

UNIVERSIDAD ADOLFO IBAÑEZ  
FACULTAD DE CIENCIAS Y TECNOLOGÍA  
Escuela de Ingeniería Industrial

MODELACIÓN DE UN SISTEMA ECONÓMICO  
BASADO EN CONCEPTOS HAYEKIANOS  
USANDO REDES NEURONALES ARTIFICIALES

IAN DEREK FARREN GAUSE

TESIS PARA OPTAR AL GRADO DE  
**MAGISTER EN CIENCIAS DE LA INGENIERÍA**  
Y AL TÍTULO DE  
**INGENIERO CIVIL INDUSTRIAL**

2005

UNIVERSIDAD ADOLFO IBAÑEZ  
FACULTAD DE CIENCIAS Y TECNOLOGÍA  
Escuela de Ingeniería Industrial

MODELACIÓN DE UN SISTEMA ECONÓMICO  
BASADO EN CONCEPTOS HAYEKIANOS  
USANDO REDES NEURONALES ARTIFICIALES

IAN DEREK FARREN GAUSE

Profesor Guía: Filadelfo de Mateo Gómez, Ph.D.  
Profesores Comisión: Filadelfo de Mateo Gómez, Ph.D.  
José Maldifassi Pohlhammer, Ph.D.  
Marcelo Villena Chamorro, Ph.D.

TESIS PARA OPTAR AL GRADO DE  
**MAGISTER EN CIENCIAS DE LA INGENIERÍA**  
Y AL TÍTULO DE  
**INGENIERO CIVIL INDUSTRIAL**

2005

# RESUMEN

Todo sistema económico es complejo. Los integrantes del sistema refuerzan sus decisiones en función de señales que entrega el mercado, lo que crea un proceso de retroalimentación en el sistema, mostrando respuestas no lineales. Por esta razón, es muy difícil hacer pronósticos en un sistema económico.

Si bien históricamente se han usado procesos estocásticos para desarrollar pronósticos en los sistemas económicos, esta metodología solo reproduce, *ex post*, un esquema de correlaciones entre las variables que fueron consideradas y que generalmente no modelan los mecanismos que condicionan esas relaciones.

En este trabajo se presenta un modelo que tiene una representación explícita, *ex ante*, de los procesos que trabajan en el mundo social; en especial en el proceso de creación de empresas.

Para desarrollar el modelo se usa la Simulación Basada en Agentes. El proceso de negociación entre agentes, fundamento del acto económico, está modelado en base a la teoría de equilibrio económico y del conocimiento desarrollada por Friedrich August von Hayek. Se usaron Redes Neuronales Artificiales para representar esta teoría en el modelo. El modelo considera un mercado de bienes, un mercado del trabajo y una política de quiebra de las empresas no solventes. No hay entidades controladoras, por lo que el modelo muestra un sistema capitalista puro. Inicialmente los agentes son productores independientes que con el pasar del tiempo podrían emplearse o formar su propia empresa contratando a otros agentes.

El modelo pone de manifiesto fenómenos emergentes de crecimiento económico cíclico, y una distribución más equitativa de la riqueza, a medida que la economía se desarrolla.

Además, se desarrolla un experimento creando una entidad que cobra impuestos y los distribuye entre los agentes, por partes iguales. Se presenta resultados para cuatro tasas de impuesto, y la tasa que optimiza la generación de riqueza.

A mis padres,  
por su apoyo y ejemplo.

# AGRADECIMIENTOS

A mi mentor Sr. Filadelfo de Mateo, por su gran ayuda y entrega de conocimientos, incluso en temas que van más allá del alcance de esta tesis.

Al profesor Pietro Terna de la Universidad de Torino, Italia, por introducirme al mundo de la Simulación Basada en Agentes.

Al profesor Paul Johnson de la Universidad de Kansas, Estados Unidos, y miembro del Grupo de Desarrollo Swarm, por su disposición a aclarar dudas referentes a la programación en Swarm.

A mi familia, por los valores que me entregaron.

A mi polola María Ignacia Alicera, por su comprensión y motivación.

A mi compañera de curso Silvana Sariego, por crear los mejores apuntes de clases y compartirlos con sus compañeros.

A Loretto de la Cruz, por su paciencia y ayuda.

# ÍNDICE

<b>CAPÍTULO 1: INTRODUCCIÓN .....</b>	<b>1</b>
<b>CAPÍTULO 2: REDES NEURONALES ARTIFICIALES.....</b>	<b>3</b>
2.1. INTRODUCCIÓN .....	3
2.2. LA ESTRUCTURA DE REDES FEED-FORWARD .....	4
2.2.1. LA RED FEED-FORWARD DE TRES CAPAS TOTALMENTE INTERCONECTADAS CON APRENDIZAJE BACKPROPAGATION .....	5
2.3. EL APRENDIZAJE EN REDES NEURONALES.....	10
2.3.1. PRINCIPALES CARACTERÍSTICAS DEL APRENDIZAJE .....	10
2.3.2. EL ALGORITMO BACKPROPAGATION .....	11
2.3.3. COMENTARIOS GENERALES ACERCA DEL BACKPROPAGATION .....	15
2.4. LA TÉCNICA DE ENTRENAMIENTO CRUZADO .....	19
<b>CAPÍTULO 3: SIMULACIÓN Y LAS CIENCIAS SOCIALES .....</b>	<b>24</b>
3.1. SIMULACIÓN VS. MODELOS ESTADÍSTICOS.....	24
3.2. SIMULANDO SOCIEDADES HUMANAS .....	25
3.3. SISTEMAS COMPLEJOS.....	26
3.3.1. FENÓMENOS EMERGENTES .....	28
3.4. VALIDACIÓN DE SIMULADORES BASADOS EN AGENTES .....	28
<b>CAPÍTULO 4: HAYEK Y SU TEORÍA DE EQUILIBRIO ECONÓMICO.....</b>	<b>30</b>
4.1. IMPLICANCIAS DEL EQUILIBRIO .....	30
4.2. EQUILIBRIO COMO CONSISTENCIA DE PLANES .....	32
4.3. LA TENDENCIA AL EQUILIBRIO.....	33
<b>CAPÍTULO 5: EL MODELO MECHA .....</b>	<b>35</b>

<b>5.1.</b>	<b>INTRODUCCIÓN .....</b>	<b>35</b>
<b>5.2.</b>	<b>LOS AGENTES.....</b>	<b>36</b>
<b>5.3.</b>	<b>PRECIOS Y CANTIDADES A OFRECER Y DEMANDAR .....</b>	<b>36</b>
5.3.1.	INTRODUCCIÓN	36
5.3.2.	ESTRUCTURA DE LA RED NEURONAL USADA	37
5.3.3.	COMENTARIOS	48
<b>5.4.</b>	<b>TIPO DE BIEN A PRODUCIR.....</b>	<b>49</b>
<b>5.5.</b>	<b>CREACIÓN DE EMPRESAS.....</b>	<b>52</b>
<b>5.6.</b>	<b>DINERO Y CRÉDITO .....</b>	<b>54</b>
<b>5.7.</b>	<b>DETERMINACIÓN DE SUELDOS .....</b>	<b>55</b>
<b>5.8.</b>	<b>DECISIÓN DE COMPRA.....</b>	<b>57</b>
<b>5.9.</b>	<b>QUIEBRAS .....</b>	<b>59</b>
<b>5.10.</b>	<b>ACCIONES DE MECHA .....</b>	<b>60</b>
5.10.1.	ACCIONES DIARIAS	61
5.10.2.	ACCIONES MENSUALES	62
<b>5.11.</b>	<b>FENÓMENOS EMERGENTES DEL MODELO.....</b>	<b>63</b>
5.11.1.	CRECIMIENTO	63
5.11.2.	CICLOS	64
5.11.3.	EQUIDAD	67
<b>5.12.</b>	<b>COMENTARIOS SOBRE LA VALIDACIÓN OPERACIONAL DEL MODELO .....</b>	<b>69</b>
<b>5.13.</b>	<b>EXPERIMENTO: REDISTRIBUCIÓN DE LA RIQUEZA.....</b>	<b>70</b>
<b><u>CAPÍTULO 6: TRABAJO FUTURO.....</u></b>		<b><u>75</u></b>
<b><u>CAPÍTULO 7: CONCLUSIONES .....</u></b>		<b><u>76</u></b>
<b><u>ANEXO.....</u></b>		<b><u>78</u></b>
<b><u>REFERENCIAS.....</u></b>		<b><u>80</u></b>

# Capítulo 1:

## INTRODUCCIÓN

La unión entre técnica de entrenamiento cruzado para redes neuronales usada por Farren (2004) y la teoría de consistencia de planes en el equilibrio económico creada por Hayek (1937) fue la fuente de inspiración para el desarrollo de esta tesis. La técnica de entrenamiento cruzado es la herramienta ideal para entrenar a una red neuronal que modela a un agente que intenta planear en forma consistente.

El propósito de esta tesis es modelar y validar un mercado de bienes usando las teorías de Hayek. Se extendió el modelo de forma que simula un sistema complejo, incorporándole un mercado del trabajo y un sistema de quiebra de empresas. El mecanismo que permite el crecimiento económico es la creación de firmas por empresarios que logran coordinar los esfuerzos de trabajadores heterogéneos para satisfacer la demanda de consumo de un bien. El empresario contrata trabajadores y dirige el actuar de cada uno de ellos. El beneficio del empresario, es el derecho sobre el residuo de los ingresos una vez que se paga los sueldos. El propósito del simulador es mostrar la relación que existe entre el crecimiento económico y la creación de empresas. Además, se profundiza en la relación crecimiento y equidad económica.

El modelo resultante de esta tesis se denomina MECHA (Modelo Económico Hayekiano) y representa a un sistema cerrado, por lo que no existe ingreso de dinero externo al sistema. Una institución central lleva un registro mensual de las cuentas de los agentes, declarando en quiebra a agentes con deudas que exceden un nivel preestablecido. La pérdida económica por esta quiebra es absorbida por los acreedores del agente quebrado.

Es importante destacar que el modelo MECHA intenta explicar fenómenos emergentes del sistema, y, si bien es posible realizar pronósticos con el mismo, ésta no es su función primera.



## CAPÍTULO 1: INTRODUCCIÓN

Dada la alta complejidad del modelo, éste fue creado usando simulación basada en agentes. Para generar MECHA, se utilizó las librerías de Swarm-2.2, y fue escrito en lenguaje Objective-C.<sup>1</sup>

Esta tesis se divide en siete capítulos, siendo éste el primero. En los capítulos dos, tres y cuatro se entregan las bases conceptuales sobre las cuales se creó el modelo. Estas son: Redes Neuronales Artificiales, algunos conceptos sobre Simulación de sistemas sociales y la teoría Hayekiana de equilibrio económico. En el quinto capítulo se presenta MECHA. Como el modelo tiene bastantes actividades, se decidió partir explicando las actividades básicas (transacción entre agentes) e ir agregando las demás actividades (selección de bienes, creación de empresa, quiebras, etc.) una a una de forma de hacer más fácil la comprensión para el lector. Finalmente, se analizan los resultados del modelo y se realiza un experimento con tasas de impuesto. En el capítulo seis se entrega la intención acerca del trabajo futuro y en el capítulo siete vemos las conclusiones de esta tesis.

---

<sup>1</sup> Swarm es un conjunto de librerías creadas por el grupo de desarrollo Swarm (SDG). Este grupo tiene como propósito promover la simulación basada en agentes.

## Capítulo 2:

# REDES NEURONALES ARTIFICIALES

## 2.1. Introducción

El origen de las redes neuronales artificiales (RNA) viene del trabajo de Hebb (1949), quien propuso una ley de aprendizaje que se transformó en la antecesora de las técnicas actuales de entrenamiento de redes neuronales. Posteriormente, Rosenblatt (1959) y Widrow y Hoff (1960) continuaron con su desarrollo. Después de este brillante inicio, la comunidad científica perdió interés en las RNA debido al resultado de Minsky y Papert (1969), quienes dieron a conocer las limitaciones de las arquitecturas existentes en esa época. El problema era que las redes simples (redes sin capa oculta), no eran capaces de solucionar problemas no separables linealmente, como el problema lógico O-exclusiva, XOR. Una solución fue ofrecida por muchos autores en forma independiente. Las principales contribuciones fueron de Werbos (1974), Parker (1985), y Le Cun (1985), quien propuso una arquitectura diferente y una regla de aprendizaje supervisado, la regla Generalizada Delta, que podía ser usada en redes complejas. El trabajo de Rumelhart y McClelland (1986) popularizó esa regla de aprendizaje y sugirió muchas aplicaciones. Investigaciones paralelas profundizaron en el estudio de reglas de aprendizaje no supervisadas para redes neuronales, Kohonen (1984) y Grossberg (1982). Más recientemente, White (1992) ha iluminado la relación entre las redes neuronales y otros campos como la estadística y la econometría.

Con el tiempo han emergido muchas arquitecturas que se diferencian en su estructura y su metodología de aprendizaje. Algunas de ellas son muy generales, como redes feed-forward con Backpropagation y mapas auto-organizativos de Kohonen, mientras que otros son más específicos, como las redes ART, Carpenter y Grossberg (1986, 1987), y redes counter-propagation Hecht-Nielsen (1987).

Como en esta tesis sólo usamos redes Feed-forward con Backpropagation, sólo se entregará una descripción de este tipo de redes.

## 2.2. La estructura de redes Feed-forward

Las redes neuronales artificiales son estructuras, inspiradas en la biología, que transforman los datos que entran a ellas en datos de salida. La transformación es producida por elementos de procesamiento llamados neuronas artificiales o neurodes (Caudill y Butler, 1990), las que están interconectadas para formar una red dividida en capas (generalmente tres): La capa de entrada que recibe los datos, la capa de salida que entrega los datos transformados y una o más capas ocultas que conectan las capas de entrada y de salida. Las propiedades básicas de una red neuronal, independientemente de su estructura específica (número de capas, número de neuronas, etc) es la siguiente:

- a) Aprendizaje: la capacidad de la red de adaptar su comportamiento al ambiente, o en otras palabras, de autónomamente representar la transformación de datos de entrada en datos de salida en función de una serie de ejemplos.
- b) Generalización: la habilidad de reaccionar en forma coherente a datos de entrada imperfectos o no vistos durante el aprendizaje.
- c) Degradación suave: la alteración o eliminación de algunos elementos de la red no le impide trabajar; sólo induce una leve disminución en el desempeño.

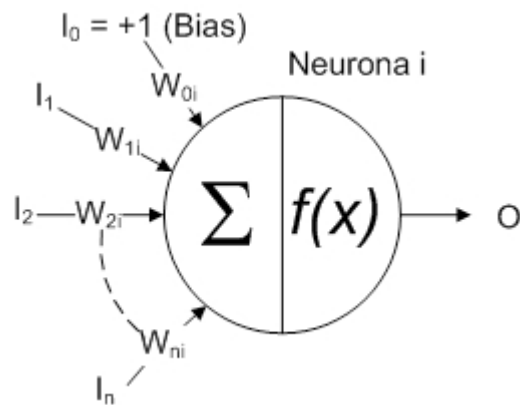
## 2.2.1. La red Feed-forward de tres capas totalmente interconectadas con aprendizaje Backpropagation

Los elementos de esta arquitectura son las neuronas y las interconexiones entre ellas.

### a) Descripción de la neurona

La neurona artificial es un elemento de procesamiento. En la Figura 1.1 se ve que:

1. Recibe las señales  $I_0 \cdot w_{0i}$ ,  $I_1 \cdot w_{1i}$ , ...  $I_k \cdot w_{ki}$ .
2. Suma las señales y genera una salida  $net_i = \sum I_j \cdot w_{ji}$  (el significado de  $w_{ji}$  será explicado mas adelante).
3. Entrega una señal de salida  $O_i$ , esta señal es una transformación de  $net_i$  por medio de una función de activación  $f$ , tal que  $O_i = f(net_i)$ .



$$net_i = \sum_{j=0}^n I_j \cdot w_{ji}$$

Figura 1.1

En forma análoga a las neuronas naturales, la función de activación tiene dos características principales: la existencia de un umbral, y límites superior e inferior. Una neurona natural es activada solo cuando su entrada alcanza un cierto umbral  $\theta$ , y su nivel de activación nunca sobrepasa un nivel de saturación dado. Con la intención de aplicar los algoritmos de aprendizaje más usados, basados en función de activación derivable, la función umbral es generalmente aproximada por una función tal como logística o la arctan. En otras aplicaciones se usa una función lineal. La Figura 1. 2 muestra la función umbral y logística.

