exposed and expected through varied exposure regimes. This simplification of problem motivated the corresponding cross-cutting cellular-based clonal selection concerns in the systems by reducing all information to single antigen. One may consider the alternative of such a minimisation in a less constrained colour space realisations of the the exposure problems. For example, an infection of antigen may involve sets of antigen that are similar (intra-infection), although distinctly different between the sets (inter-infection). This could be realised through minor variations for intra-infection colour space patterns (similar colour), and major variations for inter-infection Colour Space Pattern centroids (inter-infection). This same example scales to the habitat level, with sets of sets of antigen, where intra-habitat infections may be more similar than inter-habitat infections. These examples highlight what the bracketed cardinality provided, specifically exemplar patterns (set size of one with no variations) for each infection and habitat with promoted inter-trigger dissimilarity (minimum hamming distance). This chosen level of cardinality bracketing had the effect of removing the degeneracy, (specifically the redundancy of the triggers of adaptation) requiring an immune response for each antigenic trigger in the antigenic environments. The alternatives provide an opportunity for re-introducing such degeneracy although at the cost of the specificity of a system in an environment (generalised response against variations for each antigenic determinant at the lowest level).

Exposure Regimes

Antigenic exposure controlled the piece-wise revelation of information to the systems as well as the piece-wise spatial and temporal expectation of future exposure. In the context of forced response to a fixed information environment, the exposure regimes managed the complexity of the problems to be addressed by the systems. The complexity of exposure scaled with the increase in the level of abstraction starting with exposures, multiple exposures and multiple antigen in the definition of the exposure paradigm (Section 4.2.2), and the raise in the level of abstraction of such exposures to include spatial concerns across multiple points of exposure in infection exposures (Section 5.3.2) and habitat exposures (Section 6.3.2). This section considers the bracketing of exposures with regard to (1) the spatial points of exposure, and (2) the temporal concerns of the expectation of exposure.

Table 7.5 considers the bracketing of the spatial concerns of an exposure, in particular the triggers for adaptation in terms of the scope of the points of exposure in a system (1 or many). A point of exposure in a system is unit of adaptation, therefore given the chosen cardinality bracketing in both systems and environments, there is always a many-to-many relationship between the *units* and *triggers* of adaptation at the focused level of detail in the hierarchy. At the antigen level, the scope of the system (all points of exposure) is

Paradigm	Problem	D	A	I	B	E
<i>Cellular</i>	AEP	N	N	1	1	1
	ACSP	1	N	1	1	1
	DCSP	N	N	1	1	1
<i>Tissue</i>	IAEP	N	N	1-N	1	1
	ICSP	1	N	1-N	1	1
<i>Host</i>	HAEP	N	N	1	1-N	1
	HCSP	1	N	1	1-N	1

Table 7.5: Summary of the bracketing of spatial exposure across the abstract and realised antigenic exposure problems.

always exposed. This has the effect of requiring a mechanism to constrain a polyclonal activation (shown to be strong selection in Section 4.4.3). The reason for this design decision was the relative small scale of a single repertoire of cells in the context of multiple of such repertoires (tissues) and multiple such hosts of tissues. The same strong selection was required at the determinant level, and given the antigen focus at all levels, this had the effect of requiring explicit aggregation of response toward holistic solution to address the resultant polyclonal response. This unbounded antigen-cell exposure and explicit strong selection and aggregation is bracketed in the tissue and host levels in the introduction of discrete Antigenic Exposure Regimes (AER) providing deterministic and probabilistic patterns of interaction between the units and triggers of adaptation (referred to as $1 - N$ in the table). This had the effect of requiring the systems under such regimes to manage the spatial organisation (localisation and dissemination) of acquired information to best address the patterned interaction. One may consider the alternatives of such bracketing of spatial exposure, for example the spatial selection of cells in a repertoire considered in spatial up-front selection in Section 4.5. The symmetrical exposure regime at the tissue (STER) and at the host (SHER) levels provide a counter example of the effects of simplifying exposure complexity to that of system-wide with respect to the points of exposure of a system.

Paradigm	Problem	D	A	I	B	E
<i>Cellular</i>	AEP	N	N	1	1	1
	ACSP	N	N	1	1	1
	DCSP	N	N	1	1	1
<i>Tissue</i>	IAEP	N	N	N	1	1
	ICSP	N	N	N	1	1
<i>Host</i>	HAEP	N	N	N	N	1
	HCSP	N	N	N	N	1

Table 7.6: Summary of the bracketing of temporal exposure across the principle abstract and realised antigenic exposure problems.

Spatial exposures without a time component are a compression of the patterns of exposure over the scope of the units of adaptation. Therefore, temporal exposures can be considered an opposite case of the compression of spatial location and a focus on

the systems expectation of information on a timeline. Table 7.6 summarises the temporal expectations of the abstract and realised exposure problems with regard to the triggers for adaptation. The table clearly shows that the temporal expectations of exposure extended backward through the hierarchy from the focus of each tier. This simple but important perspective on the antigenic exposure paradigm highlights the motivating complexity for the corresponding Hierarchical Clonal Selection Framework, specifically the compounded increase in temporal expectation of information. This table also highlights the need for bracketing the information cardinality of antigenic environments across the hierarchy to bound the complexity of the effects of the compounded temporal expectations.

7.3.3 Summary

The integration and analysis of the antigenic exposure paradigm across the three levels of abstraction in the context of a compressed clonal selection system provided a context for both (1) retrospectively explaining generalised design and methodology principles for the of the problems that motivated the investigation of clonal selection algorithms, and (2) promoted the consideration of permutations and relaxation of cardinality, spatial, and temporal concerns of antigenic exposure. This section reviews the explanatory and predictive power of the Hierarchical Antigenic Exposure Framework, as follows:

1. *Explanatory Power*

- (a) *Forced Response*: The strong bracketing of the cardinality of antigen across the hierarchy resulted in a forced response expectation from systems with regard to the trigger for adaptation. Such forced response may be relaxed by loosening the bracketed cardinality constraints at a given tier in the hierarchy, in particular for the infection and habitat levels.
- (b) *Compounded Expectation*: The focus of the investigation using the antigenic exposure paradigm was on compounded expectation with varied spatial scope of exposure at the tier of interest, and strong bracketing of the spatial scope at tiers above and below the tier of interest.

2. *Predictive Power*

- (a) *Cardinality Complexity*: An alternative to the design of clonal selection systems in response to exposure complexity, is that the design of such systems can be motivated by variations in the cardinality complexity instead and/or combinations of cardinality and exposure complexity. Specific examples of cardinality complexity include sets of antigen in each antigenic infection, and sets of infections in each antigenic habitat.
- (b) *Compounded Spatial Scope*: The spatial scope of exposure was constrained at tiers of the hierarchy above and below a focused level, the relaxation of this bracketed spatial exposure will provide an additional axis of problem complexity to explore in addition to compounded expectation.

7.4 Integrated Hierarchical Framework

7.4.1 Frameworks

Given an assessment of integrated hierarchical clonal selection and antigenic exposure frameworks, this section considers the aggregation of the two frameworks towards a unified abstraction and framework for clonal selection in the context of an antigenic environment. This is a natural integration as the abstract and realised problems and algorithms in both sides of the integration were designed as a natural fit for each other. This integration is referred to as the *Integrated Hierarchical Clonal Selection Framework (IHCSA)* to differentiate it from the concerns of either independent framework, and to highlight the strong bottom-up clonal selection focus. As such, the focus of the framework is the hierarchical integration of the elaborated concerns of both the system and the environment, depicted in Figure 7.5

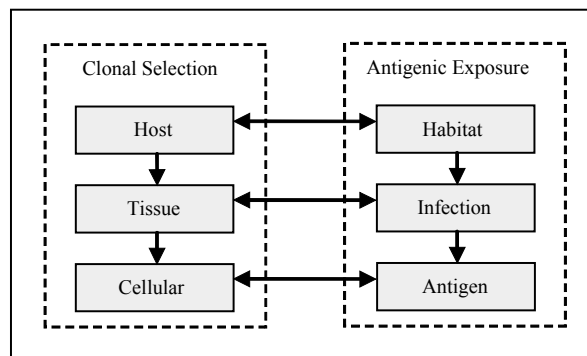


Figure 7.5: Depiction of the focus on system and environment in the Integrated Hierarchical Clonal Selection Framework (IHCSF).

7.4.2 Bracketing Analysis

This section focuses on the bracketing of concerns that have an effect both on the system and environment hierarchies. Specifically, this section is concerned with the matching or integrated cardinality across the hierarchy, and the effect of variations of this matching. Also considered in this section is the extension of the framework through the integration of additional tiers to the existing framework.

Integrated Cardinality

The focus of this subsection is the bracketing of the cardinality of both the system and environment hierarchies as they pertain to the interaction between the hierarchies. Specifically, the interaction of the triggers for adaptation with the structures under adaptation at both the scope of their intended focus (for example cells and antigen in the cellular level), and scopes above and below that focus.

Table 7.7 summarises the intended or so-called ideal integrated cardinality cases for the environmental triggers and system units of adaptation. The table clearly shows the

	<i>c</i>	<i>C</i>	<i>T</i>	<i>H</i>	<i>P</i>
<i>D</i>	1	N	N	N	N
<i>A</i>		1	N	N	N
<i>I</i>			1	N	N
<i>B</i>				1	N
<i>E</i>					1

Table 7.7: Summary of the suitability match between clonal selection and antigenic exposure components

response suitability in terms of the one-to-one relationships (upper limit) and the general capability in terms of the one-to-many relationship of system structures to environmental triggers. It should be pointed out that the upper limit (one-to-one) of the suitability between the frameworks was not investigated throughout this work as an initial assessment may suggest. For example the cellular paradigm considered a ‘*tissue of cells*’ against an ‘*infection of antigen*’ rather than a cell against an antigen or the mismatch case of a ‘*tissue of cells*’ against a single antigen. Table 7.8 elaborates on the ideal cases, and summarises the components of each hierarchy and indicates the scope of concerns of each paradigm, and the cases that were embodied in algorithms that were investigated (bold). Paradigms are marked with the first letter of the structures under adaptation, with the division of Cellular into Degenerate (D) and Cellular (C), and Host into Host (H) and Generational (G). Perspectives or tiers are defined by the interaction of the designated triggers and structures for each paradigm, for example cells (C) and antigen (A) in the cellular paradigm. The italicised cases demonstrate those cases that could be addressed from tiers above and below a given tier.

	c		C		T		H		P	
	1	<i>N</i>	1	<i>N</i>	1	<i>N</i>	1	<i>N</i>	1	<i>N</i>
D	1	D	D							
	<i>N</i>		D	<i>C</i>	<i>C</i>					
A	1		<i>D</i>	<i>C</i>	<i>C</i>					
	<i>N</i>				C	<i>T</i>	<i>T</i>			
I	1			<i>C</i>	<i>T</i>	<i>T</i>				
	<i>N</i>					T	<i>H</i>	<i>H</i>		
B	1					<i>T</i>	<i>H</i>	<i>H</i>		
	<i>N</i>							H	<i>G</i>	<i>G</i>
E	1							<i>H</i>	<i>G</i>	G
	<i>N</i>									<i>G</i>

Table 7.8: Summary of the component cardinality for both the clonal selection and antigenic exposure models.

An important point highlighted by this table are the overlapping concerns between the paradigms. For example *N* cells may be regarded as a single tissue (minimal cardinality tissue model), both of which can address *N* antigen or 1 infection (minimum cardinality infection model). The designed subsumption inherent in both independent and the inte-

grated frameworks promote this *maximum-minimum principle* at the transitions between tiers in the hierarchy. This principle allows similar concerns to be investigated from two different perspectives. For example in the cellular-tissue case, one can investigate the the systems composition and capability of a repertoire in response to a set of N antigen with explicit aggregation of the best matching cells for each antigen from the cellular perspective, or the response capabilities of a tissue against a single antigen from the tissue perspective. Both examples provide a different context (different level of abstraction with regard to the acquired immune system) that may provide varied strategies for considering the same problem. The table clearly shows such cases were investigated using the maximum of the previous tier rather than the minimum of the given tier. The highlighting of investigated cases clearly shows the scope of future work with regard to integrated cardinality cases. An interesting point is that the generational case did not follow the trend and investigate multiple environments with multiple populations. Rather, the EI-HCSA adapted multiple populations against a single antigenic environment (set of habitats). The table also provides a clear indication of those cases delineated to specific scopes of interest that were not investigated, and those cases that could not be investigated. For example, the maximum trigger for a given tier (N) with the minimum structures of adaptation (1) is an infeasible combination defining the upper limit on cardinality complexity for each system.

Figure 7.6 depicts the integrated cardinality bracketing in a hierarchical way to enforce the subsumed responsibility across the tiers, providing a standard representation of the frameworks and basis for extrapolating from the integrated relationships. The figure is information dense capturing intra- and inter-tier relationships for both the system and environment frameworks, and clearly highlighting the important maximum-minimum principle at tier-boundaries (diagonal lines). An interesting observation is that of the relationship between the cells and antigen, showing that a cell and a receptor do not differ in terms of their information content, and that a cells receptor is specialised for a given determinant on an antigen (base principles of the clonal selection theory). This is important given the compression of this relationship (bracketing) of one-antigen-to-one-determinant (and thus one receptor or cell) throughout the majority of the investigations. Only the DCCSA considered the relaxation of this bracketing, allowing cells (receptors) to be selected for by the determinants of an antigen. This specialisation was classified as apart of the cellular paradigm, although if such specialisations are elaborated and investigated (specifically the one-to-many relationship between an antigen and receptors as in the DCCSA in Section 4.4.3), one may refer to this as the *Receptor Clonal Selection Paradigm (RCSP)* to clearly distinguish it from the cellular paradigm. The figure also suggests at tiers above and below the the host and tissue tiers respectively, outside of the scope of what was investigated. These extrapolated additional tiers are considered in the next section.

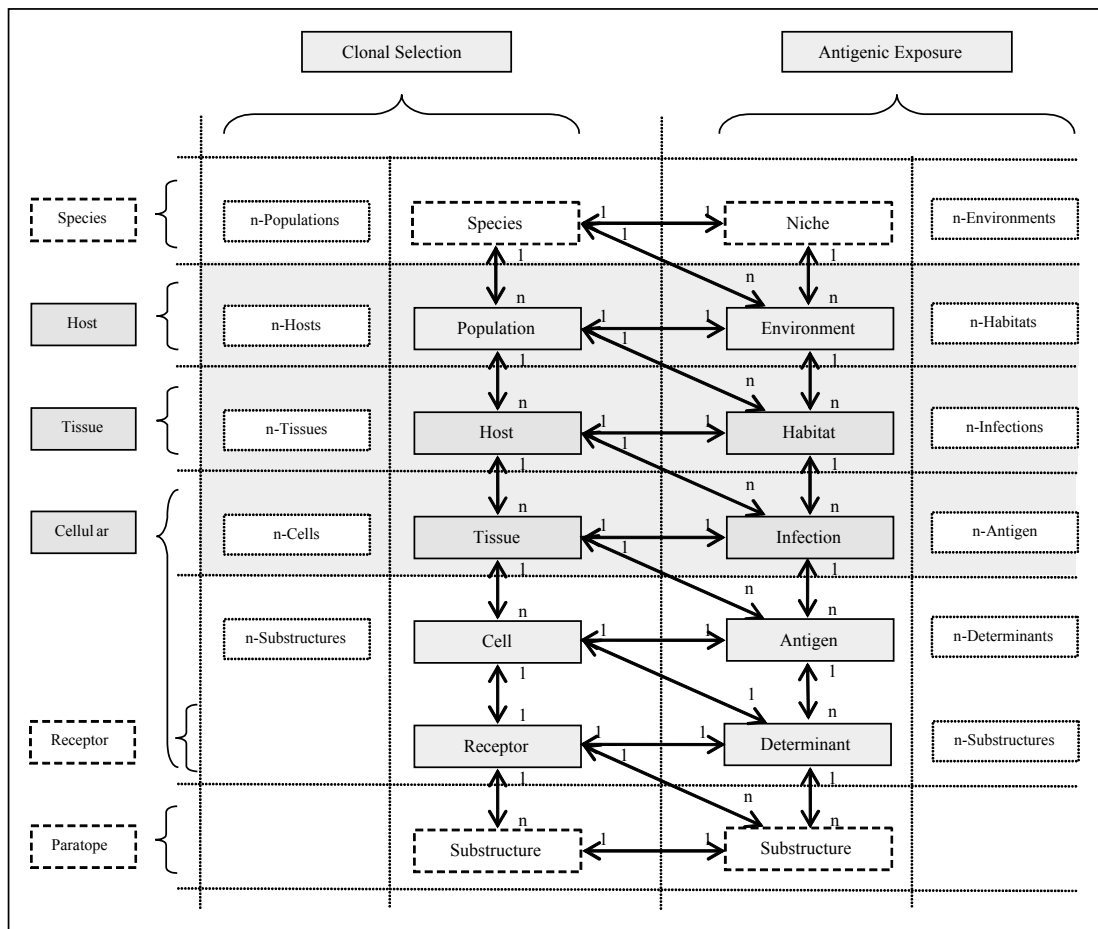


Figure 7.6: Depiction of the integrated hierarchical clonal selection and hierarchical antigenic exposure frameworks, showing subsumed relationships and component cardinalities across the investigated paradigms.

Additional Tiers

A point stressed in the integration of the three investigated clonal selection paradigms in this Chapter has been that each represents a single perspective on the application of clonal selection as an adaptive strategy. Explicitly selected in the standard form depiction of the integration in Figure 7.6 was the capability for additional vertical tiers above and below the host and cellular tiers respectively, as well as the potential for integration of additional perspectives of a given tier. This section extrapolates examples of such additional perspectives on clonal selection and integrates them into the hierarchical framework of such perspectives.

Population Clonal Selection A perspective beyond that of a population of hosts and an environment of habitats may be that of a *species* and an ecological *niche*, where a set of populations interacts with a set of environments called the *Population Clonal Selection Paradigm (PCSP)*. Like the generational variation of the host paradigm, the population paradigm involves maintenance of populations as the structures of adaptation, although the separation between the populations is spatial rather than temporal. As a tier

beyond the host level, it subsumes the concerns of the host level including the generational turn-over, introducing inter-population interaction concerns such as those from population genetics (such as migration), and epidemiology (such as the spread of disease).

Paratope Clonal Selection Given that the clonal selection theory describes the observations of immune cells, a level below the cellular level may not be a reasonable proposition. One may consider the development of a single receptor of sub-structures for a single antigenic determinant of sub-structures as the maximum in complexity of this preceding tier. The scale of the process is that of the tertiary confirmation of the immunoglobulin proteins and related immunochemical forces of their interaction with complementary structures of the determinants on macro-molecules and proteins, governed by the sequence amino acids that make up both molecules. This tier is referred to as the *Paratope Clonal Selection Paradigm (ACSP)*¹ and may be concerned with the specific chemical basis for immunoglobulin synthesis in cells, the valance of immunoglobulin molecules, and the physical forces that govern the molecular interactions and resultant affinity and avidity for determinants.

7.4.3 Summary

This final hierarchical integration represents the pinnacle of the abstractions presented. The integration and analysis of the integration framework provided a context for both (1) retrospectively explaining generalised design and methodology principles that motivated the investigation of clonal selection algorithms, and (2) promoted the consideration of unrealised permutations and relaxation of cardinality and additional tiers of the framework. This section reviews the explanatory and predictive power of the Hierarchical Integrated Clonal Selection Framework, as follows:

1. *Explanatory Power*

- (a) *Framework as a Map*: The integrated framework facilitates the *explicit* specification of immunological and antigenic concerns in terms of their scope of influence relative to each other.
- (b) *Maximum-Minimum*: Tier transitions are defined by the maximum cardinality complexity of one tier that is reduced to the minimum cardinality complexity of the next tier. Such boundaries allow the same information content (problem) to be addressed from the perspectives of the tiers above and below the boundary, revealing insights and alternative solution strategies.
- (c) *Simplest Case*: The investigations considered the average case of each paradigm, specifically the many-to-many relationship between triggers for adaptation and units of selection and adaptation.

2. *Predictive Power*

- (a) *Varied Interpretation*: The backward competence of a paradigm with its subsumed levels suggests at non-contiguous interpreted cases of cardinality and

¹Paratope was chosen as the term for a sub-structure of a receptor molecule for lack of a better term.

interaction along both the system and environmental axis. The realisation of such models facilitate the use of tier-specific mechanisms (algorithms) on backwardly-compatible (subsumed) problem domains.

- (b) *Additional Tiers*: The integration of the frameworks facilitated the elaboration of the cellular and host tiers to include concerns beyond the scope of these two limits of the framework.
- (c) *Additional Mapping*: The extensibility of the framework suggests at the potential for mapping all manner of immunological and antigenic concerns into the appropriate scope of the framework and defining relationships in terms of cardinality and interaction of structures and triggers. For example, the Negative Selection, Danger Theory, and Immune Network Artificial Immune System paradigms.

7.5 Methodology and Hierarchical Framework

The integrated Hierarchical Clonal Selection Framework was not conceived independently of the tiers of the framework, nor were the tiers conceived independently of each other, rather they were systematically co-developed. This section considers (1) an overview of the methodology used to develop and investigate the clonal selection framework, (2) the application of the framework as a template for investigation in the broader field of Artificial Immune Systems, and (3) the framework as a template for investigation in the separate but related Computational Intelligence sub-fields.

7.5.1 Partition-Reduction Methodology

Methodology Overview

In considering the applicability of the chosen hierarchy for clonal selection and the methodology for applying such a methodology, it is critical that the motivation and adopted methodology be clearly elucidated. The general systems theoretic approach used involved three important considerations (1) the distillation of the core computational principles and strategy from the clonal selection theory, (2) the division of the clonal selection concerns into that of *system* and *environment*, and (3) the reduction of the holistic acquired immune system to that of the cellular concerns of the clonal selection theory. These steps are summarised as follows:

1. *Metaphor*: Identify the relevant observations, theory, and abstractions of the selected metaphor.
2. *Strategy*: Distil the metaphor to the core computational principles and information processing strategy.
3. *Partition*: Divide the concerns from the principles and strategy into a framework of *systems* and *environments*.

4. *Reduction*: Relate the components and their interactions at the lowest and highest levels (within scope of the metaphor).

The distillation of clonal selection theory resulted in an adaptive strategy concerned with information acquisition, the concerns of which were partitioned into the *immune system* and the *antigenic environment*. The reduction from immune system to the cellular concerns of the theory focused on the three different and inter-related *functional organisations* of cells under clonal selection: the simplest case of antigens and cells in a small group (*cellular*), many groups of cells using structural concerns to influence antigenic interactions (*tissue*), and groups of immune systems influencing the transmission of disease and evolving (*host*). The selection of the levels in the reduction was also generally motivated based on the *units of selection* consideration from selectionist theory, providing a cellular basis in the reduction given that the lymphocyte is the unit of selection and adaptation in the clonal selection theory. Importantly, the general functional reduction was taken a step further by considering the cardinality of the units of selection and their interaction within each tier, and focusing on the limits of each.

Relation to Methodologies

The adopted methodology is comprised of concerns of the three reviewed approaches for investigating biologically inspired computation, immune information processing, and adaptive systems in Section 2.7. This section briefly highlights the selected aspects of the motivating methodologies and how they contributed to the integrated adopted methodology.

The general approach was a realisation of the five-step conceptual framework from biologically-inspired algorithms (2.7.1). Specifically (1) the selection of the immune system as the motivating metaphor, (2) the review of immunochemistry and immunophysiology (probes) as a basis for information processing principles, (3) the investigation of models as abstraction of the information processing properties from models, (4) the application and verification of the computational models (considered next in Chapter 8), and (5) resulting computational tools. The adopted methodology drew heavily from this approach, in particular with regard to the general process for realising viable computational tools, and with regard to the strong focus on the transition from probes (*metaphor*) to information processing models that were investigated (results of the reductions). The immune information processing approach (Section 2.7.1) provided a specific context to motivate the conceptual framework, in particular to focus on the information processing concerns of immunology. Goldberg's so-called *small models* methodology for adaptive systems (Section 2.7.3) provided a strong influence in the adopted methodology, specifically focusing the modelling effort on the decomposition of general information processing to separate sub-problems (*reduction*) and the ultimate re-integration (the IHCSF). Along with the reductionist consideration, Goldberg's methodology promoted a strong focus on the bracketing of high-order phenomena which was realised in particular with regard to the cardinality and interactions between structures. It is important to clarify that the *small models* methodology anticipates system modelling is achieved through the use of mathematical methods. The approach taken in this work involved the use of algorithmic

computational modelling and simulation. An additional alternative modelling approach that could have been used is diagrammatic object and state modelling, such as the Unified Modelling Language (UML).

Alternative Interpretations

The specific reduction was an clonal selection-specific interpretation of the immune system based on the functional aggregation of cells at three intuitive levels of detail. This suggests that alternative cellular-based reductions may be considered, as well as reductions not based on the units of selection and adaptation in the clonal selection theory. An additional consideration for a reduction may include an antibody protein focus perhaps including the concerns of other macromolecules that influence the behaviour of the immune system such as cytokines and cellular receptors for such self-molecules. This would provide a focus on the interaction of components for signalling rather than adaptation. Alternatively, one may focus on the division of the immune system into adaptive and innate, and focus on the structures and functions within each ‘sub-system’ promoting a stronger consideration of the concerns of the innate immune system and its interrelated processes with the acquired immune system. The viability of a given reduction may be assessed based on the insights it provides regarding the computational principles under consideration. As such, an alternative already considered is the ‘*immune system* \leftrightarrow *cellular*’ reduction common across paradigms in AIS, suggesting that the contributions of the IHCSF may be the host and tissue considerations and the levels of comparison they provide.

7.5.2 Hierarchical Artificial Immune Systems

The field of Artificial Immune Systems is concerned with the investigation of computational tools inspired by the structure and function of the immune system (see Section 2.3.2). Given the pervasiveness of molecular and cellular concerns in immunology and the follow-on effects of this focus into Artificial Immune Systems, one may consider the application of the proposed Integrated Hierarchical Clonal Selection Framework as a general tool in AIS. Specifically, the hierarchical perspective may be used as a framework to focus effort for the integration existing seeming disparate computational methods inspired by immunology, and potentially highlighting new opportunities for investigating AIS.

Negative Selection

The negative selection paradigm exploits the clonal selection process for adaptation-based information acquisition although focus on the specific concerns of using pattern recognition to discriminate self from nonself and the preparation of such pattern detectors (see Section 2.4.2). The first important consideration is that of negative selection as adaptive strategy (modelling in the complement space with centralised preparation and with decentralised opportunistic specialisation using clonal selection). The strategy focus of the paradigm highlights the constraints under which the strategy must operate (satisfice) within the immune system at varied levels of abstraction. Such consideration may

provide insight to advance the paradigm beyond the current highly-constrained cellular perspective, and the limitations of such approach in practical application [374].

Scope	Negative Selection
<i>Cellular</i>	Centralised detector generation and application.
<i>Tissue</i>	Centralised/Decentralised detector generation and decentralised application.
<i>Host</i>	Controlled antigenic exposure, biased restarts of environmental perspective, and adaptation of generation and application structures and processes.

Table 7.9: Summary of the Negative Selection AIS paradigm mapped onto the three tiers of the IHCSF.

Table 7.9 provides a summary of the mapping of the negative selection paradigm onto the IHCSF. Given the close ties with clonal selection, the negative selection paradigm has generally the same concerns. Specifically, effective adaptation in the cellular tier, and effective localisation and dissemination in the tissue and host tiers. Also consistent with the clonal selection paradigm is the antigenic environment, where the information in the environment encapsulates the scope of patterns that must be differentiated from self patterns. The important differences between the paradigms is the strong focus on cellular preparation required in negative selection that may be achieved locally or globally. For example, the corpus of self-patterns may be centrally managed at the cellular scope providing a focus on effective non-self pattern generation, and the general application of such detectors. At the tissue scope, the preparation of the detectors may be centralised in a *thymus model* where a specific centrally connected tissue is designated the role of managing the self-pattern corpus and detector generation, with system-wide decentralised application and localised refinement of the recirculated detectors. Alternatively the management of the corpus may decentralised and distributed throughout the tissues of the system for localised (specialised) preparation and dissemination, likely a more scalable configuration, particularly if the self corpus is accumulated in an opportunistic manner. These concerns were considered by Hofmeyr in his dissertation (see Section 2.5.2) and provide a important research agenda for negative selection *at the tissue scope of the integrated hierarchy*. The host has similar concerns as that of the tissue, although with the important shift in focus from a strongly integrated series of perspectives to strongly independent holistic perspectives with weak integration. Pattern generation is centralised within each perspective although some control is provided over the exposure to information, the restarting of perspectives with biasing, and the long-term adaptation of the structure and processes of detector preparation and application.

Immune Network

The immune network paradigm is concerned with intra-receptor interactions with and without the presence of antigen and was proposed to be a specialised case of clonal selection at the cellular level (Section 4.7). This paradigm was investigated (1) briefly, and (2) only at the cellular level, suggesting an opportunity for further elaboration and integration with the scope of artificial immune network approaches. The information processing

strategy distilled from the theory was that of repertoire maintenance using intra-cellular interactions. Such interactions form networks of cellular relationships which may be considered intra-repertoire higher order structures, that may be exploited for storing more complex information.

Scope	Immune Network
<i>Cellular</i>	Formation and maintenance of intra-repertoire relationships.
<i>Tissue</i>	Interaction and integration of inter-tissue cell-based relationships.
<i>Host</i>	Generational refinement of structures as well as the processes and structures used to form and maintain such structures.

Table 7.10: Summary of the Immune Network AIS paradigm mapped onto the three tiers of the IHCSF.

Table 7.10 provides a summary of the mapping of the immune network strategy to the three tiers of the IHCSF. The antigenic environment involves the scope of the standard clonal selection antigenic environment at each tier in the framework, in addition to the scope of intra-repertoire interactions that may result. As these relationships are scaled-up from the cellular level, they may be considered to occur between tissues, and between hosts, with the latter providing a truly abstract proposition for the theory. From the tissue perspective, the localisation and dissemination of acquired information extends from information about antigenic infections, to information about the relationships within each tissue. The dissemination of cells involved in inter-cell relationships provides an opportunity for the interaction and ultimate integration of localised relationships between tissues, providing new contexts for response, and likely initially disrupting the maintenance of localised relationships. The host perspective, in particular the generational specialisation provides a context for considering the biasing (so-called ‘seeding’) of cellular relationships, and ongoing refinement of these relationships over extended time periods across multiple host-lineage in a population using the maternal immunity approach. The generational specialisation also provides an opportunity for adapting the processes and structures used in the formation and maintenance of relationships, allowing potentially for the development of new relationship types (structures).

Other

The generality of the reduction of functional immunology from the scope of a holistic immune system to the theory-laden cellular level may provide a powerful abstraction for many unrealised and/or budding Artificial Immune Systems. Examples may include, danger theory, innate immunity, cytokine molecules, and the large number of cell types and their interactions. The so-called power of the abstraction comes from (1) the context it provides for framing the information processing qualities of an immunological structure or process, (2) the strong focus on identifying and bracketing of complexity in resultant models promoting clear schedule for experimentation and elaboration, and finally, (3) a context for the integration of findings into a coherent model of the motivating paradigm.

7.5.3 Hierarchical Framework Methodology

This section considers the exploitation of the proposed methodology in different but related biologically-inspired Computational Intelligence sub-fields. In considering this question, a brief assessment of two general fields are considered, specifically (1) *Evolutionary Computation* and (2) *Connectionist Intelligence*. The fields are considered in the context of their *metaphor, strategy, partitioning* into system and environment, and *reduction*. Given the enormity of the research into the selected Computational Intelligence sub-fields the hierarchical phrasing of the three adaptive strategies is not likely to reveal any insights into the respective paradigms, rather the expectation is that the reductionist methodology may provide a framework for better integrating findings from existing approaches.

Evolutionary Computation

The metaphor for the well studied evolutionary computation field is neo-Darwinian evolution, taking into account the genetic concerns of micro- and macroevolution. The distilled computational principles include selection and descent with modification typically via the recombination and mutation of structures. The computational strategy may be summarised as adaptation toward relative improvement under the constraints of an environment (see Section 3.4.1). The summation of the adaptive strategy highlighted the intuitive separation of the population of structures under adaptation, and a environment responsible for defining the scope of ‘fitness’, typically apportioned directly in terms of quantitative scalar values that are interpreted by the adaptive strategy as the proportionate allocation of available resources. Table 7.11 provides a summary of an interpreted reduction of evolution from the species level to the DNA level, constraining both the general adaptive strategy and fitness-assigning environment. The reduction is based on the units of selection and adaptation using the same general approach as was adopted in the reduction of clonal selection. The chosen units of selection include genomes, populations, and species, which forms an Integrated Hierarchical Evolutionary Framework.

Tier	Environment	System
<i>Genome</i>	Habitat of Localities.	<i>Population of Genomes</i> , adapting in context of the scope of information being adapted and its mapping to an assessable form.
<i>Population</i>	Niche of Habitats.	<i>Species of Populations</i> , localising and disseminating adapted genomes.
<i>Species</i>	Ecology of Niches.	<i>Ecosystem of Species</i> , adapting and competing for ecological niches.

Table 7.11: Summary of an interpretation of the field of Evolutionary Computation in terms of the hierarchical reductionist methodology used to investigate clonal selection.

