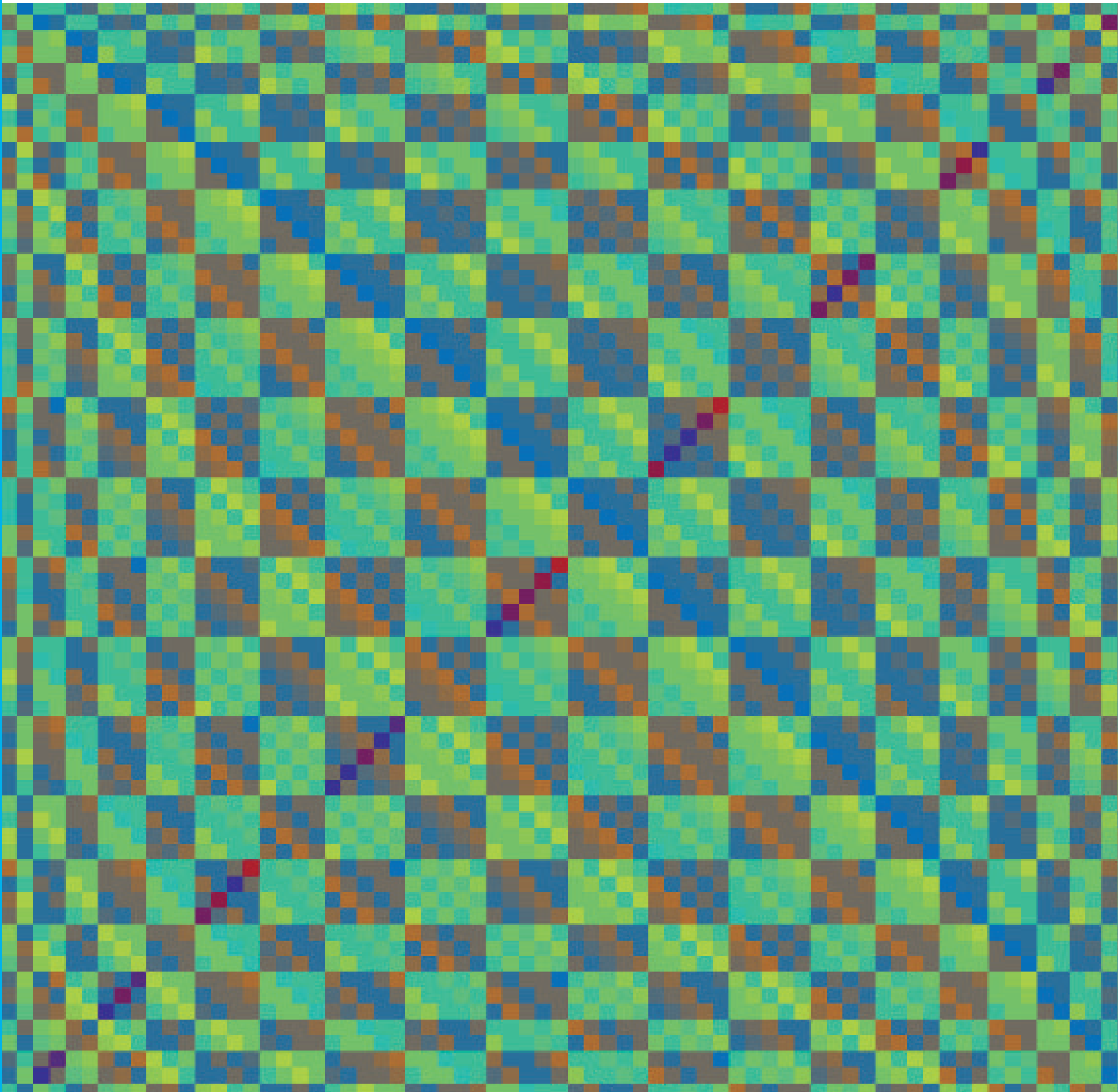
EUROPEAN
SCIENCE
FOUNDATION

Automata: from Mathematics to Applications (AutoMathA)

An ESF Standing Committee for Physical and Engineering Sciences (PESC) Programme



Automata theory (AT) is one of the longest established areas in computer science. Over the past few years AT has not only developed in many different directions but has also evolved in an exciting way at several levels: the exploration of specific new models and applications has at the same time stimulated a variety of deep mathematical theories.