Cuando usamos una función de activación como la logística, el umbral de ésta puede ser modelado introduciendo un bias entre las entradas a la neurona. Esto es equivalente a considerar otra señal de entrada con un valor constante  $\theta$ . De hecho, en la Figura 1.1 existen  $k$  señales de entrada variables y una señal de entrada fija  $I_0$  que representa el bias.

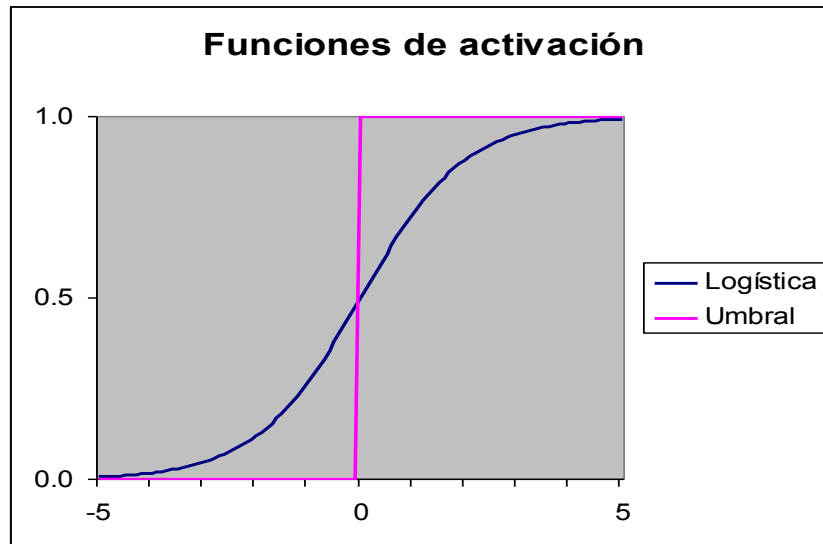


Figura 1.2

El efecto de las señales de entrada en la neurona no es homogéneo, depende de la fuerza de su conexión con la fuente de cada señal. Esta conexión es representada por un peso  $w_{ij}$ , que va desde la neurona origen  $i$  a la neurona destino  $j$ . El valor absoluto del peso representa la fuerza de la conexión, mientras que su signo corresponde a la naturaleza de la conexión: un signo positivo (negativo) crea una señal excitativa (inhibidora).

*b) La estructura de la red*

Dentro de la red, las neuronas están organizadas en tres capas:

1. Una capa de entrada, que recibe señales del exterior. Las neuronas de entrada generalmente no procesan la señal que a ellas llega, solamente la transmiten a la siguiente capa de la red. El número de neuronas de entrada es específica para cada problema y depende de la cantidad y redes tipo de información que es procesada por la red.
2. Una o más capas intermedias (u ocultas) cuyas neuronas procesan según lo descrito en (a). El número de capas ocultas determina la complejidad de la transformación hecha por la red: aunque la mayoría de las aplicaciones están basadas en redes neuronales con una capa oculta, problemas más complejos pueden requerir dos o tres capas. Más adelante se explicará en mayor detalle esta idea.
3. Una capa de salida que entrega señales al exterior una vez que la red ha procesado los datos de entrada. Las neuronas de salida realizan la transformación de señal descrita en (a).

La forma en que las neuronas están conectadas determina la estructura de la red. Una de las más usadas es la llamada red neuronal totalmente conectada, donde todas las neuronas de cada capa están conectadas con (y sólo con) todas las neuronas de la capa siguiente: neuronas de entrada con neuronas ocultas, neuronas ocultas con neuronas de salida. Dentro de una capa, no existe conexión entre una neurona y otra. Cuando la dirección del flujo de información va solo desde la capa de entrada a la capa oculta y desde la capa oculta a la capa de salida, la red es llamada una red feed-forward.

Existen otras arquitecturas que se diferencian de la ya descrita aceptando conexiones de neuronas de la misma capa o un flujo de información hacia atrás, como es el caso en las redes recurrentes, Jordan (1986). En esta tesis sólo se usa redes feed-forward de dos y tres capas, totalmente conectadas.

En esta tesis se usó redes neuronales totalmente conectadas y redes neuronales perceptron, que son un caso particular de las redes totalmente conectadas, donde la red no tiene capa oculta. La red preceptron simplemente recibe señales, las pondera según sus pesos, suma estas ponderaciones y entrega el resultado. Esta red es simplemente una función lineal.

El mecanismo es descrito en el siguiente ejemplo:

En la red de la Figura 1.3, las neuronas están agrupadas en tres capas: de entrada, oculta y de salida. Existen cuatro neuronas de entrada (llamadas 1, 2, 3, 4), tres neuronas ocultas (5, 6, 7) y dos neuronas de salida (8, 9). Podemos decir que la estructura de esta red es totalmente interconectada, ya que cada neurona en la capa de entrada está conectada con cada neurona en la capa oculta, y a su vez cada neurona de la capa oculta está conectada con cada neurona de la capa de salida. Cada conexión tiene un peso asociado  $w_{ij}$  que representa la fuerza y naturaleza de la conexión entre la neurona  $i$  a la neurona  $j$ . La información recibida por las neuronas fluye solo  $\epsilon$ , desde la capa de entrada a la capa oculta, y desde la capa oculta a la capa de salida; sin retroalimentación. No se considera bias por motivo de simplicidad.

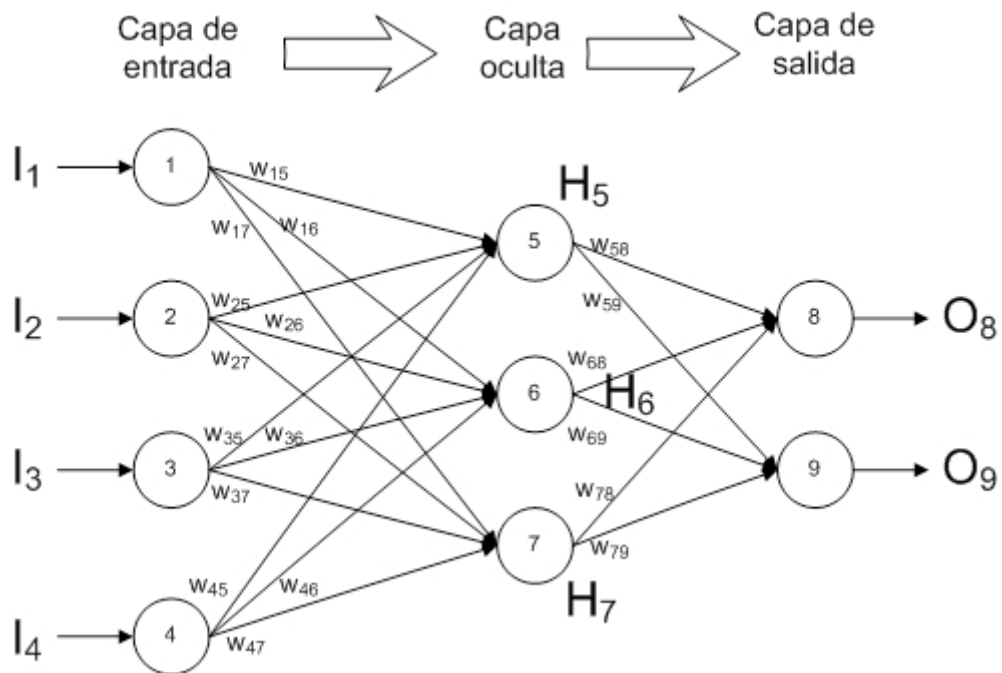


Figura 1.3

Si las cuatro entradas tienen valores  $I_1, I_2, I_3$  y  $I_4$ , la señal entregada a la neurona 5 es igual a:

$$net_5 = I_1 \cdot w_{15} + I_2 \cdot w_{25} + I_3 \cdot w_{35} + I_4 \cdot w_{45} \quad \text{Ecuación 1.1}$$

Donde  $net_i$  es la entrada a la  $i$ -ésima neurona. La neurona 5 transforma esta combinación lineal de entradas mediante la función de activación logística y entrega una salida igual a:

$$H_5 = (1 + \exp(-net_5))^{-1} \quad \text{Ecuación 1.2}$$

Esta salida en conjunto con la salida de las neuronas 6 y 7, forman una combinación lineal que es entregada a las neuronas 8 y 9. La señal recibida por la neurona 8 es:

$$net_8 = H_5 \cdot w_{58} + H_6 \cdot w_{68} + H_7 \cdot w_{78} \quad \text{Ecuación 1.3}$$

Esta neurona es nuevamente transformada por la función logística para entregar la salida final de la red:

$$O_8 = (1 + \exp(-net_8))^{-1} \quad \text{Ecuación 1.4}$$

Es posible utilizar notación matricial para una descripción más clara. Si  $A$  y  $B$  son matrices que tienen los pesos desde la capa de entrada a la capa oculta y desde la capa oculta a la capa de salida respectivamente,  $O$  es el vector de salida de la red,  $I$  es el vector de entrada y  $f$  es la función de activación (definida sobre los lectores apropiados), entonces la red puede ser descrita como:

$$O = f(f(I \cdot A) \cdot B) \quad \text{Ecuación 1.5}$$

Para el ejemplo de la Figura 1.3, las dos matrices son respectivamente:

$$A(4,3) = \begin{bmatrix} w_{15} & w_{16} & w_{17} \\ w_{25} & w_{26} & w_{27} \\ w_{35} & w_{36} & w_{37} \\ w_{45} & w_{46} & w_{47} \end{bmatrix} \quad B(3,2) = \begin{bmatrix} w_{58} & w_{59} \\ w_{68} & w_{69} \\ w_{78} & w_{79} \end{bmatrix} \quad \text{Ecuación 1.6}$$



Los pesos son los únicos elementos variables dentro de la red. Por esta razón, para una estructura dada, el comportamiento de la red depende estrictamente de los pesos que ésta posee.

White (1991) mostró en una serie de estudios que las redes neuronales artificiales que poseen capa oculta son equivalentes a modelos no lineales. Además, este tipo de redes tiene la habilidad de aproximar bastante bien cualquier función continua, por esta razón han sido llamadas aproximadores universales, Hornik *et al.*, (1989).

## 2.3. El aprendizaje en redes neuronales

### 2.3.1. Principales características del aprendizaje

Dada la estructura elegida por el usuario de la red, la intención del aprendizaje es “calibrar” la red obteniendo pesos cuyos valores transformen los datos de entrada en los datos de salida deseados.<sup>2</sup> En entrenamientos supervisados, ésta transformación es aprendida mediante la repetitiva presentación de una serie de ejemplos o patterns, cada uno de ellos compuesto por un vector de entrada y un vector de salida. Este tipo de entrenamiento es supervisado en el sentido que la salida de la red es comparada con un objetivo conocido de forma de definir un error y modificar los pesos existentes para lograr un mejor desempeño. Más adelante se describe un tipo de entrenamiento no supervisado (entrenamiento cruzado), y se compara con el entrenamiento supervisado.

La serie de ejemplos puede ser representada por un conjunto de observaciones de una variable en diferente periodo de tiempo (series de tiempo), o por una serie de observaciones en diferentes unidades al mismo tiempo (cross-section, Tally).

El método más usado para entrenar redes neuronales es el Backpropagation, Rumelhart y McClelland (1986). Este método permite a la red elegir sus pesos de forma de minimizar una función desempeño definida sobre la salida de la red y sus objetivos. La

---

<sup>2</sup> Más adelante se profundizará en la selección de estructuras o arquitecturas de una RNA.

función desempeño generalmente usada es el cuadrado de la suma de los errores, donde esta suma es tomada con respecto al número de patterns de entrenamiento y el número de neuronas de salida.<sup>3</sup> Puede ser necesaria otras funciones objetivo para problemas específicos, por ejemplo cuando el objetivo está compuesto por una secuencia binaria de 0 y 1. El error para pattern  $t$  y la neurona de salida  $j$  es la diferencia entre el objetivo  $T_{tj}$  y la salida de la red  $O_{tj}$ . De ahora en adelante  $W$  es el vector que contiene los pesos de la red.

### 2.3.2. El algoritmo Backpropagation

Considere un problema con  $T$  patterns, y una estructura de red que tiene  $K$  neuronas de entrada,  $H$  ocultas y  $J$  de salida (sólo una capa oculta). La función objetivo del problema, tal como fue definido en la sección anterior, es:

$$E(W) = \frac{1}{2} \sum_{t=1}^T \sum_{j=1}^J (T_{tj} - O_{tj})^2 \quad \text{Ecuación 1.7}$$

Donde la multiplicación por  $\frac{1}{2}$  es introducida sólo por conveniencia.

Las etapas del algoritmo son:

1. El vector  $W$  de pesos es iniciado con valores generados desde una distribución normal cuyo rango depende del usuario, pero generalmente es entre  $-0,5$  y  $+0,5$ .
2. El pattern  $t$  es presentado a la red. Este pattern es propagado hacia delante (feed-forward) en la red de la siguiente manera:

---

<sup>3</sup> Se le llama pattern a cada ejemplo de la serie con la cual se entrena a la red en forma supervisada.

$$net_{th} = w_{0h} + I_{t1} \cdot w_{1h} + I_{t2} \cdot w_{2h} + \dots + I_{tK} \cdot w_{Kh}$$

$$h = 1, 2, \dots, H$$

Ecuación 1.8

$$H_{th} = (1 + \exp(-net_{th}))^{-1}$$

$$h = 1, 2, \dots, H$$

Ecuación 1.9

$$net_{tj} = w_{0j} + H_{t1} \cdot w_{1j} + H_{t2} \cdot w_{2j} + \dots + H_{tH} \cdot w_{Hj}$$

$$j = 1, 2, \dots, J$$

Ecuación 1.10

$$O_{tj} = (1 + \exp(-net_{tj}))^{-1}$$

$$j = 1, 2, \dots, J$$

Ecuación 1.11

La Ecuación 1.8 describe como las señales exógenas  $I_t$  son transmitidas a las neuronas ocultas. La Ecuación 1.9 describe la transformación no lineal realizada por las neuronas ocultas mediante la función logística. La Ecuación 1.10 muestra como la red computa las combinaciones lineales que salen de las neuronas ocultas, las cuales son transmitidas a las neuronas de salida. La Ecuación 1.11 muestra la transformación realizada por cada neurona de salida a la señal que recibe.

- Una serie de  $J$  errores ( $T_{tj} - O_{tj}$ ) es usada para calcular  $E_t(W)$ , un error global para el pattern  $t$ :

$$E_t(W) = \frac{1}{2} \sum_{j=1}^J (T_{tj} - O_{tj})^2$$

Ecuación 1.12

Para hacer la notación más simple, de ahora en adelante ignoraremos la dependencia de los errores del vector de pesos  $W$  y expresaremos  $E_t(W)$  como  $E_t$ .

El término de error es usado para modificar los pesos que conectan las capas.

4. La regla general para la modificación de pesos es:

$$\Delta_t w_{ij} = -\alpha \frac{\partial E_t}{\partial w_{ij}} \quad \text{Ecuación 1.13}$$

Donde  $\alpha$  es la tasa de aprendizaje, un coeficiente que regula la velocidad de aprendizaje.