The *genome level* encompasses the scope of standard evolutionary approaches, and is bounded to a single population of individuals each with a distinct genome. The environment at this level is concerned with the relative difference between members of the population based on the information content of their genome. It is concerned with the

tools of DNA, information decomposition from holistic genome to functional gene, and ultimately to the base pairs of the representation. This is an important consideration as it provides a bracketing of information representation complexity, including but not limited to the concerns of the scope of the information mapped and assessable in the phenotype, intra-gene relationships (gene or allele linkage), and other representation complexities. The mapping of genotype to phenotype provides another critical axis of complexity for bracketing including concerns of development and time-scope of assessment. As such the mapping may vary between direct assessment of genome, unicellular and multicellular organisms, to the full concerns of developmental biology and lifetime-based assessment of the resulting organism. To summarise, the important axis for bracketing include: (1) *information representation* such as one-to-many relationships in base-pairs, genes, and the number of and number of sets of chromosomes (ploidy); and (2) *information mapping* such as the transformation and time-horizon on the assessment of the genotype and potentially resulting organism.

The *population level* shifts the concerns from that of organisms in a population, to a set of discrete populations in a species. The environment is scaled in turn, considering the relative improvements of populations through the aggregation of the relative improvements of the genomes in each population. The environment is generally the same for all populations, limiting the divergence of populations into different species. This level is reminiscent of hierarchical genetic algorithms (for example see [319, 119]) and parallel evolutionary algorithms (for example [68]), with concerns of inter-population interaction in the form of migrations. Examples of inter-population interactions provides an important axis for complexity for bracketing beyond the cardinality of genomes in each repertoire. This constraint is lifted in the *species level* that raises the concerns from that of populations in a species, to that of a species in an ecosystem. The environment provides a variety of ecological niches that may be exploited and competed for by the species in the ecosystem. This introduces the concerns of the divergence of species called *speciation* (for example see speciation genetic algorithms [123]), and the invasion of sudden exploitation of underexploited ecological niches *adaptive radiations* (for example see niching genetic algorithms [279]). Both the interaction between species, and the cardinality of ecological niches provide contexts of complexity that may be bracketed for investigation, with interesting combinations such as species themselves representing a part of an ecological niche in a predator-prey relationship.

The hierarchical framing and relationships between the concerns of evolution highlight the specific bracketed configurations chosen and propagated for sub-fields of the paradigm. For example in the case of niching genetic algorithms and speciation genetic algorithms the effects are realised within a single population. This observation does not violate the constraints of the proposed framework, rather suggests a cross-hierarchical bracketing where the diversity concerns of multiple populations or multiple species are compressed to that of genomes within a population. This form of bracketing is similar to that used in the HAEF with antigen from the cellular to the host levels, although in this case is used on the system side to bring features from the higher levels of abstraction down to the lowest level

of abstraction by compressing the cardinality of the units of adaptation and selection. The compression of cardinality is the basis of population genetics that considers gene frequencies, flows and drifts in populations and species of individuals. The contribution of the framework is subtle, as it does not suggest perspectives of the strategy unconsidered by the field. Rather, the contribution is that the focus of each tier provides a context for bracketing the complexity of the strategy at different levels of abstraction. This focused bracketing highlights the axis of complexity that may be bounded for investigation, and more importantly it suggests at the effects of the constraint by considering the extremes of the alternatives.

Connectionist Intelligence

The metaphor for the field of connectionist intelligence are brain cells called neurons that ultimately provide the biological basis of human-level intelligence when arranged into vast interconnected networks in the central nervous system. The distilled computational principles include collections of neurons with synapses that connect the neurons into a network structure, and the propagation of signals through the network. Propagation of signals involves sensory (input) neurons which propagate received signals through the network in parallel. Signal propagation is governed via the aggregation of inbound signals to a given neuron in the network which may or may not activate and transfer (forward propagate) the received signal. The computational strategy involves the adaptation of the ways signals are propagated via adjusting the weights between the neurons. The strategy is divided into the central nervous system of neurons and their interconnections and the environment that provides the stimuli to which the system receives and adjusts its behaviour. Table 7.12 provides a summary of an interpreted reduction of connectionist intelligence from the holistic nervous system to the neuron level, constraining both the general adaptive strategy and the sensory input environment. The reduction is based on levels of abstraction regarding the aggregations of brain cells which provide interesting constraints on the information processing that may occur including a network of cells referred to as a neuronal group that responds to a habitat of stimulus, groups of networked cells referred to as a host (nervous system), and finally a population of hosts with nervous systems.

Tier	Environment	System
<i>Cellular</i>	Stimulus of Inputs.	<i>Group of Cells</i> , providing a focus on the organisation and interaction of neuronal cells.
<i>Group</i>	Habitat of Stimulus.	<i>Host of Groups</i> , focusing on the interaction between groups of cells such as the dissemination of information. An alternative name for a group may be a cortex.
<i>Host</i>	Environment of Habitat.	<i>Population of Hosts</i> , focusing on the interaction of hosts each with their own nervous system.

Table 7.12: Summary of an interpretation of the field of Connectionist Intelligence in terms of the hierarchical reductionist methodology used to investigate clonal selection.

The *cellular level* is representative of the standard forms of Artificial Neural Networks (ANN) (for example see [334]). The environment at this level involves a stream of input patterns referred to as a *stimulus* (for example a series of multivariate vectors). Therefore, the scope of supervised and unsupervised learning schedules for adapting weights, as well as recurrent and forward propagating structures belong to the cellular level. Bracketing involves bounds on the cardinality of the units of adaptation (neurons), and their interactions (synapses). Much like the clonal selection of immune system cells, the information content from the environment is exposure-based, although in this case governed by the sensors that perceive and communicate the information to the neural network. In addition to the trigger-based adaptation of intra-network connections, recurrent connections provide a source of self-stimulation within the structure, much like the immune network paradigm.

The *group level* is concerned with multiple functional groups of neurons that collectively may form a complete nervous system. As such, the functional units are exposed to an array of stimuli called a *habitat*. This provides a scale above the standard neural network models, and considers multiple network approaches. There has been much work on such models in the field not limited to hierarchical neural networks [34], and modular and ensemble neural networks [224, 357], and many others. The rise in the level of abstraction may also provide opportunities for exploiting varied neuronal cell and connection types including relay, motor, and sensory cells as well as the variety of neural receptors and transmitters. The axes of complexity that may be bracketed include the number of functional groups of cells as well as the amount of interaction between the groups. The *host level* is concerned with populations of interacting nervous systems that respond to an environment of habitats of stimuli mediated by the each host's perception. The environment includes the stimuli that may be provided by other nervous systems as well as self-generated stimuli, and some control over preferential sensing from the environment. The scope, sources, and spatial temporal properties are expected to provide the the greatest source of complexity at this level, requiring careful bracketing of the environmental information, habitat-mediated perception of information, and the interaction control exerted by nervous systems. The population perspective permits the consideration of a generational population structure that may adapt the structure and processes that govern the propagation and manipulation of signals in networks of neurons. This is related to the field of evolutionary artificial neural networks (for example see [440]), that is concerned with the application of evolutionary processes to the concerns of neural networks at the neuron level such as the architecture, the connections weight, and the learning rules used to adjust the connection weights.

Other

The methodology, specifically the *division* and *reduction* concerns result in a clear hierarchical system-environment perspective of a biological inspired computation paradigm provides common ground for comparing and contrasting such paradigms at multiple levels of abstraction. From the perspective of clonal selection, a brief interpretation of the

related fields of evolutionary computation and connectionist intelligence provides a information processing context (more than superficial similarity) for directly relating models and algorithms at the same level of abstraction (for example the relationship of ensemble neural networks (mixture of experts) and island population genetic algorithms with the tissue clonal selection paradigm), for exploiting the findings of such related approaches, and identification of relevant and important application problem domains.

A clear difficulty in this approach is the elucidation of a discrete environment that has meaning at each level of abstraction. For the connectionist paradigm in particular, such environments may be conceptually intuitive, although difficult to articulate concisely resulting in arbitrary description. The pattern of this mapping of external concerns to an environment follows a clear trend across clonal selection, evolution and connectionist paradigms, which is the bracketing of a general environmental concern at the highest level (antigenic, ecological, and sensory environments) to the scale of a specific functional group of interest in the system. The comparison may be strength through the phrasing of additional Computational Intelligence paradigms with regard to the division and reduction aspects of the methodology. For example one may consider the following Swarm Intelligence paradigms: Particle Swarm Optimisation with reduction based on hierarchical groups of swarming units, and Ant Colony Optimisation based on hierarchical groups of foraging units.

7.5.4 Summary

The section summarises the generalised principles deduced from the consideration of the adopted methodology, the application of the resulting framework to the broader field of Artificial Immune Systems, and the general applicability of the methodology to related Computational Intelligence sub-fields.

1. *Methodology*

- (a) *Four Step Process*: The four-step process (*metaphor, strategy, division, reduction*) provides a clonal-selection focused integration and augmentation to the three reviewed standard methodologies (see Section 2.7), promoting the concerns of abstraction, modularity, extensibility, bracketing, and ultimate reintegration.
- (b) *Alternatives*: The specifics of cell-based reduction of the immune system (the cellular, tissue, and host tiers) is one interpretation of such an reduction that was demonstrated to be useful for the investigation of clonal selection. Alternative reductions exist both with regard to the specific aggregations of cells, and with regard to the use of alternative criterion from which provide a basis for reduction.

2. *Broader Applicability*

- (a) *Beyond Cells*: The framework with its basis in aggregation of cells, provides an intuitive and modular perspective on the immune system providing a context

for considering the scaling concerns of an array of cellular and molecular-based computational metaphors from immunology.

- (b) *Transferable Findings*: Given the central role clonal selection plays with regard to the negative selection and immune network paradigms, the findings from the investigation clonal selection are expected to be generally transferable, specifically with regard to the localisation and dissemination concerns investigated at the tissue and host scales.

3. *General Applicability*

- (a) *Cross-Framework Perspective*: The brief interpretation of evolutionary computation and connectionist intelligence in the context of the hierarchical framework provided a common ground for comparison at multiple levels of abstraction that revealed related *algorithms* and *problem domains* that may be exploited and integrated in the clonal selection framework.
- (b) *Hybrid Approaches*: The modularity of algorithms across paradigm hierarchies, and the control over constraints provided by bracketing of complexity facilitates the interchangeability of algorithms between frameworks, for example, the use of a genome-level GA or neuron-level ANN at the cellular level in the IHCSF when considering tissue or host level.

7.6 Chapter Summary

This section considers the broader principles and findings from the three proposed hierarchical frameworks and the adopted systematic methodology that ultimately developed the frameworks.

7.6.1 Frameworks Overview

The *Hierarchical Clonal Selection Framework* provided the integration of the cellular, tissue, and host perspectives of the clonal selection adaptive strategy, and the suppression of antigenic exposures to that of a generic exposure environment. The framework highlighted the fact that each tier imposes a distinct set of constraints on the cellular-focused strategy, where bracketing provides a tool to manage subsumed complexity of the interactions. It further suggested that the unbracketed complexity across the hierarchy results in compounded interaction effects, revealing a broader hierarchical system and promotes consideration of additional perspectives with which to constrain the cellular-focused strategy. The *Hierarchical Antigenic Exposure Framework* provided the integration of the arbitrary defined antigen, infection, and habitat perspectives of the piece-wise exposure of domain information to a generic (suppressed) trigger-response based clonal selection system. The framework showed that the expected capability of a system is compounded as additional perspectives are considered, where strong bracketing of environmental information content bounds such complexity at the expense of forced response to each exposure. The piece-wise exposure of an information environment may be investigated (and bracketed)

on three important axes of complexity: the cardinality or amount of information, the spatial and temporal patterns by which the information is exposed, and the propagation of both of these concerns down through the hierarchy.

The *Integrated Hierarchical Clonal Selection Framework* was concerned with the integration and interaction between the hierarchical clonal selection and antigenic exposure frameworks, with a full elaboration of the system and environmental concerns. Importantly, the framework highlighted the fact that the maximum in bracketable complexity of a given tier represents the minimum in bracketable complexity for the following tier in the hierarchy (the maximum-minimum principle). The non-contiguous and strongly-bracketed perspectives of the framework provide opportunities for further investigation, as does the consideration of tiers above and beyond those considered in the framework and the elaboration of those perspectives already considered. The adopted *Methodology* provided a general cellular focused approach for the systematic distillation of an interesting information processing metaphor, and controlled steps of elaboration, reduction, and further elaboration. Broadly, the artefact of the applied methodology (the IHCSF) may be a useful tool for elaborating existing AIS frameworks beyond the constraints of a cellular perspective. Generally, the application of the methodology to existing Computational Intelligence sub-fields provides both a context for comparison with superficially similar approaches toward integration (hybridisation) and racing.

7.6.2 Integration

The proposition and analysis of the hierarchical frameworks provided both insights into the design decisions, boundaries of concern, and limitations of the investigated clonal selection models and algorithms from across the three tiers. Further, the analysis also provided motivation for the investigation of the myriad of variations facilitated by the strongly applied complexity bracketing. The application of the adopted methodology and framework to related Computational Intelligence fields suggested at the potential insights that may be provided in assessing proposed approaches against state of the art. This suggestion motivates the next chapter the focus of which is dominated by the applicability of the clonal selection models and algorithms from across the three tiers in the specific context of Function Optimisation and Function Approximation problem domains.

Chapter 8

Suitability and Application

8.1 Chapter Overview

The previous chapters proposed and investigated models and algorithms inspired by clonal selection with little regard for their applicability as Computational Intelligence tools. Before applicability is considered, it is important to outline a systematic methodology to define suitability not just for the application of a technique to a problem, but also for comparison of a technique to other similar techniques as well as the transferability of results on one problem to other similar problems. Section 8.2 considers the suitability of application of the clonal selection algorithms. A simple feature-based suitability methodology is defined and is applied to the base algorithms from across the three paradigms both eliciting general features and projecting those features onto general attributes of computational problems. Two problem formalisms that embody many of the most difficult problems faced by Artificial and Computational Intelligence are reviewed: Function Optimisation in Section 8.3 and Function Approximation in Section 8.4. Each formalism is described in terms of (1) its general properties, (2) a set of specialised sub-problems, and (3) select approaches from the fields of AI and CI that both address problem instances that fit into the formalism and exhibit some superficially similar features with clonal selection algorithms from across the hierarchical framework. These problem examples provide a tangible framing for the features of the clonal selection algorithms, which are systematically mapped onto specific and *suitable* sub-problems from the respective problem domains.

8.2 Suitability of Application

From a problem-solving perspective, the tools that emerge from the field of Computational Intelligence are generally assessed with regard to their utility as *efficiently* or *effectively* solving problems. An important lesson from the ‘no-free-lunch theorem’ was to *bound claims of applicability* (Section 2.7.2). An approach toward this end is to consider the suitability of a given strategy with regard to the feature overlap with the attributes of a given problem domain. From a Computational Intelligence perspective, one may consider the architecture, processes, and constraints of a given strategy as the features of an ap-

proach. This section lists such features of the approaches from the Cellular, Tissue, and Host Paradigms and projects their features onto general capability with regard to problem features to which the approaches may be suitable. These so-called general features provide the foundation for further consideration of the application of the clonal selection algorithms from across the hierarchical framework. System features are discussed in the context of the general intention and observed behaviours. Paradigms are discussed in a compressed form, not taking into consideration the subsumed properties of the hierarchical framework.

8.2.1 Considering Suitability

In this work, the suitability of the application of an ‘approach’ to a ‘problem’ takes into considerations concerns such as the *appropriateness* (can the approach address the problem), the *feasibility* (available resources and related efficiency concerns), and the *flexibility* (ability to address unexpected or unintended effects). This section summarises a general methodology toward addressing the problem of suitability in the context of Computational Intelligence tools. This methodology involves (1) the systematic elicitation of system and problem features, and (2) the consideration of the overlap of problem-problem, algorithm-algorithm, and problem-algorithm overlap of feature sets.

Systematic Feature Elicitation

A *feature* of a system (tool, strategy, model) or a problem is a distinctive element or property that may be used to differentiate it from similar and/or related cases. Examples may include functional concerns such as: processes, data structures, architectures, and constraints, as well as emergent concerns that may have a more subjective quality such as general behaviours, organisations, and higher-order structures. The process of the elicitation of features may be taken from a system or problem perspective as follows:

- *System Perspective*: This requires a strong focus on the lower level functional elements and investigations that work toward correlating specific controlled organisations towards predictable emergent behaviours. An example is the central focus of this work in Chapters 4 through 7.
- *Problem Perspective*: May require both a generalisation of the specific case to the general problem case, as well as a functional or logical decomposition into constituent parts.

Problem *generalisation* and *functional decomposition* are important and well used patterns for problem solving in the broader fields of Artificial Intelligence and Machine Learning as the promotion of simplification and modularity can reduce the cost and complexity of achieving solutions [342, 58].

Feature Overlap

Overlap in elicited features may be considered from three important perspectives: *between systems*, *between problems*, and *between a system and a problem* (depicted in Figure 8.1). Further, such overlap may be considered at different levels of detail with regard to generalised problem solving strategies and problem definitions. These overlap cases are considered as follows:

- *System Overlap*: Defines the suitability of comparing one system to another, referred to as *comparability*. For example, systems may be considered for the same general problems and compared in terms of theoretical or empirical capability, the results of which may only be meaningful if the systems are significantly similar to each other as assessed in terms of feature overlap.
- *Problem Overlap*: Defines the suitability of comparing one problem to another, referred to as *transferability*. From a systems focus for example, transferability refers to the capability of a technique on a given problem to be transferred to another problem, the result of which is only meaningful if there is a strong overlap between the problems under consideration.
- *System-Problem Overlap*: Defines the suitability of a system on a given problem, referred to as *applicability*. For example, a system is considered suitable for a given problem if it has a significant overlap in capabilities with the requirements of the problem definition.

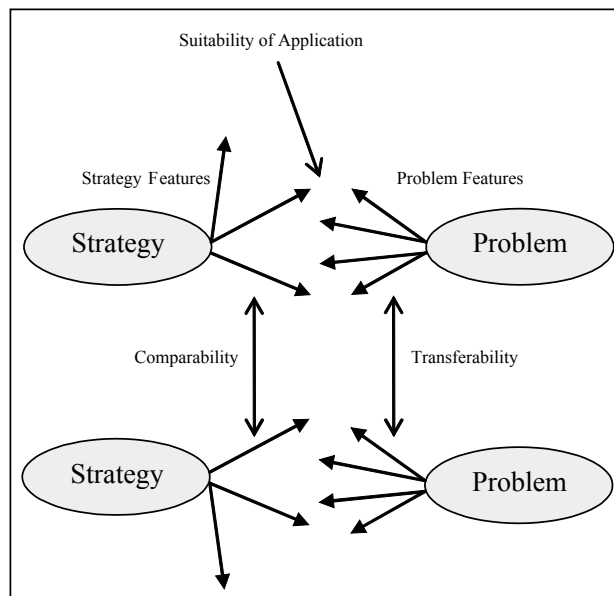


Figure 8.1: Depiction that the suitability of application of a strategy to a problem is defined by the overlap in their features, as well as the comparability with other strategies, and transferability of results to other problems.

Such mappings are expected to have noise given the subjective assessment and/or complexity required in both the elicitation and consideration overlap of the of features,

the noisiest of which is expected to be the mapping between systems and problems. The mapping of salient features of algorithms and problems was proposed as an important reconciliation of the ‘no-free-lunch theorem’ by Wolpert and Macready [436], although the important difference of this approach is that the system and algorithm are given prior to the assessment. In [435], Wolpert and Macready specifically propose the elicitation of the features from a problem first perspective, for which specialised algorithms can be defined. Therefore, this methodology of suitability may be considered a generalisation of this reconciliation suitable for the altered ‘Computational Intelligence’ (strategy first) perspective on Artificial Intelligence.

8.2.2 Cellular Algorithms as Adaptive Strategies

This section reviews the general properties of cellular algorithms as adaptive strategies, and considers the approach that such algorithms may be used for inductive modelling with different perspectives of a given problem. Finally, the general cellular algorithm features are listed, and projected onto suitable general problem features.

Cellular Algorithms Overview

The cellular algorithms are *adaptive* which is interpreted as their general capability of obtaining characteristics that improve the systems relative performance in an environment. This adaptive behaviour is achieved through a *selectionist process* of iterative selection and descent with modification. The discrete cell-based architecture is inherently *parallel* allowing for concurrent selection processes, and is *robust* providing redundancy of information and flexibility in terms of resource allocation. The method of acquiring information is called *inductive learning* (learning from example), where the approach uses the implicit assumption that specific examples are representative of the broader information content of the environment, specifically with regard to anticipated need. Generally, cellular approaches maintain a population of samples that provide both a representation for acquired information, and the basis for further induction. This method of simultaneously improving information and optimising decisions is called the *k*-armed bandit (two-armed and multi-armed bandit) problem from the field of statistical decision making [338] (for a contemporary treatment see [39]). This class of problem has had a long tradition of as a formalism for considering genetic algorithms and niching variants with regard to the adaptive processes capability of the ‘automatic’ allocation of resources proportional to expected payoff [173].

The acquired information is generally *approximate*, and is done so using a *stochastic method*. The general method is called Monte Carlo in which randomness is exploited to provide good average performance, quickly, and with a low chance of the worst case performance. Such approaches are suited to problems with many coupled degrees of freedom, for example large high-dimensional spaces. The selection method by which induction occurs may be modelled as a series of parallel random walks that exploit gradients in the underlying cost surface (directly, without derivatives), and as such is known as a Markov chain Monte Carlo method [25, 76] (sampling from a target distribution function using a

Markov chain mechanism). This highlights that the robustness of the approach extends to the induction process itself with regard to an improving approximation in the face of potentially incomplete, incorrect and inconsistent sampled data. Finally, the induction is performed in a piece-wise or *online* manner satisfying the concerns of *real-time* data, with potentially *dynamic* changes that may be benign or *adversarial* with regard to the internal model.

Inductive Model Generation

This section considers the scope of the inductive learning and resultant model in the context of a given problem, and more specifically the information made available by the problem. Three general classes of problem modelling are considered, as follows:

- *Independent Models*: The system models the entire scope of the problem irrespective of any functional decompositions that may be available. This approach may be referred to as holistic or naïve models.
- *Independent Sub-Models*: The system models the problem in terms of clearly defined sub-problems that are logically and/or functionally independent with regard to the desired solution. This approach may be referred to as partitioned, symbolic, or disjunctive models.
- *Dependent Partial-Models*: The system models the problem in terms of unclear or implicit sub-problems that are logically and/or functionally dependent with regard to the solution to the problem. This approach may be referred to as sub-symbolic, or conjunctive models.

These three model types are used to assess both the Cellular Algorithms, as well as the Tissue and Host Algorithms in the sections that follow. Figure 8.2 depicts the relationship between the cellular algorithms and the three modelling types in the context of problem instances that support such modelling. The figure clearly shows that cellular algorithms model all three induction problems using a single repertoire of cells. The implication of such an approach is that specific mechanisms may be required to facilitate the integration of independent information in the case of sub-models, and inter-dependent information in the case of partial models.

The three inductive modelling types may be related to the Antigen Colour Space Problems (ACSP) investigated in Chapter 4. The exposure of a single antigen ($N_{antigen} = 1$) to the system may be considered an independent model, whereas the multiple antigen ($N_{antigen} = N$) and explicit aggregation of solutions may be considered the modelling of multiple independent sub-problems. Importantly, given that the ACSP definition did explicitly partition the antigenic set, although did not define the ordering of the exposure of patterns to the system, the ($N_{antigen} = N$) problem case may be considered a holistic problem with unknown sub-problems and therefore given this unknown must be modelled using partial-dependent sub-models (problem-side aggregation of responses). If the ordering of antigenic exposures was known to the system, then explicit sub-repertoires could be

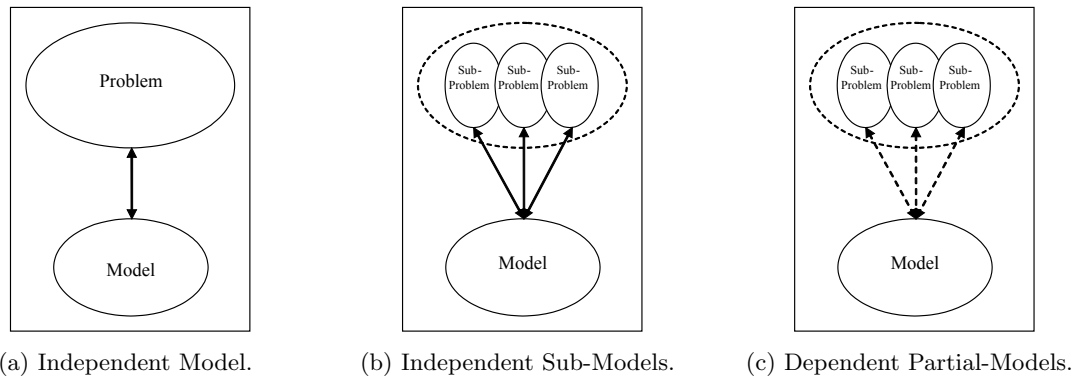


Figure 8.2: Depiction of the cellular algorithm approach to inductive learning with holistic, sub-, and partial-models of a given problem.

defined and specialised for each pattern (system-side aggregation of responses). Finally, the modelling of colour space patterns using colour components (in the case of DCCSA) may be considered the modelling of the ACSP using independent cellular components (one-cell equals one-component), although the modelling of sub-symbolic determinants (such as partial strings) would have provided an example of dependent partial modelling of single antigen.

Feature Summary

This section summarises the important descriptive features of Cellular Clonal Selection Algorithms, and projects their features onto general features of problems to which this class of algorithm may be suitable for application.

Salient System Features The following provides a list of the five features identified as being distinctive in characterising the general class of Cellular Clonal Selection Algorithms.

- *Selectionist Adaptation*: Relative improvement of the system’s decision making capability in the context of a problem via a process of selection and decent with modification to iteratively acquire and refine acquired information.
- *Inductive Learning*: Information is acquired and generalised from discrete and specific examples provided by the problem.
- *Stochastic Process*: The interaction with the problem and the resultant adaptation have an inherent element of randomness that promotes approximation of acquired information, general robustness of the system to noise, and flexibility of the system to unexpected events.
- *Population-based*: Maintenance of a diverse although clustered set (inherent redundancy) of pre-committed structures by the system, each of which may interact with the problem potentially in a concurrent manner (inherent parallelism).

- *Pattern Recognition*: The acquisition of information and overt decision making are coupled and triggered via a pattern recognition mechanism and response to information from the problem.

Projected Problem Features The following provides a projection of the characterising features of Cellular Clonal Selection Algorithms onto general problem features to which such algorithms are expected to be suitable.

- *Discrete*: The problem is comprised of discrete packets information to which it is in the systems best interest to respond.
- *Online*: Continuous decision making and improvement are desired given the variable and potential dynamic or adversarial manner in which information is made available to the system.
- *Unknowns*: Many of the parameters and/or constraints of the problem are known, although important elements regarding the information content and/or ordering of the exposure of information remain unknown to the system.
- *Noise*: The information signal may be disrupted, intermittent, and/or noisy requiring a level of robustness of interpretation.
- *Regularities*: Problem information contains regularities or patterns that must be exploited for effective progress to be made by a system.

8.2.3 Tissue Algorithms as Data Fusion

This section reviews the general properties of tissue algorithms as an adaptive data fusion approach, and considers the approach that such algorithms are useful for the inductive modelling using different perspectives of a given problem. Finally, the general tissue algorithm features are listed, and projected onto suitable general problem features.

Tissue Algorithms Overview

Tissue algorithms are *explicitly decentralised*, with each tissue providing an adaptive process making local decisions and improvements, the effects of which are *implicitly integrated* into holistic understanding of the information environment. The induction of the information content and anticipated exposure is extended across each point of exposure providing a *spatial acquisition and anticipation*, that trades-off concerns of localisation and dissemination. The tissue architecture is pre-determined, with *tightly coupled* and *fixed neighbourhood relationships* between tissues. Tissues themselves are not redundant, nor flexible in their topology (additions, modifications, deletions), although each tissues cellular information content is redundant, allowing general or specific loss of random sets of cells. The information composition across tissues is generally heterogeneous given the specialisation of each tissue (given an asymmetric information exposure pattern), although

the general capability at any tissue in the system is equivalent (*heterogeneous composition, homogeneous capability*). This equivalence in capability is provided by the continuous dissemination (high frequency, generally low amplitude) of the underlying system information, and the localised as well as opportunistic way in which such information is exploited. The general method is called *multisensor data fusion* (information fusion or distributed sensing) whereby the modularity of information acquisition is exploited and combined for holistic application providing an improved model over that of a series of independent models [195]. The modularity may be imposed requiring explicit decomposition or partitioning of an information environment using prior information. Alternatively, the modularity of the information and its integration with the system may be an inherent feature of the environment. In this latter case, the interaction with the environment governed by the tissue exposure regime may or may not be known, making information acquisition and application *opportunistic*. Integration is achieved by the continuous recirculation and mixing of acquired information, suggesting that the application of the integrated information may also be piece-wise or modular like the acquisition of the information (so-called piecewise acquisition and application from the system perspective, or exposure and expectation from the environment perspective).

Inductive Model Generation

The tissue algorithms may be considered in the context of the three generative inductive models, specifically independent models and sub-models, and dependent partial models. Figure 8.3 depicts the general relationship between the tissue algorithms and the three modelling types in the context of problem instances that support such modelling. The architecture of tissue algorithms forces an explicit partitioning of a given problem information space whether such partitioning exists or not. The partitioning of the Infection Antigenic Exposure Problem (IAEP) via the Tissue Exposure Regimes provides a basis for considering the suitability of the tissue model against each modelling type (Chapter 5). In the case of a holistic modelling, one may consider a symmetric or point exposure regime in which case the persistent unbiased recirculation generally provides a disruptive influence to such modelling. Similarly, one may consider an asymmetric exposure regime representative of an explicit decomposition of the problem, where persistent unbiased recirculation has the same disruptive effects. Specifically, the results suggest, that although the tissue algorithms can construct such inductive models (innate capability), *tissue algorithms may not be suitable for constructing holistic or sub-models unless there is an expectation that dissemination will be beneficial*, for example in the problem definition, or if the intra-tissue CCSA is capable of exploiting such information¹. This finding is further supported by analogous findings from related approaches in Section 8.3.2 and Section 8.4.2.

Importantly, the functional decomposition was not included in the IAEP definition, therefore, the tissue algorithms may be considered to have constructed partial models for the unknown sub-problems in the problem definition (TER's). As such, the aggregation

¹This latter example was demonstrated to not be the case for in the use of RCCSA as an intra-tissue algorithm.

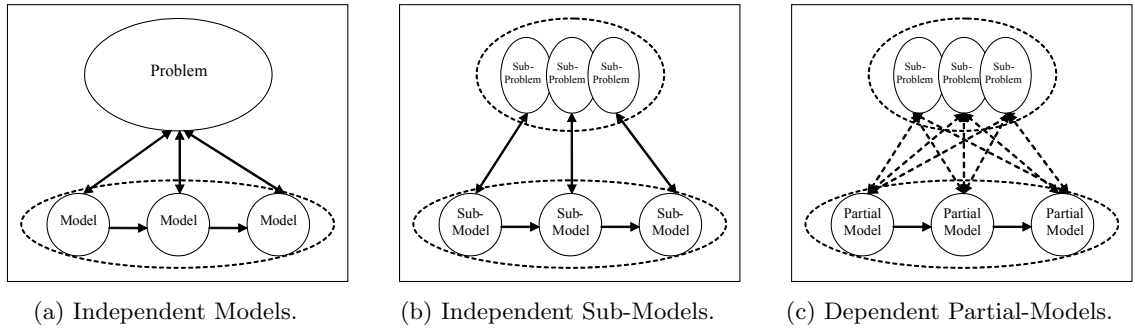


Figure 8.3: Depiction of the tissue algorithm approach to inductive learning with holistic, sub-, and partial-models of a given problem.

of results for a specific TCSA across the TER's provides a general indication of algorithm capability. The aggregation of results clearly highlight that MTCSA is more suitable for per-exposure responses, and that the *recirculation algorithms may be more suitable for holistic solution aggregation at any given single tissue under an unknown exposure regime*, likely at any given time under such a regime. More specifically, this suggests the suitability of recirculation tissue approaches are with the redundancy of acquired information to data loss and impromptu localised solution aggregation. Beyond this general suitability, the results also suggest the specific suitability of recirculation methods under a random exposure regime compared to the non-recirculation approach. Although the exploratory results from Chapter 5 are not conclusive, they suggest strong constraints on the suitability of applicability of recirculation-based Tissue Clonal Selection Algorithms.

Feature Summary

This section summarise the important descriptive features of Tissue Clonal Selection Algorithms, and projects the features onto general features of problems to which this class of algorithm may be suitable for application.

Salient System Features The following provides a list of the five features identified as being distinctive in characterising the general class of (recirculation-based) Tissue Clonal Selection Algorithms.

- *Distributed Model*: The population-based model is distributed across a series of discrete and communicating sub-populations.
- *Decentralised Adaptation*: Acquisition of information and decision making occurs locally at each distributed point of the system providing a natural mapping for problems that can be functionally or logically decomposed.
- *Tight Coupling*: Sub-populations are organised in a spatial topology with fixed neighbourhood relationships that generally communicate a low amplitude of information with a high frequency.

- *Data Fusion*: Information acquired in localised regions of the structure is disseminated and made available across the system, providing an integrated although distributed holistic model.
- *Opportunistic*: The general availability of acquired information and decision making capability facilitates the opportunistic exploitation of variably accessible problem information or information processing capability.

Projected Problem Features The following provides a projection of the characterising features of Tissue Clonal Selection Algorithms onto general problem features to which such algorithms are expected to be suitable.