Standard applications of AT include pattern matching, syntax analysis and software verification. In recent years, novel applications of automata-theoretic concepts have emerged from biology, physics, cognitive sciences, neurosciences, control, tomography, linguistics, mathematics, etc., while developments in information technology have increased the need for formally based design and verification methods to cope with emerging technical needs such as network security, mobile intelligent devices and high performance computing.

At the same time, the mathematical foundations of AT rely on more and more advanced areas of mathematics. While in the early 1960s only elementary graph theory and combinatorics were required, new tools from non-commutative algebra (semigroups, semirings and formal power series), logic, probability theory and symbolic dynamics have been successively introduced and the latest developments borrow ideas from topology and geometry. Both trends have enhanced the role of fundamental research in AT and the importance of closer interaction between theoretical and applied scientists.

This multidisciplinary programme lies at the crossroads of mathematics, theoretical computer science and applications. By setting up a framework where new applications of AT and theoretical insights can be communicated and shared by an open and qualified group of European scientists, this programme will catalyse progress in both directions.

The running period of the ESF AutoMathA programme is from June 2005 to May 2010.

Cover:
Representation of the action of the Thue-Morse operator on the cyclic group of order 70

Scientific background

Automata theory (AT) was born some fifty years ago, with the seminal work of Kleene, who first formalised the early attempts of McCulloch and Pitts, and was originally motivated by the study of neural networks. For many years, its main applications have been computer design, compilation of programming languages and pattern matching. But over the last fifteen years, applications of AT have considerably diversified and encompass – in addition to the classical ones – the following: system analysis and verification, natural language processing, information retrieval, data compression and control of discrete event systems.

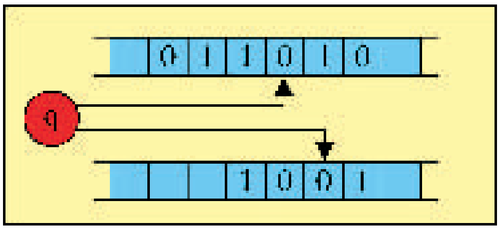
New challenges for AT include biology, molecular and cell models, mobile computing, quantum computing, tomography, picture processing or multisensor integration.

For most applications, one needs first to identify or conceive a suitable automata model and then to devise adapted automata theoretic algorithms to solve computational problems efficiently. This explains the large variety of automata models that have been studied and the need for unifying theories.

One of the main reasons for the early and steady success of automata and language theory was the availability of strong theoretical results supporting the sequential model of computation. Unsurprisingly, fundamental results are even more necessary after the emergence of the new applications listed above. The time is now right to address these new theoretical challenges, with the aim of carefully extending the theoretical base in the directions called for by the emerging applications.

Automata models

Over the last few decades, several new types of automata have been introduced that extend finite automata in various directions.



An early model, the Turing machine

Automata-based formalisms are used as models for various kinds of systems such as concurrent and distributed systems, real-time systems, embedded software systems, etc. Some proposals already proved their value, while others are under scrutiny. Hence singling out the important automata models for the intended applications is a central question.

Automata models are specified by four basic components: input, output, storage device and acceptance mode.