5.

a) Cuando la regla general entregada en la Ecuación 1.13 es usada para modificar los pesos que conectan las capas ocultas y de salida, el error global es derivado con respecto a los pesos que conectan la capa oculta y de salida. Las derivadas pueden ser expresadas en forma general mediante la notación:

$$\frac{\partial E_t}{\partial w_{ij}} = \frac{\partial E_t}{\partial O_{ij}} \cdot \frac{\partial O_{ij}}{\partial w_{ij}} \quad \text{Ecuación 1.14}$$

$$i = 0, 1, \dots, H$$

$$j = 1, 2, \dots, J$$

El índice  $i$  parte en 0 ya que considera la presencia de un bias (y su peso asociado) para cada neurona de salida. Ahora consideremos los dos factores de la parte derecha de la Ecuación 1.14. De la Ecuación 1.12, el primer factor es igual a:

$$\frac{\partial E_t}{\partial O_{ij}} = -(T_{ij} - O_{ij}) \quad \text{Ecuación 1.15}$$

De la Ecuación 1.10 y 1.11 el segundo factor es igual a:

$$\frac{\partial O_{ij}}{\partial w_{ij}} = \frac{\partial O_{ij}}{\partial net_{ij}} \cdot \frac{\partial net_{ij}}{\partial w_{ij}} = O_{ij}(1 - O_{ij})H_{ij} \quad \text{Ecuación 1.16}$$

Esta expresión particularmente simple para segundo factor es debido al uso de la función logística.

- b) Cuando la regla general dada la Ecuación 1.13 es usada para modificar los pesos que conectan las capas de entrada y oculta, el error global es derivado con respecto a los pesos que conectan las capas de entrada y oculta. Las derivadas pueden ser expresadas en forma general mediante la notación:

$$\frac{\partial E_t}{\partial w_{ij}} = \frac{\partial E_t}{\partial net_{ij}} \cdot \frac{\partial net_{ij}}{\partial w_{ij}}$$

Ecuación 1.17

$$i = 0, 1, \dots, K$$

$$j = 1, 2, \dots, H$$

El índice  $i$  parte en 0 para considerar la presencia de un bias (y sus pesos asociados) para cada neurona oculta. Ahora consideramos los dos factores de la parte derecha de la Ecuación 1.17. De la Ecuación 1.9, el primer factor es igual a:

$$\frac{\partial E_t}{\partial net_{ij}} = H_{ij}(1 - H_{ij})$$

Ecuación 1.18

El segundo factor es igual a:

$$\frac{\partial net_{ij}}{\partial w_{ij}} = I_{ij} \sum_{k=1}^J (T_{tk} - O_{tk}) O_{tk} (1 - O_{tk}) w_{jk}$$

Ecuación 1.19

6. El ciclo es repetido desde el paso (2) usando un nuevo pattern hasta que todos los patterns han sido examinados por la red (índice  $t$  va desde 1 hasta  $T$ ).
7. Los cuadrados de los  $T$  errores son sumados de forma de obtener un error global sobre todos los patterns.
8. Los pasos 2 a 7 son repetidos hasta que el error global alcance un valor especificado por el usuario.

### 2.3.3. Comentarios generales acerca del Backpropagation

#### a) Tasa de aprendizaje

La Ecuación 1.13 es derivada de un algoritmo general de gradiente descendente. En general el coeficiente que transforma la información de la primera derivada a cambios en los parámetros estimados, varía con el tiempo, aunque la tasa de aprendizaje  $\alpha$  sea constante. Generalmente este coeficiente tiene un valor pequeño para asegurar que la red alcance una solución. Un valor pequeño aumenta el número de iteraciones necesarias para obtener una solución. La tasa de aprendizaje también puede ser modificada durante el aprendizaje. White (1991) muestra que disminuir  $\alpha$  a una tasa determinada es necesario para asegurar la convergencia del algoritmo.

#### b) Momentum

Frecuentemente la Ecuación 1. 13 es modificada a:

$$\Delta_t w_{ij} = -\alpha \frac{\partial E_t}{\partial w_{ij}} + \beta \cdot \Delta_{t-1} w_{ij} \quad \text{Ecuación 1.20}$$

Donde  $\beta$  es el momentum.<sup>4</sup> Con este cambio se introduce un mayor grado de persistencia en la modificación de los pesos, ya que cada cambio depende del cambio anterior. Cuando incluimos este término, el valor de  $\alpha$  puede ser más alto para tener una convergencia más rápida, ya que  $\beta$  permitirá mayor estabilidad en el proceso de búsqueda. Incluir el momentum tiene la utilidad de evitar oscilaciones excesivas de los pesos, y puede ser una aproximación a un gradiente conjugado más general, Battiti (1992).

---

<sup>4</sup> Momentum es un término utilizado en física para describir la masa en movimiento de un objeto. A mayor velocidad o masa del objeto en movimiento, mayor es su momentum y más energía se necesita para detenerlo. A  $\beta$  se le llama momentum ya que es un factor que involucra la iteración anterior del algoritmo, por lo que entrega una mayor “inercia” a la convergencia de  $w_{ij}$ . Podemos observar en la Ecuación 1.20 que  $w_{ij}$  tenderá a moverse en la misma dirección, a menos que esa dirección sea opuesta a la dirección del término gradiente.

c) *Backpropagation y otros algoritmos*

El Backpropagation puede ser considerado como un caso especial dentro de un grupo de algoritmos generales usados en problemas no lineales, como son Newton-Raphson o Berndt. Estos algoritmos son más complicados desde un punto de vista computacional debido a que usan muchos pesos y podrían tener una convergencia muy lenta; por esta razón estos algoritmos no tienen un uso práctico en problemas basados en redes neuronales.

d) *Mínimos locales*

El Backpropagation es una técnica que realiza una búsqueda local en forma determinística, basada en un gradiente descendente, y por lo mismo el algoritmo podría terminar en un mínimo local de la función de error. Una red neuronal que entra en mínimos locales durante su aprendizaje mostrará un mal desempeño en términos de aprendizaje y de sus capacidades de generalización, ya que su entrenamiento no es el óptimo. Como las redes neuronales son generalmente funciones no lineales y muy complicadas de los pesos, es natural que muchos mínimos locales coexistan. Existen diferentes reglas empíricas para escapar de un mínimo local:

- Incluir ruido aleatorio en los patterns;
- Reducir la tasa de aprendizaje;
- Usar momentum;
- Agregar más neuronas ocultas;
- Cambiar los valores iniciales de los pesos;

Otra alternativa es usar algoritmos genéticos para entrenar la red, pero esto se escapa del alcance de esta tesis.<sup>5</sup>

---

<sup>5</sup> Un algoritmo genético imita el proceso evolutivo que tienen los genes mediante sucesivos cruces entre soluciones (genes) que evolucionan a una solución final.

*e) Condiciones iniciales*

El punto anterior sugiere la importancia de las condiciones iniciales. Generalmente los investigadores prueban la robustez de una solución iniciando el algoritmo con distintos pesos. Así ven si la solución cambia considerablemente en función de los pesos iniciales del algoritmo. También es posible usar la salida del algoritmo Backpropagation como valor inicial para una aplicación del modelo no lineal de mínimos cuadrados Newton-Raphson, White (1992).

*f) Umbral de las neuronas y bias*

Podría considerarse que el umbral  $\theta$  de las diferentes neuronas es un parámetro, sin embargo este valor es producto del entrenamiento de la red. Para cada neurona oculta y de salida, el umbral es representado por el valor del peso que conecta la entrada del bias,  $I_0$  (generalmente su valor es uno), con la neurona: de hecho este valor es sumado a la entrada de la neurona (Figura 1.1, Ecuaciones 1.1 y 1.3). La iniciación aleatoria de pesos y el posterior aprendizaje de los mismos incluyen a los pesos que conectan los bias.

*g) Aprendizaje en línea y fuera de línea*

Los autores del Backpropagation distinguen entre un aprendizaje en el cual los pesos son cambiados después de la presentación de cada pattern (entrenamiento en línea), y un aprendizaje en el cual los pesos son cambiados luego de la presentación de todos los patterns de la serie de entrenamiento.

Ambos tipos de aprendizaje tienen sus ventajas y desventajas (ver ejemplos en Battiti, 1992) Generalmente, el mismo factor es una ventaja y desventaja al mismo tiempo. Por ejemplo, el aprendizaje en línea posee una aleatoriedad que puede ser útil para escapar de mínimos locales, pero por la misma razón se podría escapar de un buen mínimo local que merecía ser explorado. En esta tesis se entrenó las redes mediante un aprendizaje en línea.



### *h) Calibrar parámetros*

Los parámetros principales del algoritmo Backpropagation son: la tasa de aprendizaje y el momentum (o su tasa de cambio, si es que varían durante el aprendizaje), y la distribución de los pesos aleatorios iniciales. El valor de estos parámetros debe ser determinado en forma empírica, ya que no existen reglas para encontrar sus valores óptimos. De hecho este es un proceso que requiere bastante tiempo, ya que una buena serie de parámetros para un problema específico sólo puede ser encontrada luego de varias pruebas.

### *i) Selección de arquitecturas*

La selección de una arquitectura no es un problema simple. Suazo (2004) ) creó un algoritmo que selecciona una arquitectura para una red feed-forward de una capa oculta dado un problema específico. Si bien el algoritmo de Suazo no considera la posibilidad de una red con más de una capa oculta, existe evidencia que muestra que redes feed-forward de una capa oculta son aproximadores universales (por ejemplo Hornik *et al.*, 1989 y 1990). En general, mientras mayor es el número de neuronas ocultas, más hábil es la red para resolver problemas específicos mediante una serie de ejemplos. Sin embargo, si existe mucho ruido entre los datos, las neuronas extras se ajustarán a este ruido y por esta razón la red no entregará datos correctos fuera de los ejemplos. Así, puede existir un problema de sobre ajuste para redes complejas vs. sub ajuste para redes muy simples. White (1988) entrega tasas de crecimiento para la complejidad de una red que indican los peligros de un sobreajuste y un subajuste a la serie de entrenamiento.

### *j) Interpretación de los pesos*

En general es imposible asignar una interpretación especificada a cada peso de la red; a veces es útil realizar una inferencia estadística a una serie de pesos, por ejemplo realizar un test de no linealidad (ver por ejemplo Lee *et al.*, 1989) o estudiar causalidades.

*k) Criterios de evaluación*

El desempeño de la red después de haber sido entrenada se determina mediante un test basado en algún indicador estadístico como por ejemplo el coeficiente de determinación entre los objetivos y las salidas de la red. Esta evaluación es realizada de dos formas: basada en los datos usados durante el entrenamiento para tener una idea de la eficacia del aprendizaje, o en datos que no pertenecen a la serie de entrenamiento para ver la capacidad de generalización de la red.

*l) Representación del conocimiento*

El conocimiento adquirido durante entrenamiento es distribuido en el sentido que cada pattern crea una actividad sobre todas las neuronas. De esta forma cada neurona está involucrada en la representación de muchos conceptos. Toda salida de la red depende de todas y cada una de las neuronas de la misma. Esta característica entrega robustez a la red.

## 2.4. La técnica de entrenamiento cruzado

El entrenamiento cruzado en RNA es un tipo de aprendizaje no supervisado creado por Farren (2004).

En este trabajo Farren desarrolló una técnica de entrenamiento cruzado que considera la consistencia de planes de cada agente, esto es, la relación entre efecto deseado y efecto real. Así, cada agente aprende a coordinar su accionar con el efecto real que este accionar tendrá y por lo tanto, aprende a planificar.

Para desarrollar este modelo se estableció la siguiente hipótesis general: un agente, actuando en un ambiente económico, debe desarrollar y adaptar su capacidad de evaluar, en forma coherente:

1. Lo que tiene que hacer para obtener un resultado específico.
2. Pronósticos sobre las consecuencias de sus actos.

Esta hipótesis se mantiene para el caso en que el agente interactúa con otros agentes.

Para un observador externo, los agentes aparentemente operan sobre la base de objetivos y planes. Obviamente, los agentes no poseen la capacidad de producir tales decisiones, las que son invenciones del observador. Es interesante destacar que en el mundo real es posible ver la misma situación. Si observamos ciertas decisiones de seres humanos en el mundo real, podríamos deducir la existencia de un metódico plan cuando en realidad las decisiones fueron tomadas en función de, por ejemplo, emociones.

El nombre aprendizaje cruzado, representa la técnica usada para entrenar la red neuronal que representa las decisiones tomadas por cada agente manteniendo una consistencia interna. Dada la hipótesis general, el comportamiento de cada agente produce resultados que pueden ser separados en dos tipos:

1. Actos a realizar.
2. Pronóstico de los efectos de esos actos.

Los objetivos necesarios para entrenar las salidas de la red que representan los actos a realizar y los necesarios para entrenar las salidas de la red que representan el pronóstico de esos actos están contruidos en forma cruzada. Los objetivos de los actos a realizar están contruidos en función de los pronósticos de los efectos de los actos, de forma de desarrollar la capacidad de decidir acciones que produzcan efectos cercanos a los resultados esperados. A su vez, los objetivos de los efectos pronosticados por el agente están contruidos en función de los actos realizados por el agente, de forma de desarrollar la capacidad de estimar los efectos de los actos que el mismo agente está decidiendo. La base de este método de aprendizaje es aprender haciendo, es decir, si bien el algoritmo de aprendizaje es Backpropagation, es sobre la base de aprender y actuar; es necesaria la acción para producir la información necesaria para entrenar la red que simula al agente. Un entrenamiento tradicional no sería útil en este caso, ya que las reglas que relacionan las

entradas y las salidas de la red tienen que ser “descubiertas” por medio de experimentos realizados por la misma red.

El aprendizaje es realizado mediante cuatro pasos por iteración. Si consideramos cada iteración como un día en la vida de nuestros agentes, cada agente tiene una experiencia de negociación diaria, de la cual aprende en cuatro pasos que estiman la salida de la red y propaga el error cuadrático medio a través de la misma, determinando los pesos de la red. Los pasos de aprendizaje son:

1. Salidas de la red: Acciones a realizar y efectos de estas acciones
2. Objetivos para los efectos: Aparecen de las acciones a realizar. De esta forma los efectos esperados de los actos realizados son cada vez más parecidos a los efectos reales de los actos realizados.
3. Objetivos para los actos: El error encontrado entre los objetivos para los efectos y los efectos pronosticados por la red en (2) son el punto de partida para determinar los objetivos de los actos. Así, los actos evolucionan en el sentido de los efectos pronosticados. Este error es compartido en forma aleatoria por los diversos actos.
4. Backpropagation: Los pesos de la red se corrigen de forma de minimizar los errores y así conseguir pronósticos de actos más cercanos a los efectos reales de los actos y actuar de forma más coherente con los pronósticos esperados de esos actos.