- *Discrete Decomposition*: In addition to facilitating discrete interaction with the system, the problem should be able to be functionally or logically decomposed into sub- or partial problems which are naturally or explicitly partitioned across points of contact with the system.
- *Online*: It is desirable for a holistic system to achieve continuous improvement against the problem, although less sensitive with regard to behaviour across specific spatial interactions.
- *Spatial-Temporal Regularities*: The notions of noisy information signals and regularities extend beyond sub-problems, and extend across the spatially discrete interactions with the system (notions of Tissue Exposure Regimes).
- *On Demand*: Information and/or decision making capability acquired by the system holistically is required at a single given point of the system in an as needed manner (impromptu and point-wise solution aggregation).
- *Hazardous Environment*: Addressing the problem requires some level of system robustness with regard to the loss, corruption, or failure of general (random sample from all tissues) or specifically-localised (content of random tissue or information in transit) sub-sets of information acquired by the system.

8.2.4 Host Algorithms as Mixtures of Experts

This section reviews the general properties of host algorithms as an adaptive mixture of experts approach, and considers the application of such algorithms for the use in inductive modelling, exploiting different perspectives of a given problem. Finally, the general host algorithm features are listed, and projected onto suitable general problem features.

Host Algorithms Overview

Like tissue algorithms, host algorithms are comprised of a set of *decentralised processes making localised decisions*, that in aggregation represent the systems information composition and general capability. Unlike the tissue paradigm, hosts are loosely coupled,

without fixed neighbourhood relationships. The localised adaptive processes are *quasi-independent* and their interactions are impromptu, opportunistic, and generally peer-to-peer. Their quasi-independence is reflected in their perspectives of the information environment, where interactions do not promote an integration of perspectives, rather are *low-bandwidth dissemination of information* (low frequency, generally high amplitude of cells) from which, in aggregation the population is expected to benefit. As such information dissemination is selective and organised although impromptu given the generally unstructured nature of the population of hosts. The general method is called *mixture of experts* (ensemble methods) where systems exploit the modularity of the perspectives of the information environment by specialising in separate although potentially overlapping perspectives [310, 327]. As with the data fusion perspective of the tissue paradigm, the modularity of the environment may be imposed through explicit decomposition or partitioning, or may be an inherent feature of the environment. Each host generates a model of its perspective of the information environment through the cellular and tissues process of induction via adaptation. The generational host algorithms relate to the general method called *multiple restarts* (iterative restarts or multiple runs) as this generative process is periodically reinitialised for a different although potentially biased set of starting conditions that influence the inductive model generation process [225, 69].

Inductive Model Generation

The host algorithms may be considered in the context of the three generative inductive models, specifically independent models and sub-models, and dependent partial models. Figure 8.4 provides a depiction of general host algorithms with regard to the three inductive model types. The figures show the explicit partitioning of the model, as was the case with the tissue algorithms. Also as with the tissue algorithms, spatial exposures regimes were the focus of the capabilities of the host algorithms embodied in the Habitat Antigenic Exposure Problem (HAEP). Model generation can be considered with regard to the population-based and generation-based approaches separately. In the case of a holistic and sub-model generation the non-interaction MP-HCSA demonstrated its suitability with regard to exposure-wise aggregation of results (system error), and its lack of suitability for impromptu point-wise aggregation of results in the same manner as the MTCSA. Interestingly, both the Random-Pairing Pathogen Transmission HCSA and the Small Sample SI-HCAS demonstrated a trend of decreased point-wise aggregation error and in the latter case system error. This highlights the *suitability of population-based HCSA with small unstructured dissemination of acquired information or triggers not only for holistic (SHER) and sub-models (AHER), but also for partial models (across unknown HER's)*, in particular, SI-HCSA-SS was competitive with MP-HCSA under the AHER and SHER exposure regimes.

The generational approach of host populations provides capability for restarting the model generation process that may remove specific training biases by using varied starting conditions. Further, the adaptation of the model generation process itself in the EI-HCSA provides capability for tuning the process toward improved resultant models. Although

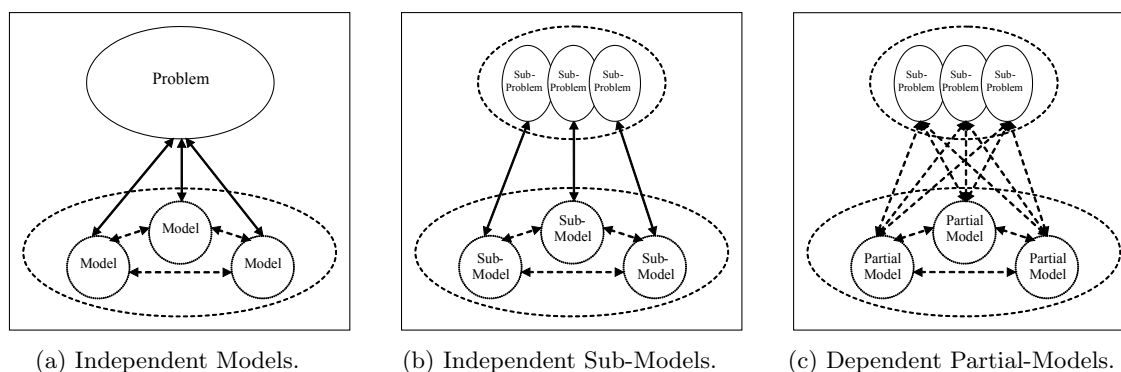


Figure 8.4: Depiction of the host algorithm approach to inductive learning with holistic, sub-, and partial-models of a given problem.

capable, neither approach was demonstrated suitable for such application from the results of the limited exploratory experimentation, likely given the specifics of the configurations used.

Feature Summary

This section summarise the important descriptive features of Host Clonal Selection Algorithms, and projects their features onto general problem features to which this class of algorithm may be suitable for application.

Salient System Features The following provides a list of the five features identified as being distinctive in characterising the general class of Host Clonal Selection Algorithms.

- *Distributed Model*: The generation, storage and maintenance of the inductive model is distributed across a set of discrete and communicating sub-systems, where the acquisition of information and decision making is decentralised across the discrete contacts with the problem promoting a natural mapping for partitioned or partitionable problems.
- *Loose Coupling*: Sub-systems are not organised into any explicit spatial topology, therefore neighbourhood interactions are impromptu, generally involving a small number of sub-systems communicating a moderate amplitude of information with a low frequency.
- *Mixture of Experts*: Sub-systems become specialists in potentially overlapping partitions of the problem space, from which infrequent intra-population communication benefits the quasi-independent partial and/or sub-models.
- *Multiple Restarts*: The generational concerns of the model provide an iterative restart mechanism providing multiple model generational trials that may be selectively biased.

- *Adaptation of Process*: The evolutionary generational concerns of the model provide an iterative adaptation mechanism for improving the model generation process with regard to a specific problem ‘environment’.

Projected Problem Features The following provides a projection of the characterising features of Host Clonal Selection Algorithms onto general problem features to which such algorithms are expected to be suitable.

- *Discrete Partitions*: The problem must facilitate discrete interaction with the sub-systems, and likely be comprised of or partitioned into multiple sub- or partial problems to which each sub-system may specialise.
- *Online Improvement*: It is desirable for the system to promote continuous iterative improvement holistically (across sub-systems) likely in response to varying information or exposure patterns, although there is less of a concern for per sub-system improvement.
- *On Demand*: Access to the acquired information and/or acquired decision making capability from across all sub-systems may be required holistically at a single point of interaction with the problem in an impromptu manner.
- *Unclear Constraints*: The constraints of inductive model generation may be unknown on unclear requiring multiple trials for generating models, and/or a meta-process to improve the model generation process.
- *Hazardous Environment*: The computational environment in which the sub-systems must operate to address the problem may be hazardous, resulting in the loss of whole sub-systems and/or sub-sets of acquired information.

8.3 Function Optimisation

Real-world optimisation problems and generalisations thereof can be drawn from most fields of science, engineering, and information technology (for a sample see [16, 402]). Importantly, optimisation problems have had a long tradition in the fields of Artificial Intelligence and Computational Intelligence in motivating basic research into new problem solving techniques, and for investigating and verifying systemic behaviour against benchmark problem instances. This section considers the *Function Optimisation Formalism* and related specialisations as a general motivating problem for demonstrating the suitability of the applicability of clonal selection algorithms from across the hierarchical framework. This is achieved firstly with a review of the problem formalism, nomenclature, and related sub-fields. Selected Artificial and Computational Intelligence research is reviewed that demonstrate feature overlap with concerns from at least one of the three clonal selection paradigms, highlighting relevant problem features that such algorithms may exploit. Finally, the Cellular, Tissue, and Host Clonal Selection Paradigms are mapped onto general cases of Function Optimisation problems to which they are suited.

8.3.1 Function Optimisation Overview

Problem Description

Optimisation (in the mathematical sense), is defined as the search for a combination of parameters commonly referred to as *decision variables* ($x = \{x_1, x_2, x_3, \dots, x_n\}$) which minimise or maximise some ordinal quantity (c) (typically a scalar called a score or cost) assigned by an *objective function* or *cost function* (f), under a set of constraints ($g = \{g_1, g_2, g_3, \dots, g_n\}$). For example, a general minimisation case would be as follows: $f(x) \leq f(x'), \forall x_i \in x$. Constraints may provide boundaries on decision variables (for example in a real-value hypercube \mathbb{R}^n), or may generally define regions of feasibility and infeasibility in the decision variable space or cost space. In applied mathematics the field may be referred to as *Mathematical Programming*. More generally the field may be referred to as *Global* or *Function Optimisation* given the focus on the objective function (for more general information on optimisation, see [222, 330, 53]).

Sub-Fields of Study

The study of optimisation is comprised of many specialised sub-fields, generally based on an overlapping taxonomy that focuses on the principle concerns in the general formalism. For example, with regard to the decision variables, one may consider univariate and multivariate optimisation problems. The type of decision variables promotes the specialities for continuous, discrete, and permutations of variables. Dependencies between decision variables under a cost function defines the fields of Linear Programming, Quadratic Programming, and Nonlinear Programming. A large class of optimisation problems can be reduced to discrete sets, which are considered in the field of Combinatorial Optimisation, to which many theoretical properties are known, most importantly that many interesting and relevant problems cannot be solved by an approach with polynomial time complexity (so-called NP-complete, for example see [311]).

The topography of the response surface for the decision variables under the cost function may be convex, which is an important class of functions to which many important theoretical findings have been made, not limited to the fact that location of the local optimal configuration also means the global optimal configuration of decisional variables has been located [53]. Many interesting and real-world optimisation problems produce cost surfaces that are non-convex or so called multi-modal² (rather than uni-modal) suggesting that there are multiple peaks and valleys. Further, many real-world optimisation problems with continuous decision variables cannot be differentiated given their complexity or limited information availability meaning that derivative-based gradient decent methods that are well understood are not applicable, requiring the use of so-called ‘direct search’ (sample or pattern-based) methods [272]. Further, real-world objective function evaluation may be noisy, non-continuous, and dynamic, and the constraints of real-world problem solving

²Taken from statistics referring to the centres of mass in distributions, although in optimisation it refers to ‘regions of interest’ in the search space, in particular valleys in minimisation, and peaks in maximisation cost surfaces.

may require a viable or approximate solution in limited time or resources, motivating the need for inductive model-generation based approaches.

Selected Examples

This section reviews a select set of approaches toward addressing optimisation problems from the field of Artificial and Computational Intelligence to both provide general insight into the state of the interaction between stochastic algorithms and the field of optimisation, and to provide a context for relating the clonal selection algorithms to optimisation problems by analogy. This section draws heavily from the field of Evolutionary Computation, Swarm Intelligence, and related Computational Intelligence sub-fields (see Section 3.4).

- *Global and Local Optimisation:* Global Optimisation refers to seeking a globally optimal structure or approximation thereof, typically in an unconstrained n -dimensional real valued space. Global is differentiated from Local Optimisation in that the latter focuses on locating an optimal structure within a constrained region of the decision variable search space, such as a single peak or valley (basin of attraction). In AI and CI literature, global optimisation problems refers to the class of optimisation problems that generally cannot be addressed through more conventional approaches such as gradient decent methods (that require derivatives), and pattern search (that can get ‘stuck’ in local optima and/or may never converge) [329, 402]. Such problems are typically addressed through a directed search strategy with stochastic elements, typically involving a population-based model in the case of global search (for example a GA, see Section 3.4.1), and a mutation hill climbing-based model for local search (see Section 3.4.2). The general strategy involves sampling broadly initially, and converging on those areas of the search space that are most likely to payoff toward the end of the search. It is common to apply a local search method to the solutions of a global search procedure as an refinement strategy.
- *Parallel Optimisation:* A natural first step in addressing difficult (large and rugged cost landscapes) is to exploit parallel and distributed hardware, to get an improved result in the same amount of time, or the same result in less time, or both [85]. Towards unifying the myriad of approaches and hardware configurations, a general consensus and taxonomy has been defined by the Parallel Evolutionary Algorithms (PEA) and Parallel Metaheuristics fields that considers the ratio of *communication* to *computation* called *granularity* [68, 13]. This taxonomy is presented concisely by Alba and Tomassini as a plot or trade-off of three concerns: (1) the number of sub-populations (models or parallel strategies working on the problem), (2) the coupling between the sub-populations (frequency and amplitude of communication), and (3) the size of the sub-populations (size or extent of the sub-models) [14]. Two important and relevant findings from the narrower field of Parallel Evolutionary Algorithms include (1) that tightly coupled (frequent migration) between coarse-grained models typically results in worse performance than a non-distributed approach [15], and (2) that loose coupling (infrequent migration) between coarse-grained models has been

consistently shown to provide a super-linear increase in performance [12, 35, 68].

- *Cooperative Search*: This is a more general approach that considers the use of multiple models that work together to address a difficult optimisation problems. Durfee, et al. consider so-called *Cooperative Distributed Problem Solving (CDPS)* in which a network of loosely coupled solvers are employed to address complex distributed problems. Specifically, such systems are desirable to match the processing capabilities of the solver with regard to the attributes of the problem. For example, a given problem may have spatially distributed, functionally distributed, or temporally distributed sub-problems to which a centralised and monolithic system may not be suitable. Lesser considers CDPS and proposes such models perform *distributed search* on dependent or independent and potentially overlapping sub-problems as a motivating perspective for conducting research into Distributed Artificial Intelligence (DAI)³ [271]. Lesser points out that in real world applications, it is hard to get a optimal mapping between the allocated resources to the needs or availability of information for a given problem, suggesting that such problems may be caused by a mismatch in processing times and/or number of sub-problems, interdependencies between sub-problems, and local experts whose expertise cannot be effectively communicated. For a more detail on the relationships between parallel and cooperative search, El-Abd and Kamel provide a rigorous taxonomy [134].
- *Hybrid Search*: Hybrid Search is a perspective on optimisation that focuses on the use of multiple and likely different approaches either sequentially (as in the canonical global and local search case), or in parallel (such as in Cooperative Search). For example in this latter case, it is common in the field of PEA to encourage different levels of exploration and and exploitation across island populations by varying the operators or operator configurations used [382, 3]. Talbi proposed a detailed 4-level taxonomy of Hybrid Metaheuristics that encompassed concerns of parallel and cooperating approaches [379]. The taxonomy encompasses parallel and cooperative considerations for optimisation and focuses on the discriminating features in the lowest level such as heterogeneity, and specialisation of approaches.
- *Functional Decomposition*: Three examples of a functional decomposition of optimisation include (1) multiple objectives, (2) multiple constraints, and (3) partitions of the decision variable search space. Multi-Objective Optimisation (MOO) is a sub-field that is concerned with the optimisation of decision variables under a set of objective functions. A solution to a MOO conventionally involves locating and returning a set configurations under the objective functions called the optimal or non-dominated Pareto set of solutions [124]. The complexity with MOO problems is in the typically unknown dependencies between decision variables across objectives, that in the case of conflicts, must be traded off (Purshouse and Fleming provide a taxonomy of such complexity [331]). Constraint Satisfaction Problem's (CSP) involve the optimisation of decision variables under a set of constraints. The principle

³This perspective provided the basis for what became the field of Multi-Agent Systems (MAS).

complexity in such problems is in locating structures that are feasible or violate the least number of constraints, potentially optimising such feasibility [399, 263]. Search Space Partitioning involves partitioning of the decision variable search space (for example see Multispace Search by Gu, et al [131, 193, 192]). This is a critical consideration given that for equal-sized dimensions bounds on parameters, the increase in decision variables results in an exponential increase in the volume of the space to search.

- *Availability Decomposition:* Optimisation problems may be partitioned by the concerns of temporal and spatial distribution of (1) information availability, and (2) computation availability. An interesting area of research regarding variable information availability for optimisation problems is called Interactive Evolutionary Computation, in which one or a collection of human operators dynamically interact with an optimisation process [378]. Example problem domains include but are not limited to computer graphics, industrial design, image processing, and drug design. There is an increasing demand to exploit clusters of heterogeneous workstations to complete large-scale distributed computation tasks like optimisation, typically in an opportunistic manner such as when individual machines are under utilised. The effect is that optimisation strategies, such as random partitioning of the search space (independent non-interacting processing) are required to take advantage of such environments for optimisation problems [351, 274].
- *Meta Optimisation:* One may optimise at a level above that considered in previous sections. Specifically, (1) the iterative generation of an inductive model called multiple restart optimisation, and (2) the optimisation of the parameters of the process that generates an inductive model of an optimisation problem. Multiple or iterative restarts involves multiple independent algorithm executions from different (random) starting conditions. It is generally considered as a method for achieving an improved result in difficult optimisation problems in which a given strategy is deceived by local or false optima [300, 225], typically requiring a restart schedule [158]. A second and well studied form of meta optimisation involves the optimisation of the search process itself. Classical examples include the self-adaptation of mutation parameters (step sizes) in the Evolutionary Strategy [333, 352] and Evolutionary Programming [144, 143] approaches. Smith and Fogarty provided a review of genetic algorithms with adaptive strategies providing a taxonomy in which the meta-adaptations are applied at one of three levels: (1) the population (adapting the overall sampling strategy), (2) the individual (adapting the creation of new samples in the decision variable space), and (3) components (modifying component contributions and/or individual step sizes as in ES and EP) [366].

8.3.2 Clonal Selection Applicability

This section considers the suitability of application of the clonal selection algorithms from across the Cellular, Tissue, and Host Paradigms to the general problem of Function Opti-

misation. Applicability is considered in the context of the algorithm features reviewed in Section 8.2 against the selected examples reviewed in the previous section.

Cellular Algorithms

A functional mapping of clonal selection onto the Function Optimisation formalism is as follows: *cells* represent the decision variables (or a mapping thereof) that define a structure, *antigen* is abstracted as the cost function and/or constraints imposed on the structure, the *repertoire* represents a clonal set of cells specialised (converged) toward addressing the antigen in response to the *affinity* assigned to structures by the cost function. This single-antigen mapping of clonal selection to the optimisation formalism was shown in in Section 3.3.2 to be the basis of the development and investigation of a variety of clonal selection algorithms not limited to CLONALG, BCA, and the IA family of algorithms where in each case a repertoire of cells was adapted to address a cost function. This mapping embodies the *affinity maturation principle* reviewed in Section 3.3.1 as a constrained clonal selection metaphor suitable for application to optimisation domains, specifically as a process directed toward the improvement of affinity (maturation) of cells in the context of an antigen.

The general cellular algorithms (CCSA and RCCSA) are suited to global and local optimisation in the case where the objective function is mapped to an affinity interaction with a single antigen. One may consider the procedure a local search when a single cell is used ($N_{cells} = 1$) and a multi-point global search procedure otherwise ($N_{cells} > 1$). As such, the procedure is generally less sophisticated than a an Evolutionary or Swarm Intelligence approach, and belongs to the general class of directed random search, and (point and/or population-based) stochastic-based hill climbing and steady-state genetic algorithm without crossover (see Section 3.4). The point-wise basis to exposures across all cellular algorithms promotes considerations of parallelism within the repertoire, perhaps exemplified in the SCCSA which strongly resembles the so-called fine-grained parallisation of genetic algorithms for multi-processor hardware (such as diffuse and cellular genetic algorithms [323]). The mediated (ECCSA) and network (NCCSA) algorithms do not have clear analogies in the field of optimisation. Although these approaches are immature, further development may reveal their capability in learning complex inter-relationships between decision variables or other functional decompositions such as objectives, constraints, and regions in the search space. Finally, one may consider the functional decompositions of optimisation problems and their mapping onto the foundational CCSA as either explicit sub-models or inferred partial models. The generality of these foundational approaches means that augmentation may be required to effectively manage explicit sub-problems such as MOO and CSP. A summary of the suitability of application for the Cellular Clonal Selection Algorithms to Function Optimisation is provided in Table 8.1.

Tissue Algorithms

The functional mapping of tissue algorithms to Function Optimisation is as follows: *tissues* are mapped onto different (logically or physically) global or local optimisation processes,

Algorithm	$N_{antigen}$	N_{cells}	Optimisation Strategy	Inductive Model
<i>MCCSA</i>	1	1	Local Search	Holistic
<i>R/CCSA</i>	1	N	Global Search	Holistic/Partial
<i>R/CCSA</i>	N	N	Decomposition Search	Sub/Partial
<i>SCCSA</i>	1	N	Parallel Global	Holistic/Partial
<i>E/NCCSA</i>	1	N	Decomposition Search	Partial
<i>E/NCCSA</i>	N	N	Decomposition Search	Sub/Partial

Table 8.1: Summary of the projected applicability of Cellular Clonal Selection Algorithms to Function Optimisation Problems.

recirculation is mapped onto inter-process communication and any overheads associated with such communication (for example network latency and/or memory requirements), *infections* may be mapped as single objective functions (single antigen), or comprised of a set of objective functions, constraints or other functional decompositions (multiple antigen). The *tissue exposure regime* may be influenced by a pre-designed arrangement of tissues (architecture and network topology), or may represent the interaction of the system with a known or unknown functional or logical decomposition of a given problem.

As considered in with cellular algorithms, a single tissue may be exploited for a global or local optimisation process. Multiple non-interacting tissues in the MTCSA may be considered parallel global search or parallel multiple restarts, whereas those algorithms that use recirculation (RTCSA, HTCSA, ITCSA) may be considered cooperative search parallel strategies. Given the high-frequency nature of recirculation, the RTCSA fits in somewhere between coarse and fine-grained parallel algorithms, with a moderate population size, moderate number of tissues, and fixed (directed cyclic graph) topology. Importantly the reviewed findings from parallel EA's suggest that such a configuration may perform as well as or worse than a non-distributed model (single population) for global optimisation. Interestingly, such behaviour (strong competition and system-wide takeover) may be desirable if parallel tissues cooperate using hybrid global optimisation strategies. For example, variations on cellular algorithms (such as a global and local variation of a given CCSA), or the holistic replacement of CCSA with other similar approaches (such as evolutionary or swarm intelligence algorithms). Finally, the tissue architecture provides a natural mapping for functional and availability decompositions of a given optimisation problem. For example, objectives, constraints, and search space partitions may be explicitly or randomly assigned to tissues of the algorithm, where recirculation promotes the cooperative concerns such as trade-off's between the sub-problems. The tissue architecture is suited to such problem descriptions, although this suitability may be further improved if the functional decomposition is motivated by an availability decomposition such as distributed information availability, event-driven (such as human interaction), or opportunistic allocation of computation time. Table 8.2 summarises the applicability of TCSA to Function Optimisation problems.

Algorithm	$N_{infection}$	$N_{tissues}$	Optimisation Strategy	Inductive Model
<i>MTCSA</i>	1	1	Global Search	Holistic
<i>MTCSA</i>	1	N	Parallel Global	Holistic
<i>MTCSA</i>	N	N	Parallel/Decomposition	Sub
<i>R/H/ITCSA</i>	1	N	Cooperative/Hybrid	Holistic
<i>R/H/ITCSA</i>	N	N	Cooperative/Decomposition	Sub/Partial

Table 8.2: Summary of the projected applicability of Tissue Clonal Selection Algorithms to Function Optimisation Problems.

Host Algorithms

The functional mapping of host algorithms to Function Optimisation is as follows: *hosts* may map onto different (logically or functionally) global or local search processes ($N_{tissues} = 1$ per H), or more interestingly onto groups of tightly-coupled clusters of global or local search processes ($N_{tissues} > 1$ per H). Similarly, a habitat may be a single objective function ($N_{infection} = 1$ per B), a set of objectives functions or constraints ($N_{infection} > 1$ per B), or a set of holistic although different function optimisation problems ($N_{habitats} > 1$). As with the mapping of the tissue paradigm, the complexity of the system-problem interaction is relegated to the (in this case host) exposure regime.

Given the bracketing of tissues, one may consider the extent of capabilities of the tissue algorithm subsumed by a host or a population of hosts, as was the case between the tissue and cellular paradigms. As such, if we ignore (constrain to $N_{tissue} = 1$) tissue cardinality, the MP-HCSA provides a set of independent global or local strategies. The SI-HCSA may be exploited for cooperative parallel global optimisation, given their strong resemblance in configuration to parallel evolutionary algorithms. Specifically, the small to moderate number of hosts, moderate to large population sizes, and low frequency, low amplitude sharing of information (decrease in frequency compared to experimentation in Chapter 6). Further, the sharing-based population host algorithms encourage an unstructured topology of impromptu host interactions, more akin to a P2P network topology than a master-server, star or other fixed topology commonly exploited by PEA's. The cooperative relationship of SI-HCSA may also be exploited for hybrid optimisation, perhaps with an increased frequency of migration, as was the case in the exploratory experimentation. As with tissue paradigm, the population-based host algorithms may exploit the natural modularity of optimisation problems in terms of their functional and information availability decompositions, although it is unclear whether the varied information sharing scheme of SI-HCSA would promote the sufficient level of integration of partial-models where required (dependent decompositions). Interesting, the THCSA, specifically the Pathogen and Vaccination variations promote the consideration of sharing the triggers of adaptation, in this case an explicitly partitioned or limited availability resources between nodes in the system. Where this mechanism is supported under the constraints of a specific problem. Such behaviour may promote an adaptive-self organisation of responsibility within the host population.

The generational host algorithms provide an clear relationship to the meta-optimisation techniques. MG-HCSA with a single host provides a model for a sequential restart strat-

Algorithm	$N_{habitats}$	N_{hosts}	Optimisation Strategy	Inductive Model
<i>MP-HCSA</i>	1	1	Global Search	Holistic
<i>MP-HCSA</i>	1	N	Parallel Global	Holistic
<i>SI-HCSA</i>	1	N	Parallel/Cooperative/Hybrid	Holistic
<i>SI/THCSA</i>	N	N	Cooperative/Decomposition	Sub/Partial
<i>MG-HCSA</i>	1/N	1	Sequential Iterative-Restarts	Holistic
<i>MG-HCSA</i>	1	1/N	Parallel Iterative-Restarts	Holistic
<i>MI-HCSA</i>	1	1/N	Multiple Restart/Meta	Holistic
<i>EI-HCSA</i>	1	N	Meta-Optimisation	Holistic

Table 8.3: Summary of the projected applicability of Host Clonal Selection Algorithms to Function Optimisation Problems.

egy, and a parallel restart strategy when configured with multiple hosts. The MI-HCSA algorithm provides a restart strategy in which the initialisation of naïve systems is biased with acquired information. This strategy may be effective in environments where the availability is subject to frequent and unexpected change to seed new instances of host processes (such as in the exploitation of idle computation time in a network environment). Finally, the suggestion of optimising the initial conditions in the EI-HCSA provides a link between the restart strategy in GHCSA and meta optimisation strategies that optimise the processes that operate on a given problem instance. Importantly the meta optimisation strategies (restarts and optimisation of process) apply not only to all cellular and tissue approaches to which the hosts subsume, but also to the capabilities of the population-based host algorithms. Table 8.3 summarises the applicability of HCSA to function optimisation problems.

8.3.3 Summary

This section summarises the insight provided into the applicability of the clonal selection algorithms from across the three paradigms generally with regard to Function Optimisation, and specifically with regard to some specialisations of the problem and techniques for addressing the problem.

1. *General Suitability*
 - (a) *Mapping*: The mapping between the general clonal selection adaptive strategy and function optimisation is natural for the specific case of blindly adapting a clonal set for a single antigen.
 - (b) *Systems*: Beyond the adaptive cellular level, the tissue and host paradigms provide a basis for decentralised and distributed optimisation, given a suitable problem mapping.
 - (c) *Environments*: Global optimisation alone is insufficient to motivate application to tissue and host paradigms, requiring functional and availability decompositions of a given optimisation problem to provide an effective mapping of the problem to the antigenic exposure paradigm.

2. *Algorithm Applicability*

- (a) *Cellular*: R/CCSA are suitable for local and global optimisation tasks, potentially for functional decompositions of optimisation problems, and SCCSA is suitable for parallelisation.
- (b) *Tissue*: MTCSA is suitable for parallel global optimisation, R/H/ITCSA are suitable for cooperative hybrid and potentially functional decomposition, and likely not suitable for cooperative global search.
- (c) *Host*: MP-HCSA is suitable for parallel global search, SI-HCSA is suitable for parallel global search and likely cooperative hybrid and functional decomposition, THCSA variants are suitable for functional decompositions and potentially cooperative parallel optimisation. MG-HCSA is suitable for multiple restarts, MI-HCSA may be suitable for biased multiple restarts, and EI-HCSA suggests at meta optimisation with further extension.

8.4 Function Approximation

The phrasing of real-world problems in the Function Approximation formalism are among the most computationally difficult considered in the broader field of Artificial Intelligence for reasons including: incomplete information, high-dimensionality, noise in the sample observations, and non-linearities in the target function. This section considers the *Function Approximation Formalism* and related specialisations as a general motivating problem for demonstrating the suitability of the applicability of clonal selection algorithms from across the hierarchical framework. Firstly the general problem is reviewed including standard nomenclature and a summary of sub- and related fields of study. Selected Artificial and Computational Intelligence research is reviewed that demonstrate feature overlap with concerns from at least one of the three clonal selection paradigms, highlighting relevant problem features that such algorithms may exploit. Finally, the Cellular, Tissue, and Host Clonal Selection Paradigms are mapped onto general cases of Function Approximation problems to which that are proposed to be well suited.

8.4.1 Function Approximation Overview

Problem Description

Function Approximation (in the mathematical sense) is an *inductive problem* of finding a function (f) that approximates a target function (g), where typically the approximated function is selected based on a sample of observations (x , also referred to as the *training set*) taken from the unknown target function. In machine learning, the function approximation formalism is used to describe general problem types commonly referred to as *pattern recognition*, such as classification, clustering, and curve fitting (so-called decision or discrimination function). Specifically, such general problem types are described in terms of approximating an unknown Probability Density Function (PDF), which underlies the relationships in the problem space, and represented to some degree in the sample data. This function approximation perspective of such problems is commonly referred to as *statistical machine learning* and/or density estimation [159, 46].

Sub-Fields of Study

The function approximation formalism can be used to phrase some of the hardest problems faced by Computer Science, and Artificial Intelligence in particular such as natural language processing and computer vision. The general process focuses on (1) the collection and preparation of the observations from the target function, (2) the selection and/or preparation of a model of the target function, and (3) the application and ongoing refinement of the prepared model. Some important problem-based sub-fields include: *Feature Selection* where a feature is considered an aggregation of attributes, where only those features that have meaning in the context of the target function are necessary to the modelling process [262, 194], *Classification* where observations are inherently organised into labelled groups (classes) and a supervised process models an underlying discrimination function to classify unobserved samples, *Clustering* where observations may be organised into inherent groups based on common features although the groups are unlabelled requiring a process to model an underlying discrimination function without corrective feedback, and *Curve or Surface Fitting* where a model is prepared that provides a ‘best-fit’ (or regression) for a set of observations that may be used for *interpolation* over known observations and *extrapolation* for observations outside what has been observed.

The field of Function Optimisation is related to Function Approximation, as many sub-problems of Function Approximation may be defined as optimisation problems. As such many of the inductive modelling paradigms are differentiated based on the representation used and/or the optimisation process used to minimise error or maximise effectiveness on a given approximation problem. The difficulty of Function Approximation problems centre around (1) the nature of the unknown relationships between attributes and features, (2) the number (dimensionality) of attributes and features, and (3) general concerns of noise in such relationships and the dynamic availability of samples from the target function. Additional difficulties include the incorporation of prior knowledge (such as imbalance in samples, incomplete information and the variable reliability of data), and problems of invariant features (such as transformation, translation, rotation, scaling and skewing of features).

Selected Examples

This section reviews a select set of approaches toward addressing Function Approximation problems from the field of Artificial and Computational Intelligence to both provide general insight into the state of the interaction between stochastic algorithms and the field, and to provide a context for relating the clonal selection algorithms to Function Approximation problems by analogy. The review draws heavily from the fields of artificial neural networks, specifically competitive learning, as well as related inductive machine learning fields such as instance based learning.

- *Vector Quantisation*: Vector Quantisation (VQ) refers to a method of approximating a target function using a set of exemplar (prototype or codebook) vectors. The exemplars represent a discrete sub-set of the problem, generally restricted to the

features of interest using the natural representation of the observations in the problem space, typically an unconstrained n -dimensional real valued space. The VQ method provides the advantages of a non-parametric model of a target function (like instance-based and lazy learning such as k NN, see Section 3.4) using a symbolic representation that is meaningful in the domain (like tree-based approaches). The promotion of compression addresses the storage and retrieval concerns of k NN, although the selection of codebook vectors (the so-called quantisation problem) is a hard problem that is known to be NP-complete [162]. More recently Kuncheva and Bezdek have worked towards unifying quantisation methods in the application to classification problems, referring to the approaches as Nearest Prototype Classifiers (NPC) and proposing a generalised nearest prototype classifier [264, 265].