Fundamental aspects

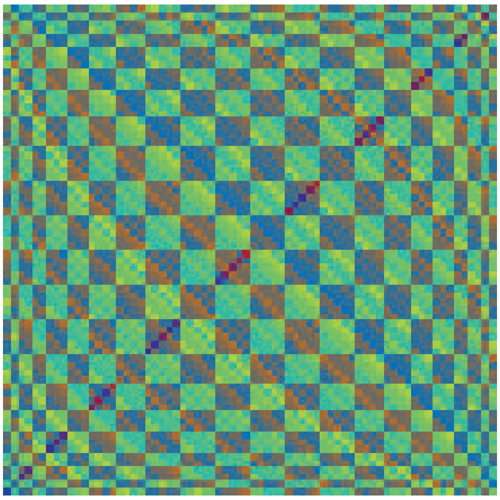
— **Mathematical foundations**

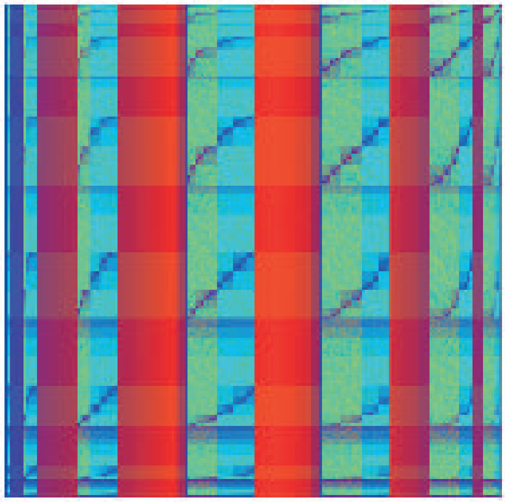
The need for strong mathematical foundations for AT goes back to the early 1960s. The first domains involved were logic, with the works of Büchi and Rabin, and semigroup theory, with the emergence of the abstract notions of recognisable and rational sets. This threefold approach (AT, logic and non-commutative algebra) is at the core of the mathematical theory of automata and examples abound to show that progress in one component deeply influences the others.

Another fundamental issue might be called ‘automatic mathematics’. The general question can be formulated as follows: what kind of mathematical objects, such as functions, relations, sets, game strategies, can be realised by finite automata?

— **Algorithms on automata and words**

As explained above, the study covers many different types of automata (working on different objects, with different acceptance conditions, presented in different ways), and their descriptive and computational complexities need investigation.





Operators acting on the cyclic group of order 70 (on the left) and on the dihedral group of order 256 (on the right)
© J. Almeida

— **Languages and formal grammars**

Within Chomsky's hierarchy, context-free grammars are the classical model for the syntax of programming and technical languages. Derived models have been introduced subsequently. Most recent developments of grammars have been motivated by scientific and technological requirements from other disciplines.
The need to define structured texts in the World Wide Web has triggered the study of specialised grammar models, based on multisort bracketing.
Other examples include modelisation of DNA sequences, and 2D pictures.