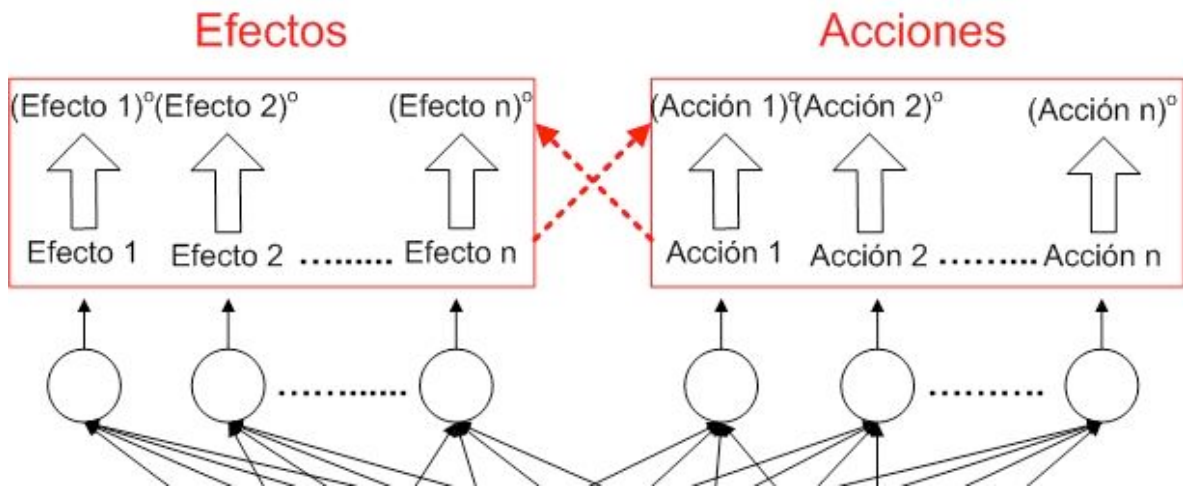


Figura 1.4

Formalmente, el entrenamiento cruzado puede ser genéricamente explicado introduciendo el efecto  $E$  que aparece de realizar dos acciones  $A_1$  y  $A_2$ .  $E$ ,  $A_1$  y  $A_2$  son salidas de la red. El objetivo para el efecto es:

$$E^o = f(A_1, A_2) \tag{Ecuación 1.21}$$

donde  $f$  es una función que representa el efecto real de las acciones  $A_1$  y  $A_2$ . La intención es encontrar una salida  $E$  (efecto pronosticado) lo más cercana posible a  $E^o$ , que es el efecto real. El error de  $E$  es:

$$e = E^o - E \tag{Ecuación 1.22}$$

Para minimizar este error, es propagado a través de toda la red. La intención ahora es encontrar los actos a realizar que sean más consistentes con los efectos pronosticados, es decir, encontrar los objetivos para los actos ( $A_1^o$ ,  $A_2^o$ ) que son coherentes con el efecto entregado por la red  $E$  para así corregir los valores  $A_1$  y  $A_2$ . Para esto definimos las funciones  $g_1$  y  $g_2$  que representan las funciones inversas de  $f$  para cada una de las variables  $A_1$  y  $A_2$ .

$$\begin{aligned} A_1 &= g_1(E, A_2) \\ A_2 &= g_2(E, A_1) \end{aligned} \tag{Ecuación 1.23}$$

Definiendo la variable aleatoria  $\phi_1 \sim U(0,1)$  (proveniente de una distribución uniforme en el intervalo  $[0,1]$ ), y  $\phi_2 = 1 - \phi_1$ , tenemos que:

$$A_1^o = g_1(E^o - e \cdot \phi_1, A_2)$$

$$A_2^o = g_2(E^o - e \cdot \phi_2, A_1)$$

Ecuación 1.24

Los errores de  $A_1$  y  $A_2$  son:

$$a_1 = A_1^o - A_1$$

$$a_2 = A_2^o - A_2$$

Ecuación 1.25

La idea es minimizar los errores  $e$ ,  $a_1$ ,  $a_2$  mediante Backpropagation y representar así a un agente internamente coherente.

## Capítulo 3:

# SIMULACIÓN Y LAS CIENCIAS SOCIALES

## 3.1. Simulación vs. Modelos Estadísticos

La simulación es un tipo particular de modelación. Un modelo es una simplificación (más pequeña, menos detallada, menos compleja, etc.) de otra estructura o sistema. Hasta ahora los modelos más relevantes de las ciencias sociales son modelos estadísticos. La construcción de los modelos de simulación y los modelos estadísticos son bastante similares. Ambos modelos abstraen la realidad, los modelos de simulación lo hacen mediante un programa computacional, mientras los modelos estadísticos crean una serie de ecuaciones en función de algunos parámetros (por ejemplo, coeficientes beta). Ambos tipos de modelos deben ser “calibrados” usando datos recolectados del sistema. Los modelos de simulación usan estos datos para determinar distribuciones, mientras que los modelos estadísticos usan estos datos para encontrar los parámetros. Una vez que los modelos han sido creados, es necesario validarlos. Ambos tipos de modelos son validados usando datos del sistema real.

Aunque pareciera haber fuertes similitudes entre los modelos de simulación y los modelos estadísticos, existen diferencias importantes. Los modelos de simulación están orientados al proceso, mientras que los modelos estadísticos intentan explicar correlaciones entre variables medidas en un instante de tiempo. Por esta razón, es de esperar que un modelo de simulación posea una representación explícita, *ex ante*, de los procesos que trabajan en el mundo social mientras que un modelo estadístico solo reproduce, *ex post*, un esquema de correlaciones entre las variables que fueron consideradas y que generalmente no modelan los mecanismos que condicionan esas relaciones.

Ambos tipos de modelos pueden ser usados tanto para explicar como para pronosticar fenómenos. Sin embargo, aunque algunos modelos tanto estadísticos como de simulación enfatizan la necesidad de hacer pronósticos, todas las simulaciones deben satisfacer ambos criterios: un buen modelo pronosticador contribuirá, en cierta medida, al entendimiento del sistema mientras que un modelo explicativo siempre será capaz de hacer pronósticos, incluso aunque éstos no sean muy precisos.

Es interesante destacar que la finalidad de la ciencia es explicar fenómenos, no pronosticarlos necesariamente. Los terremotos, la aparición de una estrella en el cielo, los fenómenos sociales, etc., son fenómenos no pronosticables aún, sin embargo sería falso afirmar que los geólogos, los astrónomos o los psicólogos no son científicos.

## 3.2. Simulando Sociedades Humanas

Las técnicas utilizadas actualmente para simular sistemas sociales fueron originalmente desarrolladas en campos como la física y la inteligencia artificial. Si bien el estudio de las ciencias sociales es distinto al de las ciencias naturales en el sentido que modelar sociedades está enfocado en temas muy distintos que los temas de interés de un modelo de, por ejemplo, el movimiento de un conjunto de partículas, estas técnicas han probado ser exitosas. Además, existen características específicas de las ciencias sociales que solo pueden ser estudiadas usando Simulaciones Basadas en Agentes.

La Simulación Basada en Agentes difiere a la Simulación tradicional en que la entidad básica del modelo es el agente, y no el proceso como lo es en la Simulación Discreta. Es decir, la Simulación Basada en Agentes intenta modelar comportamientos específicos de los individuos, mientras que la Simulación tradicional modela características promediadas de la población y el modelo simula cambios en estas características promediadas para toda la población. Como consecuencia, la Simulación Basada en Agentes es muy útil en situaciones en las que se tiene un gran número de individuos heterogéneos que se comportan de diferente manera.



Uno de los temas de investigación en la simulación social es que incluso cuando los agentes son programados con reglas muy simples, el comportamiento de estos agentes considerados como un todo puede resultar extremadamente complejo.

Prácticamente todos los métodos estadísticos convencionales para el análisis de sistemas sociales están basados sobre la suposición de una relación lineal entre las variables. Es decir, la variable dependiente es proporcional a la suma de una serie de variables independientes. Esta suposición es bastante restrictiva. Un nuevo campo interdisciplinario llamado teoría de la complejidad, Waldrop (1992) y Kauffman (1995), estudia sistemas no lineales.

### 3.3. Sistemas Complejos

El sonido es un sistema lineal, por eso podemos escuchar una voz sonando sobre una guitarra y reconocer a ambos. Aunque las ondas de sonido se mezclen, ambas mantienen sus identidades. La luz también es un sistema lineal, por eso es posible ver luz del semáforo incluso en días luminosos. Los rayos de luz operan en forma independiente, pasando unos a través de otros como si nada hubiera ahí. En cierto sentido, incluso la economía es un sistema lineal ya que todos los agentes actúan en forma independiente. Cuando un individuo compra el diario no tiene efecto sobre la decisión de otro individuo en sus compras del supermercado.

Sin embargo, muchos sistemas naturales son no lineales. Sin duda que el funcionamiento de nuestro cerebro no es lineal, aunque el sonido de una voz y el de una guitarra sean independientes cuando entran al oído, el impacto emocional que ambos sonidos tienen en la persona es mucho mayor que cada uno de ellos por separado. La economía tampoco es realmente lineal. Las millones de decisiones tomadas por los agentes sobre comprar o no comprar pueden reforzarse unas a otras, creando un boom o una

recesión económica. Este clima económico puede retroalimentar el sistema alterando las decisiones de compra.<sup>6</sup>

De hecho, exceptuando los fenómenos físicos más simples, casi todo y todos en el mundo es parte de un sistema no lineal. Un cambio insignificante en un lugar puede causar un efecto abrumador en otro. El todo es casi siempre bastante mayor a la suma de sus partes. El movimiento de las alas de una mariposa puede generar una tormenta en otro lugar (efecto mariposa).

Estos sistemas son llamados Sistemas Complejos. Desde el punto de vista de un científico o matemático, estos sistemas son muy difíciles de estudiar, ya que no les puede entender en forma analítica. Generalmente no existe un sistema de ecuaciones que pueda explicar o hacer pronósticos del sistema. La única forma efectiva de estudiar comportamientos no lineales es creando un modelo de Simulación Basada en Agentes del sistema en un computador y analizar su comportamiento. Incluso cuando el científico puede entender el sistema, éste es impredecible. Por mucho que se estudie las propiedades de la economía o del cerebro, será imposible pronosticar una recesión económica o la emoción que produciría una pieza musical.

Erróneamente, la filosofía convencional de las ciencias sociales generalmente hace una conexión entre explicar y pronosticar un fenómeno. Se tiende a asumir que la prueba sobre la validez de una teoría es que ésta haga pronósticos exitosos. Este criterio no es apropiado para teorías no lineales. La teoría de la Complejidad nos muestra que aunque entendamos completamente los factores que afectan al actuar de un individuo, no sería suficiente para pronosticar el comportamiento de un grupo de esos individuos.

---

<sup>6</sup> Recordemos la economía chilena post crisis asiática, cuando nadie quería invertir aunque ya no hubiera razones para desconfiar de la estabilidad de la economía.

### 3.3.1. Fenómenos Emergentes

El concepto de Fenómenos Emergentes es una de las ideas más importantes que nacen a partir de la teoría de la Complejidad. Los Fenómenos Emergentes ocurren cuando interacciones entre objetos de un nivel dan origen a diferentes tipos de objetos en otro nivel. Es decir, un fenómeno es emergente si requiere en su descripción conceptos que no son requeridos para describir el comportamiento de las componentes que lo subyacen. Por ejemplo, la temperatura es una propiedad emergente del movimiento de átomos. Sin embargo un átomo no tiene temperatura, un conjunto de ellos sí. Las emociones son propiedades emergentes de la interacción entre neuronas aunque las neuronas no tienen emociones.

La Simulación Basada en Agentes es una herramienta ideal para estudiar fenómenos emergentes en un sistema complejo.

## 3.4. Validación de Simuladores Basados en Agentes

Un Simulador Basado en Agentes puede llegar a ser tan complejo que no tendría sentido buscarle validez mediante métodos usados en Simulación Discreta (generalmente intervalos de confianza para la diferencia entre medias de las distribuciones real y generada). Un Simulador Basado en Agentes intenta llegar hasta el fondo del sistema en su proceso de modelación, no considera “sub-sistemas”.<sup>7</sup> La Simulación Discreta profundiza en la modelación del sistema solo hasta el punto en el cual el nivel de complejidad es muy alto, en este punto crea “sub-sistemas” y recoge datos de ellos para asignarle comportamientos mediante una distribución de probabilidad.

---

<sup>7</sup> Por esta razón los Simuladores Basados en Agentes son llamados Simuladores “Bottom-up”.

Es claro entonces que una desventaja de los Simuladores Basados en Agentes es que, dada su alta complejidad, generalmente no es posible modelar un sistema relativamente complejo con un nivel de validez similar al de otros tipos de Simuladores. Sin embargo, recordemos que los Simuladores deben ser desarrollados con un propósito o aplicación específica, y su validez es determinada con respecto a ese propósito, Sargent (1988). Este concepto tiene especial importancia en Simuladores Basados en Agentes.

## Capítulo 4:

# HAYEK Y SU TEORÍA DE EQUILIBRIO ECONÓMICO

## 4.1. Implicancias del equilibrio

*“La razón es la facultad mediante la cual el hombre identifica e integra el material entregado por sus sentidos. La razón integra las percepciones del hombre formando abstracciones o conceptos, aumentando así el conocimiento del hombre desde un nivel perceptual, que comparte con los animales, a un nivel conceptual, que solo él puede alcanzar. El método utilizado por la razón en este proceso es la lógica, y la lógica es el arte de la identificación no contradictoria” (Rand, 1984).*

Toda teoría se basa en abstracciones de la realidad. Las abstracciones muchas veces simplifican el problema en estudio mediante suposiciones. La pregunta es: ¿Qué son abstracciones permisibles y qué abstracciones llevan a una teoría no válida? ¿Cuándo las suposiciones implícitas en una abstracción son tales que el resultado no representa la realidad?

Un ejemplo. Las leyes de movimiento de Newton formaron una teoría válida hasta hoy. Sin embargo, posteriormente Einstein con su teoría de la relatividad entregó leyes de movimiento más generales que son válidas para cualquier tipo de partícula y a cualquier velocidad. Esto no significa que Newton esté equivocado, de hecho sus ecuaciones están contenidas en las leyes de Einstein. Cuando la velocidad de la partícula es baja (mucho menor que la velocidad de la luz), la diferencia entre las leyes de movimiento de Einstein y las de Newton son pequeñas. Newton entrega una excelente aproximación a un muy bajo costo de cálculo. Por esta razón, la ley de la relatividad no tiene un factor importante en la vida diaria. Para velocidades altas, las leyes de Newton no son válidas, y las de Einstein son

usadas. Podríamos decir que las leyes de Newton son un caso particular dentro de las leyes de Einstein, con más suposiciones implícitas en sus abstracciones.

En este ejemplo, claramente Newton supuso implícitamente que la partícula que él estudiaba se movería a velocidad baja. Esta suposición llevó a abstracciones erradas en sentido estricto, pero, ¿es por eso una teoría no válida? No. Un modelo debe ser desarrollado para un propósito específico (o aplicación), y su validez debe ser determinada con respecto a ese propósito, Sargent (1988). Las leyes de Newton cumplen su propósito, y si bien hay suposiciones implícitas en ellas, esta ley representa la realidad.

Para el caso del análisis de equilibrio puro, esta teoría considera observar las consecuencias de cualquier cambio económico mientras se imaginan todos los demás cambios relevantes como ausentes.<sup>8</sup> De esta forma, se puede crear una idea general de causa y efecto. El término “Ceteris Paribus” no es nuevo para nadie que tenga un conocimiento básico de economía, éste significa mantener todas las variables que no están en estudio como constantes. No existe duda de que el análisis de equilibrio puro es de gran utilidad para el aprendizaje de la economía, ya que presenta una visión muy básica del sistema económico. Por esta razón, y siguiendo el proceso de validez de Sargent, las teorías de análisis de equilibrio puro son válidas dentro de aplicaciones destinadas al aprendizaje de la economía.

Ahora, en un sentido más amplio, esta teoría no es válida para representar la realidad. Cualquier teoría económica que no considere el pasar del tiempo y sus efectos en el accionar humano y en el cambio del conocimiento, no es muy útil para ayudarnos a entender procesos económicos.<sup>9</sup>

---

<sup>8</sup> Hayek llama análisis de equilibrio puro a la teoría de equilibrio económico imperante en ese tiempo, Hayek (1937).

<sup>9</sup> Como veremos más adelante, el conocimiento o “knowledge” es usado por Hayek como la información que un individuo posee, y el uso que le da a ésta.

## 4.2. Equilibrio como consistencia de planes

Friedrich August von Hayek definió el equilibrio como una situación en la cual “los diferentes planes de acción en el tiempo que tienen los individuos que componen una sociedad son mutuamente compatibles” (Hayek, 1937). Un aspecto importante de esta definición es que desligó al concepto de equilibrio con la dimensión física de un estado de descanso o de igualdad entre fuerzas a uno basado en la mente humana. El equilibrio es concebido entonces como una situación en la cual el conocimiento individual y las expectativas, y las acciones basadas en éstas, son compatibles con la información que el individuo posee, donde la información para un individuo incluye los actos de otros individuos. No tiene sentido definir el equilibrio económico sin considerar la percepción individual, ya que después de todo estamos trabajando con acciones humanas, y esas acciones están determinadas por la percepción. Entonces, en el caso de la oferta y demanda de un mercado único y bien definido, el precio no cambiará cuando todos los individuos estén realizando exitosamente sus planes (mutuamente relacionados) de comprar y vender, pero cuando uno o más de estos planes no sean realizados exitosamente podemos esperar que sean replanteados y veremos cambios en los precios.<sup>10</sup>

En resumen, el equilibrio es una situación en la cual los planes individuales están coordinados. Cada plan puede ser ejecutado exitosamente. Los medios calzan con el fin.