- *Parallelisation*: Instance-based approaches are inherently parallel given the generally discrete independent nature in which they are used, specifically in a case or per-query manner. As such, parallel hardware can be exploited in the preparation of the corpus of prototypes (parallel preparation), and more so in the application of the corpus given its read-only usage [1, 301, 326]. With regard to vector quantisation specifically, there is an industry centred around the design and development of VQ and WTA algorithms and circuits given their usage to compress digital audio and video data [302, 312].
- *Cooperative Methods*: Classical cooperative methods in the broader field of statistical machine learning are referred to as *Ensemble Methods* [310, 327] or more recently *Multiclassifier Systems* [171]. *Boosting* is based on the principle of combining a set of quasi-independent weak learners that collectively are as effective as a single strong learner [244, 349]. The seminal approach is called Adaptive Boosting (AdaBoost) that involves the preparation of a series of classifiers, where subsequent classifiers are prepared for the observations that are misclassified by the preceding classifier models (creation of specialists) [350]. *Bootstrap Aggregation* (bagging) involves partitioning the observations into N randomly chosen subsets (with reselection), and training a different model on each [55]. Although robust to noisy datasets, the approach requires careful consideration as to the consensus mechanism between the independent models for decision making. *Stacked Generalisation* (stacking) involves creating a sequence of models of generally different types arranged into a stack, where subsequently added models generalise the behaviour (success or failure) of the model before it with the intent of correcting erroneous decision making [434, 394].
- *Functional Decomposition*: As demonstrated, it is common in ensemble methods to partition the dataset either explicitly or implicitly to improve the approximation of the underlying target function. A first important decomposition involves partitioning the problem space into sub-spaces based on the attributes, regular groups of attributes called features, and decision attributes such as class labels. A popular method for attribute-based partitioning is called the Random Subspace Method, involving the random partitioning of attributes to which specialised model is pre-

pared for each (commonly used on tree-based approaches) [209]. A related approach involves a hierarchical partitioning of attributes space into sub-vectors (sub-spaces) used to improve VQ-based compression [170]. Another important functional decomposition involves the partitioning of the set of observations. There are many ways in which observations may be divided, although common approaches include pre-processing using clustering techniques to divide the set into natural groups, additional statistical approaches that partition based on central tendency and outliers, and re-sampling methods that are required to reduce the volume of observations.

- *Availability Decomposition*: The availability observations required to address function approximation in real-world problem domains motivate the current state of the art in Distributed Data Mining (DDM or so-called Collective Data Mining), Parallel Data Mining (PDM), and Distributed Knowledge Discovery in Database (DKDD) [239]. The general information availability concerns include (1) the *intractable volume of observations*, and (2) the *spatial (geographical) and temporal distribution of information* [444]. It is common in many large real-world problems for it to be infeasible to centralise relevant observations for modelling, requiring scalable, load balancing, and incremental acquisition of information [362].
- *Meta Approximation*: The so-called ensemble or multiple-classifier methods may be considered meta approximation approaches as they are not specific to a given modelling technique. As with function optimisation, meta-approaches may be divided into restart methods and meta-learning algorithms. The use of restart methods is a standard practice for connectionist approaches, and more generally in approaches that use random starting conditions and a gradient or local search method of refinement. The method provides an opportunity for over-coming local optima in the error-responds surface, when there is an unknown time remaining until convergence [278], and can exploit parallel hardware to provide a speed advantage [47]. Ensemble methods and variants are examples of meta approximation approaches, as well as the use of consensus classifiers (gate networks in mixtures of experts) to integrate and weight the decision making properties from ensembles.

8.4.2 Clonal Selection Applicability

This section considers the suitability of the application of the clonal selection algorithms from across the Cellular, Tissue, and Host Paradigms to the general problem of Function Approximation. Applicability is considered in the context of the algorithm features reviewed in Section 8.2 against the selected examples reviewed in the previous section.

Cellular Algorithms

A generalised mapping of clonal selection onto the Function Approximation formalism is as follows: *cells* are exemplars (a.k.a. codebooks or prototypes) in the parameter space (or a mapping thereof) of the target function, *antigen* are the sample of observations from

the target function (therefore there are always N), the *repertoire* represents an approximation of the target function adapted in the presence of the sample observations, and *affinity* is calculated based on a distance metric between observations and exemplars. As considered in the review of the state of clonal selection algorithm research (Section 3.3.2), this was the perspective used in the development and application of AIRS to classification where exemplars quantised observations. This mapping was also used in the application of CLONALG to binary pattern recognition, and was the basis for the colour space problems investigated across the three clonal selection paradigms in this work. As considered in paradigms related to the field of clonal selection algorithms in Section 3.4, the algorithms competition-based approach to preparing the exemplars in response to the sample observations is related to the field of *competitive learning*, whereas the interpretation of the resulting exemplars is related to the field of *instance-based learning*, specifically k -Nearest Neighbours (k NN).

The cellular clonal selection algorithms provide a series of global optimisers suitable for vector quantisation, not unlike the use of EC, SA, and hill climbing for the same task. The general nature of VQ makes the cellular algorithms suitable for classification (Nearest Prototype Classifiers), clustering, and curve fitting. Given the competition-based interaction between exemplars and the unknown context-based relationships that such competition represents, it is reasonable to categorise the cellular clonal selection algorithms as the induction of partial-models. Although it is important to point out that it is the degree to which such relationships have meaning in the problem space that defines whether the exemplars may be considered partial (strongly related) or sub (weakly or unrelated) models of the broader function approximation problem. The iterative (per-exposure) competition based learning and decision making suggest at the general suitability of the approach to incremental and dynamic specialisations of the general problem. The spatial properties of SCCSA show similarities with projection-based clustering methods such as SOM and Neural Gas, although unlike those approaches SCCAS in its present form does not demonstrate the preservation of topological features. As with function optimisation, the role of ECCAS and NCCAS is not clear given the approaches immature although demonstrated similarities with sub-symbolic feature-based approaches such as Learning Classifier System. The VQ approach to function approximation highlighted an important perspective on negative selection algorithms and their relationship to clonal selection algorithms. Specifically, that such approaches may be considered in the context of a function approximation formalism (the standard domains of anomaly and change detection as discriminating functions), where the sub-string approaches used (hamming distance, r-chunks, and others) may be considered the simultaneous identification and extraction of features and instance-based modelling of the underlying PDF. Table 8.4 summarises the suitability of the Clonal Selection Algorithms for general variations of the Function Approximation formalism.

Tissue Algorithms

The mapping of tissue algorithms onto the Function Approximation formalism is as follows: *tissues* represent individual vector quantisation models each with a set of exemplars,

Algorithm	$N_{antigen}$	N_{cells}	Approximation Strategy	Inductive Model
<i>R/CCSA</i>	N	N	Vector Quantisation	Sub/Partial
<i>SCCSA</i>	N	N	Vector Quantisation	Partial
<i>E/NCCSA</i>	N	N	Feature Extraction	Partial

Table 8.4: Summary of the projected applicability of Cellular Clonal Selection Algorithms to Function Approximation Problems.

recirculation represents inter-process communication (parallel and/or distributed) between the quantisation models, *infections* represent logical or functional groups of observations from the target function, and the *tissue exposure regime* represents the mapping of the logical or function partitions of the observations to the individual models where the mapping may be known: designed allowing tissue-specialisation on sub-problems, or unknown: based on availability decompositions forcing tissue-specialisation on partial-problems.

The MTCSA provides a candidate for ensemble methods given the required independence between such models, although the sharing of exemplars between the ensemble in the case of recirculation algorithms likely makes a mapping for the R/H/ITCSA unsuitable for these specific cooperate approaches. Specifically, the tissue topology and the MTCSA has a strong resemblance to the ‘bagging’ and potentially the ‘boosting’ ensemble methods. More generally, the TCSA approaches may be used for functional decompositions of the problem. The persistent communication and information sharing in the recirculation algorithms strongly suggests suitability of the approaches as a distributed vector quantisation method beyond efficiency concerns of parallelisation (such as Parallel-AIRS), in particular for those real-world problems with information and communication availability constraints (considered in DDM and DKDD). This suitability is likely bounded to cases where adaptive localisation and dissemination are required, for example where there is an unknown partitioning of the observations (availability of information) and the use and/or exploitation of the decision making capability (such as use of the approximated function for discrimination) also occurs in an unknown and partitioned (localised) manner. This suitability scenario may be considered a general application-case of Hofmeyr’s proposed use of his negative selection system (later ARTIS and LISYS) in distributed anomaly detection (see Section 2.5.2), although constrained to a localised ring network topology of machines. Table 8.5 provides a summary of the suitability of tissue clonal selection algorithms to function approximation.

Algorithm	$N_{infection}$	$N_{tissues}$	Approximation Strategy	Inductive Model
<i>MTCSA</i>	N	N	Cooperative/Decomposition	Sub
<i>R/H/ITCSA</i>	N	N	Cooperative/Decomposition	Partial

Table 8.5: Summary of the projected applicability of Tissue Clonal Selection Algorithms to Function Approximation Problems.

Host Algorithms

The functional mapping of host algorithms is as follows: *hosts* may be considered collections of highly integrated vector quantisation models or large (bracketed) models, *host interaction* may be considered the inter process, inter system, or inter-network communication involving either the sharing of model information or information to which models are exposed, *habitats* represent different perspectives of the observations in terms of functional or availability decompositions, where the *host exposure regime* defines the designed or opportunistic mapping between such decompositions and the system.

The population of hosts provide the similar capabilities as the tissue algorithms, although scaled up. In particular the MP-HCSA provides a broader interpretation of ensemble methods considering the hierarchical formulation of such approaches. Importantly the SI-HCSA provides a less rigid communication mechanism between VQ models, that permits a more directed and/or rapid replication of acquired information more akin to ARTIS in a broadcast environment, and peer-to-peer propagation. This is also related to epidemic and gossip based replication of information in distributed databases which host models could adopt toward predictable efficiency gains [127, 413]. The generational algorithms exploit the restart property inherent in the global optimisation algorithm that underlies the adaptive capability of the approach. As such MI-HCSA provides general bias of ensembles and EI-HCSA provides general adaptation of the process within ensembles. Table 8.6 provides a summary of the suitability of the Host Clonal Selection Algorithms to Function Approximation.

Algorithm	$N_{habitats}$	N_{hosts}	Approximation Strategy	Inductive Model
<i>MP-HCSA</i>	N	N	Cooperative/Decomposition	Sub
<i>SI/THCSA</i>	N	N	Cooperative/Decomposition	Partial
<i>MG-HCSA</i>	N	1	Sequential Iterative-Restarts	Sub
<i>MG-HCSA</i>	N	N	Parallel Restarts	Sub
<i>MI-HCSA</i>	N	N	Multiple Restart/Meta	Holistic
<i>EI-HCSA</i>	N	N	Meta-Optimisation	Holistic

Table 8.6: Summary of the projected applicability of Host Clonal Selection Algorithms to Function Approximation Problems.

8.4.3 Summary

This section summarises the insights provided into the suitability of the applicability of the clonal selection algorithms from across the three paradigms, generally with regard to Function Approximation, and specifically with regard to some specialisations of the problem and techniques for addressing the problem.

1. General Suitability

- (a) *Mapping*: The mapping of clonal selection to vector quantisation based function approximation provides a natural fit, that may only be further improved through the parallel task of feature selection.

- (b) *Systems*: Clonal selection provides an optimisation process over the scope of managed exemplars, that provides a natural parallel realisation of the mapping between the process and function optimisation. As such, the system is always comprised of > 1 cell.
- (c) *Environments*: The set of observations from the target function and the iterative exposure and adaptation or application in response to such observations embodies the principles and provide a natural mapping for onto the antigenic exposure paradigm.

2. Algorithm Applicability

- (a) *Cellular*: R/CCSA provides a competitive learning realisation of vector quantisation applicable for compression, clustering, and nearest prototype classification.
- (b) *Tissue*: MTCSA provides a natural suitability for ensemble methods, bagging in particular, whereas the recirculation algorithms (R/H/ITCSA) are suited to availability decompositions of problems, in particular information availability concerning observations from the target function, and the unknown and point-wise application of the approximated function across the tightly-coupled deployment (DDM and DKDD).
- (c) *Host*: MP-HCAS provides a broader scoped suitability for parallel and hierarchical ensemble methods, where as the SI/THCSA are suited to less constrained topologies for DDM and KDD with both model and observation information sharing. Generation HCSA are suitable for restart and iterative tuning of the model and process respectively, both of which are concerns of the underlying stochastic-optimisation process.

8.5 Chapter Summary

This section reviews the findings from this chapter regarding the investigated clonal selection algorithms suitability and applicability, specifically in the context of Function Optimisation and Function Approximation.

8.5.1 Suitability and Applicability

General Suitability Section 8.2 considered the general suitability of the clonal selection algorithms, specifically with regard to the systematic elicitation of features from the general algorithms (cellular, tissue, host) from across the hierarchical clonal selection framework. The general suitability and projection onto problem attributes is summarised as follows:

1. *Cellular Algorithms*: Stochastic-adaptive algorithms that use a population-based pattern recognition as a method for induction and modelling of a problems. Suited to

problems that facilitate discrete interaction, that provide noisy although somewhat regular signals with unknown levels of complexity likely in an online manner.

2. *Tissue Algorithms*: Distributed model of a domain using a decentralised adaptation across multiple tightly coupled populations with high-frequency communication promoting integration of acquired information. Suited to noisy complex problems with regularities that are both spatial as well as temporal, operating in a potentially hazardous and/or opportunistic computational environment.
3. *Host Algorithms*: Distributed model of a domain using a set of loosely-coupled populations each specialising in an aspect of the problem with low frequency communication and long term strategies including model restarts and adaptation of internal processes. Suited to large and complex problems requiring online continuous improvement, likely with unclear constraints and hazards.

Problem Applicability Sections 8.3 and 8.4 considered the applicability of the clonal selection algorithms in the context of the Function Optimisation and Function Approximation problem formalisms respectively, highlighting sub-problems to which specific algorithms are suggested to be suited. The general mapping of the clonal selection paradigm and applicability to the respective problems is summarised as follows:

1. *Optimisation*: The adaptive qualities of clonal selection, specifically a clone (group) of cells in response to an antigenic stimulus provides a natural mapping to an adaptive strategy for function optimisation. This is a reduction of the capability and complexity of the general strategy, which may be further exploited through functional and/or information availability decompositions of a given problem instance, both improving the mapping and employing more available problem specific information.
2. *Approximation*: The adaptive qualities of clonal selection considered in a parallel and decentralised where a system is presented with a large set of stimuli to which to model, provides a natural mapping to a decentralised adaptation strategy for exemplar and pattern recognition based function approximation.

8.5.2 Integration

This final chapter provides closure on the adaptive and distributed conceptual models and clonal selection algorithms proposed and investigated in this work by grounding their salient features to specific problem formalisms and sub-problems. Although the claims of suitability and applicability are reasoned from the qualitative observations from previous chapters, such claims require verification on specific standard benchmark and real-world problem instances. This clear and important extension of this investigation is considered in conjunction with further extensions, limitations, and criticisms in the following Chapter that concludes this work.

Chapter 9

Conclusions

9.1 Research Goals

This work has demonstrated that the interpreted structure and function of the immune system do provide an effective motivating metaphor for the realisation of distributed clonal selection algorithms. This section reviews the contributions of this dissertation in the context of the research hypothesis, defined by five goals outlined in Section 1.3.

1. Identify a systematic methodology for realising a novel biological inspired computational framework and models. A systematic methodology was outlined in Section 2.7, comprised of (1) an immunological and information processing centric framework for progressing from metaphor, model, strategy, and algorithm, (2) an experimental framework for empirical verification and demonstration of computational models and algorithms, and (3) an adaptive algorithm centric framework for the integration of qualitative and quantitative information regarding so-called ‘conceptual machines’. This methodology was applied at three different and inter-related scales: (1) *General Metaphor*: The broad consideration of the information processing properties of the clonal selection theory, investigated in the context of three specific and distinct interpretations called paradigms. (2) *Specific Metaphors*: The consideration of computational paradigms inspired by a specific immunological metaphors that constrain the clonal selection process, which are investigated through instantiations of specific models and algorithms. (3) *Specific Algorithms*: The consideration of a specific algorithm in the context of a broader paradigm which is described in terms of its further constrained metaphor, information processing strategy, general model, algorithmic procedure instantiation, and empirical verification and demonstration of expected behaviours.

2. Identify limitations with and elaborate upon the base clonal selection paradigm. The Clonal Selection Theory was considered in depth in Section 3.2 providing a context for effectively interpreting the motivations for the state of the art in clonal selection algorithms. More importantly, the depth of immunological background provided insight into the constraints that were imposed in realising the classical and state of the art algorithms. The review of clonal selection algorithms in Section 3.3 highlighted

three main limitations, specifically: (1) *Lack of explicit consideration of adaptive qualities*: Information processing inspired by structures and functions of the clonal selection theory were traditionally considered adaptive implicitly given the theories relationship with evolutionary theory and resultant evolutionary algorithms. (2) *Lack of positioning in the broader field*: Consideration of inspired algorithms was traditionally limited to related algorithms in the field of Artificial Immune Systems with some consideration of the related approaches in the field of Evolutionary Computation, broader consideration in computational and machine intelligence was rarely considered. (3) *Lack of consideration beyond the cellular*: Clonal selection algorithms, as well as the majority of the broader Artificial Immune System algorithms are concerned primarily with the structure and function of the immune system at the cellular level, specifically given the cellular (lymphocyte immune cell) and molecular (antibody protein) focus of the principle inspiring theories.

Each of these three limitations were address in turn, specifically: Section 3.4 considered clonal selection in the broader context of Computational and Machine Intelligence, strengthening the relationship with Evolutionary Computation and Lazy Learning, and highlighting the relationship with Hill Climbing algorithms and Competitive Learning. Section 3.5 considered clonal selection as an adaptive strategy, phrasing the approach in Holland's adaptive systems formalism. Section 3.6 outlined an agenda for the investigation and elaboration of the broader clonal selection paradigm, including: the *Cellular Clonal Selection* perspective involving the investigation of the classical paradigm pursued in Chapter 4, the new *Tissue Clonal Selection* perspective investigated in Chapter 5, and the new *Host Clonal Selection* perspective investigated in Chapter 6.

3. Identify immunological structures and/or functions which clearly exhibit distributed information processing. An important motivation for the project was the identified promise of decentralised, autonomous, and distributed information processing from in the broader field of Artificial Immune Systems reviewed in Section 2.5. The review demonstrated that although such approaches had been proposed, few investigated the intersection of clonal selection and distributed information processing explicitly. Section 3.6 identified two distinct perspectives of clonal selection that exhibited distributed information processing properties: clonal selection across the tissues of a host organism, and clonal selection across the hosts in a population of organisms. The *Tissue Clonal Selection Paradigm* was investigated in Chapter 5, involving a review of the motivating metaphor in the physiology of the lymphatic system and the behaviour of migrating lymphocytes. The underlying strategy of Tissue Clonal Selection Algorithms was defined as the localised although decentralised acquisition and application of information with continuous dissemination toward general system capability. Toward this end, investigation of tissue algorithms focused on the localisation and dissemination of information in unknown distributed exposure scenarios. The *Host Clonal Selection Paradigm* was investigated in Chapter 6, providing a review of the motivating metaphor in the immunisation and evolutionary pressures on the immune system. The underlying trend of Host Clonal Selection algorithms was defined as similar to that of Tissue Algorithms involving the localisation and dissemination of information, with the increase in localisation providing

host-specialisation, and the variation of disseminated giving hosts control over what information is disseminated, and when. Toward this end, investigation of host algorithms also focused on the information acquisition and application capabilities under a range of unknown distributed exposure scenarios.

4. Study the behaviours, capabilities, and limitations of devised computational models. A model problem domain and problem instances were defined in Section 4.2.3 called the Colour Space Domain suitable for testing and assessing clonal selection algorithms with properties of optimisation and pattern recognition. Each clonal selection paradigm was phrased in terms of (1) abstract models that outlined the motivating structural and/or functional metaphor and general computational concerns (Sections 4.2, 5.3, and 6.3 for the cellular, tissue, and host paradigms respectively), and (2) realised algorithms, problems, and measures that outlined the specific system behavioural and capability expectations (Sections 4.3, 5.4, and 6.4 for the cellular, tissue, and host paradigms respectively). Clonal Selection Algorithm behavioural capabilities and expectations were systematically demonstrated and verified empirically with specific algorithm instantiations on specific problem instance scenarios across the three paradigms. Importantly, the verification process bounded the claims of expected capability, limiting them to specific algorithms and/or problem scenarios. The suitability of application of the general paradigms was assessed in Chapter 8. This involved a distillation of each approach into distinctive information processing features based on observed and verified behaviours (Section 8.2). The information processing features were then systematically mapped onto specialised cases of Function Optimisation (Section 8.3.2) and Function Approximation (Section 8.4.2) problem domains.

5. Propose an integrated clonal selection framework that unites the base and distributed variants. The clonal selection paradigm was partitioned in section Section 3.6 into that of ‘*systems*’ and ‘*environments*’, where clonal selection governed immune systems were mapped to ‘systems’, and the scope of antigen and pathogen were mapped to ‘environments’ (elaborated initially in Section 4.2.2). The clonal selection (systems) partition was reduced into the three distinct paradigms, each of which was investigated. Chapter 7 considered the reconstruction of the three paradigms and the system-environment partition into a hierarchical framework. A Hierarchical Clonal Selection Framework was defined in Section 7.2 that aggregated the clonal selection algorithms from across the three paradigms, highlighting the specific constraints imposed on the quintessential information processing procedure. A Hierarchical Antigenic Exposure Framework was defined in Section 7.3 that aggregated the contrived antigenic environments from across the three paradigms, highlighting the commonality of piece-wise information exposure, and the common source of complexity in the cardinality and spatial temporal patterns of exposure, which are compounded across the paradigms. The frameworks were combined into an Integrated Hierarchical Clonal Selection Framework in Section 7.4, uniting the system and environmental concerns, suggesting at additional perspectives, and accounting for the bracketing of complexity in each specific clonal selection paradigm. Finally,

the adopted partition-reduction methodology was reviewed and its application considered in the broader context in the field of Artificial Immune Systems, and as a general tool for investigating related Computational Intelligence fields of study (Section 7.5).

9.2 Contributions

This section reviews the contributions as standalone conceptual and computational artefacts. Artefacts are defined as discrete and differentiable objects produced and qualified within this investigation that may draw from and used beyond the work.

9.2.1 Computational Artefacts

A computational artefact contribution is a defined formalism that may be exploited beyond this investigation for computation, including tools such as metaheuristics or solvers and benchmark problems.

1. *Cellular Clonal Selection*: A generic paradigm for unifying existing cellular-based clonal selection algorithms (Chapter 4), and promoting the development of novel algorithms based on inter-cellular interactions, specifically the Spatial Clonal Selection Algorithm (Section 4.5), Mediated Clonal Selection Algorithm (Section 4.6), and the integration of the network metaphor embodied in the Network Clonal Selection Algorithm (Section 4.7). These algorithms are generally suited to global function optimisation and vector quantisation based function approximation using a holistic inductive model of a given problem (Sections 8.3.2 and 8.4.2).
2. *Tissue Clonal Selection*: A distributed clonal selection paradigm inspired by the lymphatic system and lymphocyte recirculation (Chapter 5), including the following specific decentralised algorithms: the Recirculation (Section 5.5), Homing (Section 5.6), and Inflammation (Section 5.7) Tissue Clonal Selection Algorithms. These algorithms are generally suited to cooperative function optimisation and approximation with hybrid approaches and/or partial inductive models based on functional decompositions of the problem (Sections 8.3.2 and 8.4.2).
3. *Host Clonal Selection*: A distributed clonal selection paradigm inspired by populations of organisms with immune systems undergoing immunisation and evolution (Chapter 6), including the following specific decentralised algorithms: Transmission (Section 6.5) and Shared Immunity (Section 6.6) Population Host Clonal Selection Algorithms generally suited to parallel and cooperative function optimisation and ensemble function approximation, and the Maternal (Section 6.7) and Evolved Immunity (Section 6.8) Generational Host Clonal Selection Algorithms generally suited to multiple restart and adaptation of process for both optimisation and function approximation problems (Sections 8.3.2 and 8.4.2).

9.2.2 Conceptual Artefacts

A conceptual artefact contribution is a defined formalism that may be exploited beyond this investigation for structuring practices, including methodologies and frameworks.

1. *Systematic Methodology*: An integrated set of practices derived from best practices for realising biologically inspired computational intelligence models from metaphor to computational artefact, including an immunological-centric conceptual framework (Section 2.7.1), an experimental framework (Section 2.7.2), an engineering-centric integration framework (Section 2.7.3), a partition and reduction framework for investigation a specific information processing strategy (Sections 3.6 and 7.5.1), and a feature-centric suitability framework for assessing applicability, comparability, and transferability (Section 8.2.1).
2. *Hierarchical Framework*: A conceptual scaffold were the capabilities and complexity of a given constrained perspective of the information processing principle is subsumed by the successive level (Chapter 7), used to integrate clonal selection algorithms (Section 7.2), antigenic exposure problems (Section 7.3), integration of these two frameworks into a unified hierarchical model of clonal selection (Section 7.4), with suggested broader applicability in for other immunological-based computational intelligence approaches (Sections 7.5.2).

9.3 Limitations

Limitations are observed as decision points with alternative options, each of which represents a trade-off in terms of the contributions made from the investigation in the context of the hypothesis and research goals. This section critically assesses the limitations of this work, specifically in the context of design and methodological constraints imposed on the investigation, as well as the alternatives, benefits, and restrictions they provide. The limitations of three specific decision points are considered: (1) the method of biological interpretation and transition to computational intelligence algorithm, (2) the focus in the chosen method of algorithm development, and (3) the methodology adopted to verify and demonstrate expected behaviours and capabilities of developed algorithms.

9.3.1 Biological Interpretation

The motivating immunological metaphor for the Cellular, Tissue, and Host Clonal Selection Paradigms was based on a qualitative interpretation of the structure and function of the acquired immune system at different scales resulting in generalised models and algorithms. A problem with this approach is that the behavioural expectations of realised models was also qualitative, and therefore difficult to assess and verify. An alternative approach which is common in the field of Artificial Immune Systems is to use *mathematical models* of immunological function as a basis for computational intelligence models and algorithms, for example the shape-space and affinity landscape geometric formalisms that

underlie the principles of much of the field (Section 3.2.2). This alternative reduces up-front effort by exploiting interpretations made by theoretical immunologists and produces specific rather than general approaches with a clear mathematical foundation. The strong specialisation of such approaches likely limits the broader applicability of the approaches, and the inspiring models are likely bounded by immunological dogma (for example self-nonsel self discrimination) which may or may not be a useful perspective in computational intelligence modelling and algorithm development. A specific and relevant example from this work is the tissue paradigm which suggests at multiple tissues communicating in series (general), whereas the alternative may be to base the organisation of tissues on a specific model of the lymphatic system [19], or a communication mechanism based on mathematical models of lymphocyte trafficking [371, 370, 369]. The qualitative interpretation approach taken in this work was chosen as it facilitates the conception of broader fields of study (paradigms) which may be useful in and of themselves, and which may be specialised at a later time with restrictive mathematical models.

9.3.2 Approach Development

The development of the clonal selection algorithms was focused on the strategy and the constraints to its information processing properties resulting in the broadly conceived algorithms and paradigms for grouping algorithms without an explicit problem domain application. A standard alternative is to take a problem centric perspective of the development of approaches, and specialise an existing strategy or paradigm for a given problem domain or instance. The benefit of this alternative is that it results in specific and comparable findings for a specialised algorithm on a known difficult problem class, directly contributing to an assessment of the relative utility of the approach. The strong problem focus of this alternative in turn imposed strong limitations on the broader applicability of an approach given its specialisation to the features and available information regarding the problem. A specific example from his work is the foundational cellular clonal selection algorithms developed in the context of global function optimisation (such as BCA, the IA family, and to a lesser extent CLONALG). Although this class of algorithm is well suited for this application, the single-antigen perspective of the strategy ignores the potential application of the approach to function approximation, and has the compound effect of bounding the applicability of algorithms that subsume the cellular level¹. The strategy focus in the development of approaches was chosen in the spirit of the goals of Computational Intelligence toward investigating generalised unconventional methods of computation, specifically bound to the information processing qualities of clonal selection. The generality of the developed approaches may be bound at a later time through specialisations of the algorithms in the context of known problem instances to which they are highly suited.

¹The field of Evolutionary Computation is generally bound by the first application of the approach to function optimisation [121], that in turn has generally limited subsumed approaches such as Parallel Evolutionary Algorithms.

9.3.3 Verification and Demonstration

The verification of the behaviours and capabilities of developed modes as achieved empirically using experimentation with general algorithm instances on a generalised and trivial pattern recognition problem domain. A problem with this approach of verification and demonstration is that it is qualitative and inaccurate, specifically in the context of Goldberg's economy of modelling in Section 2.7.3. An alternative is to demonstrate the capability and verify behaviour using mathematical modelling tools. Analytical tools would provide generalised descriptions of relationships and behaviours that would be more accurate although a more restricted interpretation of the motivating information processing properties. An alternative to the development of costly mathematical models is the approach outlined by Goldberg's is to use a spectrum of modelling types in addition to empirical assessment and statistical analysis, specifically facet-wise and dimensional models. The qualitative empirical approach was chosen primarily given the speed both in implementation and in obtaining results. Empirical demonstration required implementation providing a grounding of conceptual models to the specific concerns of a functioning information processing system. The broader inaccuracy of qualitative observations was address through the use of multiple problem scenarios providing calibration and cross-correlation of behaviour, and through the use of statistical tools.

9.4 Further Research

This section proposes recommendations for further research and improvements on the work from the investigation. Extensions are partitioned as follows: (1) *framework* in which additional related metaphors are proposed as a basis for new paradigms within the integrated hierarchical clonal selection framework, (2) *paradigms* in which additional and related constrained metaphors are considered within the context of each clonal selection paradigm as the basis for new algorithms and architectures, (3) and *algorithms* in which elaborations and improvements on algorithms proposed in this work are considered. Many of the proposed extensions, particularly the alternate and additional motivating metaphors were preliminary investigated as apart of this work, although were not integrated into the framework given the constraints imposed on the project.

9.4.1 Framework Extensions

Integrated frameworks were considered in Chapter 7 which proposed a hierarchical clonal selection, antigenic exposure, and integrated clonal selection frameworks. The integration of the three clonal selection paradigms highlighted the following extensions that may be pursued beyond this work:

1. *Additional Configurations*: The investigation of clonal selection systems beyond the simplest configuration (such as minimum component or interaction cardinality) in what was referred to as varied interpretation of clonal selection models, as highlighted by the bracketing analysis of the integrated framework (Section 7.4.2). This

investigation would immediately provide new interpretations of clonal selection algorithms from across the three paradigms, and reveal insights into the behaviour and application of clonal selection algorithms covered by the framework.

2. *Additional Tiers*: The elicitation and investigation of systems motivated by constraints on the information processing properties of clonal selection *below the cellular level* (such as the molecular concerns of epitopes and paratopes) and *above the host level* (such as interacting species of organisms) (Section 7.4.2). These investigations would reveal insights into the scope of sub-cellular and super-host bracketing already imposed by the algorithms covered by the framework, as well as provide a basis for new constraints and interpretations of clonal selection as a basis for Computational Intelligence algorithms.
3. *Additional Mappings*: The investigation of the clonal selection centric acquired immune system models across the hierarchies biased by one or more of the related computational paradigms (such as negative selection, immune network, or danger theory) as remarked in the bracketing analysis (Section 7.4.3), and explicitly considered with regard to the integrated framework in Section 7.5.2. This investigation would not only provide insights into the approaches covered by the framework, it would explicitly merge related immunological computational paradigms immediately providing a firm basis of models, architectures, and algorithms to be exploited by the integrated paradigm.

In addition to improving and augmenting the integrated clonal selection framework, additional holistic metaphors may be incorporated toward providing further general insights into the computational strategy, as well as additional metaphors for information processing, as follows:

1. *Asexual Evolution*: Clonal Selection Theory describes a process of asexual evolution, akin to the evolution of micro-organisms not limited to bacteria, virus, unicellular organisms, as well as multi-cellular asexual reproduces. This obvious connection was made by Burnet in his monograph on the theory where he described the relationship of the principles of clonal selection with bacteria and cancer cells [64]. Much is known regarding the genetic and adaptive models of asexual evolution, for example advantageous and deleterious mutation rates in populations [169, 340], which may be interpreted and integrated into an improved computational model of B-cell receptor evolution in the cellular paradigm. This fundamental work would not only provide insights into the configuration and application of cellular algorithms, given the cellular basis of all three paradigms (and the high-level constraints imposed on the cellular mechanisms) it would provide insights for all clonal selection based approaches.
2. *Pathogen Models*: The antigenic exposure paradigm was defined as an abstract problem for clonal selection systems encapsulating the properties of all self and foreign

antigen (defined in Section 4.2.2, elaborated in Sections 5.3.2 and 6.3.2, and integrated in Section 7.3). As such, the paradigm received little elaboration beyond definition providing a large opportunity for the investigation of pathogenic models. Pathogen models may be elicited and investigated as either models for problems to be solved by clonal selection systems, or as distinct computational models for problem solving in and of themselves. Three clear motivating metaphors for the latter case are as follows: (1) *Pathogen Evolution*: Suggested in Section 6.3.2 as providing a model for dynamic problems in which a system and problem co-adapt in response to each other based on reviewed parasitism in Section 6.2.2, (2) *Intra-Host Dynamics*: That describe the strategies and behaviours of a pathogen after it enters a host organism including evasive and adaptive qualities [26, 161] (3) *Intra-Population Dynamics*: Also suggested in Section 6.3.2 and providing part of the motivation for system-controlled transmission in the PT-HCSA (Section 6.5)².