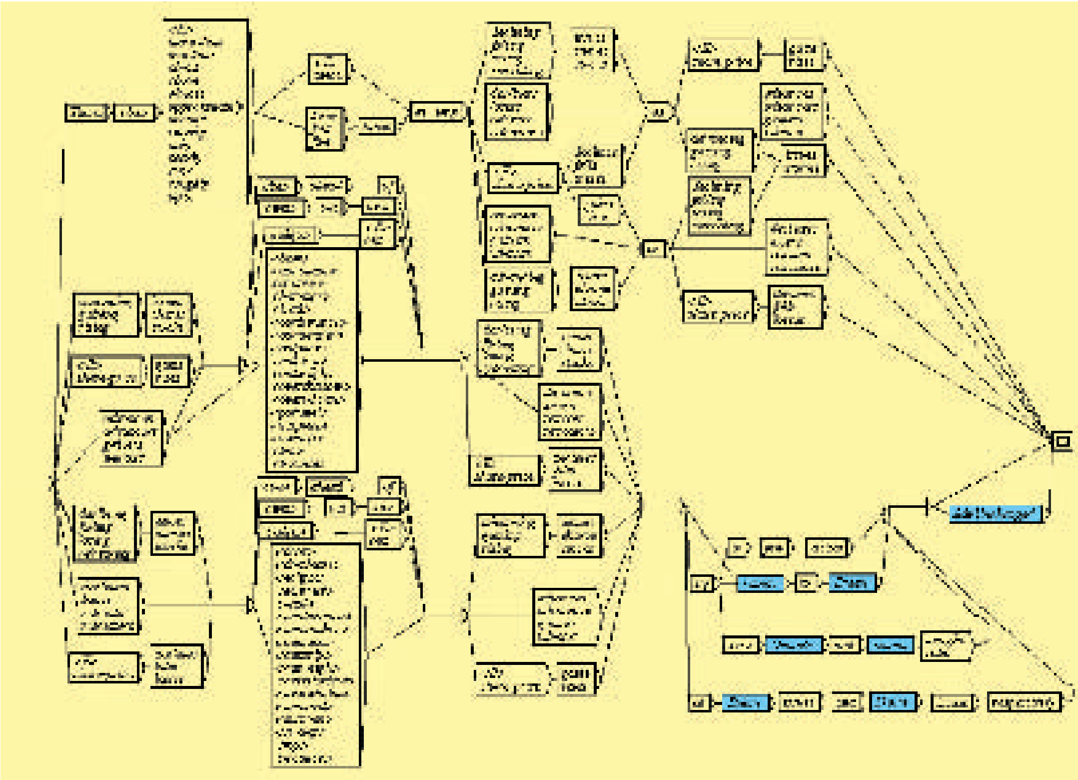
— **Symbolic dynamics and coding**

Dynamical systems originally arose as a model for physical phenomena such as the motion of planets or the behaviour of molecules in a gas. Symbolic dynamics can be viewed, in a sense, as dynamical systems in which space and time have been discretised.
The main open problems of this theory are perhaps the isomorphism problem for subshifts of finite type and the characterisation of conjugate subshifts. The automata theoretic approach to this problem has already led to promising partial results and is being pursued. In particular, the links between coverings of finite automata and conjugacy should shed light both on algorithms used in AT and on methods used in symbolic dynamics.
The road-colouring problem, another famous open problem in symbolic dynamics, is actually a problem on synchronising automata. This problem is studied in connection with another open problem on synchronising automata, namely Cerny's conjecture.

Applications

— **System analysis and verification**

It is widely acknowledged that to ensure reliability of computer systems, development methods must include automatic verification and design techniques that allow the detection of defective behaviours of the system.
Software design and verification requires the study of structural properties of automata. New paradigms are needed to handle concurrency, timing constraints, quantitative reasoning, data manipulation, dynamism, parameterisation, etc.
Game problems in an automata-theoretic setting is a powerful framework for controlling a system and force it to satisfy some given property against any environment.
Model-checking is a method for formally verifying finite-state concurrent systems. Specifications are expressed as logic formulae, and symbolic algorithms are used to check if the specification holds or not. Games, temporal logics and alternating automata provide efficient tools for model-checking.
The behaviour of nets, communicating automata etc., is best explained by mathematical structures with an explicit notion of concurrency (for example asynchronous automata, branching automata or Message Sequence Charts).
Configurations of infinite-state systems can be encoded by more complex automata called 'finite state transducers' on words or on trees.



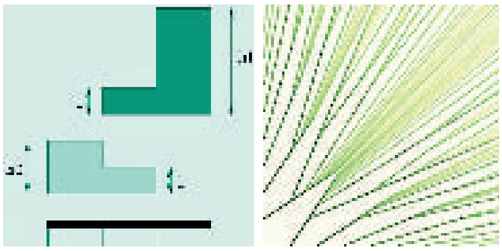
Stock Exchange-related sentences can be represented by an automaton.

Natural language processing, text and hypertext processing, information retrieval, data compression

Other applications concern semistructured information; that is, information whose structure is not a priori established, such as texts, hypertexts or web pages. Novel databases and search engines face the problem of representing and efficiently querying such data, naturally represented by graphs. These graphs may be interpreted as automata, and queries as logical formulae, reducing a query to an instance of a model-checking problem.