Hayek hizo una importante diferenciación entre el equilibrio individual y el equilibrio del sistema.<sup>11</sup>

*"Desde hace mucho tiempo siento que el concepto de equilibrio y los métodos que utilizamos en el análisis puro tienen un sentido claro sólo cuando son utilizados en el análisis del actuar de una sola persona y que pasamos a una esfera distinta e introducimos un nuevo elemento de carácter totalmente distinto cuando lo aplicamos*

---

<sup>10</sup> Es posible concebir una situación de equilibrio “estadístico”, donde los errores individuales se compensan y el precio no es alterado. En esta situación, aunque los planes no son compatibles, existe un equilibrio como igualdad de fuerzas. Los planes individuales son exitosos “en promedio” (Hayek, 1937).

<sup>11</sup> El término sistema es usado como un medio que posee a más de un individuo. Este puede ser un sistema económico completo o un subsistema de él (un mercado, por ejemplo).

*a la explicación de las interacciones de un número de individuos diferentes" (Hayek, 1937).*

Si tomamos a un individuo aislado, se puede hablar de equilibrio individual cuando las acciones del individuo están en equilibrio, es decir, éstas son parte de un plan. Vemos entonces que el concepto de equilibrio individual implica actuar en forma racional. De aquí se puede sacar dos conclusiones:

1. Dado que el éxito de todo plan depende en parte de la información a que el individuo tiene acceso, y uso que da a ésta, vemos que el conocimiento es de vital importancia en el equilibrio individual. El conocimiento es el entendimiento que de la realidad tiene un individuo. Esto considera: características psicológicas (inteligencia, emociones, gustos, educación, etc.) y acceso a información. Es muy probable que en algunos aspectos, el conocimiento de distintos individuos respecto de una misma circunstancia sea distinto. Es más, estos conocimientos pueden ser inconsistentes entre ellos.
2. Como las relaciones de equilibrio existen cuando las acciones sucesivas de una persona son parte del mismo plan, cualquier cambio en el conocimiento de la persona la llevaría a alterar su plan y a alterar las relaciones de equilibrio entre sus acciones tomadas antes del cambio de conocimiento y las acciones tomadas después. En otras palabras, las relaciones de equilibrio existen sólo durante el periodo en el cual el pronóstico de las acciones tomadas es correcto. Como el equilibrio es una relación entre acciones, y las acciones son ejecutadas sucesivamente en el tiempo, obviamente el paso del tiempo es esencial para dar al concepto de equilibrio algún significado.

### 4.3. La tendencia al equilibrio

Para Hayek, el equilibrio no es un estado que en la realidad existe, aunque bajo el concepto lógico existe. Él estaba más interesado en el estudio de una tendencia hacia el



equilibrio en términos de un mayor grado de coordinación de planes. El conocimiento y los planes de los distintos miembros de una sociedad serán cada vez más y más compatibles; en otras palabras, las expectativas de la gente y particularmente de los empresarios, serán más y más correctas. Así, vemos que la tendencia al equilibrio es fácilmente identificable en el mundo real, eliminando la abstracción de la definición inicial y entregando un significado claro al sentido común.

## Capítulo 5:

# EL MODELO MECHA

## 5.1. Introducción

MECHA es un modelo de alta complejidad en la interacción de los agentes, por lo que se usó la simulación basada en agentes para crearlo.<sup>12</sup>

El propósito del modelo es estudiar el crecimiento económico y su relación con la creación de empresas.

Este modelo tiene como unidad básica a cada individuo integrante del sistema económico. Cada individuo o agente posee una ubicación determinada en una grilla y tiene la capacidad de consumir y de producir uno o más bienes de un total de cuatro tipos de bienes que posee el mercado. La ubicación de los productores no cambia, sólo tratan de satisfacer la demanda de los consumidores que se mueven a través de la grilla. Por consecuencia, todo productor tiene una dirección determinada por dos coordenadas fijas y tiene además a ocho productores vecinos.

En un inicio, los agentes son productores independientes. Con el pasar del tiempo éstos podrían emplearse o formar su propia empresa. Los empresarios pueden contratar a los agentes vecinos. Los agentes tienen la capacidad de proponer precios, cantidades y el tipo de bien que desean consumir o producir.

MECHA consta de un mercado de bienes, un mercado del trabajo y una política de quiebra. Además, existen reglas en el mercado que serán descritas más adelante.

---

<sup>12</sup> MECHA es una sigla que reúne en una palabra las principales características del modelo, es decir, **M**odelo **E**conómico **H**ayekiano.

## 5.2. Los agentes

MECHA representa a cada agente como una unidad capaz de aprender en función de sus experiencias y la información a la cual tiene acceso. Los agentes determinan el precio, la cantidad y el tipo de bien a demandar u ofrecer en función del precio, cantidad y bien demandado u ofrecido en periodos anteriores y el éxito que con éstos tuvo.

Para modelar este sistema de aprendizaje se utiliza redes neuronales. Cada agente tiene una red neuronal que controla sus decisiones relacionadas con precios y cantidades a demandar, otra red que controla sus decisiones en cuanto a precios y cantidades a ofrecer y una última red que controla el tipo de bien a ofrecer.

Los productores son inicialmente unidades independientes que producen en función de su capacidad. Con el tiempo éstos pueden transformarse en empresarios y contratar a otros agentes, empleados, o bien, seguir independientes.

## 5.3. Precios y Cantidades a Ofrecer y Demandar

### 5.3.1. Introducción

Siguiendo la línea de sentido de individualismo y orden económico propuesto por Hayek (1949), en esta tesis se modela un mercado formado por agentes que se comportan en forma estrictamente individualista, sin las simplificaciones de un mercado tipo “remate”, en el cual los agentes simplemente ignoran el pasar del tiempo y su efecto en su accionar (toda acción lleva implícita una reacción, que es pronosticada por el individuo). Bajo el enfoque de Hayek, todo plan puede ser realizado sólo si el agente puede realizar pronósticos sobre el efecto que tendrán sus acciones. Recordemos que un plan es un conjunto de acciones a seguir para lograr un efecto deseado, Hayek (1937). El plan puede ser más o menos exitoso, dependiendo de la diferencia entre el efecto deseado y el efecto real de las acciones del individuo.

Según Hayek (1937), en la medida que los agentes aprendan a realizar planes satisfactorios, el sistema tenderá a un equilibrio. Recordemos que un plan es satisfactorio sólo si es consistente a todos los demás planes de los otros individuos del sistema.

MECHA se basa en las ideas de Hayek. Por esta razón se decidió crear un sub-modelo explicativo que nos haga más claras estas ideas.

Para modelar la toma de decisiones de los agentes se usó redes neuronales feed-forward, con una capa oculta y un número arbitrario de neuronas ocultas.<sup>13</sup> La red es entrenada con la técnica de entrenamiento cruzado. Las salidas de la parte derecha de las redes entregarán las decisiones de precio y cantidad a consumir u ofrecer (acciones). En la parte izquierda, las redes entregarán el pronóstico hecho por la misma red acerca del gasto (efecto). Para que la red realice pronósticos válidos, la relación  $\frac{acción}{efecto}$  dada por  $\frac{precio \cdot cantidad}{gasto}$  debe ser igual a uno.

La idea es ver qué propiedades emergen del modelo. Según Hayek un modelo de este tipo debería llegar a un equilibrio. No es posible anticipar el punto de equilibrio en este modelo mediante un estudio del mismo.

### 5.3.2. Estructura de la red neuronal usada

A continuación, se entrega una migración paso a paso desde la teoría de Hayek sobre la consistencia de planes, Hayek (1937), a un modelo de red neuronal con entrenamiento cruzado.

- a) El primer paso consiste en modelar la toma de decisiones de un individuo aislado siguiendo la idea de equilibrio individual entregada por Hayek.

---

<sup>13</sup> La arquitectura óptima para una red que simula el comportamiento humano es un problema interesante, pero se escapa del alcance de esta tesis.

- b) El segundo paso consiste en insertar al individuo ya creado en un sistema siguiendo la idea de equilibrio global entregada por Hayek.

## a) Modelo de las decisiones tomadas por un individuo aislado

Volviendo al concepto de equilibrio económico entregado por Hayek, específicamente a su idea de equilibrio individual, podemos crear una red neuronal feed-forward que modele las decisiones a tomar por un individuo aislado que busca un equilibrio. Según Hayek (1937) todo individuo tenderá al equilibrio en la medida que pueda realizar planes exitosos, o sea, que pueda pronosticar correctamente el efecto que sus acciones tendrán.

### a.1) Descripción del modelo

Para modelar al consumidor que toma decisiones sobre la base de planes, se usó una red neuronal de tres entradas: El gasto del periodo anterior, el precio propuesto en el periodo anterior y la cantidad propuesta en el periodo anterior. Las salidas de la red son las mismas, pero para el periodo actual, esto es: Gasto, precio y cantidad para este periodo.

Vemos que la primera salida (Gasto) es el efecto que la red entrega para las acciones de la segunda y tercera salida (precio y cantidad). A su vez, la segunda y tercera salidas son las acciones que entrega la red, las cuales se espera que tengan el efecto entregado por la primera salida.

Es claro que el individuo decide en función de la experiencia que tuvo en periodos anteriores.<sup>14</sup>

---

<sup>14</sup> El número de periodos que usa el individuo como experiencia para tomar decisiones, depende del algoritmo de aprendizaje a usar, más adelante veremos esto.

La idea es que el individuo sea consistente en sus planes, esta consistencia se da cuando sus planes son exitosos, y al estar aislado, solo necesita que las acciones que realiza tengan los efectos que pensó que tendrían.

La Figura 2.1 muestra la red con sus respectivos objetivos de entrenamiento por salida. En la primera parte de esta tesis se presenta la técnica de entrenamiento cruzado.

En la medida que la red aprende a pronosticar el efecto de sus acciones, el individuo se aproxima al equilibrio.

El entrenamiento de la red se realiza sobre la base de un número variable de patterns, dependiendo del tipo de aprendizaje que se desee: aprendizaje a corto plazo o a largo plazo. La red aprende por defecto en base a la última serie de datos de salida (los datos del periodo anterior), pero el modelo permite cambiar el parámetro del número de patterns en la serie de entrenamiento ( $n$ ) de forma de usar los datos que van desde el último periodo hasta  $n$  periodos atrás. Así, se puede estudiar la diferencia entre un agente que actúa a corto plazo, adaptándose a la situación local, a mediano plazo, donde el agente empieza a desarrollar la capacidad de actuar en forma global o a largo plazo.

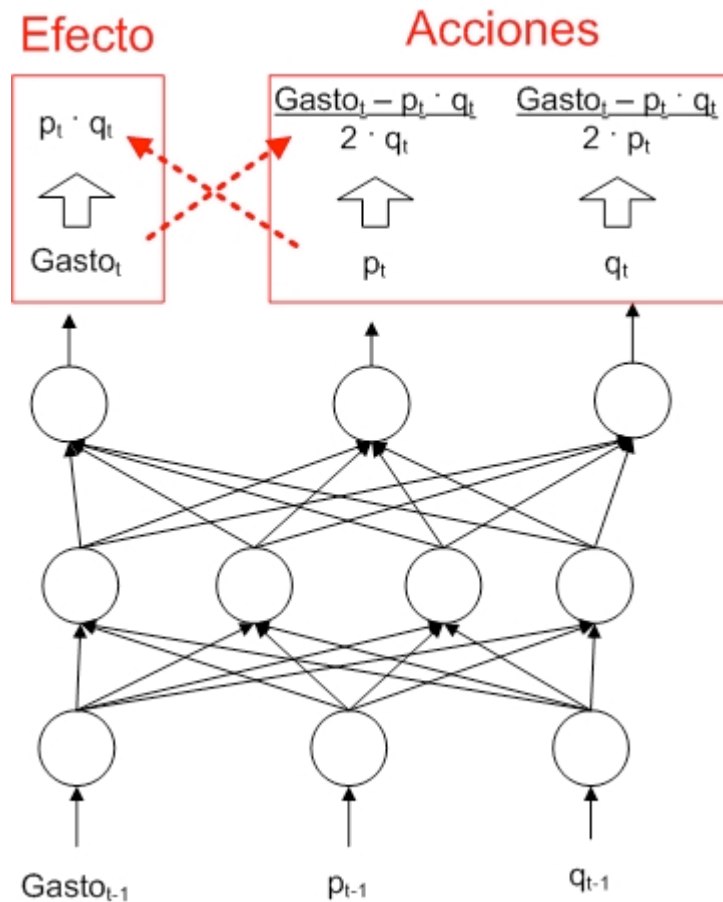


Figura 2.1

### a.2) Resultados

En la primera corrida del modelo, el agente aprende de los datos del periodo anterior únicamente. Vemos en la Figura 2.2 que el agente logra una consistencia interna completa cerca de la iteración número 70. Es interesante resaltar la simpleza de las series. El agente sólo lo aumenta o disminuye el gasto en función de lo que pasó en el periodo anterior hasta lograr el equilibrio acción/efecto.

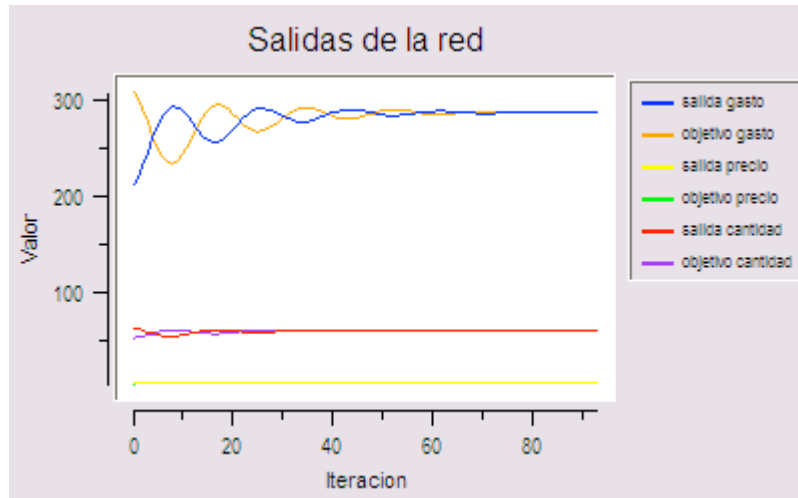


Figura 2.2

La segunda corrida del modelo considera a un agente un tanto más cauto en la toma de decisiones. Este agente usará los datos de diez periodos atrás para entrenar su red. En la Figura 2.3 podemos ver que el agente toma decisiones más complejas, tiene un plan que va más allá de lo obvio. Logra el equilibrio cerca de la iteración 30.

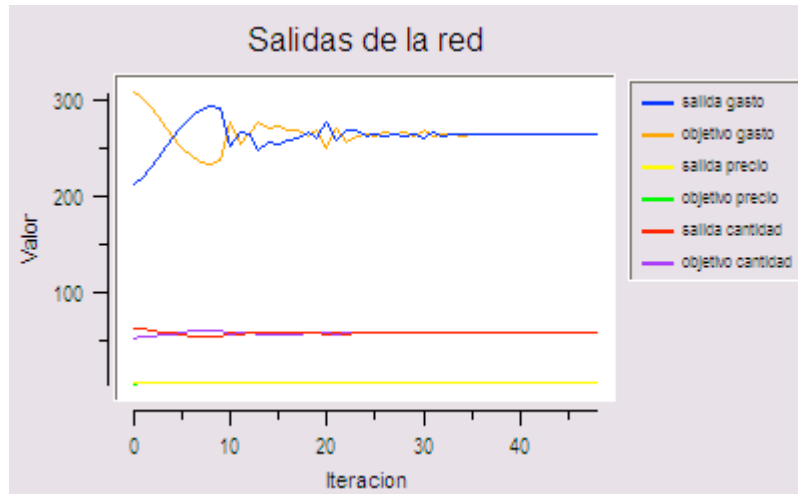


Figura 2.3

En la Figura 2.4 vemos las acciones de un agente que usa los últimos 50 datos para entrenar su red, y los efectos. Si bien el agente logra un equilibrio en la iteración 70, este equilibrio no es tan robusto como en los casos anteriores.