3. *Cancer*: Uniting asexual somatic evolution and an elaborated self-antigen exposure paradigm, cancer provides a motivation for a computational paradigm that may be investigated independently or as a complement clonal selection. Cancer provides a co-evolutionary problem model for clonal selection in which the accumulation of mutations is exploited for survival of tumour cells. The acquired immune system must discriminate the endogenous antigen and (at the very least) counter the adaptations to surface structures. More interestingly, cancer also provides an intra-host Darwinian microcosm [409, 56, 410], that can aggressively recruit resources such as cells and nutrient sources, disseminate throughout the host (such as the metastasis of malignant tumour cells), and in some cases can be transmitted between hosts [299].

9.4.2 Paradigm Extensions

This section enumerates the three clonal selection paradigms and highlights elaborations of proposed models as additional motivating metaphors for new computational models, not specific to any given algorithm within the paradigms.

1. *Cellular Paradigm*: Likely the most important of the three paradigms, the cellular paradigm may be elaborated in three specific directions: (1) lymphocyte life cycle and types, and (2) regulatory mechanisms. The model of lymphocytes as atomic data points is a dramatic simplification of both the complexity of the life cycle of immune cells, as well as the diverse variety of cells that are involved in the immune response. The computational properties of cellular life cycles introduces task-division and developmental mechanism (Section 5.2.2), whereas the exploitation of the diverse set of immune cell types further promotes task division and more importantly interaction and collaboration between cells. The regulation of lymphocytes under clonal selection in which the cellular repertoire is continually acquiring information

²There is much work on epidemiology, specifically adaptive transmission [50], and adaptive virulence [335], to motivate this paradigm.

and expanding, presents problems of cell population size regulation whilst maintaining an effective immunological memory. More elaborate homeostasis mechanisms should be investigated toward providing improved flexibility in cellular algorithms, in particular with regard to promoting decentralised rather than centralised resource allocation in response to antigenic exposures within a given repertoire. Such approaches may be motivated by the variety of mathematical models that describe lymphocyte homeostasis [290, 87].

2. *Tissue Paradigm*: Two general areas for elaboration of the tissue paradigm include (1) the investigation of the variety of lymphocyte trafficking behaviours and the decentralised receptor centric mechanisms that influence them (Section 5.2.2) and (2) the tissues of the lymphatic system and the behavioural restrictions and antigenic interactions promoted by the different tissue structures and arrangements (reviewed in Section 5.2.1, with models proposed in Section 5.3.3). A final additional motivating metaphor is the emerging field of *Lymphocyte Population Biology*, in which the behaviours and interactions of immune cells are studied using the tools of population genetics and ecology, driven more recently by Freitas Antonio and colleagues at the Pasteur Institute in France [154, 165, 4]. This aggregated perspective of lymphocyte behaviours introduces interesting models of lymphocyte niche exploitation and regulatory concerns of density dependent competition.
3. *Host Paradigm*: The two principle areas of the host paradigm present a large capacity for elaboration given the generality of base motivations, specifically (1) host population immunisation (reviewed in Section 6.2.1), and (2) host immune system evolution (reviewed in Section 6.2.2). Host vaccination provides a promising area for elaboration, specifically the computational concerns of vaccine identification and transmission regimes in the context of the benefits of population coverage called ‘herd immunity’. Such investigations are expected to provide models for information dissemination in distributed databases such as database replication [127], and peer-to-peer content dissemination networks [413] suggested in Section 8.4.2. The evolution of the immune system is a broader motivating metaphor, although further research may consider the following two promising cases: (1) the adaptation of innate immunity where a selective process motivates the integration of lifetime acquired traits into the genome (the intersection of lifetime and evolutionary learning considered in Section 6.3.1), and (2) the adaptation of the affinity maturation process beyond the base repertoire where the mechanisms used in lifetime are also evolved.

9.4.3 Algorithm Extensions

The investigation of algorithms in this work was broad in coverage although shallow in depth, providing plenty of opportunity for further ‘basic science’. Two general classes of fundamental research are highlighted regarding the proposed algorithms: (1) application of algorithms as outlined in Chapter 8, and (2) investigation into the predictable elicit-

tion of information processing capabilities not limited to parameter studies, specialisation, and additional empirical benchmarking. This section highlights specific cases of the second class of fundamental research for specific algorithms across the three clonal selection paradigms.

1. *Cellular Algorithms*: Three specific cellular algorithms are in need of elaboration to capitalise on their expected potential: (1) degenerate, (2) mediated, and (3) network. Cohen's notions of emergent specificity through degenerate pattern recognition cells in different contexts (reviewed in Section 4.4.3) provides the potential for a powerful sub-symbolic population-based adaptive paradigm, requiring investigation into viable representation and aggregation mechanisms. The two-signal model of lymphocyte activation provides an important capability for the system to decide to which signals to respond and not respond. The mediated algorithm requires further investigation into mapping mechanisms for affinity-based activation potentials between cell populations (see Section 4.6). Finally, the immune network paradigm has motivated many immuno-cognitive models, although the specific approach suggested in this work was an explicit pattern recognition method akin to Learning Classifier Systems, requiring much work into representation and mapping schemes (Section 4.7). An opportunistic improvement on the cell-cell interaction models would be to adopt a LCS architecture, representation, and credit assignment mechanism and bias the approaches by immune network behaviours.
2. *Tissue Algorithms*: The tissue algorithms, specifically the use of recirculation and mechanisms to impose localisation biases on recirculation (such as homing and inflammation) would benefit from analytical analysis. Specifically, (1) probabilistic models of information availability across a given structure and how such models are effected by localisation biases, and (2) probabilistic models of information dissemination based on varied exposure models. These models would motivate predictable configurations (such as migration or imprinting rates) for eliciting well defined system capability under different exposure scenarios supporting empirical findings. Further, an investigation into the field of *Spatial-Temporal Learning* and the relationship of Tissue (and Host) Algorithms would provide a deeper theoretical understanding of the fundamental problem and recirculation-based strategy solution of systemic immunity within a host [354].
3. *Host Algorithms*: Mathematical models of vaccination and virus transmission are available with which to bias the configuration of shared immunity and transmission host clonal selection algorithms. Such models would provide predictable intra-population interactions allowing focus of these algorithms to shift from base interaction mechanism toward more interesting population-based information acquisition and application scenarios (mixtures of experts and ensembles). A similar situation exists with the generational host algorithms that must overcome the fundamentals predictable inter-generational interaction before investigating the interesting and desirable effects of cross-generational information acquisition. This highlights that

the compounded complexity effect of the host level of algorithms that are both dependent on predictable algorithms which they subsume, as well as the predictable foundational behaviours (such as interaction) of the paradigm.

9.5 Final Words

This work has demonstrated that a holistic perspective of the immune system as a computational metaphor provides a fertile ground for *adaptive* and *distributed* Artificial Immune Systems, and Clonal Selection Algorithms specifically. The culmination of this work advocates a broader perspective of the acquired immune system beyond a classical cellular focus as a viable path toward a so-called second generation of Artificial Immune Systems. This path not only facilitates, but requires the currently accepted approach of turning back to the motivating metaphor as a strategy for developing systems that incorporate more detailed information processing models of immunology. The computational focus of immunophysiology and the integrated framework presented in this work provide a general scaffold for such work, the ongoing investigation of which promotes a multifaceted Computational Intelligence research agenda of (1) basic (foundational) research, (2) exploratory research, and (3) application-focused research.

Appendix A

Tables of Empirical Results

A.1 Tissue Clonal Selection

Tables of results from the empirical investigations undertaken in Chapter 5.

A.1.1 Recirculation Empirical Study

Table A.1 summarises the results from the empirical study in Section 5.5.3 investigating the Recirculation Tissue Clonal Selection Algorithm (RTCSA). The non-parametric Mann-Whitney U statistical test was calculated pair-wise for all algorithms on each TER group.

Problem	System	HD		HE		ATD		ATE	
		\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
ATER	MTCSA	1610.68	78.407	0.013	0.006	85.903	0.92	0.155	0.011
ATER	RTCSA-SML	733.841	57.614	0.131	0.021	78.82	2.812	0.148	0.017
ATER	RTCSA-MED	608.013	45.143	0.149	0.024	76.792	3.723	0.153	0.018
ATER	RTCSA-LRG	562.136	54.663	0.152	0.025	76.969	4.213	0.152	0.017
Significant		True		True ^a		True ^b		False	
OTER	MTCSA	1708.99	89.685	0.033	0.01	84.394	1.151	0.152	0.01
OTER	RTCSA-SML	690.006	51.553	0.122	0.022	78.947	4.497	0.137	0.021
OTER	RTCSA-MED	602.637	38.677	0.141	0.024	77.86	4.731	0.145	0.018
OTER	RTCSA-LRG	561.076	47.781	0.15	0.021	78.593	4.583	0.145	0.02
Significant		True		True ^c		True ^d		False ^e	
PTER	MTCSA	876.922	40.518	0.093	0.026	92.366	0.442	0.158	0.01
PTER	RTCSA-SML	690.785	22.011	0.148	0.019	93.192	0.182	0.172	0.018
PTER	RTCSA-MED	635.74	21.472	0.162	0.024	93.515	0.129	0.173	0.018
PTER	RTCSA-LRG	572.749	14.306	0.167	0.021	93.826	0.118	0.169	0.017
Significant		True		True ^f		True		True ^g	
RTER	MTCSA	2313.88	100.596	0.121	0.021	74.7	1.777	0.123	0.01
RTER	RTCSA-SML	602.977	71.514	0.11	0.018	76.113	4.058	0.11	0.016
RTER	RTCSA-MED	561.37	55.816	0.109	0.015	76.277	4.166	0.115	0.017
RTER	RTCSA-LRG	503.919	65.54	0.114	0.024	75.347	4.988	0.111	0.019
Significant		True		False		False ^h		False ⁱ	
STER	MTCSA	2142.95	160.562	0.026	0.009	76.972	2.421	0.093	0.017
STER	RTCSA-SML	936.657	125.483	0.046	0.013	79.098	2.812	0.087	0.021
STER	RTCSA-MED	672.651	72.745	0.065	0.019	76.659	3.648	0.098	0.024
STER	RTCSA-LRG	579.055	62.934	0.07	0.019	76.445	3.952	0.094	0.023
Significant		True		True ^j		True ^k		False ^l	

^aFalse for RTCSA-MED and RTCSA-LRG

^bFalse for RTCSA-MED and RTCSA-LRG

^cFalse for RTCSA-MED and RTCSA-LRG

^dFalse for RTCSA-SML and RTCSA-MED, False for RTCSA-SML and RTCSA-LRG, False for RTCSA-MED and RTCSA-LRG

^eTrue for MTCSA and RTCSA-SML

^fFalse for RTCSA-MED and RTCSA-LRG

^gFalse for RTCSA-SML and RTCSA-MED, False for RTCSA-SML and RTCSA-LRG, False for RTCSA-MED and RTCSA-LRG

^hTrue for MTCSA and RTCSA-SML, True for MTCSA and RTCSA-MED

ⁱTrue for MTCSA and RTCSA-SML, True for MTCSA and RTCSA-LRG

^jFalse for RTCSA-MED and RTCSA-LRG

^kFalse for MTCSA and RTCSA-MED, False for MTCSA and RTCSA-LRG, False for RTCSA-MED and RTCSA-LRG

^lTrue for RTCSA-SML and RTCSA-MED

Table A.1: Summary of results from an empirical study into the RTCSA compared to the MTCSA.

Table A.2 summarises the results of the RTCSA for the analysis of the approach in Section 5.5.3.

System	Problem	HD		HE		ATD		ATE	
		\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
TCSA	TER								
MTCSA	ATER	1610.68	78.407	0.013	0.006	85.903	0.92	0.155	0.011
MTCSA	OTER	1708.99	89.685	0.033	0.01	84.394	1.151	0.152	0.01
MTCSA	PTER	876.922	40.518	0.093	0.026	92.366	0.442	0.158	0.01
MTCSA	RTER	2313.88	100.596	0.121	0.021	74.7	1.777	0.123	0.01
MTCSA	STER	2142.95	160.562	0.026	0.009	76.972	2.421	0.093	0.017
Significant		True		True		True		True ^a	
RTCSA-LRG	ATER	562.136	54.663	0.152	0.025	76.969	4.213	0.152	0.017
RTCSA-LRG	OTER	561.076	47.781	0.15	0.021	78.593	4.583	0.145	0.02
RTCSA-LRG	PTER	572.749	14.306	0.167	0.021	93.826	0.118	0.169	0.017
RTCSA-LRG	RTER	503.919	65.54	0.114	0.024	75.347	4.988	0.111	0.019
RTCSA-LRG	STER	579.055	62.934	0.07	0.019	76.445	3.952	0.094	0.023
Significant		False ^b		True ^c		True ^d		True ^e	

^aFalse for OTER and ATER, PTER and ATER

^bTrue for RTER and ATER, RTER and OTER, RTER and PTER, STER and RTER

^cFalse for ATER and OTER

^dFalse for ATER and OTER, RTER and ATER, STER and ATER, STER and OTER, STER and RTER

^eFalse for OTER and ATER

Table A.2: Summary of results from an empirical study into the RTCSA comparing TER's with the MTCSA local and RTCSA-LRG global information management strategies.

A.1.2 Homing Empirical Study

Table A.3 summarises the results from the empirical study in Section 5.6.3 investigating the Homing Tissue Clonal Selection Algorithm (HTCSA). The non-parametric Mann-Whitney U statistical test was calculated pair-wise for all algorithms on each TER group.

Problem	System	HD		HE		ATD		ATE	
		\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
ATER	MTCSA	1610.68	78.407	0.013	0.006	85.903	0.92	0.155	0.011
ATER	RTCSA-S	733.841	57.614	0.131	0.021	78.82	2.812	0.148	0.017
ATER	HTCSA-S	747.071	52.578	0.099	0.019	81.631	3.123	0.142	0.016
ATER	HTCSA-M	816.023	53.374	0.084	0.019	83.046	2.228	0.15	0.018
ATER	HTCSA-L	816.023	53.374	0.084	0.019	83.046	2.228	0.15	0.018
Significant		True ^a		True ^b		True ^c		False ^d	
OTER	MTCSA	1708.99	89.685	0.033	0.01	84.394	1.151	0.152	0.01
OTER	RTCSA-S	690.006	51.553	0.122	0.022	78.947	4.497	0.137	0.021
OTER	HTCSA-S	737.646	81.211	0.099	0.02	79.893	3.451	0.136	0.018
OTER	HTCSA-M	756.301	64.802	0.1	0.019	81.077	3.568	0.146	0.015
OTER	HTCSA-L	756.301	64.802	0.1	0.019	81.077	3.568	0.146	0.015
Significant		True ^e		True ^f		True ^g		False ^h	
PTER	MTCSA	876.922	40.518	0.093	0.026	92.366	0.442	0.158	0.01
PTER	RTCSA-S	690.785	22.011	0.148	0.019	93.192	0.182	0.172	0.018
PTER	HTCSA-S	678.451	20.88	0.136	0.019	93.232	0.137	0.169	0.016
PTER	HTCSA-M	685.395	27.472	0.145	0.017	93.206	0.161	0.17	0.018
PTER	HTCSA-L	685.395	27.472	0.145	0.017	93.206	0.161	0.17	0.018
Significant		True		True		False ⁱ		False ^j	
RTER	MTCSA	2313.88	100.596	0.121	0.021	74.7	1.777	0.123	0.01
RTER	RTCSA-S	602.977	71.514	0.11	0.018	76.113	4.058	0.11	0.016
RTER	HTCSA-S	635.615	52.025	0.109	0.021	76.828	3.359	0.113	0.017
RTER	HTCSA-M	673.063	58.672	0.114	0.022	75.595	3.745	0.114	0.018
RTER	HTCSA-L	673.063	58.672	0.114	0.022	75.595	3.745	0.114	0.018
Significant		True ^k		False ^l		False ^m		False ⁿ	
STER	MTCSA	2142.95	160.562	0.026	0.009	76.972	2.421	0.093	0.017
STER	RTCSA-S	936.657	125.483	0.046	0.013	79.098	2.812	0.087	0.021
STER	HTCSA-S	984.195	130.37	0.045	0.014	79.29	2.245	0.086	0.023
STER	HTCSA-M	995.975	114.778	0.046	0.013	79.731	2.071	0.09	0.022
STER	HTCSA-L	995.975	114.778	0.046	0.013	79.731	2.071	0.09	0.022
Significant		False ^o		False ^p		False ^q		False	

^aFalse for RTCSA-S and HTCSA-S, HTCSA-M and HTCSA-L

^bFalse for HTCSA-M and HTCSA-L

^cFalse for HTCSA-S and HTCSA-M, HTCSA-S and HTCSA-L, HTCSA-M and HTCSA-L

^dTrue for MTCSA and HTCSA-S

^eFalse for HTCSA-S and HTCSA-M, HTCSA-S and HTCSA-L, HTCSA-M and HTCSA-L

^fFalse for RTCSA-S and HTCSA-M, HTCSA-S and HTCSA-L, HTCSA-M and HTCSA-L

^gFalse for RTCSA-S and HTCSA-S, HTCSA-S and HTCSA-M, HTCSA-S and HTCSA-L, HTCSA-M and HTCSA-L

^hTrue for MTCSA and RTCSA-S, MTCSA and HTCSA-S, HTCSA-S and HTCSA-M, HTCSA-S and HTCSA-L

ⁱTrue for MTCSA and RTCSA-S, MTCSA and HTCSA-S, MTCSA and HTCSA-M, MTCSA and HTCSA-L

^jTrue for MTCSA and RTCSA-S, MTCSA and HTCSA-S, MTCSA and HTCSA-M, MTCSA and HTCSA-L

^kFalse for RTCSA-S and HTCSA-S, HTCSA-M and HTCSA-L

^lTrue for MTCSA and HTCSA-S

^mTrue for MTCSA and RTCSA-S, MTCSA and HTCSA-S

ⁿTrue for MTCSA and RTCSA-S, MTCSA and HTCSA-S, MTCSA and HTCSA-M, MTCSA and HTCSA-L

^oTrue for MTCSA and RTCSA-S, MTCSA and HTCSA-S, MTCSA and HTCSA-M, MTCSA and HTCSA-L

^pTrue for MTCSA and RTCSA-S, MTCSA and HTCSA-S, MTCSA and HTCSA-M, MTCSA and HTCSA-L

^qTrue for MTCSA and RTCSA-S, MTCSA and HTCSA-S, MTCSA and HTCSA-M, MTCSA and HTCSA-L

Table A.3: Summary of results from an empirical study into the HTCSA.

A.1.3 Inflammation Empirical Study

Table A.4 summarises the results from the empirical study in Section 5.7.3 investigating the Inflammation Tissue Clonal Selection Algorithm (ITCSA). The non-parametric Mann-Whitney U statistical test was calculated pair-wise for all algorithms on each TER group.

Problem	System	HD		HE		ATD		ATE	
<i>TER</i>	<i>TCSA</i>	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
ATER	MTCSA	1610.68	78.407	0.013	0.006	85.903	0.92	0.155	0.011
ATER	RTCSA-S	733.841	57.614	0.131	0.021	78.82	2.812	0.148	0.017
ATER	ITCSA-N	1885.64	81.789	0.014	0.007	86.034	0.846	0.149	0.01
ATER	ITCSA-S	800.373	70.917	0.122	0.027	80.34	3.392	0.138	0.019
Significant		True		True		True ^a		True	
OTER	MTCSA	1708.99	89.685	0.033	0.01	84.394	1.151	0.152	0.01
OTER	RTCSA-S	690.006	51.553	0.122	0.022	78.947	4.497	0.137	0.021
OTER	ITCSA-N	4298.85	145.475	0.061	0.026	36.523	5.666	0.223	0.015
OTER	ITCSA-S	977.768	151.773	0.125	0.022	58.817	11.181	0.16	0.027
Significant		True		True ^b		True		True ^c	
PTER	MTCSA	876.922	40.518	0.093	0.026	92.366	0.442	0.158	0.01
PTER	RTCSA-S	690.785	22.011	0.148	0.019	93.192	0.182	0.172	0.018
PTER	ITCSA-N	1186.91	88.006	0.151	0.032	89.802	0.958	0.164	0.011
PTER	ITCSA-S	769.419	37.061	0.13	0.023	93.063	0.24	0.169	0.017
Significant		True		True ^d		True		True ^e	
RTER	MTCSA	2313.88	100.596	0.121	0.021	74.7	1.777	0.123	0.01
RTER	RTCSA-S	602.977	71.514	0.11	0.018	76.113	4.058	0.11	0.016
RTER	ITCSA-N	3300.08	188.667	0.132	0.02	64.262	3.485	0.128	0.009
RTER	ITCSA-S	742.331	84.362	0.091	0.016	73.855	4.428	0.093	0.012
Significant		True		True ^f		True ^g		True	
STER	MTCSA	2142.95	160.562	0.026	0.009	76.972	2.421	0.093	0.017
STER	RTCSA-S	936.657	125.483	0.046	0.013	79.098	2.812	0.087	0.021
STER	ITCSA-N	3711.46	130.433	0.066	0.013	51.85	3.256	0.149	0.022
STER	ITCSA-S	1331.72	189.05	0.066	0.014	64.593	4.886	0.123	0.02
Significant		True		True ^h		True		True ⁱ	

^aFalse for MTCSA and ITCSA-N

^bFalse for RTCSA-S and ITCSA-S

^cFalse for MTCSA and ITCSA-S

^dFalse for RTCSA-S and ITCSA-N

^eFalse for RTCSA-S and ITCSA-N, False for RTCSA-S and ITCSA-S, False for ITCSA-N and ITCSA-S

^fFalse for MTCSA and RTCSA-S

^gFalse for MTCSA and ITCSA-S

^hFalse for ITCSA-N and ITCSA-S

ⁱFalse for MTCSA and RTCSA-S

Table A.4: Summary of results from an empirical study into the ITCSA.

A.2 Host Clonal Selection

Tables of results from the empirical investigations undertaken in Chapter 6.

A.2.1 Transmission Empirical Study

Table A.5 summarises the results from the empirical study in Section 6.5.2 investigating the Transmission Host Clonal Selection Algorithm (THCSA). The non-parametric Mann-

Whitney U statistical test was calculated pair-wise for all algorithms on each HER group.

Problem	System	PH		PE		AHD		AHE	
		\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
HER	HCSA	1606.24	57.906	0.013	0.007	85.97	0.702	0.157	0.01
AHER	MHCSA	2257.49	92.704	0.028	0.01	75.558	1.552	0.116	0.011
AHER	PT-RP	2007.71	68.019	0.018	0.008	79.905	1.037	0.138	0.011
AHER	PT-SP	2218.74	103.049	0.023	0.01	76.433	1.72	0.127	0.011
AHER	VT-L	1952.49	45.303	0.015	0.008	80.94	0.692	0.141	0.011
AHER	VT-S	True ^a		True ^b		True		True ^c	
OHER	MHCSA	1792.98	99.787	0.031	0.011	83.271	1.398	0.152	0.009
OHER	PT-RP	2269.86	87.319	0.042	0.015	75.496	1.48	0.122	0.011
OHER	PT-SP	2098.97	74.909	0.043	0.015	78.482	1.156	0.134	0.01
OHER	VT-L	2254.70	88.167	0.037	0.013	75.797	1.508	0.126	0.012
OHER	VT-S	2015.66	95.136	0.038	0.011	79.928	1.476	0.138	0.009
OHER	Significant	True ^d		False ^e		True ^f		True ^g	
PHER	MHCSA	894.093	40.821	0.102	0.03	92.315	0.493	0.159	0.009
PHER	PT-RP	1716.58	88.816	0.102	0.03	84.24	1.175	0.147	0.01
PHER	PT-SP	1251.29	50.971	0.102	0.03	88.887	0.64	0.152	0.009
PHER	VT-L	2193.93	94.627	0.102	0.03	76.966	1.621	0.131	0.011
PHER	VT-S	1716.78	67.258	0.102	0.03	84.268	0.955	0.147	0.011
PHER	Significant	True ^h		False		True ⁱ		True ^j	
RHER	MHCSA	2341.24	72.031	0.122	0.023	74.238	1.264	0.123	0.009
RHER	PT-RP	2353.58	106.255	0.118	0.023	73.869	1.878	0.117	0.011
RHER	PT-SP	2368.57	141.087	0.112	0.019	73.619	2.647	0.116	0.011
RHER	VT-L	2369.23	118.525	0.123	0.022	73.59	2.095	0.118	0.012
RHER	VT-S	2364.53	115.166	0.125	0.018	73.853	2.108	0.124	0.009
RHER	Significant	False		False ^k		False		False ^l	
SHER	MHCSA	2144.72	172.247	0.028	0.012	76.929	2.717	0.094	0.019
SHER	PT-RP	2150.92	187.501	0.029	0.009	76.78	2.886	0.093	0.019
SHER	PT-SP	2161.80	194.783	0.029	0.011	76.603	3.084	0.093	0.018
SHER	VT-L	2164.46	178.009	0.026	0.009	76.583	2.847	0.094	0.018
SHER	VT-S	2162.90	173.729	0.026	0.011	76.697	2.77	0.093	0.018
SHER	Significant	False		False		False		False	

^aFalse for PT-RP and VT-L

^bFalse for MHCSA and VT-S, PT-RP and VT-L, PT-SP and VT-S

^cFalse for PT-SP and VT-S

^dFalse for PT-RP and VT-L

^eTrue for MHCSA and PT-RP, MHCSA and PT-SP, MHCSA and VT-S

^fFalse for PT-RP and VT-L

^gFalse for PT-RP and VT-L, PT-SP and VT-S

^hFalse for PT-RP and VT-S

ⁱFalse for PT-RP and VT-S

^jFalse for PT-RP and PT-SP, PT-RP and VT-S, PT-SP and VT-S

^kTrue for PT-SP and VT-L, PT-SP and VT-S

^lTrue for MHCSA and PT-SP, PT-RP and VT-S, PT-SP and VT-S

Table A.5: Summary of results from an empirical study into the THCSA compared to MP-HCSA.

Table A.6 summarises the pathogen transmission and vaccination transmission results from an empirical study into the THCSA from Section 6.5.2.

Problem	System	PH		PE		AHD		AHE	
		\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
AHER	PT-RP	2257.49	92.704	0.028	0.01	75.558	1.552	0.116	0.011
AHER	VT-L	2218.74	103.049	0.023	0.01	76.433	1.72	0.127	0.011
Significant		False		False		True		True	
OHER	PT-RP	2269.86	87.319	0.042	0.015	75.496	1.48	0.122	0.011
OHER	VT-L	2254.70	88.167	0.037	0.013	75.797	1.508	0.126	0.012
Significant		False		False		False		False	
PHER	PT-RP	1716.58	88.816	0.102	0.03	84.24	1.175	0.147	0.01
PHER	VT-L	2193.93	94.627	0.102	0.03	76.966	1.621	0.131	0.011
Significant		True		False		True		True	
RHER	PT-RP	2353.58	106.255	0.118	0.023	73.869	1.878	0.117	0.011
RHER	VT-L	2369.23	118.525	0.123	0.022	73.59	2.095	0.118	0.012
Significant		False		False		False		False	
SHER	PT-RP	2150.92	187.501	0.029	0.009	76.78	2.886	0.093	0.019
SHER	VT-L	2164.46	178.009	0.026	0.009	76.583	2.847	0.094	0.018
Significant		False		False		False		False	

Table A.6: Summary of the pathogen transmission and vaccination transmission results from an empirical study into the THCSA.

A.2.2 Shared Immunity Empirical Study

Table A.7 summarises the results from the empirical study in Section 6.6.2 investigating the Shared Immunity Host Clonal Selection Algorithm (SI-HCSA). The non-parametric Mann-Whitney U statistical test was calculated pair-wise for all algorithms on each HER group.

Problem	System	PH		PE		AHD		AHE	
		\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
HER	HCSA	1606.24	57.906	0.013	0.007	85.97	0.702	0.157	0.01
AHER	MHCSA	1021.50	53.326	0.012	0.005	88.691	0.593	0.082	0.011
AHER	SI-SS	660.944	45.926	0.03	0.01	88.719	0.893	0.086	0.011
AHER	SI-SL	679.249	48.4	0.028	0.012	88.249	1.101	0.086	0.017
AHER	SI-LS	418.271	35.337	0.076	0.013	84.116	2.518	0.104	0.012
Significant		True ^a		True ^b		True ^c		True ^d	
OHER	MHCSA	1792.98	99.787	0.031	0.011	83.271	1.398	0.152	0.009
OHER	SI-SS	944.546	65.987	0.027	0.01	88.584	0.739	0.093	0.013
OHER	SI-SL	634.095	50.014	0.053	0.014	87.804	0.986	0.092	0.016
OHER	SI-LS	611.977	43.081	0.041	0.014	87.712	1.252	0.086	0.013
OHER	SI-LL	426.655	44.262	0.095	0.019	82.11	2.352	0.116	0.016
Significant		True ^e		True ^f		True ^g		True ^h	
PHER	MHCSA	894.093	40.821	0.102	0.03	92.315	0.493	0.159	0.009
PHER	SI-SS	718.195	51.235	0.105	0.023	91.069	0.647	0.124	0.01
PHER	SI-SL	500.736	34.383	0.112	0.016	88.74	1.034	0.131	0.014
PHER	SI-LS	522.83	37.182	0.108	0.017	88.611	1.09	0.125	0.01
PHER	SI-LL	402.229	39.746	0.118	0.015	82.375	2.624	0.131	0.012
Significant		True		False ⁱ		True ^j		True ^k	
RHER	MHCSA	2341.24	72.031	0.122	0.023	74.238	1.264	0.123	0.009
RHER	SI-SS	792.363	65.438	0.094	0.014	84.998	1.899	0.094	0.012
RHER	SI-SL	530.103	49.515	0.107	0.018	83.07	2.296	0.105	0.014
RHER	SI-LS	556.441	57.438	0.106	0.016	83.359	2.092	0.105	0.012
RHER	SI-LL	420.333	43.22	0.121	0.017	79.294	3.281	0.126	0.013
Significant		True ^l		True ^m		True ⁿ		True ^o	
SHER	MHCSA	2144.72	172.247	0.028	0.012	76.929	2.717	0.094	0.019
SHER	SI-SS	1393.50	158.49	0.027	0.01	83.148	1.765	0.072	0.019
SHER	SI-SL	852.855	107.04	0.034	0.013	83.743	1.163	0.068	0.016
SHER	SI-LS	911.251	119.976	0.033	0.012	84.085	1.771	0.07	0.018
SHER	SI-LL	564.659	52.838	0.055	0.011	81.828	2.587	0.087	0.014
Significant		True ^p		True ^q		True ^r		True ^s	

^aFalse for SI-SL and SI-LS

^bFalse for MHCSA and SI-SS, SI-SL and SI-LS

^cFalse for SI-SS and SI-SL, SI-SS and SI-LS, SI-SL and SI-LS

^dFalse for SI-SS and SI-SL, SI-SS and SI-LS, SI-SL and SI-LS

^eFalse for SI-SL and SI-LS

^fFalse for MHCSA and SI-SS

^gFalse for SI-SL and SI-LS

^hFalse for SI-SS and SI-SL, SI-SS and SI-LS, SI-SL and SI-LS

ⁱTrue for MHCSA and SI-LL, SI-SS and SI-SL, SI-SS and SI-LL, SI-LS and SI-LL

^jFalse for SI-SL and SI-LS

^kFalse for SI-SS and SI-LS, SI-SL and SI-LL, SI-LS and SI-LL

^lFalse for SI-SL and SI-LS

^mFalse for MHCSA and SI-LL, SI-SL and SI-LS

ⁿFalse for SI-SL and SI-LS

^oFalse for MHCSA and SI-LL, SI-SL and SI-LS

^pFalse for SI-SL and SI-LS

^qFalse for MHCSA and SI-SS, MHCSA and SI-LS, SI-SS and SI-LS, SI-SL and SI-LS

^rFalse for SI-SS and SI-SL, SI-SS and SI-LL, SI-SL and SI-LS

^sFalse for MHCSA and SI-LL, SI-SS and SI-SL, SI-SS and SI-LS, SI-SL and SI-LS

Table A.7: Summary of results from an empirical study into the SI-HCSA compared to MP-HCSA.

A.2.3 Maternal Immunity Empirical Study

Table A.8 summarises the results from the empirical study in Section 6.7.2 investigating the Maternal Immunity Host Clonal Selection Algorithm (MI-HCSA). The non-parametric Mann-Whitney U statistical test was calculated pair-wise for all algorithms on each HER group.