Data compression algorithms and systems also use automata intensively. This occurs during the two main steps of these methods: modelling and coding.

Almost all of the most efficient data-compression methods, such as the Ziv-Lempel text compression variants, consider that dictionaries implemented by automata store repetitive patterns. A more recent technique, not yet fully explored, uses automata to store the set of minimal words that do not occur in the source text, called its ‘antidictionary’. This notion makes sense for developing text-compression software only because the whole chain of treatment is based on automata.



How to optimally pack Tetris heaps (left).
The cells in the phase diagram (right) show the heights of the pieces for which the optimal arrangements are the same.

Control and discrete event systems

How to evaluate the amount of parallelism in a computer architecture? How to maximise the throughput of a production line? These are the sort of questions that it is important to address in the context of artificially made systems such as communication networks, transportation networks or manufacturing systems. The term ‘discrete event systems’ (DES) was coined in the 1980s to describe an approach to modelling these systems that is event-driven rather than time-driven and that uses discrete state variables.

A first approach to studying a DES is concerned with the precise ordering of the events. A typical objective is to build a controller that allows and disallows transitions between states of the system in order to guarantee a certain desired behaviour. This approach is based on automata and formal languages.
A recent and innovative development is the max-plus approach: the max operation represents sequential dependencies between events, and the plus operation corresponds to the execution times of the events. This new approach involves the use of several additional mathematical tools such as exotic semirings, formal power series and probability theory. Extending the results of the logical model to this richer setting is an important and challenging objective.

New challenges

New challenges include biology (molecular and cell models), mobile computing, quantum computing, tomography, picture processing and multisensor integration.

Activities

The programme includes the following planned activities:

– short-term visits/exchanges among the programme participants;
– organisation of workshops for programme participants, to allow the dissemination of early research results and experiences;
– sponsoring of conferences in the area of AutoMathA;
– organisation of schools on the subjects covered by the programme;
– organisation of open workshops in the area of AutoMathA;
– setting up a comprehensive Internet research dissemination channel and publication activities.

More information at: www.esf.org/automatha

Funding

ESF scientific programmes are principally financed by the Foundation's Member Organisations on an *à la carte* basis. AutoMathA is supported by:
Fonds zur Förderung der wissenschaftlichen Forschung, Austria; Fonds National de la Recherche Scientifique, Belgium; Akademie věd České republiky, Czech Republic; Grantová agentura České republiky, Czech Republic; Forskningsrådet for Natur og Univers, Denmark; Suomen Akatemia/Finlands Akademi; Finland; Centre National de la Recherche Scientifique, France; Deutsche Forschungsgemeinschaft, Germany; Magyar Tudomanyos Akademia, Hungary; the Israeli Academy of Sciences and Humanities, Israel; the CNR Instituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni, Torino, Italy; Polska Akademia Nauk, Poland; Fundação para e Ciência e a Tecnologia, Portugal; Slovenská Akadémia Vied, Slovak Republic; Schweizerischer Nationalfonds zur Förderung der wissenschaftlichen Forschung/Fonds National Suisse de la Recherche Scientifique, Switzerland.

AutoMathA Steering Committee

Professor Jean-Eric Pin (Chair)
Laboratoire d'Informatique Algorithmique
LIAFA
Université Paris Denis Diderot et CNRS
Case 7014
2, place Jussieu
75251 Paris Cedex 05
France
Tel: +33 1 44 27 68 45
Fax: +33 1 44 27 68 49
E-mail: Jean-Eric.Pin@liafa.jussieu.fr

Professor Jorge Almeida
Centro de Matemática
Departamento de Matemática Pura
Faculdade de Ciências
Universidade do Porto
Rua de Campo Alegre, 687
4169-007 Porto
Portugal
Tel: +351 22 0100732
Fax: +351 22 0100708
E-mail: jalmeida@fc.up.pt