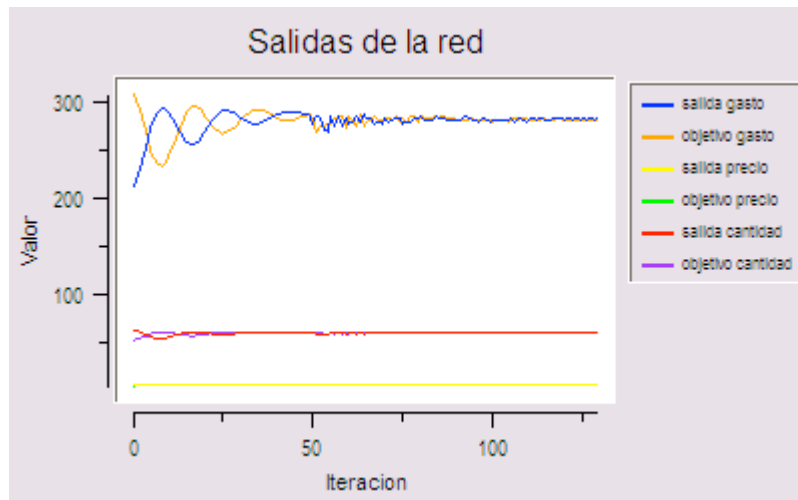


Figura 2.4

## b) Modelo de las decisiones tomadas por un individuo no aislado

En este caso, las decisiones del individuo son más complejas. El correcto pronóstico del efecto de las acciones ahora depende también de lo que aprende del comportamiento de los demás individuos del sistema. Si bien, ahora el individuo puede pronosticar perfectamente el efecto que tendrían sus acciones, es posible que sus acciones no sean consistentes con el mercado. En otras palabras, el consumidor puede pronosticar bien el gasto que tendrá con un precio y cantidad determinados, pero nadie le asegura que a ese precio puede comprar en el mercado. Así, vemos que tal como dijo Hayek (1937), la información en base a la cual el individuo toma sus decisiones ahora debiera incluir datos sobre los demás individuos (ambiente).

### b.1) Descripción del modelo

En cada periodo un consumidor busca un productor. Si el precio que ofrece el productor es menor o igual al que demanda el consumidor, se produce el intercambio.

- Para los agentes consumidores la red tiene las siguientes variables de entrada:
  1.  $Gasto_{t-1}$ : es el gasto en el que incurrió el consumidor en el periodo anterior.
  2.  $p_{t-1}$ : el precio que pagó el consumidor por el bien en el periodo anterior.
  3.  $q_{t-1}$ : la cantidad del bien que compró el consumidor en el periodo anterior.
  
- Para los agentes productores la red tiene las siguientes variables de entrada:
  1.  $Ingreso_{t-1}$ : el ingreso que percibió el productor en el periodo anterior.
  2.  $p_{t-1}$ : el precio al que vendió el bien el productor en el periodo anterior.
  3.  $q_{t-1}$ : la cantidad del bien que vendió el productor en el periodo anterior.
  
- Las salidas de las redes que simulan a los consumidores, y sus respectivos objetivos de entrenamiento son:
  1.  $P_{Real_t}$ : Es el pronóstico realizado por el consumidor para el precio de intercambio. El objetivo del entrenamiento de esta salida es el precio propuesto por el productor, en caso que hubiese habido intercambio en el periodo anterior, o bien, el precio propuesto por este mismo consumidor en caso que no hubiese habido intercambio en el periodo anterior.
  2.  $Q_{Real_t}$ : El pronóstico realizado por el consumidor para la cantidad a intercambiar. El objetivo de entrenamiento de esta salida es la cantidad real intercambiada, pudiendo ésta ser igual a cero.
  3.  $Gasto_t$ : El pronóstico del efecto que producirá sus acciones en el gasto al que incurrirá. El objetivo de entrenamiento es el efecto real que tienen sus acciones en el gasto.
  4.  $p_t$ : El precio propuesto por el consumidor. Esta es una acción. El objetivo de entrenamiento proviene del gasto, así se entrena esta acción de forma de coordinar la relación acción/efecto.

5.  $q_t$ : La cantidad propuesta por el consumidor. Es una acción. El objetivo de entrenamiento proviene del gasto (ver  $p_t$ ).
- Las salidas de las redes que simulan a los productores, y sus respectivos objetivos de entrenamiento son:
    1.  $P_{Real}_t$ : Es el pronóstico realizado por el productor para el precio de intercambio. El objetivo de entrenamiento de esta salida es el precio propuesto por el consumidor en caso de que hubiese habido intercambio en el periodo anterior, o bien, el precio propuesto por este mismo productor en caso que no hubiese habido intercambio en el periodo anterior (similar al caso del consumidor).
    2.  $Q_{Real}_t$ : El pronóstico realizado por el consumidor para la cantidad a intercambiar. El objetivo de entrenamiento de esta salida es la cantidad real intercambiada, pudiendo ésta ser igual a cero.
    3. Ingreso: El pronóstico del efecto que producirá sus acciones al ingreso que tendrá. El objetivo de entrenamiento es el efecto real que tienen sus acciones en el ingreso.
    4.  $p_t$ : El precio propuesto por el productor. Esta es una acción. El objetivo de entrenamiento proviene del ingreso, así se entrena esta acción de forma de coordinar la relación acción/efecto.
    5.  $q_t$ : La cantidad propuesta por el productor. Es una acción. El objetivo de entrenamiento proviene del ingreso (ver  $p_t$ ).

Para un consumidor, la red se muestra en la Figura 2.5, donde  $p_{real}_t^o$ ,  $q_{real}_t^o$ ,  $gasto_t^o$ ,  $p_t^o$  y  $q_t^o$  son los objetivos de entrenamiento para las respectivas salidas. Si consideramos que  $p_t^p$  y  $q_t^p$  son el precio y la cantidad propuestos por el productor con el cual nuestro consumidor está transando, podemos ver los objetivos de entrenamiento en la tabla 2.1.

Para la implementación del modelo, se utilizó las librerías Swarm-2.2. Al igual que en el modelo anterior, los agentes pueden aprender a actuar a corto plazo o largo plazo (y todos los puntos intermedios).

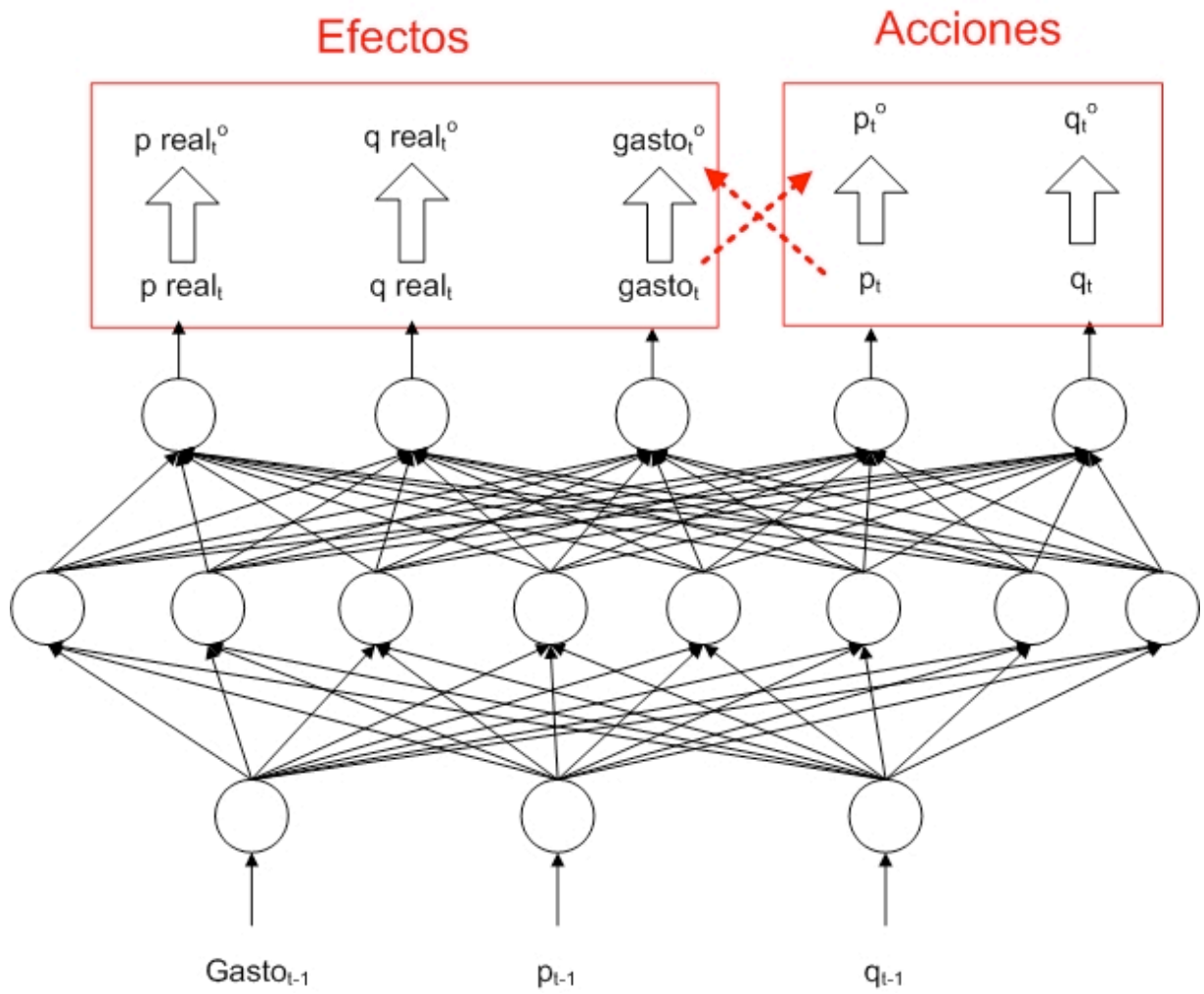


Figura 2.5

$p_{real_t^o}$	$p_{real_t^o} = \min(p_t, p_t^p)$
$q_{real_t^o}$	$if(p_t \geq p_t^p)$ $q_{real_t^o} = \min(q_t, q_t^p)$ $else$ $q_{real_t^o} = 0$
$gasto_t^o$	$gasto_t^o = p_t \cdot q_t$

$p_t^o$	$corrección1 = \frac{gasto_t - gasto_t^o}{2 \cdot q_t}$ <p>if (<math>q_{real_t^o} = 0</math>)  <math>corrección2 = 0,15</math>  else  <math>corrección2 = -0,15</math></p> $p_t^o = p_t + \max(abs(corrección1), abs(corrección2))$
$q_t^o$	$corrección1 = \frac{gasto_t - gasto_t^o}{2 \cdot p_t}$ $p_t^o = p_t + corrección1$

Tabla 2.1

### b.2) Resultados

El modelo fue probado con 10 consumidores y 10 productores.

Con agentes que aprenden a corto plazo, es decir, se adecuan al mercado según sus experiencias recién pasadas (en el periodo anterior), emerge un equilibrio espontáneo en el precio promedio del mercado (ver Figura 2.6). El equilibrio se logra cerca de la iteración 120.

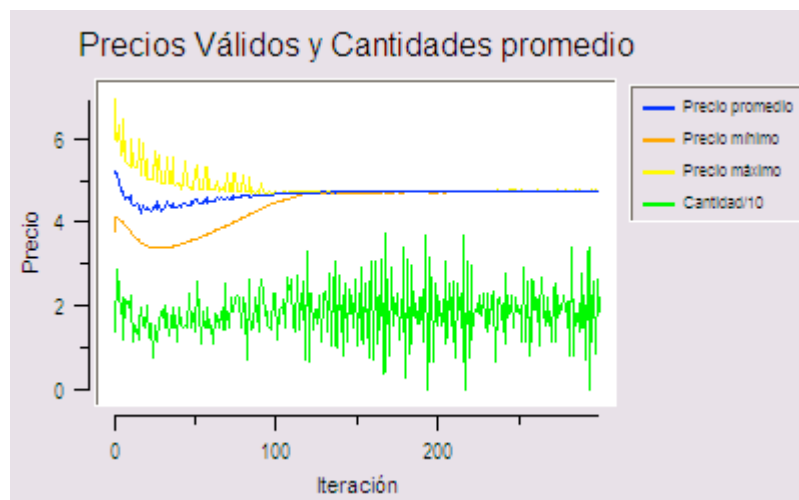


Figura 2.6

Vemos en la Figura 2.7 que para un mercado compuesto por agentes que se entrenan con los datos de los últimos 10 periodos, los precios se demoran alrededor de 7000 iteraciones en llegar a un equilibrio. Hasta la iteración 2500 aprox. algunos agentes siguen planes distintos a los de la mayoría.

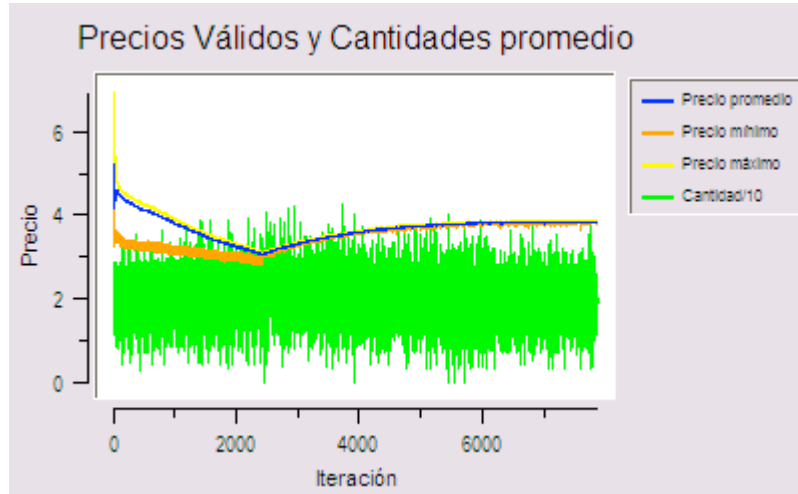


Figura 2.7

Ahora, con agentes que aprenden usando la información de los últimos 50 periodos, ver Figura 2.8, un equilibrio de precios espontáneo emerge muy rápidamente. Sin embargo existen agentes que se diferencian en su comportamiento. Siguiendo la tendencia de la Figura 2.6 y la Figura 2.7, probablemente antes de la iteración número 100.000 todos los agentes tendrán planes similares.

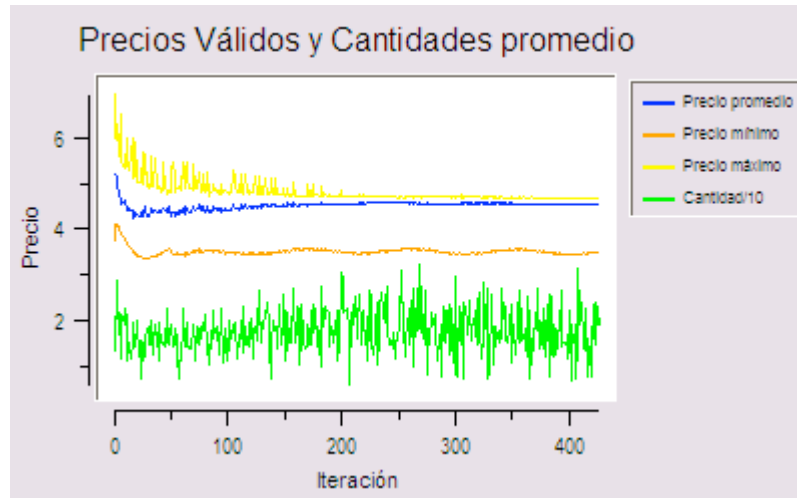


Figura 2.8

### 5.3.3. Comentarios

Para el modelo de un individuo aislado, *“las relaciones de equilibrio existen sólo durante el periodo en el cual el pronóstico de las acciones tomadas es correcto”*, Hayek (1937), confirmando así la teoría de Hayek sobre el equilibrio en individuos aislados.