Problem	System	PH		PE		AHD		AHE	
<i>HER</i>	<i>HCSA</i>	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
AHER	MG-HCSA	1270.97	78.43	0.106	0.021	89.685	0.78	0.162	0.009
AHER	MIHCSA-S	1216.84	38.157	0.106	0.02	90.254	0.351	0.159	0.009
AHER	MIHCSA-M	1233.45	57.015	0.104	0.019	90.036	0.531	0.162	0.009
AHER	MIHCSA-L	1334.18	54.323	0.058	0.016	89.042	0.533	0.157	0.007
Significant		True ^a		True ^b		True ^c		False ^d	
OHER	MG-HCSA	1260.92	92.225	0.121	0.02	89.618	0.877	0.159	0.009
OHER	MIHCSA-S	1253.80	59.516	0.118	0.019	89.694	0.606	0.162	0.009
OHER	MIHCSA-M	1330.90	60.777	0.107	0.017	88.938	0.644	0.16	0.009
OHER	MIHCSA-L	1392.42	70.266	0.087	0.018	88.271	0.726	0.157	0.009
Significant		True ^e		True		True ^f		False ^g	
PHER	MG-HCSA	882.648	43.358	0.125	0.016	92.239	0.492	0.161	0.008
PHER	MIHCSA-S	890.847	26.554	0.122	0.022	92.24	0.303	0.16	0.008
PHER	MIHCSA-M	890.237	42.345	0.122	0.019	92.299	0.451	0.161	0.008
PHER	MIHCSA-L	979.421	33.188	0.112	0.019	91.631	0.363	0.161	0.009
Significant		True ^h		True ⁱ		True ^j		False	
RHER	MG-HCSA	1590.41	66.83	0.157	0.02	85.945	0.811	0.153	0.01
RHER	MIHCSA-S	1592.31	72.526	0.153	0.019	85.877	0.882	0.151	0.01
RHER	MIHCSA-M	1618.38	69.943	0.151	0.016	85.521	0.876	0.146	0.009
RHER	MIHCSA-L	1717.02	70.612	0.145	0.019	84.245	0.911	0.144	0.01
Significant		True ^k		False ^l		True ^m		True	
SHER	MG-HCSA	2321.94	100.304	0.047	0.01	74.558	1.826	0.123	0.011
SHER	MIHCSA-S	2314.24	132.046	0.045	0.009	74.572	2.239	0.121	0.01
SHER	MIHCSA-M	2328.53	122.946	0.043	0.007	74.419	2.161	0.12	0.011
SHER	MIHCSA-L	2275.90	102.101	0.044	0.01	75.279	1.804	0.115	0.012
Significant		False		False ⁿ		False		False ^o	

^aFalse for MIHCSA-S and MIHCSA-M

^bFalse for MG-HCSA and MIHCSA-S, MG-HCSA and MIHCSA-M, MIHCSA-S and MIHCSA-M

^cFalse for MIHCSA-S and MIHCSA-M

^dTrue for MG-HCSA and MIHCSA-L

^eFalse for MG-HCSA and MIHCSA-S

^fFalse for MG-HCSA and MIHCSA-S

^gTrue for MIHCSA-S and MIHCSA-L

^hFalse for MG-HCSA and MIHCSA-S, MG-HCSA and MIHCSA-M, MIHCSA-S and MIHCSA-M

ⁱFalse for MG-HCSA and MIHCSA-S, MG-HCSA and MIHCSA-M, MIHCSA-S and MIHCSA-M

^jFalse for MG-HCSA and MIHCSA-S, MG-HCSA and MIHCSA-M, MIHCSA-S and MIHCSA-M

^kFalse for MG-HCSA and MIHCSA-S, MG-HCSA and MIHCSA-M, MIHCSA-S and MIHCSA-M

^lTrue for MG-HCSA and MIHCSA-L

^mFalse for MG-HCSA and MIHCSA-S, MG-HCSA and MIHCSA-M, MIHCSA-S and MIHCSA-M

ⁿTrue for MG-HCSA and MIHCSA-M

^oTrue for MG-HCSA and MIHCSA-L, MIHCSA-S and MIHCSA-L

Table A.8: Summary of results from an empirical study into the MI-HCSA.

A.2.4 Evolved Immunity Empirical Study

Table A.9 summarises the results from the empirical study in Section 6.8.2 investigating the Evolved Immunity Host Clonal Selection Algorithm (EI-HCSA). The non-parametric Mann-Whitney U statistical test was calculated pair-wise for all algorithms on each HER group.

Problem	System	PH		PE		AHD		AHE	
		\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
AHER	MGHCSA	1272.81	62.233	0.111	0.016	89.697	0.601	0.163	0.01
AHER	EI-HCSA	1048.32	71.445	0.112	0.019	90.22	0.561	0.165	0.014
	Significant	True		False		True		False	
OHER	MGHCSA	1319.73	75.51	0.123	0.017	89.064	0.77	0.161	0.01
OHER	EI-HCSA	1142.77	72.429	0.112	0.016	88.975	0.626	0.163	0.009
	Significant	True		True		False		False	
PHER	MGHCSA	902.305	35.86	0.118	0.016	92.149	0.399	0.161	0.009
PHER	EI-HCSA	588.025	55.818	0.126	0.021	91.98	0.678	0.164	0.012
	Significant	True		False		False		False	
RHER	MGHCSA	1605.65	65.629	0.146	0.019	85.69	0.789	0.149	0.01
RHER	EI-HCSA	1454.08	67.685	0.158	0.018	85.77	0.853	0.153	0.012
	Significant	True		True		False		False	
SHER	MGHCSA	2328.98	117.362	0.048	0.01	74.437	2.041	0.123	0.011
SHER	EI-HCSA	2272.29	105.348	0.049	0.009	74.247	1.812	0.125	0.012
	Significant	False		False		False		False	

Table A.9: Summary of results from an empirical study into the EI-HCSA.

Appendix B

Works Arising from the Study

This appendix lists all works derived from this research project in the interest of full disclosure.

B.1 Peer-Reviewed Publications

This section lists peer-reviewed publications derived from this research project.

1. J. Brownlee *IIDLE: An Immunological Inspired Distributed Learning Environment for Multiple Objective and Hybrid Optimisation*. Proceedings of the IEEE Congress in Evolutionary Computation (CEC'06); Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada. USA: IEEE Press; 2006: 507-513.

B.2 Software

This section lists software created as a result of this research project. All experimentation in this dissertation were achieved using the OAT platform with customised extensions.

1. *IIDLE: Immunologically Inspired Distributed Learning Environment*: IIDLE is based in a holistic abstraction of the acquired immune system as a spatially distributed, circulating and heterogeneous population of specialised discrete units that provide a homogeneous defence against external pathogenic material. It is these discrete, atomic, immutable and disposable units or information packets that represent the knowledge captured by the system, and are the substrate for the maintenance and learning processes. Includes Java implementation, with applets, user interface, and optimisation experiments. See [60] and download the software from <http://www.ict.swin.edu.au/personal/jbrownlee/iidle/index.html> or <http://jason.brownlee05.googlepages.com>.
2. *WEKA Classification Algorithms*: A collection of plug-in algorithms for the WEKA machine learning workbench including Artificial Neural Network (ANN) algorithms and Artificial Immune System (AIS) algorithms. Algorithms include Learning Vector Quantisation (LVQ), Perceptron, Back-Propagation, Self-Organising Map (SOM), Artificial Immune Recognition System (AIRS), and the Clonal Selection Algorithm (CLONALG). Download the software from: <http://weka.classalgos.sourceforge.net>.
3. *Optimisation Algorithm Toolkit (OAT)*: A workbench and toolkit for developing, evaluating, and playing with classical and state-of-the-art optimisation algorithms

on standard benchmark problem domains; including reference algorithm implementations (Java), graphing, visualisations and much more. Algorithms include Evolutionary, Swarm, Immune System and others. Domains include Continuous Function, Travelling Salesman, Protein Folding and others. Download the software from: <http://optalgtoolkit.sourceforge.net>.

B.3 Unpublished Technical Reports

This section lists unpublished technical reports derived from this research project, many of which were reviewed internally (Complex Intelligent Systems Laboratory). These technical reports are available on the Internet at the following URL: <http://jason.brownlee05.googlepages.com>.

B.3.1 Year 2005

1. J. Brownlee *Artificial Immune Recognition System (AIRS) - A Review and Analysis* [Technical Report]. Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2005 Jan; Technical Report ID: 1-01.
2. J. Brownlee *Clonal Selection Theory and CLONALG - The Clonal Selection Classification Algorithm (CSCA)* [Technical Report]. Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2005 Jan; Technical Report ID: 2-01.
3. J. Brownlee *Immunos-81 - The Misunderstood Artificial Immune System* [Technical Report]. Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2005 Feb; Technical Report ID: 3-01.
4. J. Brownlee *Immunological Inspired Novelty Detection* [Technical Report]. Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2005 Apr; Technical Report ID: 4-01.
5. J. Brownlee *On Biologically Inspired Computation* [Technical Report]. Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2005 Jul; Technical Report ID: 5-01.
6. J. Brownlee *Protein Folding - A Benchmark Combinatorial Optimisation Problem* [Technical Report]. Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2005 Jul; Technical Report ID: 6-01.
7. J. Brownlee *The Pole Balancing Problem - A Review of a Benchmark Control Theory Problem* [Technical Report]. Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2005 Jul; Technical Report ID: 7-01.
8. J. Brownlee *Introduction to IIDLE - The Immunological Inspired Distributed Learning Environment* [Technical Report]. Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2005 Sep; Technical Report ID: 8-01.
9. J. Brownlee *Preliminary Experiments with IIDLE* [Technical Report]. Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2005 Sep; Technical Report ID: 9-01.

10. J. Brownlee *Implementation Specification for IIDLE* [Technical Report]. Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2005 Sep; Technical Report ID: 10-01.
11. J. Brownlee *Further Preliminary Experiments with IIDLE* [Technical Report]. Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2005 Oct; Technical Report ID: 11-01.

B.3.2 Year 2006

1. D. Angus, and J. Brownlee *Discrete History Ant Systems* [Unpublished Manuscript]. Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2006.
2. J. Brownlee *A Coarse Taxonomy of Artificial Immune Systems* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2006 May; Technical Report ID: TR2006-01.

B.3.3 Year 2007

1. J. Brownlee *A Note on Research Methodology and Benchmarking Optimization Algorithms* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jan; Technical Report ID: 070125.
2. J. Brownlee *Clonal Selection Algorithms* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Feb; Technical Report ID: 070209A.
3. J. Brownlee *A Review of the Clonal Selection Theory of Acquired Immunity* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Feb; Technical Report ID: 070223A.
4. J. Brownlee *A Series of Adaptive Models Inspired by the Acquired Immune System* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Feb; Technical Report ID: 070227A.
5. J. Brownlee *Complex Adaptive Systems* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Mar; Technical Report ID: 070302A.
6. J. Brownlee *Satisficing, Optimization, and Adaptive Systems* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Mar; Technical Report ID: 070305A.
7. J. Brownlee *The 'shape-space' and 'affinity landscape' Immunological Paradigms* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Mar; Technical Report ID: 070310A.

8. J. Brownlee *The Physiology of Lymphocyte Migration* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Mar; Technical Report ID: 070316A.
9. J. Brownlee *A Series of Discrete Repertoire Models Inspired by Lymphocyte Migration* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Mar; Technical Report ID: 070320A.
10. J. Brownlee *An Adaptive Systems Formalism* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Mar; Technical Report ID: 070320A
11. J. Brownlee *'Colour Space': A Contrived Problem Domain for Investigating Adaptive Models* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Mar; Technical Report ID: 070322A.
12. J. Brownlee *'Small Models': A Methodology for Designing and Investigating Adaptive Systems* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Mar; Technical Report ID: 070326A.
13. J. Brownlee *Autonomous Distributed Control in the Immune System Using Diffuse Feedback* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Mar; Technical Report ID: 070329A.
14. J. Brownlee *Artificial Immune System Thesis Bibliography* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Mar; Technical Report ID: 070330A.
15. J. Brownlee *Experimental Scenarios in Colour Space* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Apr; Technical Report ID: 070402A
16. J. Brownlee *Populations of Interacting Immune Systems: Evolution and Immunization* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Apr; Technical Report ID: 070410A.
17. J. Brownlee *A Series of Multiple Immune System Adaptive Models Inspired by Evolution and Immunization* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Apr; Technical Report ID: 070413A
18. J. Brownlee *Realizing Elementary Clonal Selection Algorithms* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Apr; Technical Report: 070418A

19. J. Brownlee *The 'Pathogenic Exposure' Paradigm* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Apr; Technical Report ID: 070422A
20. J. Brownlee *Antigen-Antibody Interaction* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Apr; Technical Report ID: 070427A
21. J. Brownlee *Realizing Elementary Discrete Repertoire Clonal Selection Algorithms* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Apr; Technical Report ID: 070430A
22. J. Brownlee *Cognition, Immunology, and the Cognitive Paradigm* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 May; Technical Report: 070504A.
23. J. Brownlee *Lazy and Competitive Learning* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 May; Technical Report: 070508A.
24. J. Brownlee *Learning Classifier Systems* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 May; Technical Report: 070514A.
25. J. Brownlee *Towards Unification...* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 May; Technical Report: 070516A
26. J. Brownlee *A Clonal Selection Algorithm and Extensions* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 May; Technical Report: 070521A.
27. J. Brownlee *A T-Cell Mediated Clonal Selection Algorithm* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 May; Technical Report: 070523A
28. J. Brownlee *An Intra-Repertoire Recognition Algorithm* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 May; Technical Report: 070525A
29. J. Brownlee *A Spatial Clonal Selection Algorithm* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 May; Technical Report: 070531A
30. J. Brownlee *Optimizing and Matching Bitstrings* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jun; Technical Report: 070604A.

31. J. Brownlee *A Lymphocyte Homing Algorithm* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jun; Technical Report: 070607A.
32. J. Brownlee *A Lymphoid Tissue Model and Lymphocyte Recruitment Algorithms* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jun; Technical Report: 070609A.
33. J. Brownlee *A Hierarchical Framework of the Acquired Immune System* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jun; Technical Report: 070613A.
34. J. Brownlee *Darwinism and Selectionist Theories* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jun; Technical Report: 070616A.
35. J. Brownlee *Information Processing with a Lymphoid Tissue Architecture* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jun; Technical Report: 070619A.
36. J. Brownlee *A Population-Based Clonal Selection Algorithm and Extensions* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jun; Technical Report: 070621A.
37. J. Brownlee *A Generational Population-Based Clonal Selection Algorithm and Extensions* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jun; Technical Report: 070623A.
38. J. Brownlee *Models, Algorithms, and the Hierarchical Acquired Immune System Framework* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jun; Technical Report: 070625A.
39. J. Brownlee *Integration of the Pathogenic Exposure Paradigm and the Hierarchical Immune System Framework* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jun; Technical Report: 070626A.
40. J. Brownlee *Towards General Information Processing Models of the Acquired Immune System* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jun; Technical Report: 070627A.
41. J. Brownlee *A Review of the Immunological Inspired Distributed Learning Environment* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jul; Technical Report: 070702A.

42. J. Brownlee *Convergence and Dominance in the Minimal Clonal Selection Algorithm* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jul; Technical Report: 070706A.
43. J. Brownlee *Immunology as a Metaphor for Adaptive and Distributed Information Processing* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Jul; Technical Report: 070714A.
44. J. Brownlee *Debrief of PhD Project Practices* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Dec; Technical Report: 20071204-A.
45. J. Brownlee *Unexplored Territory: Seeds for Future Research Investigations* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Dec; Technical Report: 20071210-A.
46. J. Brownlee *OAT HowTo: High-level Domain, Problem, and Algorithm Implementation* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Dec; Technical Report: 20071218A.
47. J. Brownlee *OAT: The Optimization Algorithm Toolkit* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Dec; Technical Report: 20071220A.
48. J. Brownlee *OAT: A Quick Start Guide* [Technical Report]. Victoria, Australia: Complex Intelligent Systems Laboratory (CIS), Centre for Information Technology Research (CITR), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology; 2007 Dec; Guide: 20071220A.

Bibliography

- [1] A. Aamodt and E. Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *Artificial Intelligence Communications*, 7(1):39–59–, 1994.
- [2] S Aaronson. Np-complete problems and physical reality. *ACM SIGACT News (COLUMN: Complexity theory)*, 36(1):30–52–, 2005.
- [3] P. Adamidis, P. Adamidis, and V. Petridis. Co-operating populations with different evolution behaviours. In V. Petridis, editor, *Proc. IEEE International Conference on Evolutionary Computation*, pages 188–191, 1996.
- [4] F. Agens, M. M. Rosado, and A. A. Freitas. Peripheral b cell survival. *Cell Mol Life Sci*, 57(8-9):1220–1228, Aug 2000.
- [5] D. W. Aha. Case-based learning algorithms. In *Proceedings of the 1991 DARPA Case-Based Reasoning Workshop*, pages 147–158–. Morgan Kaufmann, 1991.
- [6] D. W. Aha. Artificial intelligence review. *Editorial: Lazy Learning*, 11(1-5):7–10–, February 1997.
- [7] D. W. Aha. Feature weighting for lazy learning algorithms. In H. Motoda, editor, *Feature Extraction, Construction and Selection: A Data Mining Perspective*, pages –. Kluwer, Norwell MA, USA, 1998.
- [8] D. W. Aha, D. Kibler, and M. K. Albert. Instance-based learning algorithms. *Machine Learning*, 6(1):37–66–, January 1991.
- [9] S. C. Ahalt, A. K. Krishnamurthy, P. Chen, and D. E. Melton. Competitive learning algorithms for vector quantization. *Neural Networks*, 3(3):277–290–, 1990.
- [10] E. Ahmed, A. S. Elgazzar, and A. S. Hegazi. An overview of complex adaptive systems. *ARXIV, eprint arXiv:nlin/0506059*, pages –, July 2005.
- [11] E. Ahmed and A. H. Hashish. On modelling the immune system as a complex system. *Theory in biosciences*, 124(3-4):413–418–, 2006.
- [12] E. Alba. Parallel evolutionary algorithms can achieve super-linear performance. *Information Processing Letters*, 82:7–13, 2002.
- [13] E. Alba. *Parallel Metaheuristics: A New Class of Algorithms*. John Wiley, 2005.
- [14] E. Alba and M. Tomassini. Parallelism and evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 6(5):443–462, 2002.
- [15] E. Alba and J. M. Troya. Influence of the migration policy in parallel distributed gas with structured and panmictic populations. *Applied Intelligence*, 12:163–181, 2000.

- [16] M. M. Ali, C. Storey, and A. Trn. Application of stochastic global optimization algorithms to practical problems. *Journal of Optimization Theory and Applications*, 95(3):545–563, 1997.
- [17] K. P. Anchor, P. D. Williams, G. H. Gunsch, and G. B. Lamont. The computer defense immune system: current and future research in intrusion detection. In *Proceedings of the 2002 Congress on Evolutionary Computation (CEC '02)*, volume 2, pages 1027–1032–. IEEE Press, USA, 2002.
- [18] K. P. Anchor, J. B. Zydallis, G. H. Gunsch, and G. B. Lamont. Extending the computer defense immune system: Network intrusion detection with a multiobjective evolutionary programming approach. In P. J. Bentley, editor, *First International Conference on Artificial Immune Systems (ICARIS 2002)*, pages 1221–, UK, 2002.
- [19] A. O. Anderson. Structure and organization of the lymphatic system. In Ethan M. Shevach, editor, *Immunophysiology : the role of cells and cytokines in immunity and inflammation*, pages 14–45–. Oxford University Press, New York, 1990.
- [20] A. O. Anderson and S. Shaw. Chapter 3: Lymphocyte trafficking. In T.A. Fleisher, B.D. Schwartz, W.T. Shearer, and W. Strober, editors, *Clinical Immunology, Principles and Practice*, pages 39–49–. Mosby, London, 1995.
- [21] P. W. Anderson, K. J. Arrow, and D. Pines. *Proceedings of The Santa Fe Institute Studies in the Sciences of Complexity - Economy As an Evolving Complex System*. Addison Wesley Publishing Company, USA, 1988.
- [22] R. M. Anderson. Modern vaccines: immunisation and herd immunity. *The Lancet*, 335(8690):641–646–, March 1990.
- [23] R. M. Anderson and R. M. May. Vaccination and herd immunity to infectious diseases. *Nature*, 318:323–329–, November 1985.
- [24] P. S. Andrews and J. Timmis. Inspiration for the next generation of artificial immune systems. In Marcin L. Pilat, Peter J. Bentley, and Jonathan Timmis, editors, *Lecture Notes in Computer Science*, pages 126–138–. Springer-Verlag, 2005.
- [25] C. Andrieu, N. de Freitas, A. Doucet, and M. I. Jordan. An introduction to mcmc for machine learning. *Machine Learning*, 50:5–43, 2003.
- [26] J. B. Andr, J. B. Ferdy, and B. Godelle. Within-host parasite dynamics, emerging trade-off, and evolution of virulence with immune system. *Evolution*, 57(7):1489–1497, 2003.
- [27] A.M. Anile, V. Cutello, G. Narzisi, G. Nicosia, and S. Spinella. Lipschitzian pattern search and immunological algorithm with quasi-newton method for the protein folding problem: An innovative multistage approach. In *Lecture Notes in Computer Science*, pages 307–323–. Springer, 2006.
- [28] W. Brian Arthur. Introduction: Process and emergence in the economy. In Steven Durlauf and David A. Lane, editors, *The Economy as an Evolving Complex System II*, volume Volume XXVII, pages 1–04–. Addison-Wesley Pub. Co, Reading, Mass, USA, 1997.
- [29] C. G. Atkeson, A. W. Moore, and S. Schaal. Locally weighted learning. *Artificial Intelligence Review*, 11(1-5):11–73–, February 1997.

- [30] T. Bäck. Optimal mutation rates in genetic search. In *Proceedings of the Fifth International Conference on Genetic Algorithms*, pages 2–9–. Morgan Kaufman, San Mateo, CA, USA, 1993.
- [31] T. Back, D. B Fogel, and Z. Michalwicz. *Evolutionary Computation 1 - Basic Algorithms and Operators*. Institute of Physics (IoP) Publishing, Bristol, UK, 2000.
- [32] T. Back, D. B Fogel, and Z. Michalwicz. *Evolutionary Computation 2 - Advanced Algorithms and Operators*. Institute of Physics (IoP) Publishing, Bristol, UK, 2000.
- [33] R. Barr, B. Golden, J. Kelly, M. Rescende, and W. Stewart. Designing and reporting on computational experiments with heuristic methods. *Journal of Heuristics*, 1:9–32–, 1995.
- [34] S. Behnke. *Hierarchical Neural Networks for Image Interpretation*. Springer, 2003.
- [35] T. C. Belding. The distributed genetic algorithm revisited. In *Proceedings of the 6th International Conference on Genetic Algorithms*, 1995.
- [36] P. J. Bentley, J. Greensmith, and S. Ujjin. Two ways to grow tissue for artificial immune systems. In Marcin L. Pilat, Peter J. Bentley, and Jonathan Timmis, editors, *Lecture Notes in Computer Science*, pages 139–. Springer, Berlin / Heidelberg, Germany, 2005.
- [37] C. Berek, A. Berger, and M. Apel. Maturation of the immune response in germinal centers. *Cell*, 67(6):1121–1129–, 1991.
- [38] C. Berek and M. Ziegner. The maturation of the immune response. *Immunology Today*, 14(8):400–404–, 1993.
- [39] D. Bergemann and J. Valimaki. Bandit problems. Cowles Foundation Discussion Papers 1551, Cowles Foundation, Yale University, January 2006.
- [40] H. Bersini and F. Varela. Hints for adaptive problem solving gleaned from immune networks. In *Lecture Notes In Computer Science*, pages 343–354–. Springer-Verlag, London, UK, 1990.
- [41] K. Beyer, J. Goldstein, R. Ramakrishnan, and U. Shaft. When is "nearest neighbor" meaningful? In *Lecture Notes in Computer Science*, pages 217–, Springer Berlin / Heidelberg, 1998.
- [42] J. Bienenstock and A. D. Befus. Mucosal immunology. *Immunology*, 41(2):249–270–, 1980.
- [43] M. Birattari, P. Balaprakash, and M. Dorigo. Aco/f-race: Ant colony optimization and racing techniques for combinatorial optimization under uncertainty. In M. Gendreau, P. Greistorfer, W. J. Gutjahr, R. F. Hartl, and M. Reimann, editors, *MIC 2005: The Sixth Metaheuristics International Conference*, pages 107–112–, 2005.
- [44] M. Birattari and M. Dorigo. How to assess and report the performance of a stochastic algorithm on a benchmark problem: Mean or best result on a number of runs? Technical report, IRIDIA, Universit Libre de Bruxelles, Brussels, Belgium, 2005.
- [45] M. Birattari, T. Sttzle, L. Paquete, and K. Varrentropp. A racing algorithm for configuring metaheuristics. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 11–18–. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2002.

- [46] C. M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.
- [47] A. Di Blas, A. Jagota, and R. Hughey. Optimizing neural networks on simd parallel computers. *Parallel Computing*, 31:97–115, 2005.
- [48] C. Blum and A. Roli. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys (CSUR)*, 35(3):268–308–, September 2003.
- [49] L. Boggess and J. S. Hamaker. The effect of irrelevant features on airs, an artificial immune-based classifier. In *Intelligent Engineering Systems through Artificial Neural Networks (ANNIE)*, pages 219–224–, 2003.
- [50] S. Bonhoeffer and M. A. Nowak. Intra-host versus inter-host selection: viral strategies of immune function impairment. *Proc Natl Acad Sci U S A*, 91(17):8062–8066, Aug 1994.
- [51] L. Bottou and V. Vapnik. Local learning algorithms. *Neural Computation*, 4(6):888–900–, November 1992.
- [52] J. Y. L. Boudec and S. Sarafijanovic. An artificial immune system approach to misbehavior detection in mobile ad hoc networks. In Masayuki Murata and Naoki Wakamiya, editors, *Lecture Notes in Computer Science*, pages 396–. Springer, Berlin / Heidelberg, Germany, 2004.
- [53] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [54] P. Brandtzaeg. Mucosal immunity: integration between mother and the breast-fed infant. *Vaccine*, 21(24):3382–3388–, 2003.
- [55] L. Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, 1996.
- [56] J. Breivik. Commentary: cancer–evolution within. *Int J Epidemiol*, 35(5):1161–1162, Oct 2006.
- [57] P. Bretscher and M Cohn. A theory of self-nonsel self discrimination. *Science*, 169(950):1042–109–, 1970.
- [58] R. Brooks. A robust layered control system for a mobile robot. *IEEE Journal Of Robotics And Automation*, 2(1):14–23, 1986.
- [59] D. S. Broomhead and David Lowe. Radial basis functions, multi-variable functional interpolation and adaptive networks. Technical report, Royal Signals and Radar Establishment Malvern, England, 1988.
- [60] J. Brownlee. IIDLE: An immunological inspired distributed learning environment for multiple objective and hybrid optimisation. In *Proceedings of the IEEE Congress in Evolutionary Computation (CEC’06)*, pages 507–513. IEEE Press, 2006.
- [61] M. D. Buhmann. Radial basis functions. *Acta Numerica*, 9(00):1–38–, 2000.
- [62] P. Bull, A. Knowles, G. Tedesco, and A. Hone. Diophantine benchmarks for the b-cell algorithm. In *Proceedings 5th International Conference on Artificial Immune Systems (ICARIS 2006)*, pages –, 2006.

- [63] F. M. Burnet. A modification of jernes theory of antibody production using the concept of clonal selection. *Australian Journal of Science*, 20,:67–69–, 1957.
- [64] F. M. Burnet. *The clonal selection theory of acquired immunity*. Vanderbilt University Press, Nashville, Tennessee, U.S.A., 1959.
- [65] F. M. Burnet. Clonal selection and after. *Theoretical Immunology*, pages 63–85–, 1978.
- [66] E. C. Butcher and L. J. Picker. Lymphocyte homing and homeostasis. *Science*, 272(5258):60–67–, April 1996.
- [67] E. C. Butcher, M. Williams, K. Youngman, L. Rott, and M. Briskin. Lymphocyte trafficking and regional immunity. *Advances in Immunology*, 72:209–253–, 1999.
- [68] E. Cant-Paz. *Efficient and Accurate Parallel Genetic Algorithms*. Kluwer Academic Publishers (Springer), 2000.
- [69] E. Cant-Paz and D. E. Goldberg. Are multiple runs of genetic algorithms better than one? In *Genetic and Evolutionary Computation, GECCO 2003*, volume 2723 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2003.
- [70] J. Carneiro and J. Stewart. Rethinking "shape space": evidence from simulated docking suggests that steric shape complementarity is not limiting for antibody-antigen recognition and idiotypic interactions. *Journal of theoretical biology*, 169(4):391–402–, 1994.
- [71] G. A. Carpenter and S. Grossberg. The art of adaptive pattern recognition by a self-organizing neural network. *Computer*, 21(3):77–88–, March 1988.
- [72] D. J. Cavicchio. *Adaptive search using simulated evolution*. PhD thesis, Computer and Communication Sciences, University of Michigan, USA, 1970.
- [73] D. J. Cavicchio. Reproductive adaptive plans. In *Proceedings of the ACM annual conference - Volume 1*, volume 1, pages 60–70–. ACM Press, New York, NY, USA, 1972.
- [74] K Chan, A Saltelli, and S Tarantola. Sensitivity analysis of model output: variance-based methods make the difference. In *Proceedings of the 29th conference on Winter simulation (Winter Simulation Conference)*, pages 261–268–. ACM Press, New York, NY, USA, 1997.
- [75] M. Chiarandini, D. Basso, and T. Sttzle. Statistical methods for the comparison of stochastic optimizers. In Michel Gendreau, Peter Greistorfer, Walter J. Gutjahr, Richard F. Hartl, and Marc Reimann, editors, *MIC2005: Proceedings of the 6th Metaheuristics International Conference*, pages 189–196–, 2005.
- [76] E. Clark, A. Hone, and J. Timmis. A markov chain model of the b-cell algorithm. In *Lecture Notes in Computer Science*, pages 318–330–. Springer, Berlin / Heidelberg, 2005.
- [77] C. A. Coello Coello and N. C. Cortes. An approach to solve multiobjective optimization problems based on an artificial immune system. In P. J. Bentley, editor, *First International Conference on Artificial Immune Systems (ICARIS)*, pages 212–221–, 2002.

- [78] C. A. Coello Coello and N. C. Cortes. Hybridizing a genetic algorithm with an artificial immune system for global optimization. *Engineering Optimization*, 36(5):607–634–, October 2004.
- [79] C. A. Coello Coello and N. C. Cortes. A parallel implementation of an artificial immune system to handle constraints in genetic algorithms: Preliminary results. In *Congress on Evolutionary Computation (CEC'2002)*, volume 1, pages 819–824–, USA, 2002. IEEE Computer Society.
- [80] I. R. Cohen. Antigenic mimicry, clonal selection and autoimmunity. *Journal of Autoimmunity*, 16(3):337–340–, 2001.
- [81] I. R. Cohen. *Tending Adam's Garden: Evolving the Cognitive Immune Self*. Academic Press, New York, 2001.
- [82] I. R. Cohen, U. Hershberg, and S. Solomon. Antigen-receptor degeneracy and immunological paradigms. *Molecular Immunology*, 40(14-15):993–996–, 2004.
- [83] N. C. Cortes and C. A. Coello Coello. Multiobjective optimization using ideas from the clonal selection principle. In *Lecture Notes in Computer Science*, pages 158–170–, Berlin / Heidelberg, 2003. Springer.
- [84] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27–, January 1967.
- [85] T. G. Crainic and N. Hail. Parallel metaheuristics applications. In *Parallel Metaheuristics*. John Wiley & Sons, Inc., 2005.
- [86] E. Crist and A. I. Tauber. Selfhood, immunity, and the biological imagination: The thought of frank macfarlane burnet. *Biology and Philosophy*, 15(4):509–533–, September 2000.
- [87] J. Cunliffe. Tissue homeostasis and immunity—more on models. *Scand J Immunol*, 64(3):172–176, Sep 2006.
- [88] V. Cutello, D. Lee, S. Leone, G. Nicosia, and M. Pavone. Clonal selection algorithm with dynamic population size for bimodal search spaces. In *Lecture Notes in Computer Science*, pages 949–958–. Springer, 2006.
- [89] V. Cutello, D. Lee, G. Nicosia, M. Pavone, and I. Prizzi. Aligning multiple protein sequences by hybrid clonal selection algorithm with insert-remove-gaps and block-shuffling operators. In *Lecture Notes in Computer Science*, pages 321–334–. Springer, Berlin / Heidelberg, 2006.
- [90] V. Cutello, G. Morelli, G. Nicosia, and M. Pavone. Immune algorithms with aging operators for the string folding problem and the protein folding problem. In Jens Gottlieb, editor, *Lecture Notes in Computer Science*, pages 80–90–, Berlin / Heidelberg, Germany, 2005. Springer.
- [91] V. Cutello, G. Narzisi, and G. Nicosia. A multi-objective evolutionary approach to the protein structure prediction problem. *Journal of the Royal Society Interface*, 3(6):139–151–, 2006.
- [92] V. Cutello, G. Narzisi, G. Nicosia, and M. Pavone. Clonal selection algorithms: A comparative case study using effective mutation potentials. In Marcin L. Pilat, Peter J. Bentley, and Jonathan Timmis, editors, *LNCS*, pages 13–28–, Germany, 2005. Springer.