Professor Véronique Bruyère
Institute of Computer Science
University of Mons-Hainaut
Le Pentagone
6 Avenue du Champ de Mars
7000 Mons
Belgium
Tel: +32 65373444
Fax: + 32 65373459
E-mail: Veronique.Bruyere@umh.ac.be

Professor Stefano Crespi-Reghizzi
Dipartimento di Elettronica e Informazione
Politecnico di Milano
Piazza Leonardo da Vinci, 32
Milano
Italy
Tel: +39 02 23 99 35 18
Fax: +39 02 23 99 34 11
E-mail: crespi@elet.polimi.it

Professor Søren Eilers
Department of Mathematics
University of Copenhagen
Universitetsparken 5
2100 Copenhagen Ø
Denmark
Tel: +45 35320757
Fax: +45 35320704
E-mail: eilers@math.ku.dk

Professor Zoltan Esik
Institute of Computer Science
University of Szeged
PO Box 652
6701 Szeged
Hungary
Tel: +36 62 544 289
Fax: +36 62 420 292
E-mail: ze@inf.u-szedeg.hu

Professor Jozef Gruska
Faculty of Informatics
Masaryk University
Botanická 68a
602 00 Brno
Czech Republic
Tel: +420 549 49 4592
Fax: +420 549 49 1820
E-mail: gruska@fi.muni.cz

Professor Tatiana Jajcayova
Institute of informatics
Faculty of Mathematics, Physics
and Informatics
Comenius University
Mlynska dolina
842 15 Bratislava
Slovak Republic
Tel: +421 7 602 95 278
Fax: +421 7 654 25 882
E-mail: jajcayova@fmph.uniba.sk

Professor Juhani Karhumaki
Department of Mathematics
University of Turku
20014 Turku
Finland
Tel: +358 2 333 5613
Fax: +358 2 333 6595
E-mail: karhumak@cs.utu.fi

Professor Andrzej Kisielewicz
Institute of Mathematics
Wrocław University
Pl. Grunwaldzki 2/4
50-384 Wrocław
Poland
Tel: +48 71 375 7467
Fax: +48 71 375 7429
E-mail: kisiel@math.uni.wroc.pl

Professor Werner Kuich
Institut für Algebra und
Computermathematik
Technische Universität Wien
Wiedner Hauptstrasse 8-10
1040 Wien
Austria
Tel: +43 1 58801 11820
Fax: +43 1 58801 11892
E-mail: kuich@pop.tuwien.ac.at

Professor Stuart W. Margolis
Department of Informatics
Bar Ilan University
52900 Ramat Gan
Israel
Tel: +972 3 531 7608
Fax: +972 3 535 3325
E-mail: margolis@math.biu.ac.il

Professor Wolfgang Thomas
Fachgruppe Informatik
RWTH Aachen
Ahornstrasse 55
52056 Aachen
Germany
Tel: +49 241 80 21700
Fax: +49 241 80 22215
E-mail: thomas@informatik.rwth-aachen.de

Professor Jacques Duparc
École des HEC
Université de Lausanne
1015 Dorigny
Switzerland
Tel: +41 21 692 35 88
Fax: +41 21 692 35 85
E-mail: jacques.duparc@unil.ch

ESF Liaison

Dr. Patricia Arsene
Science

Ms. Catherine Werner
Administration

Physical and Engineering
Sciences Unit (PESC)
European Science Foundation
1 quai Lezay-Marnésia
BP 90015
Tel: +33 (0)3 88 76 71 28
Fax: +33 (0)3 88 37 05 32
E-mail: cwerner@esf.org

For the latest information on this programme consult the AutoMathA home page:
www.esf.org/automatha

EUROPEAN
SCIENCE
FOUNDATION

European Science Foundation
1 quai Lezay-Marnésia, BP 90015
67080 Strasbourg cedex, France
Tel: +33 (0)3 88 76 71 00
Fax: +33 (0)3 88 37 05 32
www.esf.org

January 2006 – Print run: 1500