Para el modelo multiagente, las relaciones de equilibrio existen durante el periodo en el cual *“los diferentes planes de acción en el tiempo que tienen los individuos que componen una sociedad son mutuamente compatibles”* (Hayek, 1937) confirmando así la teoría de Hayek sobre el equilibrio en sistemas económicos.

Dado que de las pruebas de ambos modelos emergió la propiedad de equilibrio, podemos validar estos modelos como simuladores de un sistema económico Hayekiano compuesto por un solo individuo, para el caso del primer modelo, y de múltiples individuos, para el caso del segundo modelo.

Para modelar las decisiones de precio y cantidad a comprar o vender de cada individuo en esta tesis, se utilizó el modelo del sistema Hayekiano multiagente ya presentado. Los parámetros del modelo que se usó consideran una red neuronal con ocho neuronas

ocultas.<sup>15</sup> Además, el propósito de esta tesis es simular agentes que deciden sobre la base a su experiencia, y claramente la experiencia no puede basarse únicamente en el periodo anterior. Al igual que los seres humanos, se intentó usar la mayor cantidad de periodos anteriores para formar la historia de experiencias a la cual tiene acceso el agente de forma de lograr un aprendizaje a largo plazo. En condiciones normales, nadie planea en función de lo que pasó solo el día anterior. El modelo de esta tesis considera los últimos 15 patterns. Si bien se intentó experimentar con aprendizajes a mayor plazo, no fue posible por limitaciones computacionales.

## 5.4. Tipo de bien a producir

Los consumidores expresan su intención de compra en uno de los cuatro bienes transables, o en términos de la red neuronal, una de las cuatro “señales de demanda”. Las cuatro señales posibles son: (0, 0), (0, 1), (1, 0), (1, 1). El modelo asume que el consumidor decide comprar siempre que la respuesta que éste reciba del productor satisfaga la función lógica O-exclusiva XOR.

X	Y	XOR(X, Y)
0	0	0
0	1	1
1	0	1
1	1	0

Es importante destacar que la función lógica O-exclusiva, a diferencia de las funciones AND, OR y NOT, no es linealmente separable. Esto significa que los datos 1 no pueden ser separados linealmente de los datos 0 en un gráfico de dos dimensiones como el siguiente:

---

<sup>15</sup> La mejor arquitectura para una red que simula el comportamiento humano es un problema interesante, pero se escapa del alcance de esta tesis.



1	1	0
0	0	1
<b>XOR</b>	0	1

Cada agente produce el bien determinado por una red neuronal perceptron, con dos entradas y una salida (Rosenblatt, 1959). Las entradas de la red reciben la señal de demanda del consumidor, y la salida entrega la respuesta del productor pudiendo ser ésta satisfactoria (verdadero) o no satisfactoria (falso).

La función de activación de la neurona que posee cada red  $f(z)$  es la función umbral, tal que:

$$if(z \leq 0) f(z) = 0$$

$$if(z > 0) f(z) = 1$$

Las Figuras 2.9 y 2.10 muestran dos ejemplos de este tipo de redes.

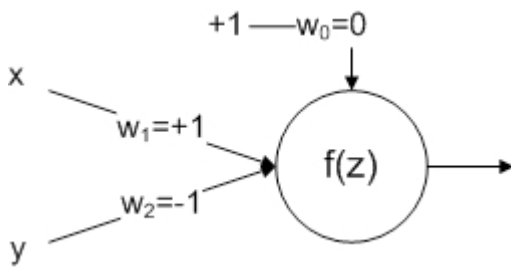


Figura 2.9

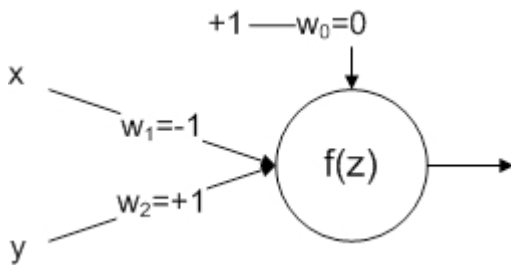


Figura 2.10

x	y	Salida Red
0	0	0
0	1	0
1	0	1
1	1	0

x	y	Salida Red
0	0	0
0	1	1
1	0	0
1	1	0

Las entradas a la red son ponderadas, y para la suma de estas ponderaciones se determina si el umbral es alcanzado. Si el umbral es alcanzado, la salida de la red es 1, en otro caso, la salida es 0. Además, la red tiene un peso externo adicional que pondera al bias.

Se utilizó este tipo de red por su incapacidad de resolver problemas a base de patterns que no sean linealmente separables (Minsky y Papert, 1969). Por esta razón, es imposible que una red perceptron pueda aprender la función XOR.

Esta red es simplemente una función lineal  $f(x, y) = w_0 + x \cdot w_1 + y \cdot w_2$ , donde  $x$  e  $y$  son las dos entradas a la red y  $w_0$ ,  $w_1$  y  $w_2$  son los pesos de la red que conectan el bias, la entrada  $x$  y la entrada  $y$ , respectivamente. Para que la red pueda aprender a responder correctamente a todas las instancias de la función XOR, el sistema de ecuaciones:

$$w_0 + 0 \cdot w_1 + 0 \cdot w_2 < 1$$

$$w_0 + 1 \cdot w_1 + 1 \cdot w_2 < 1$$

$$w_0 + 1 \cdot w_1 + 0 \cdot w_2 \geq 1$$

$$w_0 + 0 \cdot w_1 + 1 \cdot w_2 \geq 1$$

debería entregar los pesos correspondiente a la red entrenada. Pero este sistema no tiene solución.

Si consideramos que  $f(x, y) = 0$ , vemos proyectado en el plano  $(x, y)$  la línea  $f(x) = -\frac{x \cdot w_1 + w_0}{w_2}$ . Esta línea divide el plano  $(x, y)$  entre los puntos proyectados de  $f(x, y) > 0$  y los proyectados de  $f(x, y) < 0$ . Claramente entonces, estas redes solo pueden generar una función  $f(x, y)$  cuya proyección en el plano  $(x, y)$ ,  $f(x)$ , es una línea o sea, solo pueden aprender problemas linealmente separables. Cada red, por si sola, no puede simular la función XOR. No hay forma de que una línea pueda dividir las instancias positivas de las negativas de la función XOR.

La función XOR fue elegida por ser el generador más simple de un problema no separable linealmente.

La red es entrenada usando la Regla Delta creada por Widrow y Hoff (1960), que es simplemente un caso particular del algoritmo Backpropagation, donde, si llamamos  $t$  al objetivo de entrenamiento de la red (resultado de la función XOR),  $\Delta w_i = t \cdot x_i$ .

## 5.5. Creación de Empresas

La unión de dos o más agentes crea una empresa. La unión entre agentes tiene dos ventajas simultáneas:

1. Con la unión de dos o más productores, el jefe recibe todas las señales de demanda que recibiría estando solo, más las que recibe por medio de sus trabajadores. Esto da la oportunidad al jefe de entrenar más rápidamente la red neuronal de precios y cantidades y de reaccionar más rápidamente ante cambios en el ambiente. En la línea de la teoría de Hayek, esto significa un mayor acceso a información (o conocimiento).
2. Dada la simplicidad de la red perceptron que determina qué bien producir, los productores no serán capaces de responder afirmativamente las cuatro señales de demanda existentes. La red se entrenará periodo a periodo de forma de adaptarse a la demanda.

Este aprendizaje representa fielmente la “propensión natural al trueque e intercambio” del ser humano descrita por Adam Smith (1776).

Esta propensión es modelada en aún mayor profundidad. Cuando un productor es incapaz de satisfacer cierta señal de demanda, éste pregunta a cada uno de sus ocho vecinos qué habrían respondido en caso de ellos haber recibido la misma señal. Cada vez que un vecino da la respuesta satisfactoria según la función XOR, la confianza del productor con ese vecino se incrementa. Así, un índice de confianza para cada vecino es actualizado cada vez que el productor recibe una señal de demanda. Cuando el índice de confianza con un vecino es mayor que cierto límite, el productor independiente considerará convertirse en un empresario, tratando de contratar al vecino confiable. Si la contratación es exitosa, el

nuevo empresario estará a cargo de coordinar su respuesta a la señal de demanda más la respuesta de su nuevo trabajador.

En términos de la red neuronal, una contratación significa agregarle un nuevo enlace a la red del productor contratante que lo una con la salida de la red del productor contratado, creando así un algoritmo constructivo, Auer, Burgsteiner y Maass (2001).

La Figura 2.11 muestra una empresa formada por dos agentes que logró una situación eficaz. Llamamos eficaz a la empresa capaz de responder exitosamente a los cuatro tipos de señales de demanda. Una empresa que no ha alcanzado la eficacia, continuará su proceso de aprendizaje en función de los errores que cometa en su evolución. Esto implica que el proceso de contratación podría volver a suceder en periodos posteriores hasta que la empresa alcance una situación de eficacia tal que sea capaz de responder satisfactoriamente a todas las señales de demanda.

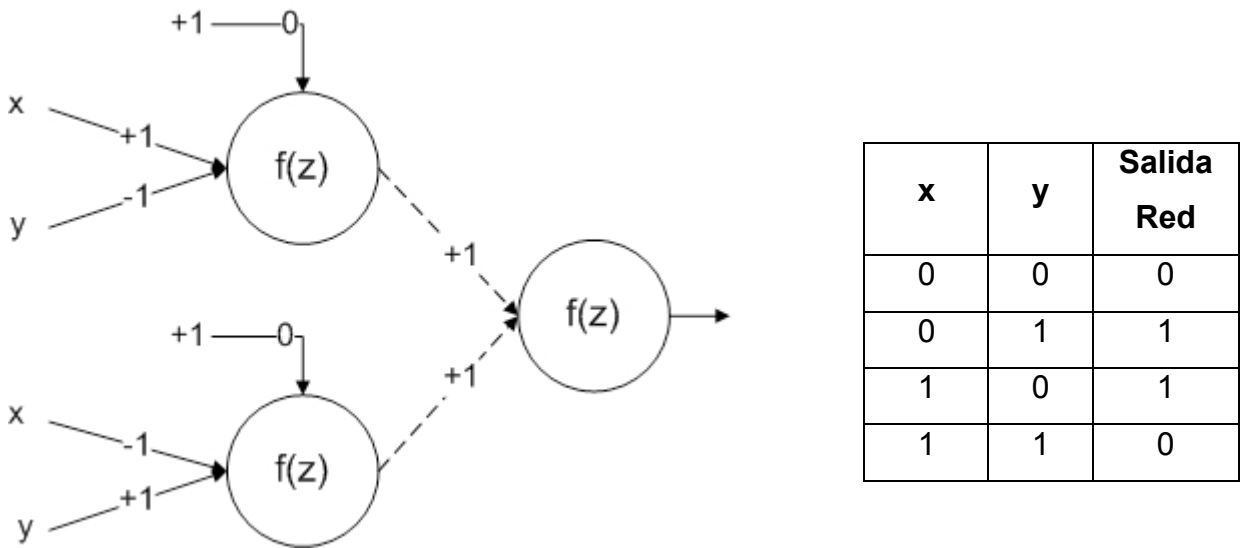


Figura 2.11

Una vez que se crea una empresa, la red del empresario tiene un enlace más (y su respectivo peso  $w$ ) para ser entrenado. Así el empresario es el encargado de “dirigir” el comportamiento de la empresa.

Una empresa eficaz puede estar compuesta de dos hasta nueve miembros. El tamaño de la empresa tiene ventajas y desventajas. La ventaja de ser una empresa pequeña es que el producto será compartido por una menor cantidad de agentes. La desventaja de una empresa pequeña es que el “área de marketing” es menor, es decir, dado que la empresa ocupa una menor área en la grilla, ésta tiene una menor probabilidad de recibir consumidores. La ventaja de ser una empresa grande es que el “área de marketing” es mayor. La desventaja de ser una empresa grande es que esta empresa tendrá que dividir el mismo producto entre una mayor cantidad de agentes.

Las empresas eficaces no son eternas. Estas pueden quebrar o ser disueltas mediante una tasa de mutación de empresas incluida en el modelo.

## 5.6. Dinero y Crédito

La parte monetaria de la economía es modelada como un sistema de crédito. El sistema de crédito fue elegido por su simplicidad y su correspondencia con la realidad. Todos los agentes inician su “vida” con una riqueza igual a cero. En cada transacción el productor aumenta su riqueza, y el consumidor disminuye la misma. Esto significa que para que un agente tenga un superávit monetario necesariamente debe haber otro agente que posea un déficit monetario. Cómo los agentes reaccionan ante superávits y déficits es crucial con respecto a las propiedades macroeconómicas del modelo. Estos son aspectos de los sistemas económicos que son muy difíciles de estudiar sin el uso de la simulación basada en agentes.

Los agentes se endeudan como productores y como consumidores. Para que la producción se realice, debe haber una demanda por consumo, y esto requiere que los consumidores estén dispuestos a endeudarse antes de recibir algún tipo de ingreso en forma de sueldos (para trabajadores) o utilidades (para independientes o empresarios). Más adelante, el dinero gastado por los consumidores es devuelto en forma de sueldo o utilidades, pero no hay garantía de que cada consumidor reciba exactamente lo que gastó.

## 5.7. Determinación de Sueldos

Cuando un productor independiente decide convertirse en empresario, o cuando un empresario que aún no ha formado una empresa eficaz decide contratar más empleados, debe que convencer al(los) vecino(s) elegido(s) a trabajar para él. Cada agente guarda registro de sus ingresos del mes anterior. Dado que para realizar una venta se requiere que el productor cumpla con las exigencias de precio y tipo de bien determinadas por el consumidor, el ingreso que el productor tuvo en el mes anterior es un índice que mide el nivel de eficacia de cada productor.

Para el caso de un productor independiente, este potencial empresario ofrece pagar como sueldo a su vecino independiente el ingreso que él tuvo en el periodo anterior dividido por dos. Si se trata de un empresario que decide aumentar el tamaño de su empresa, ofrece al potencial trabajador su ingreso en el periodo anterior dividido por el número de trabajadores actuales más uno. De esta forma el empresario gana en “área de marketing” y en know how, ya que nuevos integrantes en la empresa significa una empresa más grande y con mayores probabilidades de recibir un cliente dentro de la grilla. Además cada nuevo empleado trae consigo su red neuronal, la que se suma a la red de la empresa y crea así una red más flexible a cambios en la demanda.<sup>16</sup> Así, vemos que el modelo está diseñado de forma que los empresarios distribuyan sus ingresos en partes iguales con sus trabajadores (como la oferta del sueldo es en base a los ingresos del mes anterior, el empresario puede realmente pagar a sus trabajadores menos o más de lo que él recibe dependiendo de los ingresos del mes actual, sin embargo la intención del empresario es una distribución equitativa).

Este diseño es bastante razonable. El agente que quiere contratar considera su experiencia como normal. Por esta razón, para el empresario una unidad extra de producción y de marketing puede llevar a ingresos promedios equivalentes a los ingresos promedios actuales de la empresa.

---

<sup>16</sup> Ver “Creación de empresas”.

Sin duda la proposición del empresario es generosa, ya que ofrece al trabajador un sueldo que corresponde aproximadamente a una división equitativa del ingreso esperado. Por esta razón, el modelo tiene este sistema de división del ingreso como un parámetro, que puede ser cambiado.

Si el sueldo ofrecido por el empresario o independiente es mayor o igual al ingreso que el candidato tuvo en el periodo anterior, éste aceptará la oferta de trabajo. El candidato acepta una oferta que es igual a su ingreso como independiente ya que el modelo supone que los agentes poseen una aversión al riesgo, característica natural de la mayoría de los seres humanos.

Si el sueldo ofrecido por el empresario o independiente es menor al ingreso del candidato en el periodo anterior, éste no aceptará la oferta de trabajo. En este caso el empresario o independiente intentará contratar al mismo productor independiente candidato dentro de un mes, ofreciéndole un nuevo sueldo que depende del éxito que la empresa tuvo durante el mes. Durante este mes, el candidato puede tener menos fortuna y sus ingresos pueden disminuir, por lo que éste aceptará ser contratado por el mismo sueldo que rechazó en el periodo anterior.