- [93] V. Cutello, G. Narzisi, G. Nicosia, and M. Pavone. An immunological algorithm for global numerical optimization. In *LNCS*, pages 284–295–, 2005.
- [94] V. Cutello, G. Narzisi, G. Nicosia, M. Pavone, and G. Sorace. How to escape traps using clonal selection algorithms. In *The First International Conference on Informatics in Control, Automation and Robotics, ICINCO 2004*, pages 322–326–. INSTICC Press, 2004.
- [95] V. Cutello and G. Nicosia. An immunological approach to combinatorial optimization problems. In J.C. Riquelme and M. Toro, editors, *Lecture Notes In Computer Science*, pages 361–370–, London, UK, 2002. Springer-Verlag.
- [96] V. Cutello and G. Nicosia. Multiple learning using immune algorithms. In *Proceedings of 4th International Conference on Recent Advances in Soft Computing, RASC 2002*, pages 102–107–, 2002.
- [97] V. Cutello and G. Nicosia. Chapter vi. the clonal selection principle for in silico and in vivo computing. In Fernando J. Von Zuben, editor, *Recent Developments in Biologically Inspired Computing*, pages 104–146–, Hershey, London, Melbourne, Singapore, 2005. Idea Group Publishing.
- [98] V. Cutello and G. Nicosia. A clonal selection algorithm for coloring, hitting set and satisfiability problems. In *Lecture Notes in Computer Science*, pages 324–337–. Springer, 2006.
- [99] V. Cutello, G. Nicosia, and E. Pavia. A parallel immune algorithm for global optimization. In *Series on Advances in Soft Computing*, pages 467–475–. Springer-Verlag, 2006.
- [100] V. Cutello, G. Nicosia, and M. Pavone. A hybrid immune algorithm with information gain for the graph coloring problem. In *Lecture Notes in Computer Science*, pages 171–182–, Berlin / Heidelberg, 2003. Springer.
- [101] V. Cutello, G. Nicosia, and M. Pavone. Exploring the capability of immune algorithms: A characterization of hypermutation operators. In Vincenzo Cutello, Peter J. Bentley, and Jon Timmis, editors, *Lecture Notes in Computer Science*, pages 263–276–, Berlin / Heidelberg, 2004. Springer.
- [102] V. Cutello, G. Nicosia, and M. Pavone. An immune algorithm with hypermacromutations for the dill’s 2d hydrophobic-hydrophilic model. In *Congress on Evolutionary Computation, (CEC2004)*, volume 1, pages 1074–1080–, USA, 2004. IEEE Computer Society.
- [103] V. Cutello, G. Nicosia, M. Pavone, and G. Narzisi. Real coded clonal selection algorithm for unconstrained global optimization using a hybrid inversely proportional hypermutation operator. In *ACM*, volume 2, pages 950–954–. ACM Press, New York, NY, USA, 2006.
- [104] V. Cutello, G. Nicosia, M. Pavone, and J. Timmis. An immune algorithm for protein structure prediction on lattice models. *IEEE Transactions on Evolutionary Computation*, pages –, 2006.
- [105] N. Danilova. The evolution of immune mechanisms. *Journal of experimental zoology. Part B. Molecular and developmental evolution*, 306(6):496–520–, 2006.

- [106] D. Dasgupta. Chapter ten: Information processing in the immune system. In M. Dorigo and F. Glover, editors, *New Ideas in Optimization*, pages 203–216. McGraw-Hill, England, 1999.
- [107] D. Dasgupta and H. Brian. Mobile security agents for network traffic analysis. *DARPA Information Survivability Conference and Exposition (DISCEX II'01)*, 2:1332–, 2001.
- [108] L. N. de Castro and J. Timmis. An artificial immune network for multimodal function optimization. In *Proceedings of the 2002 Congress on Evolutionary Computation (CEC '02)*, volume 1, pages 699–704–. IEEE Computer Society, USA, 2002.
- [109] L. N. de Castro and J. Timmis. *Artificial Immune Systems: A new computational intelligence approach*. Springer-Verlag, Great Britain, 2002.
- [110] L. N. de Castro and J. Timmis. Artificial immune systems: A novel paradigm to pattern recognition. In L. Alonso and C. Fyfe, editors, *Artificial Neural Networks in Pattern Recognition*, pages 67–84–. University of Paisley, UK, 2002.
- [111] L. N. de Castro and J. Timmis. Artificial immune systems as a novel soft computing paradigm. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, 7(8):526–544–, August 2003.
- [112] L. N. de Castro and F. J. Von Zuben. Artificial immune systems - part i: Basic theory and applications. Technical report, Department of Computer Engineering and Industrial Automation, School of Electrical and Computer Engineering, State University of Campinas, Brazil, December 1999.
- [113] L. N. de Castro and F. J. Von Zuben. Artificial immune systems: Part ii: A survey of applications. Technical report, Department of Computer Engineering and Industrial Automation, School of Electrical and Computer Engineering, State University of Campinas, Brazil, February 2000.
- [114] L. N. de Castro and F. J. Von Zuben. The clonal selection algorithm with engineering applications. In David E. Goldberg, Erick Cant-Paz, Lee Spector, Ian C. Parmee, and Hans-Georg Beyer, editors, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '00), Workshop on Artificial Immune Systems and Their Applications*, pages 36–37–. Morgan Kaufmann, 2000.
- [115] L. N. de Castro and F. J. Von Zuben. An evolutionary immune network for data clustering. In *Proceedings Sixth Brazilian Symposium on Neural Networks*, pages 84–89–. IEEE Computer Society, USA, 2000.
- [116] L. N. de Castro and F. J. Von Zuben. Chapter xii: ainet: An artificial immune network for data analysis. In R. A. Sarker and C. S. Newton, editors, *Data Mining: A Heuristic Approach*, pages 231–259–. Idea Group Publishing, USA, 2001.
- [117] L. N. de Castro and F. J. Von Zuben. Learning and optimization using the clonal selection principle. *IEEE Transactions on Evolutionary Computation*, 6(3):239–251, June 2002.
- [118] L. N. de Castro and F. J. Von Zuben. Chapter 1: From biologically inspired computing to natural computing. In Fernando J. Von Zuben, editor, *Recent developments in biologically inspired computing*, pages –. Idea Group, Hershey, PA, 2005.

- [119] E. D. de Jong, D. Thierens, and R. A. Watson. Hierarchical genetic algorithms. In *Parallel Problem Solving from Nature - PPSN VIII*, Lecture Notes in Computer Science, pages 232–241. Springer, 2004.
- [120] K. A. De Jong. *An Analysis of the Behavior of a Class of Genetic Adaptive Systems*. PhD thesis, Computer and Communication Sciences, University of Michigan, USA, 1975.
- [121] K. A. De Jong. Genetic algorithms are not function optimizers. In *Proceedings of the Second Workshop on Foundations of Genetic Algorithms (FOGA)*, pages 5–17–. Morgan Kaufmann, 1992.
- [122] S. Fazekas de St.Groth and R. G. Webster. Disquisitions on original antigenic sin. i. evidence in man. *The Journal of Experimental Medicine*, 124:331–345–, 1966.
- [123] K. Deb. Chapter 14 - speciation methods. In David B Fogel and Zbigniew Michalwicz, editors, *Evolutionary Computation 2 - Advanced Algorithms and Operations*, pages 93–100–. Institute of Physics (IOP) Publishing Ltd., UK, 2000.
- [124] K. Deb. *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley and Sons, 2001.
- [125] K. Deb and S. Agrawal. Understanding interactions among genetic algorithm parameters. In Colin R. Reeves, editor, *Proceedings of the Fifth Workshop on Foundations of Genetic Algorithms (FOGA)*, pages 265–286–. Morgan Kaufmann, 1999.
- [126] K. Deb and David E. Goldberg. An investigation of niche and species formation in genetic function optimization. In *Proceedings of the Third International Conference on Genetic Algorithms*, pages 42–50–. Morgan Kaufmann, 1989.
- [127] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, and D. Terry. Epidemic algorithms for replicated database maintenance. In *Proceedings of the sixth annual ACM Symposium on Principles of distributed computing*, 1987.
- [128] D. DeSieno. Adding a conscience to competitive learning. In *IEEE International Conference on Neural Networks*, volume 1, pages 117–124–. IEEE Press, USA, 1988.
- [129] P. D’haeseleer. An immunological approach to change detection: theoretical results. In *Proceedings of the 9th IEEE Computer Security Foundations Workshop*, pages 18–26–. IEEE Computer Society, Washington, DC, USA, 1996.
- [130] P. D’haeseleer, S. Forrest, and P. Helman. An immunological approach to change detection: algorithms, analysis and implications. In *IEEE Symposium on Security and Privacy*, pages 110–119–. IEEE Computer Society, Washington, DC, USA, 1996.
- [131] B. Du, B. Du, J. Gu, W. Wang, and D.H.K. Tsang. Multispace search for minimizing the maximum nodal degree. In Jun Gu, editor, *Proc. Sixth International Conference on Computer Communications and Networks*, pages 364–367, 1997.
- [132] A. E. Eiben and M. Jelasity. A critical note on experimental research methodology in ec. In *Proceedings of the 2002 Congress on Evolutionary Computation (CEC ’02)*, volume 1, pages 582–587–. IEEE Press, USA, 2002.
- [133] A. E. Eiben and C. A. Schippers. On evolutionary exploration and exploitation. *Fundamenta Informaticae*, 35(1-4):35–50–, August 1998.

- [134] M. El-Abd and M. Kamel. A taxonomy of cooperative search algorithms. In *Hybrid Metaheuristics*, 2005.
- [135] A. P. Engelbrecht. *Computational Intelligence: An Introduction*. J. Wiley & Sons, 2002.
- [136] J. Englund, W. P. Glezen, and P. A. Piedra. Maternal immunization against viral disease. *Vaccine*, 16(14-15):1456–1463–, 1998.
- [137] J. D. Farmer, S. A. Kauffman, N. H. Packard, and A. S. Perelson. Adaptive dynamic networks as models for the immune system and autocatalytic sets. *Annals of the New York Academy of Sciences*, 504(1):118131–, 1986.
- [138] J. D. Farmer, N. H. Packard, and Alan S. Perelson. The immune system, adaptation, and machine learning. *Physica D*, 22:187–204–, 1986.
- [139] D. Ferber. The education of t cells. *Science: IMMUNOLOGY*, 316(5822):191–193–, April 2007.
- [140] P. E. M. Fine. Herd immunity: History, theory, practice. *Oxford Journals, Epidemiologic Reviews*, 15(2):265–302–, 1993.
- [141] T. C. Fogarty. Varying the probability of mutation in the genetic algorithm. In *Proceedings of the third international conference on Genetic algorithms*, pages 104–109–. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1989.
- [142] D. B. Fogel. *Evolutionary Computation : Toward a New Philosophy of Machine Intelligence*. John Wiley & Sons, Inc., New Jersey, USA, 1995.
- [143] D.B. Fogel, L.J. Fogel, and J.W. Atmar. Meta-evolutionary programming. In *Conference Record of the Twenty-Fifth Asilomar Conference on Signals, Systems and Computers*, pages 540–545, 4–6 Nov. 1991.
- [144] L. J. Fogel, A. J. Owens, and M. J. Walsh. *Artificial Intelligence Through Simulated Evolution*. Wiley Publishing, 1966.
- [145] N. Forbes. Biologically inspired computing. *Computing in Science and Engineering*, 2(6):83–87–, 2000.
- [146] N. Forbes. *Imitation of life: how biology is inspiring computing*. Mit Press, Cambridge, Massachusetts, 2004.
- [147] S. Forrest and S. A. Hofmeyr. Immunology as information processing. In *Design Principles for the Immune System and Other Distributed Autonomous Systems*, pages 361–388–. Oxford University Press, New York, 2001.
- [148] S. Forrest, S. A. Hofmeyr, and A. Somayaji. Computer immunology. *Communications of the ACM*, 40(10):88–96–, October 1997.
- [149] S. Forrest, S. A. Hofmeyr, A. Somayaji, and T. A. Longstaff. A sense of self for unix processes. In *Proceedings of the 1996 IEEE Symposium on Security and Privacy*, pages 120–128–. IEEE Computer Society, 1996.
- [150] S. Forrest, B. Javornik, R. E. Smith, and A. S. Perelson. Using genetic algorithms to explore pattern recognition in the immune system. *Evolutionary Computation*, 1(3):191–211–, 1993.

- [151] S. Forrest, A. S. Perelson, L. Allen, and R. Cherukuri. Self-nonsel self discrimination in a computer. In *Proceedings of the 1992 IEEE Symposium on Security and Privacy*, pages 202–212–. IEEE Computer Society Press, Washington, DC, USA, 1994.
- [152] T. Francis Jr. On the doctrine of original antigenic sin. *Proceedings of the American Philosophical Society*, 104(6):572–578–, December 1960.
- [153] O. Francois and C. Lavergne. Design of evolutionary algorithms - a statistical perspective. *IEEE Transactions on Evolutionary Computation*, 5(2):129–148–, April 2001.
- [154] A. A. Freitas and B. Rocha. Population biology of lymphocytes: the flight for survival. *Annu Rev Immunol*, 18:83–111, 2000.
- [155] A. A. Freitas and J. Timmis. Revisiting the foundations of artificial immune systems: A problem-oriented perspective. In Peter Bentley and Emma Hart, editors, *Lecture Notes in Computer Science*, pages 229–241–. Springer, Berlin, Heidelberg, 2003.
- [156] Y. X. Fu and D. D. Chaplin. Development and maturation of secondary lymphoid tissues. *Annual review of immunology*, 17:399–433–, 1999.
- [157] T. Fukuda, K. Mori, and M. Tsukiyama. Immune networks using genetic algorithm for adaptive production scheduling. *15th IFAC World Congress*, 3:57–60–, 1993.
- [158] A. S. Fukunaga. Restart scheduling for genetic algorithms. In *Parallel Problem Solving from Nature PPSN V*, 1998.
- [159] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, 1990.
- [160] N. Ganguly, G. Canright, and A. Deutsch. Design of an efficient search algorithm for p2p networks using concepts from natural immune systems. In *Proceedings of the 8th International Conference on Parallel Problem Solving from Nature (PPSN VIII)*, pages –, 2004.
- [161] V. V. Ganusov, C. T. Bergstrom, and R. Antia. Within-host population dynamics and the evolution of microparasites in a heterogeneous host population. *Evolution*, 56(2):213–223, 2002.
- [162] M. Garey, D. Johnson, and H. Witsenhausen. The complexity of the generalized lloyd - max problem (corresp.). *IEEE Transactions on Information Theory*, 28(2):255–256, Mar 1982.
- [163] S. M. Garrett. Parameter-free, adaptive clonal selection. In *Congress on Evolutionary Computing (CEC 2004)*, volume 1, pages 1052–1058–. IEEE Press, USA, 2004.
- [164] S. M. Garrett. How do we evaluate artificial immune systems? *Evolutionary Computation*, 13(2):145–177–, 2005.
- [165] E. Gaudin, M. Rosado, F. Agenes, A. Mclean, and A. A. Freitas. B-cell homeostasis, competition, resources, and positive selection by self-antigens. *Immunological Reviews*, 197(1):102–, February 2004.
- [166] M. Gell-Mann. Complex adaptive systems. In David Pines and David Meltzer, editors, *Complexity: metaphors, models, and reality*, pages 17–45–. Addison-Wesley, USA, 1994.

- [167] A. J. T. George and D. Gray. Receptor editing during affinity maturation. *Immunology Today*, 20(4):196–, April 1999.
- [168] A. J. T. George and D. Gray. Jumping or walking: which is better? *Immunology Today*, 21(1):55–, 2000.
- [169] P. J. Gerrish and R. E. Lenski. The fate of competing beneficial mutations in an asexual population. *Genetica*, 102-103(1-6):127–144, 1998.
- [170] A. Gersho, A. Gersho, and Y. Shoham. Hierarchical vector quantization of speech with dynamic codebook allocation. In Y. Shoham, editor, *Proc. IEEE International Conference on ICASSP '84. Acoustics, Speech, and Signal Processing*, volume 9, pages 416–419, 1984.
- [171] J. Ghosh. Multiclassifier systems: Back to the future. In *Proceedings of the Third International Workshop on Multiple Classifier Systems*, 2002.
- [172] W. P. Glezen and M. Alpers. Maternal immunization. *Clinical infectious diseases, International Conference on Acute Respiratory Infections*, 28(2):188–239–, 1999.
- [173] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison Wesley Publishing Company, Inc., USA, Canada, 1989.
- [174] D. E. Goldberg. From genetic and evolutionary optimization to the design of conceptual machines. *Evolutionary Optimization*, 1(1):1–12–, 1999.
- [175] D. E. Goldberg. The design of innovating machines: A fundamental discipline for a postmodern systems engineering. In *Engineering Systems Symposium*, pages –. MIT Engineering Systems Division, USA, 2004.
- [176] D. E. Goldberg and J. R. Genetic algorithms with sharing for multimodal function optimization. In *Proceedings of the Second International Conference on Genetic Algorithms on Genetic algorithms and their application*, pages 41–49–. Lawrence Erlbaum Associates, Inc., Mahwah, NJ, USA, 1987.
- [177] D. Edward Goldberg. *The Design of Innovation: Lessons from and for Competent Genetic Algorithms*. Kluwer Academic Publishers, USA, 2002.
- [178] D. Goodman, L. Boggess, and A. Watkins. Artificial immune system classification of multiple-class problems. In A. Buczak, J. Ghosh, M. Embrechts, O. Ersoy, and S Kercel, editors, *Fuzzy Logic, Evolutionary Programming, Complex Systems and Artificial Life*, volume 12, pages 179–184–, New York, 2002. ASME Press.
- [179] D. Goodman, L. Boggess, and A. Watkins. An investigation into the source of power for airs, an artificial immune classification system. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN'03)*, pages 1678–1683–, Portland, Oregon, USA, 2003.
- [180] D. E. Goodman and L. C. Boggess. The role of hypothesis filtering in (airs), an artificial immune classifier. In Anna L. Buczak, Joydeep Ghosh, Mark J. Embrechts, and Okan Ersoy, editors, *Proceedings of the Artificial Neural Networks in Engineering Conference, 2003.*, volume 13, pages 243–248–, New York, 2004. ASME Press.
- [181] C. C. Goodnow. Chance encounters and organized rendezvous. *Immunological Reviews*, 156:5–10–, 1997.

- [182] H. J. Gould, P. Takhar, H. E. Harries, S. R. Durham, and C. J. Corrigan. Germinal-centre reactions in allergic inflammation. *Trends in Immunology*, 27(10):446–452–, October 2006.
- [183] J. L. Gowans. The recirculation of lymphocytes from blood to lymph in the rat. *American journal of physiologic imaging*, 146(1):54–69–, 1959.
- [184] J. L. Gowans and E. J. Knight. The route of re-circulation of lymphocytes in the rat. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 159(975):257–282–, 1964.
- [185] J. Greensmith. *New Frontiers For An Artificial Immune System*. PhD thesis, University of Leeds, UK, 2003.
- [186] J. Greensmith and S. Cayzer. An artificial immune system approach to semantic document classification. In Peter Bentley and Emma Hart, editors, *Lecture Notes in Computer Science*, pages 136–146–. Springer, Berlin / Heidelberg, 2003.
- [187] A. Grilo, A. Caetano, and A. Rosa. Immune system simulation through a complex adaptive system model. In M. Koppen, S. Ovaska, T. Furuhashi, and F. Homann, editors, *Springer Engineering Series, Soft Computing in Industrial Applications - Recent Advances*, pages 681–704–. Springer, 2002.
- [188] S. Grossberg. Neural expectation: cerebellar and retinal analogs of cells fired by learnable or unlearned pattern classes. *Biological Cybernetics*, 10(1):49–57–, January 1972.
- [189] S. Grossberg. Adaptive pattern classification and universal recoding: I. parallel development and coding of neural feature detectors. *Biological Cybernetics*, 23(3):121–134–, September 1976.
- [190] S. Grossberg. Adaptive pattern classification and universal recoding: Ii. feedback, expectation, olfaction, illusions. *Biological Cybernetics*, 23(4):187–202–, December 1976.
- [191] S. Grossberg. Competitive learning: from interactive activation to adaptive resonance. *Cognitive Science*, 11:23–63–, 1988.
- [192] J. Gu. Multispace search: A new optimization approach. In *Algorithms and Computation*, 1994.
- [193] J. Gu. Multispace search for satisfiability and np-hard problems. In *Satisfiability Problem: Theory and Applications : DIMACS Workshop*, 1997.
- [194] I Guyon and A Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, 2003.
- [195] D. L. Hall and S. A. H. McMullen. *Mathematical Techniques in Multisensor Data Fusion*. Artech House, 2004.
- [196] L. A. Hanson, M. Korotkova, and E. Telemo. Breast-feeding, infant formulas, and the immune system. *Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology*, 90(6 Suppl 3):59–63–, 2003.
- [197] P. K. Harmer, P. D. Williams, G. H. Gunsch, and G. B. Lamont. An artificial immune system architecture for computer security applications. *IEEE Transactions on Evolutionary Computation*, 6(3):252–280–, June 2002.

- [198] W Harriman, H Volk, N Defranoux, and M Wabl. Immunoglobulin class switch recombination. *Annual Review of Immunology*, 11:361–384–, April 1993.
- [199] E. Hart and P. Ross. Clustering moving data with a modified immune algorithm. In J. Gottlieb, P.L. Lanzi, R.E. Smith, S. Cagnoni, E. Hart, G.R. Raidl, and H. Ti-jink, editors, *Lecture Notes in Computer Science*, pages 394–. Springer, Berlin / Heidelberg, 2001.
- [200] E. Hart and P. Ross. Exploiting the analogy between the immune system and sparse distributed memories. *Genetic Programming and Evolvable Machines*, 4(4):333–358–, December 2003.
- [201] E. Hart and J. Timmis. Application areas of ais: The past, the present and the future. In Marcin L. Pilat, Peter J. Bentley, and Jonathan Timmis, editors, *Lecture Notes in Computer Science*, pages 483–. Springer, Berlin / Heidelberg, Germany, 2005.
- [202] R. Hecht-Nielsen. Counterpropagation networks. In *Proceedings of the IEEE International Conference on Neural Networks*, pages –. IEEE Press, USA, 1987.
- [203] R. Hecht-Nielsen. Counterpropagation networks. *Applied Optics*, 26(23):4979–4984–, 1987.
- [204] S. M. Hedrick. The acquired immune system: a vantage from beneath. *Immunity*, 21(5):607–615–, 2004.
- [205] U. Hershberg, S. Solomon, and Irun R. Cohen. What is the basis of the immune systems specificity? In *The MIRIAM Project Series*, pages 377–384–. ESCULAPIO Pub. Co., Bologna, Italy, 2003.
- [206] J. Hesser and R. Mnner. Towards an optimal mutation probability for genetic algorithms. In *Lecture Notes In Computer Science*, pages 23–32–. Springer-Verlag, London, UK, 1991.
- [207] R. Hightower, S. Forrest, and A. S. Perelson. The baldwin effect in the immune system: Learning by somatic hypermutation. In *Adaptive Individuals in Evolving Populations: Models and Algorithms*, pages 159–167–. Addison Wesley, Reading, MA. USA, 1995.
- [208] R. R. Hightower, S. Forrest, and A. S. Perelson. The evolution of emergent organization in immune system gene libraries. In *Proceedings of the 6th International Conference on Genetic Algorithms*, pages 344–350–, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc.
- [209] T. K. Ho. The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:832–844, 1998.
- [210] G. W. Hoffmann. A neural network model based on the analogy with the immune system. *Journal of Theoretical Immunology*, 122(1):33–67–, September 1986.
- [211] S. Hofmeyr and S. Forrest. Immunity by design: An artificial immune system. In Jason Daida, Agoston E. Eiben, Max H. Garzon, Vasant Honavar, Mark Jakiela, and Robert E. Smith, editors, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*, volume 2, pages 1289–1296–. Morgan-Kaufmann, Orlando, Florida, USA, 1999.

- [212] S. A. Hofmeyr. *An Immunological Model of Distributed Detection and its Application to Computer Security*. PhD thesis, Department of Computer Sciences, University of New Mexico, USA, 1999.
- [213] S. A. Hofmeyr. An interpretative introduction to the immune system. In I. R. Cohen L. A. Segel, editor, *Design Principles for Immune Systems and Othe Distributed Autonomous Systes*, volume Santa Fe Institute, Studies in the Sciences of Complexity, pages 3–26. Oxford University Press, 2001.
- [214] S. A. Hofmeyr and S. Forrest. Architecture for an artificial immune system. *Evolutionary Computation*, 8(4):443–473–, December 2000.
- [215] S. A. Hofmeyr, S. Forrest, and A. Somayaji. Intrusion detection using sequences of system calls. *Journal of Computer Security*, 6(3):151–180–, 1998.
- [216] J. H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, volume 2nd Edition. MIT Press, USA, 1992.
- [217] J. H. Holland. Complex adaptive systems. *Daedalus*, 121(1):17–30–, 1992.
- [218] J. H. Holland. *Hidden Order: How Adaptation Builds Complexity*. Addison Wesley Publishing Company, USA, 1995.
- [219] A. Hone and J. Kelsey. Optima, extrema, and artificial immune systems. In Vincenzo Cutello, Peter J. Bentley, and Jon Timmis, editors, *Lecture Notes in Computer Science*, pages 80–90–. Springer, Berlin / Heidelberg, 2004.
- [220] T. Honkela. *Self-Organizing Maps in Natural Language Processing*. PhD thesis, Neural Networks Research Centre, Helsinki University of Technology, Finland, 1997.
- [221] J. N. Hooker. Testing heuristics: We have it all wrong. *Journal of Heuristics*, 1(1):33–42–, September 1995.
- [222] R. Horst, P. M. Pardalos, and N. V. Thoai. *Introduction to Global Optimization*. Kluwer Academic Publishers, 2nd edition, 2000.
- [223] N. Hozumi and S. Tonegawa. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *National Academy of Sciences*, 73(10):3628–3632–, 1976.
- [224] T. Hrycej. *Modular Learning in Neural Networks A Modularized Approach to Neural*. Wiley, 1992.
- [225] X. Hu, R. Shonkwiler, and M. Spruill. Random restarts in global optimization. Technical Report Tech Report 110592-015, School of Mathematics, Georgia Institute of Technology, January 1994.
- [226] N. Intrator and S. Edelman. Competitive learning in biological and artificial neural computation. *Trends in Cognitive Sciences*, 1(7):268–272–, October 1997.
- [227] Y. Ishida. Fully distributed diagnosis by pdp learning algorithm: towards immune network pdp model. In *IJCNN International Joint Conference on Neural Networks*, volume 1, pages 777–782–. IEEE Computer Society, USA, 1990.
- [228] Y. Ishida. Distributed and autonomous sensing based on the immune network. *Artificial Life and Robotics: Computer Science and Engineering*, 2(1):1–7–, March 1998.

- [229] Y. Ishida. *Immunity-Based Systems - A Design Perspective*. Springer-Verlag, Berlin, Germany, 2004.
- [230] A. Ishiguro, T. Kondo, Y. Watanabe, and Y. Uchikawa. Dynamic behavior arbitration of autonomous mobile robots using immune networks. In *IEEE International Conference on Evolutionary Computation*, volume 2, pages 722–727–. IEEE Press, USA, 1995.
- [231] A. Ishiguro, Y. Watanabe, T. Kondo, Y. Shirai, and Y. Uchikawa. A robot with a decentralized consensus-making mechanism based on the immune system. In *Proceedings. ISADS 97., Third International Symposium on Autonomous Decentralized Systems*, pages 231–237–. IEEE Press, USA, 1997.
- [232] N. K. Jerne. The natural-selection theory of antibody formation. *Proceedings of the National Academy of Sciences of the United States of America*, 41(11):849–857–, 1955.
- [233] N. K. Jerne. Clonal selection in a lymphocyte network. In *Cellular Selection and Regulation in the Immune Response*, volume 29 of *Society of General Physiologists Series*, pages 39–48. Raven Press, 1974.
- [234] N. K. Jerne. Towards a network theory of the immune system. *Annales d’immunologie (Annals of Immunology), Institut Pasteur (Paris, France), Societe Francaise d’Immunologie*, 125(C):373–389–, January 1974.
- [235] N. K. Jerne. Idiotypic networks and other preconceived ideas. *Immunological Reviews*, 79:5–24–, 1984.
- [236] D. S. Johnson. A theoreticians guide for experimental analysis of algorithms. In D. S. Johnson and C. C. McGeoch, editors, *Proceedings of the 5th and 6th DIMACS Implementation Challenges*, pages 215–250–. American Mathematical Society, 2002.
- [237] J. Jost. External and internal complexity of complex adaptive systems. *Theory in biosciences*, 123(1):69–88–, 2004.
- [238] J. H. Jun, D. W. Lee, and K. B. Sim. Realization of cooperative strategies and swarm behavior in distributed autonomous robotic systems using artificial immune system. In *Systems, Man, and Cybernetics, 1999. IEEE SMC '99 Conference Proceedings. 1999 IEEE International Conference on*, volume 6, pages 614–619vol.6, 12-15 Oct. 1999.
- [239] H. Kargupta and P. Chan. *Advances in Distributed and Parallel Knowledge Discovery*. AAAI Press/MIT Press, 2000.
- [240] S. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, USA, 1993.
- [241] S. Kauffman and S. Levin. Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology*, 128(1):11–45–, 1987.
- [242] S. A. Kauffman and E. D. Weinberger. The nk model of rugged fitness landscapes and its application to maturation of the immune response. *Journal of theoretical biology*, 141(2):211–245–, 1989.
- [243] S. A. Kauffman, E. D. Weinberger, and A. S. Perelson. Maturation of the immune response via adaptive walks on affinity landscapes. In *Theoretical Immunology I: Santa Fe Institute Studies in the Sciences of Complexity*, pages –. Addison Wesley, USA, 1988.

- [244] M. Kearns. Thoughts on hypothesis boosting. Unpublished manuscript, 1988.
- [245] M. A. Keller and E. R. Stiehm. Passive immunity in prevention and treatment of infectious diseases. *Clinical Microbiology Reviews*, 13(4):602–614–, 2000.
- [246] J. Kelsey and J. Timmis. Immune inspired somatic contiguous hypermutation for function optimisation. In *Lecture Notes in Computer Science*, pages 207–218–. Springer, Berlin / Heidelberg, 2003.
- [247] J. Kelsey, J. Timmis, and A. Hone. Chasing chaos. In *The 2003 Congress on Evolutionary Computation, (CEC '03)*, volume 1, pages 413–419–. IEEE Computer Society, USA, 2003.
- [248] J. O. Kephart. A biologically inspired immune system for computers. In Pattie Maes, editor, *Artificial Life IV*, pages 130–139–. MIT Press, Cambridge, Massachusetts, USA, 1994.
- [249] J. O. Kephart, G. B. Sorkin, W. C. Arnold, D. M. Chess, G. J. Tesauero, and S. R. White. Biologically inspired defences against computer viruses. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCA '95)*, pages 985–996–. Morgan Kaufmann Publishers, Inc., 1995.
- [250] J. O. Kephart, G. B. Sorkin, and M. Swimmer. An immune system for cyberspace. In *IEEE International Conference on Systems, Man, and Cybernetics, 'Computational Cybernetics and Simulation'*, volume 1, pages 879–884–. IEEE Press, USA, 1997.
- [251] J. Kim and P. Bentley. Immune memory and gene library evolution in the dynamic clonal selection algorithm. *Genetic Programming and Evolvable Machines*, 5(4):361–391–, December 2004.
- [252] J. Kim and P. J. Bentley. Negative selection and niching by an artificial immune system for network intrusion detection. In *A late-breaking paper, Genetic and Evolutionary Computation Conference (GECCO '99)*, pages –, 1999.
- [253] J. Kim and P. J. Bentley. Towards an artificial immune system for network intrusion detection: an investigation of clonal selection with a negative selection operator. In *Proceedings of the 2001 Congress on Evolutionary Computation*, volume 2, pages 1244–1252–, USA, 2001. IEEE Computer Society.
- [254] J. Kim and P. J. Bentley. Towards an artificial immune system for network intrusion detection: An investigation of dynamic clonal selection. In *Congress on Evolutionary Computation (CEC-2002)*, pages 1015–1020–, 2002.
- [255] J. W. Kim. *Integrating Artificial Immune Algorithms for Intrusion Detection*. PhD thesis, Department of Computer Science, University College London, UK, 2002.
- [256] J. Klein and N. Nikolaidis. The descent of the antibody-based immune system by gradual evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 102(1):169–174–, January 2005.
- [257] V. B. Klimovich. Actual problems of evolutionary immunology. *Journal of Evolutionary Biochemistry and Physiology*, 38(5):562–574–, September 2002.
- [258] T. Knight and J. Timmis. Aine: An immunological approach to data mining. In T. Lin and Xindon Wu, editors, *First IEEE International Conference on Data Mining (ICDM'01)*, pages 297–304–. IEEE Computer Society, USA, 2001.

- [259] T. P. Knight. *MARIA: A Multilayered Unsupervised Machine Learning Algorithm Based on the Vertebrate Immune System*. PhD thesis, The University of Kent at Canterbury, UK, 2005.
- [260] G. Kohler and C. Milstein. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, 256:495–497–, 1975.
- [261] T. Kohonen. *Self-Organizing Maps*, volume Third. Springer-Verlag, Berlin Heidelberg, 2001.
- [262] M. Kudo and J. Sklansky. Comparison of algorithms that select features for pattern classifiers. *Pattern Recognition*, 33:25–41, 2000.
- [263] V. Kumar. Algorithms for constraint-satisfaction problems : a survey. *The AI magazine (AI mag.)*, 13, 1992.
- [264] L. Kuncheva and J. Bezdek. An integrated framework for generalized nearest prototype classifier design. *Int. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6:437–457, 1998.
- [265] L.I. Kuncheva and J.C. Bezdek. Nearest prototype classification: clustering, genetic algorithms, or random search? *IEEE Transactions on Systems, Man and Cybernetics, Part C*, 28(1):160–164, 1998.
- [266] H. Y. K. Lau, V. W. K. Wong, and I. S. K. Lee. Immunity-based autonomous guided vehicles control. *Applied Soft Computing*, pages –, 2005.
- [267] H. Y.K. Lau and V. W.K. Wong. A strategic behavioral-based intelligent transport system with artificial immune system. *IEEE International Conference on Systems, Man and Cybernetics*, 4:3909–3914–, 2004.
- [268] H. Y.K. Lau and V. W.K. Wong. An immunity-based distributed multiagent-control framework. *IEEE Transactions on Systems, Man and Cybernetics, Part A*, 36(1):91–108–, January 2006.
- [269] J. Lederberg. Ontogeny of the clonal selection theory of antibody formation. reflections on darwin and ehrlich. *Annals of the New York Academy of Sciences*, 546:175–182–, 1988.
- [270] D. W. Lee, H. B. Jun, and K. B. Sim. Artificial immune system for realization of cooperative strategies and group behavior in collective autonomous mobile robots. In *Proceedings of the Forth International Symposium on Artificial Life and Robotics, AAAI*, pages 232–235–, 1999.
- [271] V.R. Lesser. An Overview of DAI: Viewing Distributed AI as Distributed Search. *Journal of Japanese Society for Artificial Intelligence-Special Issue on Distributed Artificial Intelligence*, 5(4):392–400, January 1990.
- [272] R. M. Lewis, V. T., and M. W. Trosset. Direct search methods: then and now. *Journal of Computational and Applied Mathematics*, 124:191–207, 2000.
- [273] Z. Li, C. J. Woo, M. D. Iglesias-Usse, D. Ronai, and M. D. Scharff. The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes and Development*, 18:1–11–, 2004.
- [274] P. Liu, P. Liu, and D. W. Wang. Reduction optimization in heterogeneous cluster environments. In Da-Wei Wang, editor, *Proc. 14th International Parallel and Distributed Processing Symposium IPDPS 2000*, pages 477–482, 2000.

- [275] Y. J. Liu and C. Arpin. Germinal center development. *Immunological Reviews*, 156(1):111126–, 1997.
- [276] G. F. Luger and W. A. Stubblefield. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. Addison Wesley Longman Inc, USA, 1997.
- [277] I. C. M. MacLennan. Germinal centers. *Annual review of immunology*, 12:117–139–, 1994.
- [278] M. Magdon-ismail and A. F. Atiya. The early restart algorithm. *Neural Computation*, 12:1303–1312, 2000.
- [279] S. Mahfoud. *Evolutionary Computation 2: Advanced Algorithms and Operations*, chapter Niching Methods, pages 87–92. IoP, 2000.
- [280] E. N. Marieb. *Essentials of human anatomy & physiology*. Pearson/Benjamin Cummings, San Francisco, CA, USA, 2006.
- [281] P Marrow. Nature-inspired computing technology and applications. *BT Technology Journal*, 18(4):13–23–, October 2000.
- [282] T. Martinetz. Competitive hebbian learning rule forms perfectly topology preserving maps. In *Proceedings of the International Conference on Artificial Neural Networks (ICANN93)*, pages 427–434–. Springer, Amsterdam, 1993.
- [283] T. Martinetz and K. Schulten. A "neural gas" network learns topologies. In Kai Mkisara, Olli Simula, and Jari Kangas, editors, *Artificial Neural Networks*, pages 397–402–. Elsevier, North-Holland, Amsterdam, 1991.
- [284] T. Martinetz and K. Schulten. Topology representing networks. *Neural Networks*, 7(3):507–522–, 1994.
- [285] T. M. Martinetz, S. G. Berkovich, and K. J. Schulten. 'neural-gas' network for vector quantization and its application totime-series prediction. *IEEE Transactions on Neural Networks*, 4(4):558–569–, July 1993.
- [286] G. Marwah and L. Boggess. Artificial immune systems for classification: Some issues. In *First International Conference on Artificial Immune Systems*, pages 149–153–, 2002.
- [287] P. Matzinger. Tolerance, danger, and the extended family. *Annual Review of Immunology*, 12:991–1045–, 1994.
- [288] P. Matzinger. The danger model: A renewed sense of self. *Science*, 296(5566):301–305–, 2002.
- [289] T. W. McDade and C. M. Worthman. Evolutionary process and the ecology of human immune function. *American Journal of Human Biology*, 11(6):705–717–, 1999.
- [290] A. R. McLean, M. M. Rosado, F. Agenes, R. Vasconcellos, and A. A. Freitas. Resource competition as a mechanism for b cell homeostasis. *Proc. Natl. Acad. Sci. USA, Immunology*, 94:5792–5797, May 1997.
- [291] H. Meshref and H. VanLandingham. Artificial immune systems: application to autonomous agents. In *IEEE International Conference on Systems, Man, and Cybernetics*, volume 1, pages 61–66–. IEEE Press, USA, 2000.

- [292] M. E. Meyer-Hermann and P. K. Maini. Cutting edge: Back to "one-way" germinal centers. *The Journal of Immunology*, 174:2489–2493–, 2005.
- [293] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell. *Machine Learning: An Artificial Intelligence Approach*. Morgan Kaufmann, 1983.
- [294] R. Michelan and F. J. Von Zuben. Decentralized control system for autonomous navigation based on an evolved artificial immune network. In *Proceedings of the 2002 Congress on Evolutionary Computation, (CEC '02)*, volume 2, pages 1021–1026–. IEEE Press, USA, 2002.
- [295] F. Mizessyn and Yoshiteru Ishida. Immune networks for cement plants. In *Proceedings International Symposium on Autonomous Decentralized Systems, ISADS 93*, pages 282–288–. IEEE Computer Society, USA, 1993.
- [296] P. H. Mohr, N. Ryan, and J. Timmis. Exploiting immunological properties for ubiquitous computing systems. In Vincenzo Cutello, Peter J. Bentley, and Jon Timmis, editors, *Lecture Notes in Computer Science*, pages 277–289–. Springer, Berlin / Heidelberg, Germany, 2004.
- [297] K. Mori, M. Tsukiyama, and T. Fukuda. Immune algorithm with searching diversity and its application to resource allocation problem. *The Transactions of the Institute of Electrical Engineers of Japan*, 113-C(10):872–878–, 1993.
- [298] H. Mühlenbein. How genetic algorithms really work: I. mutation and hillclimbing. In *Parallel Problem Solving from Nature 2, PPSN-II*, pages 15–26–. Elsevier, 1992.
- [299] C. Murgia, J. K Pritchard, S. Y. Kim, A. Fassati, and R. A Weiss. Clonal origin and evolution of a transmissible cancer. *Cell*, 126(3):477–487, Aug 2006.
- [300] M. Muselli. A theoretical approach to restart in global optimization. *Journal of Global Optimization*, 10:1–16, 1997.
- [301] M. V. Nagendra Prasad, V. R. Lesser, and S. E. Lander. Retrieval and reasoning in distributed case bases. *Journal of Visual Communication and Image Representation, Special Issue on Digital Libraries*, 7(1):74–87, 1996.
- [302] A. Nakada, A. Nakada, T. Shibata, M. Konda, T. Morimoto, and T. Ohmi. A fully parallel vector-quantization processor for real-time motion-picture compression. *IEEE Journal of Solid-State Circuits*, 34(6):822–830, 1999.
- [303] R. M. Nesse. Maladaptation and natural selection. *The Quarterly Review of Biology*, 80:6270–, 2005.
- [304] J. Newborough and S. Stepney. A generic framework for population-based algorithms, implemented on multiple fpgas. In Marcin L. Pilat, Peter J. Bentley, and Jonathan Timmis, editors, *Lecture Notes in Computer Science*, pages 43–55–. Springer, Berlin, Heidelberg, 2005.
- [305] G. J. V. Nossal. Negative selection of lymphocytes. *Cell*, 76(2):229–239–, 1994.
- [306] G. J. V. Nossal and J. Lederberg. Antibody production by single cells. *Nature*, 181:1419–1420–, 1958.
- [307] M. Nussenzweig. Immune receptor editing: revise and select. *Cell*, 95(7):875–878–, 1998.

- [308] W. H. Oddy. Breastfeeding protects against illness and infection in infants and children: a review of the evidence. *Breastfeeding review : professional publication of the Nursing Mothers' Association of Australia*, 9(2):11–18–, 2001.
- [309] W. L. Olszewski. The lymphatic system in body homeostasis: Physiological conditions. *Lymphatic Research and Biology*, 1(1):11–24–, January 2003.
- [310] D. Opitz and R. Maclin. Popular ensemble methods: An empirical study. *Journal of Artificial Intelligence Research*, 11:169–198, 1999.
- [311] C. H. Papadimitriou and K. Steiglitz. *Combinatorial Optimization: Algorithms and Complexity*. Courier Dover Publications, 1998.
- [312] K.K. Parhi, K.K. Parhi, F.H. Wu, and K. Genesan. Sequential and parallel neural network vector quantizers. *IEEE Transactions on Computers*, 43(1):104–109, 1994.
- [313] R. Paton. Introduction to computing with biological metaphors. In *Computing With Biological Metaphors*, pages 1–8–. Chapman & Hall, London, UK, 1994.
- [314] William E. Paul. *Immunology - Recognition and Response : Readings from Scientific American magazine / edited by William E. Paul*. Scientific American Inc., W. H. Freeman and Company, USA, 1991.
- [315] L. Pauling. A theory of the structure and process of formation of antibodies. *Journal of the American Chemical Society*, 62:2643–2657–, 1940.
- [316] G. Paun. Bio-inspired computing paradigms (natural computing). *Unconventional Programming Paradigms*, 3566:155–160–, 2005.
- [317] W. Pedrycz. *Computational Intelligence: An Introduction*. CRC Press, New York, 1997.
- [318] E. S. Peer, A. P. Engelbrecht, and F. van den Bergh. Cirq@up optibench: a statistically sound framework for benchmarking optimisation algorithms. In *The 2003 Congress on Evolutionary Computation, (CEC '03)*, volume 4, pages 2386–2392–. IEEE Press, USA, 2003.
- [319] M. Pelikan and D. E. Goldberg. Hierarchical problem solving by the bayesian optimization algorithm. IlliGAL Report No. 2000002, Illinois Genetic Algorithms Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL, 2000.
- [320] A. S. Perelson. Immune network theory. *Immunological Reviews*, 110:5–36–, August 1989.
- [321] A. S. Perelson and G. F. Oster. Theoretical studies of clonal selection: Minimal antibody repertoire size and reliability of self-non-self discrimination. *Journal of Theoretical Biology*, 81(4):645–670–, December 1979.
- [322] A. S. Perelson and G. Weisbuch. Immunology for physicists. *Reviews of Modern Physics*, 69(4):1219–1268–, 1997.
- [323] C. C. Pettey. Diffusion (cellular) models. In David B Fogel and Zbigniew Michalwicz, editors, *Evolutionary Computation 2 - Advanced Algorithms and Operations*, pages –. Institute of Physics (IOP) Publishing Ltd., UK, 2000.
- [324] L. J. Picker and E. C. Butcher. Physiological and molecular mechanisms of lymphocyte homing. *Annual Review of Immunology*, 10:561–591–, 1992.

- [325] D. M. Pierre, D. Goldman, Y. Bar-Yam, and A. S. Perelson. Somatic evolution in the immune system: the need for germinal centers for efficient affinity maturation. *Journal of Theoretical Biology*, 186(2):159–171–, 1997.
- [326] E. Plaza and J. Llus A. F. Martn. Cooperative case-based reasoning. In *Distributed Artificial Intelligence Meets Machine Learning Learning in Multi-Agent Environments*, 1997.
- [327] R. Polikar. Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine*, 6(3):21–45, 2006.
- [328] M. J. D. Powell. Radial basis functions for multivariable interpolation: a review. In *Algorithms for approximation*, pages 143–167–. Clarendon Press, New York, NY, USA, 1987.
- [329] W. L. Price. A controlled random search procedure for global optimisation. *The Computer Journal*, 20:367–370, 1977.
- [330] Computational Science Education Project. Mathematical optimization. Online, 1995.
- [331] R. C. Purshouse and P. J. Fleming. Conflict, harmony, and independence: Relationships in evolutionary multi-criterion optimisation. In *Proceedings of the Second International Conference on Evolutionary Multi-Criterion Optimization (EMO)*, pages 16–30, 2003.
- [332] K. Rajewsky. Immunology: Burnet’s unhappy hybrid. *Nature*, 394:624–625–, August 1998.
- [333] I. Rechenberg. *Evolutionstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution*. Frommann-Holzboog, 1973.
- [334] R. D. Reed and R. J. Marks II. *Neural Smithing - Supervised Learning in Feedforward Artificial Neural Networks*. MIT Press, USA, 1999.
- [335] R. R. Regoes, M. A. Nowak, and S. Bonhoeffer. Evolution of virulence in a heterogeneous host population. *Evolution Int J Org Evolution*, 54(1):64–71, Feb 2000.
- [336] L. Riechmann, M. Clark, H. Waldmann, and G. Winter. Reshaping human antibodies for therapy. *Nature*, 332:323–327–, 1988.
- [337] B. Rinkevich. Primitive immune systems: Are your ways my ways? *Immunological Reviews*, 198(1):2535–, April 2004.
- [338] H. Robbins. Some aspects of the sequential design of experiments. *Bull. Amer. Math. Soc.*, 58:527–535, 1952.
- [339] H. Rosen, G. Sanna, and C. Alfonso. Egress: a receptor-regulated step in lymphocyte trafficking. *Immunological Reviews*, 195(1):160177–, 2003.
- [340] I. M Rouzine, J. Wakeley, and J. M Coffin. The solitary wave of asexual evolution. *Proc Natl Acad Sci U S A*, 100(2):587–592, Jan 2003.
- [341] D. E. Rumelhart and D. Zipser. Feature discovery by competitive learning. *Cognitive Science*, 9(1):75–112–, 1985.
- [342] S. J. Russell. and P Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, 1995.

- [343] M. Salmi and S. Jalkanen. Lymphocyte homing to the gut: attraction, adhesion, and commitment. *Immunological Reviews*, 206(1):100–113–, 2005.
- [344] D. Salmon, S. Teret, C. MacIntyre, D. Salisbury, M. Burgess, and N. Halsey. Compulsory vaccination and conscientious or philosophical exemptions: past, present, and future. *The Lancet*, 367(9508):436–442–, 2006.
- [345] S. Sarafijanovic and J. Y. L Boudec. An artificial immune system for misbehavior detection in mobile ad-hoc networks with virtual thymus, clustering, danger signal, and memory detectors. In Vincenzo Cutello, Peter J. Bentley, and Jon Timmis, editors, *Lecture Notes in Computer Science*, pages 342–356–. Springer, Berlin / Heidelberg, Germany, 2004.
- [346] S. Sarafijanovic and J. Y. L Boudec. An artificial immune system approach with secondary response for misbehavior detection in mobile ad hoc networks. *IEEE Transactions on Neural Networks*, 16(5):1076–1087–, September 2005.
- [347] J. David Schaffer, Richard A. Caruana, Larry J. Eshelman, and Rajarshi Das. A study of control parameters affecting online performance of genetic algorithms for function optimization. In *Proceedings of the third international conference on Genetic algorithms*, pages 51–60–. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1989.
- [348] J. David Schaffer and Larry J. Eshelman. On crossover as an evolutionarily viable strategy. In Lashon B. Booker, editor, *Proceedings of the 4th International Conference on Genetic Algorithms*, pages 61–68–. Morgan Kaufmann, 1991.
- [349] R. E. Schapire. The strength of weak learnability. *Machine Learning*, 5(2):197–227, 1992.
- [350] R. E. Schapire. The boosting approach to machine learning: An overview. In D. D. Denison, M. H. Hansen, C. Holmes, B. Mallick, and B. Yu, editors, *Nonlinear Estimation and Classification*, 2003.
- [351] T. Schnekenburger. Parallel randomized algorithms in heterogeneous environments. In *Int. Conference on Systems Engineering*, 1993.
- [352] H. P. Schwefel. *Numerical optimization of computer models*. John Wiley & Sons, Inc., New York, NY, USA, 1981.
- [353] A. Seeker and A. A. Freitas. Wairs: improving classification accuracy by weighting attributes in the aircs classifier. In *Proc. IEEE Congress on Evolutionary Computation CEC 2007*, pages 3759–3765, 25–28 Sept. 2007.
- [354] L. A. Segel. Some spatio-temporal models in immunology. *International Journal of Bifurcation and Chaos*, 12(11):2343–2347, 2002.
- [355] L. A. Segel and A. S. Perelson. Computations in shape space: A new approach to immune network theory. *Theoretical Immunology, Part Two, SFI Studies in the Sciences of Complexity*, pages 377–401–, 1988.
- [356] M. Shannon and R. Mehr. Reconciling repertoire shift with affinity maturation: The role of deleterious mutations. *The Journal of Immunology*, 162:3950–3956–, 1999.
- [357] A. J. C. Sharkey. *Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems*. Springer, 1999.

- [358] A. M. Silverstein. The clonal selection theory: what it really is and why modern challenges are misplaced. *Nature Immunology*, 3:793–796–, 2002.
- [359] S. P. N. Singh and S. M. Thayer. Immunology directed methods for distributed robotics: A novel, immunity-based architecture for robust control & coordination. In *Proceedings of SPIE: Mobile Robots XVI*, pages –, 2001.
- [360] S. P. N. Singh and S. M. Thayer. Kilorobot search and rescue using an immunologically inspired approach. *Distributed Autonomous Robotic Systems*, 5:–, June 2002.
- [361] S. P. N. Singh, S. M. Thayer, and W. P. Thayer. A foundation for kilorobotic exploration. In *Proceedings of the 2002 Congress on Evolutionary Computation, CEC '02.*, volume 2, pages 1033–1038–. IEEE Press, USA, 2002.
- [362] D. Skillicorn. Strategies for parallel data mining. *IEEE Concurrency*, 7(4):26–35, 1999.
- [363] H. B. Slade and S. A. Schwartz. Mucosal immunity: The immunology of breast milk. *Journal of Allergy and Clinical Immunology*, 80(3, Part 1):348–358–, September 1987.
- [364] A. Sloman. Must intelligent systems be scruffy? In G T McKee and G C Dean, editors, *Evolving Knowledge in Natural Science and Artificial Intelligence*, pages –. Pitman, London, 1990.
- [365] D. J Smith, S. Forrest, and A. S. Perelson. Immunological memory is associative. In *ICMAS '96: Multiagent systems workshop on immunity-based systems*, pages 62–70–, United States, 1996.
- [366] J. E. Smith and T. C. Fogarty. Operator and parameter adaptation in genetic algorithms. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, 1:81–87, 1997.
- [367] W. M. Spears. Crossover or mutation? In *Proceedings of Foundations of Genetic Algorithms Workshop*, volume 2, pages 221–237–. Morgan Kaufmann, San Mateo, CA, USA, 1992.
- [368] W. M. Spears. *The Role of Mutation and Recombination in Evolutionary Algorithms*. PhD thesis, George Mason University, Fairfax, Virginia, USA, 1998.
- [369] W. Srikusalanukul, F. De Bruyne, and P. McCullagh. Modelling of peripheral lymphocyte migration: System identification approach. *Immunology and Cell Biology*, 78(3):288–, June 2000.
- [370] D. J. Stekel. The simulation of density-dependent effects in the recirculation of t lymphocytes. *Scandinavian journal of immunology*, 47(5):426–430, 1998.
- [371] D. J. Stekel, C. E. Parker, and M. A. Nowak. A model of lymphocyte recirculation. *Immunology Today*, 18(5):216–221, May 1997.
- [372] S. Stepney, R. E. Smith, J. Timmis, and A. M. Tyrrell. Towards a conceptual framework for artificial immune systems. In Vincenzo Cutello, Peter J. Bentley, and Jon Timmis, editors, *Lecture Notes in Computer Science (LNCS)*, pages 53–64–. Springer-Verlag, Germany, 2004.
- [373] S. Stepney, R. E. Smith, J. Timmis, A. M. Tyrrell, M. J. Neal, and Andrew N. W. Hone. Conceptual frameworks for artificial immune systems. *International Journal of Unconventional Computing*, 1(3):315–338–, July 2005.

- [374] T. Stibor. *On the Appropriateness of Negative Selection for Anomaly Detection and Network Intrusion Detection*. PhD thesis, Darmstadt University of Technology, Germany, 2006.
- [375] T. Stibor, J. Timmis, and C. Eckert. On the appropriateness of negative selection defined over hamming shape-space as a network intrusion detection system. In *The 2005 IEEE Congress on Evolutionary Computation, (CEC 2005)*, volume 2, pages 995–1002–. IEEE Press, USA, 2005.
- [376] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, Cambridge, Massachusetts; London, England, 1998.
- [377] M. A. Swartz. The physiology of the lymphatic system. *Advanced drug delivery reviews*, 50(1-2):3–20–, 2001.
- [378] H. Takagi. Interactive evolutionary computation: fusion of the capabilities of ec optimization and human evaluation. *Proceedings of the IEEE*, 9(89):1275–1296–, September 2001.
- [379] E. Talbi. A taxonomy of hybrid metaheuristics. *Journal of Heuristics*, 8:541–564–, 2001.
- [380] D. W. Talmage. Allergy and immunology. *Annual Review of Medicine*, 8:239–256–, 1957.
- [381] D. W. Talmage. Immunological specificity: Unique combinations of selected natural globulins provide an alternative to the classical concept. *Science*, 129:1643–1648–, 1959.
- [382] R. Tanese. Distributed genetic algorithms. In *Proceedings of the third international conference on Genetic algorithms*, 1989.
- [383] J. C. Tay and A. Jhavar. Cafiss: a complex adaptive framework for immune system simulation. In *Proceedings of the 2005 ACM symposium on Applied computing*, pages 158–164–. ACM Press, New York, NY, USA, 2005.
- [384] S. M. Thayer and S. P. N. Singh. Development of an immunology-based multi-robot coordination algorithm for exploration and mapping domains. In *IEEE/RSJ International Conference on Intelligent Robots and System*, volume 3, pages 2735–2739–. IEEE Press, USA, 2002.
- [385] P. Theodosopoulos and T. Theodosopoulos. Quantitative biology, abstract. Technical report, arXiv, arXiv.org, June 2004.
- [386] P. Theodosopoulos and T. Theodosopoulos. Robustness and evolvability of the b cell mutator mechanism. Technical report, arXiv, arXiv.org, June 2005.
- [387] G. J. Thorbecke, A. R. Amin, and V. K. Tsiagbe. Biology of germinal centers in lymphoid tissue. *The Federation of American Societies for Experimental Biology Journal (FASEB)*, 8:832–840–, 1994.
- [388] J. Timmis. Artificial immune systems - today and tomorrow. *Natural Computing*, 6(1):1–18–, March 2007.
- [389] J. Timmis, M. Amos, W. Banzhaf, and A. Tyrrell. ”going back to our roots”: second generation biocomputing. Technical Report arXiv:cs/0512071v1, arXiv, Dec 2005.

- [390] J. Timmis, C. Edmonds, and J. Kelsey. Assessing the performance of two immune inspired algorithms and a hybrid genetic algorithm for function optimisation. In *Proceedings of the Congress on Evolutionary Computation (CEC04)*, volume 1, pages 1044–1051–. IEEE Press, USA, 2004.
- [391] J. Timmis and M. Neal. Investigating the evolution and stability of a resource limited artificial immune system. In *Workshop Program*, pages 40–41–. AAAI Press, USA, 2000.
- [392] J. Timmis, M. Neal, and J. Hunt. An artificial immune system for data analysis. *Biosystems*, 55(1):143–150–, 2000.
- [393] J. Timmis and M. J. Neal. A resource limited artificial immune system for data analysis. *Knowledge Based Systems Journal: Special Issue*, 14(3-4):121–130–, 2001.
- [394] K. M. Ting and I. H. Witten. Issues in stacked generalization. *Journal of Artificial Intelligence Research*, 10:271–289, 1999.
- [395] N. Toma, S. Endo, K Yamada, and H. Miyagi. The immune distributed competitive problem solver with majorhistocompatibility complex and immune network. In *IEEE International Conference on Systems, Man, and Cybernetics*, volume 3, pages 1865–1870–. IEEE Press, USA, 2000.
- [396] S. Tonegawa. Somatic generation of antibody diversity. *Nature*, 302(5909):575–581–, April 1983.
- [397] S. Tonegawa. Somatic generation of immune diversity. *Bioscience Reports, Biomedical and Life Sciences, Nobel Lecture*, 8(1):3–26–, February 1988.
- [398] F. Trepel. Number and distribution of lymphocytes in man. a critical analysis. *Klinische Wochenschrift*, 52(11):511–515–, 1974.
- [399] E. Tsang. *Foundations of Constraint Satisfaction*. Academic Press, 1993.
- [400] J. Twycross and U. Aickelin. Towards a conceptual framework for innate immunity. In *LNCS*, pages 112–125–. Springer, Germany, 2005.
- [401] J. Twycross and U. Aickelin. Biological inspiration for artificial immune systems. In *Proceedings of the 6th International Conference on Artificial Immune Systems (ICARIS-2007)*, pages –, 2007.
- [402] A. Trn, M.M. Ali, and S. Viitanen. Stochastic global optimization: Problem classes and solution techniques. *Journal of Global Optimization*, 14:437–447, 1999.
- [403] F. J. Varela. A cognitive view of the immune system. *World Futures*, 42(1-2):31–40–, 1994.
- [404] P. A. Vargas, L. N. de Castro, R. Michelan, and F. J. Von Zuben. An immune learning classifier network for automated navigation. In Peter J. Bentley and Emma Hart, editors, *Lecture Notes in Computer Science*, pages 69–80–. Springer, Berlin Heidelberg, 2003.
- [405] P. A. Vargas, L. N. de Castro, and F. J. Von Zuben. Artificial immune systems as complex adaptive systems. In *Proc. of ICARIS (International Conference on Artificial Immune Systems)*, pages 115–123, 2002.

- [406] P. A. Vargas, L. N. de Castro, and F. J. Von Zuben. Mapping artificial immune systems into learning classifier systems. In *Lecture Notes in Computer Science*, pages 163–186–. Springer, Berlin / Heidelberg, 2003.
- [407] T. J. Vaughan, J. K. Osbourn, and P. R. Tempest. Human antibodies by design. *Nature Biotechnology*, 16:535–539–, 1998.
- [408] M. Villalobos-Arias, C. A. Coello Coello, and O. Hernandez-Lerma. Convergence analysis of a multiobjective artificial immune system algorithm. In Vincenzo Cutello, Peter J. Bentley, and Jon Timmis, editors, *Lecture Notes in Computer Science*, pages 226–235–, Berlin / Heidelberg, 2004. Springer.
- [409] P. Vineis. Cancer as an evolutionary process at the cell level: an epidemiological perspective. *Carcinogenesis*, 24(1):1–6, Jan 2003.
- [410] P. Vineis and M. Berwick. The population dynamics of cancer: a darwinian perspective. *Int J Epidemiol*, 35(5):1151–1159, Oct 2006.
- [411] U. H. von Andrian and T. R. Mempel. Homing and cellular traffic in lymph nodes. *Nature reviews. Immunology*, 3(11):867–878–, 2003.
- [412] C. von der Malsburg. Self-organization of orientation sensitive cells in the striate cortex. *Biological Cybernetics*, 14(2):85–100–, December 1973.
- [413] S. Voulgaris, M. Jelasity, and M. van Steen. A robust and scalable peer-to-peer gossiping protocol. In *Agents and Peer-to-Peer Computing*, 2005.
- [414] T. A. Waldmann. Immunotherapy: past, present and future. *Nature Medicine*, 9(3):269–277–, 2003.
- [415] J. H. Walker and S. M. Garrett. Dynamic function optimisation: Comparing the performance of clonal selection and evolutionary strategies. In *Lecture Notes in Computer Science*, pages 273–284–. Springer-Verlag, Germany, 2003.
- [416] R. Aaron Warnock, Sanaz Askari, Eugene C. Butcher, and Ulrich H. von Andrian. Molecular mechanisms of lymphocyte homing to peripheral lymph nodes. *Journal of Experimental Medicine (JEM)*, 182(2):205–216–, 1998.
- [417] A. Watkins, X. Bi, and A. Phadke. Parallelizing an immune-inspired algorithm for efficient pattern recognition. In A. Buczak, J. Ghosh, M. Embrechts, and O. Ersoy, editors, *Fuzzy Logic, Evolutionary Programming, Complex Systems and Artificial Life*, volume 13, pages 225–230–. ASME Press, New York, USA, 2003.
- [418] A. Watkins and L. Boggess. A new classifier based on resource limited artificial immune systems. In *Part of the 2002 IEEE World Congress on Computational Intelligence*, pages 1546–1551–. IEEE, May 2002.
- [419] A. Watkins and J. Timmis. Artificial immune recognition system (airs): Revisions and refinements. In P.J. Bentley, editor, *1st International Conference on Artificial Immune Systems (ICARIS2002)*, pages 173–181–. University of Kent at Canterbury Printing Unit, UK, 2002.
- [420] A. Watkins and J. Timmis. Exploiting parallelism inherent in airs, an artificial immune classifier. In V. Cutello, P. Bentley, and J. Timmis, editors, *Lecture Notes in Computer Science (LNCS)*, volume 3239, pages 427–438–. Springer-Verlag GmbH, 2004.

- [421] A. Watkins, J. Timmis, and L. Boggess. Artificial immune recognition system (airs): An immune-inspired supervised learning algorithm. *Genetic Programming and Evolvable Machines*, 5(3):291–317–, September 2004.
- [422] A. B. Watkins. Airs: A resource limited artificial immune classifier. Master’s thesis, Mississippi State University, USA, 2001.
- [423] A. B. Watkins and L. C. Boggess. A resource limited artificial immune classifier. In *Part of the 2002 IEEE World Congress on Computational Intelligence held in Honolulu*, pages 926–931–, USA, May 2002. IEEE Computer Society.
- [424] R. G. Weinand. Somatic mutation, affinity maturation and the antibody repertoire: a computer model. *Journal of Theoretical Biology*, 143(3):343–382–, 1990.
- [425] G. Weiss. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. MIT Press, 1999.
- [426] D. Wettschereck, D. W. Aha, and T. Mohri. A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. *Artificial Intelligence Review*, 11(1-5):273–314–, February 1997.
- [427] J. White and S. M. Garrett. Improved pattern recognition with artificial clonal selection? In *Lecture Notes in Computer Science*, pages 181–193–. Springer-Verlag, Berlin / Heidelberg, Germany, 2003.
- [428] R. H. White. Competitive hebbian learning. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN-91)*, volume 2, pages 949–. IEEE Press, USA, 1991.
- [429] R. H. White. Competitive hebbian learning: algorithm and demonstrations. *Neural Networks*, 5(2):261–275–, 1992.
- [430] R. H. White. Recurrent competitive hebbian learning. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, volume 4, pages –. IEEE Press, USA, 1992.
- [431] D. Whitley, S. Rana, J. Dzubera, and K E. Mathias. Evaluating evolutionary algorithms. *Artificial Intelligence - Special volume on empirical methods*, 85(1-2):245–276–, 1996.
- [432] P. D. Williams, K. P. Anchor, J. L. Bebo, G. H. Gunsch, and G. D. Lamont. Cdis: Towards a computer immune system for detecting network intrusions. In L. M and A. Wespi, editors, *Lecture Notes in Computer Science*, pages 117–133–. Springer, Berlin / Heidelberg, 2001.
- [433] I. H. Witten and E. Frank. *Data Mining: Practical machine learning tools with Java implementations*. Morgan Kaufmann, San Francisco, 2000.
- [434] D. H. Wolpert. Stacked generalization. *Neural Networks*, 5(LA-UR-90-3460):241–259, 1992.
- [435] D. H. Wolpert and W. G. Macready. No free lunch theorems for search. Technical report, Santa Fe Institute, Sante Fe, NM, USA, 1995.
- [436] D. H. Wolpert and W. G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(67):67–82–, 1997.

- [437] S. Wright. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the Sixth International Congress on Genetics*, pages 355–366–, 1932.
- [438] L J Wysocki and M L Geftter. Gene conversion and the generation of antibody diversity. *Annual Review of Biochemistry*, 58:July 1989–, July 1989.
- [439] L. Xu, A. Krzyzak, and E. Oja. Rival penalized competitive learning for clustering analysis, rbfnets, and curve detection. *IEEE Transactions on Neural Networks*, 4(4):636–649–, July 1993.
- [440] X. Yao. A review of evolutionary artificial neural networks. *International Journal of Intelligent Systems*, 8(4):539–567, 1993.
- [441] X. Yao, Y. Liu, and G. Lin. Evolutionary programming made faster. *IEEE Transactions on Evolutionary Computation*, 3(2):82–102, July 1999.
- [442] A. J. Young and J. Hay. Lymphocyte migration in the development and disease. *Seminars in Immunology*, 11(2):71–, 1999.
- [443] Y. Yu and C. Z. Hou. A clonal selection algorithm by using learning operator. In *Proceedings of 2004 International Conference on Machine Learning and Cybernetics*, volume 5, pages 2924–2929–. IEEE Press, USA, 2004.
- [444] M. J. Zaki. Parallel and distributed data mining: An introduction. In *Revised Papers from Large-Scale Parallel Data Mining, Workshop on Large-Scale Parallel KDD Systems, SIGKDD*, pages 1–23, 1999.
- [445] Y. Zhang and C. Hou. A clone selection algorithm with niching strategy inspiring by biological immune principles for change detection. In *IEEE International Symposium on Intelligent Control*, pages 1000–1005–. IEEE Press, USA, 2003.
- [446] R. M. Zinkernagel, S. Ehl, P. Aichele, S. Oehen, T. Kundig, and H. Hengartner. Antigen localisation regulates immune responses in a dose- and time-dependent fashion: a geographical view of immune reactivity. *Immunological Reviews*, 156:199–209–, 1997.