El último sueldo transado entre el empresario y sus trabajadores será el sueldo válido durante ese mes. En meses posteriores, el sueldo puede aumentar o disminuir dependiendo del éxito que tenga el empresario. Esto no es lo mismo que decir que el sueldo de los trabajadores es una porción fija del ingreso que tenga el empresario, ya que el pago de los sueldos negociados el mes anterior puede resultar tanto en pérdidas como ganancias para el empresario en el presente mes. Una vez que la empresa es eficaz, el sueldo de los trabajadores de la misma será fijado al nivel ofrecido al último trabajador contratado. Este sueldo dará utilidades al empresario dependiendo de su suerte en atraer consumidores.

## 5.8. Decisión de compra

Dado que para el modelo MECHA una iteración es un ciclo en el cual cada uno de los agentes del modelo efectúa una señal de consumo, sería poco realista considerar un “día” como una iteración. En una sociedad humana, los integrantes de ella realizan múltiples compras diarias. Más aún, sería poco realista considerar que todos los agentes realizan la misma cantidad de compras en un día. Claramente, en el mundo real unos individuos realizan más compras diarias que otros.

La cantidad de compras que realiza un individuo depende del acceso que tenga al dinero, es decir, de su riqueza. Para Clower y Johnson (1976), los individuos integrantes de una economía dan valor tanto a su riqueza como al consumo que realizan. Según ellos, ahorrar no es necesariamente una decisión para consumir mañana, sino que una decisión para aumentar la riqueza. El consumo puede ser modelado como un proceso que depende de la diferencia entre la riqueza actual y la deseada. A partir de evidencia empírica, Clower y Johnson (1976) entregan una función de consumo del siguiente tipo:

$$c = h \cdot w^b$$

Ecuación 2.1

donde  $c$  es la parte de la riqueza que se destina a consumo,  $w$  es la riqueza,  $h$  es una constante positiva y  $b$  es un número entre cero y uno. En su estimación empírica, Clower y Johnson encontraron que  $b$  tiene un valor entre 0.35 y 0.40 para Estados Unidos e Inglaterra.

Como el modelo MECHA no posee bienes de capital o un mercado financiero excepto el sistema de crédito, se debe hacer algunas modificaciones a la función de consumo de Clower y Johnson. Bruun y Luna (1999) usan en su modelo una modificación de esta función.

$$c = \frac{h}{1 + e^{\frac{-w}{b}}}$$

Ecuación 2.2



donde  $h$  y  $b$  son constantes positivas. En su modelo Bruun y Luna usaron  $h = 10$  y  $b = 25$ . La gráfica de esta función se puede ver un la Figura 2.12.

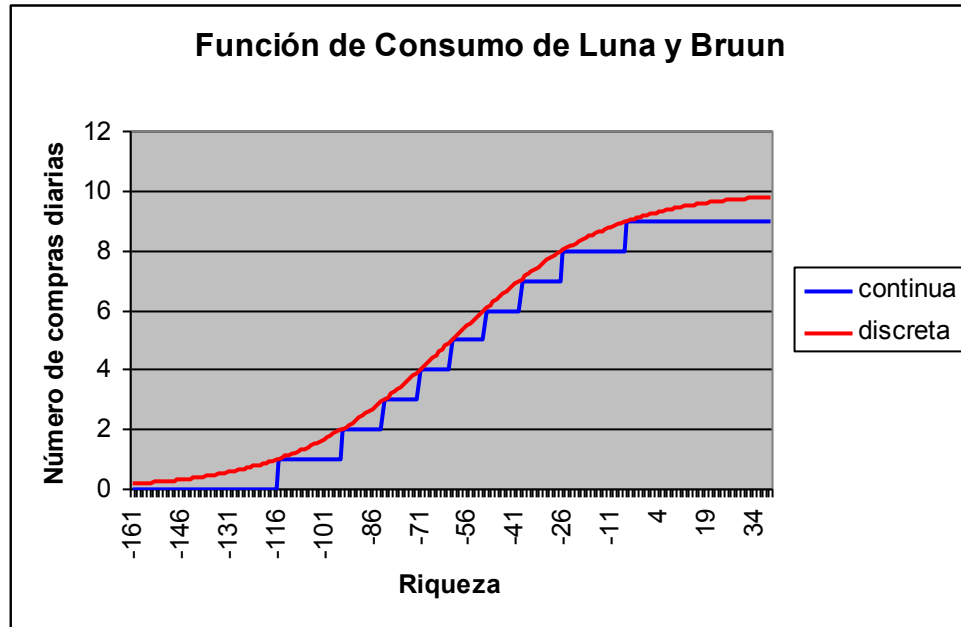


Figura 2.12

Como la riqueza total del modelo es siempre cero, esta función representa mejor la realidad del mismo, ya que permite que los agentes con riqueza negativa consuman. Sin agentes dispuestos a endeudarse para consumir, la economía del modelo nunca despegaría. Además, en su modelo Bruun y Luna utilizaron un mínimo de consumo igual a una compra, lo que se puede interpretar como un mínimo de subsistencia.

Para agentes cuyas riquezas son positivas, ambas funciones son similares. Para agentes con riqueza igual a cero o negativa, la función de Bruun y Luna asegura que el consumo no será nunca negativo.

El modelo creado en esta tesis usa la función de decisión de compra entregada por Bruun y Luna.

Además de decidir cuantas compras hacer, los consumidores deben decidir cuando moverse en busca de otro productor para tener una mayor probabilidad de satisfacer las

necesidades del próximo periodo. La ubicación de cada agente dentro de la grilla es fija. Los agentes tienen una dirección de compra, que es variable. Esta dirección de compra se puede interpretar como la dirección del “carro de compras” similar a un supermercado, la cual puede recorrer toda la grilla y encontrarse con los productores. Si las compras realizadas durante el presente día por un consumidor fueron exitosas, este consumidor no tendrá razones para abandonar a ese productor y se quedará en esa dirección de compra durante el próximo día. Si las compras realizadas por el consumidor en esa ubicación no fueron exitosas, el consumidor buscará a otro productor para el día siguiente, o sea, moverá su carro de compras a otra ubicación aleatoria. Una tasa de mutación es incluida en el modelo de forma que los consumidores puedan cambiar su ubicación de compra aunque estén sobre una firma eficaz que responda exitosamente a todas sus compras.

## 5.9. Quiebras

Para lograr una asignación eficiente de recursos en el largo plazo, MECHA considera la quiebra de empresas (o independientes) cuyas deudas son mayores a cierto límite de dinero. Mientras las empresas mantengan sus deudas (de existir) por bajo el nivel límite, éstas tendrán acceso libre a crédito. Sin un límite a las deudas, los empresarios podrían contratar trabajadores sin lograr nunca una empresa eficaz. Estos empresarios pueden limitar la formación de empresas eficaces en su “barrio”, y por lo tanto limitar el crecimiento.

Las pérdidas relacionadas a la quiebra de una empresa son absorbidas por sus acreedores, tal como es en el sistema chileno.

Cuando un agente es declarado en quiebra, éste es reinicializado, es decir, su posición financiera cambia a cero y los pesos de todas sus redes neuronales son reasignados aleatoriamente. Además la dirección de compra del agente es también aleatoriamente reasignada.

Es necesario destacar la importancia que tiene esta “actividad” en el modelo.<sup>17</sup> Las quiebras dan a la economía la posibilidad de evolucionar, sin ellas no hay motivación para que un agente se desarrolle y sea cada vez más eficaz.

## 5.10. Acciones de MECHA

Para lograr una explicación más clara de las acciones dinámicas del modelo, en las próximas páginas se presentan diagramas de actividades UML para las acciones diarias y las acciones mensuales.

---

<sup>17</sup> En el lenguaje de Swarm, actividad se le llama a cada acción que realiza el modelo.

### 5.10.1. Acciones diarias

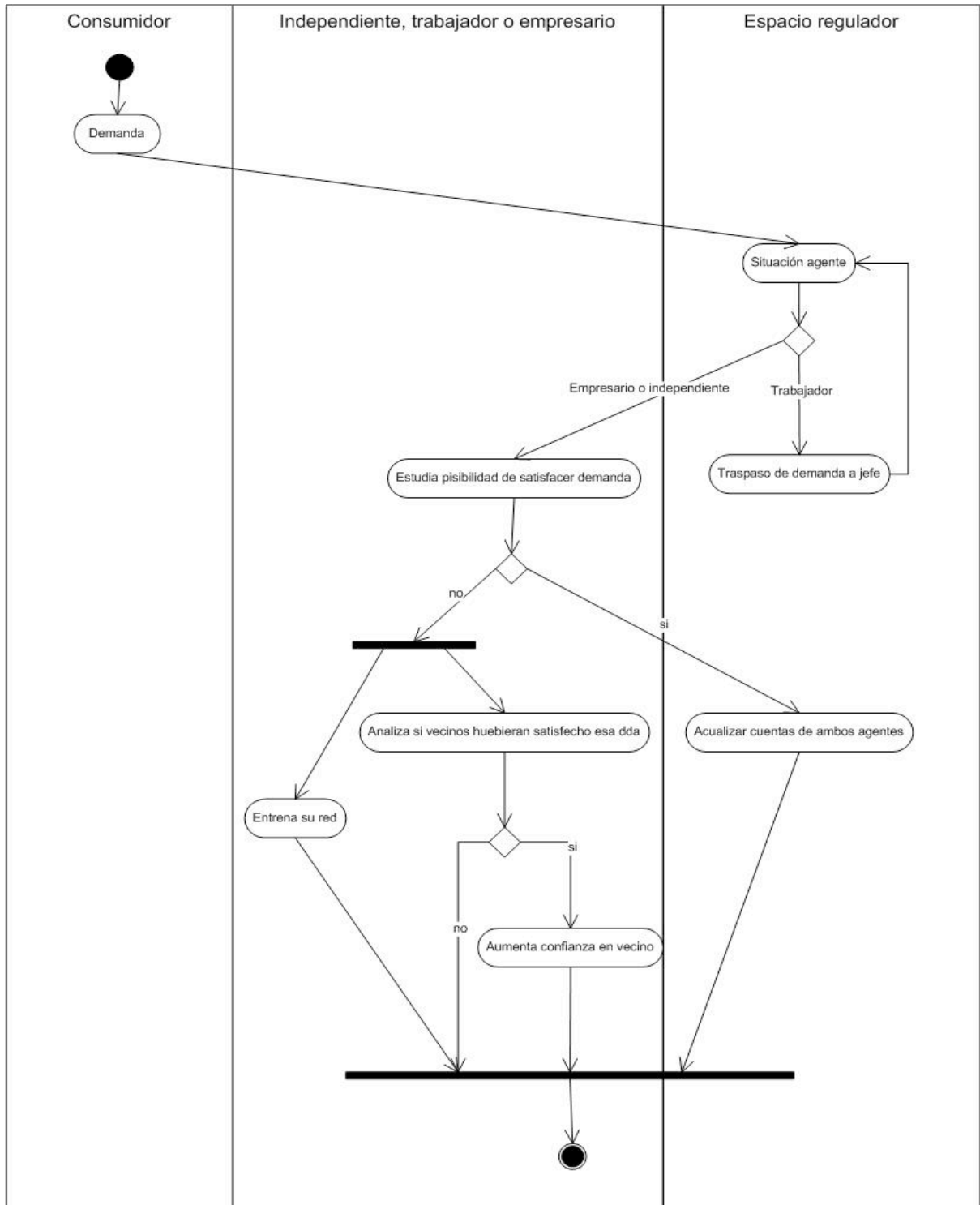


Figura 2.13

5.10.2. Acciones mensuales

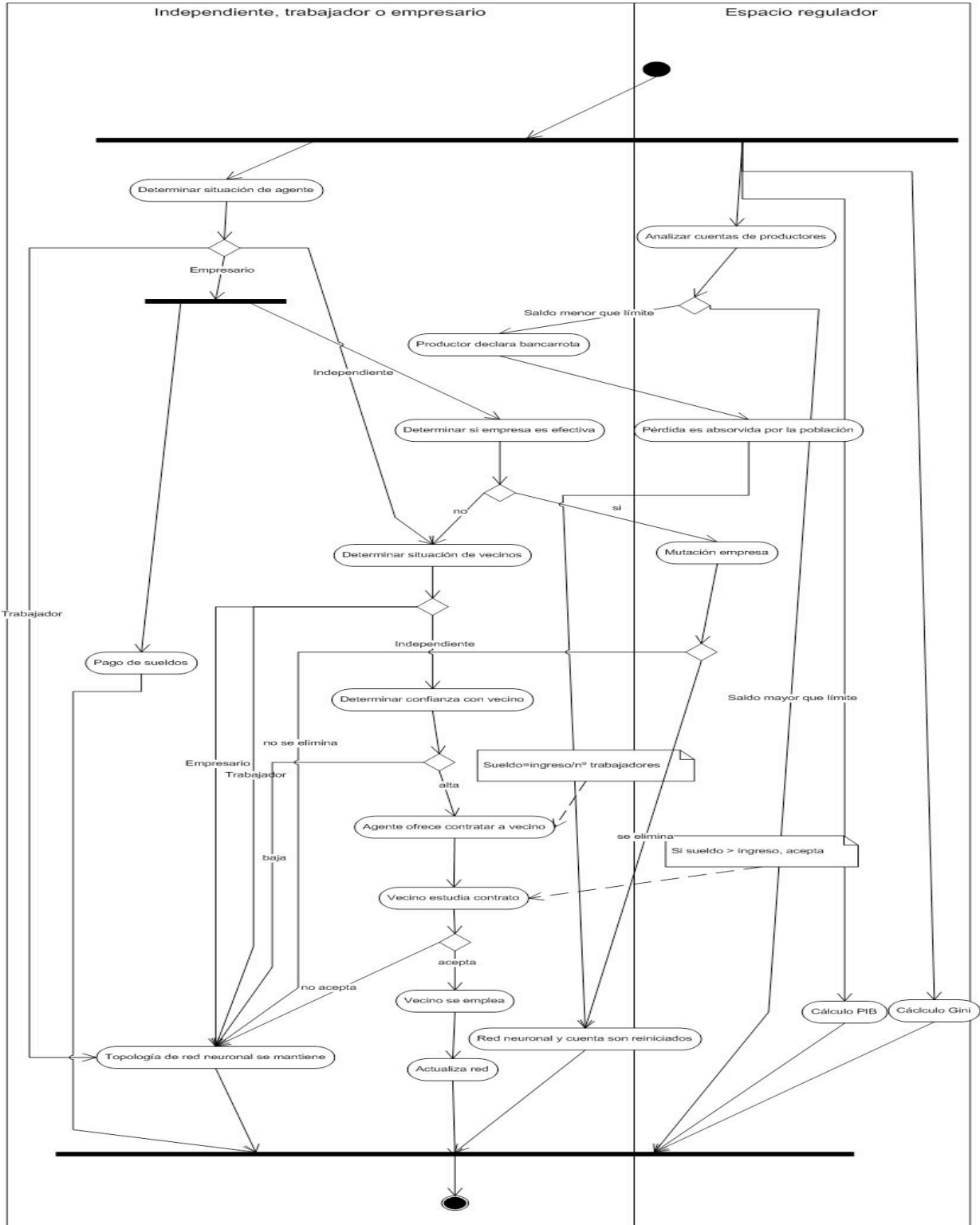


Figura 2.14

## 5.11. Fenómenos Emergentes del Modelo

### 5.11.1. Crecimiento

Del modelo MECHA emerge un crecimiento económico en el largo plazo marginalmente decreciente hasta que se logra un equilibrio en la asíntota que tiende a 2.507.552 Unidades Monetarias, ver Figura 2.15. Este límite al crecimiento es debido a que la función de demanda es marginalmente decreciente. Además, por esta misma razón el producto per cápita es independiente del número de agentes en el modelo, con un valor de aproximadamente 6.300 unidades monetarias.

En la Figura 2.16 vemos una secuencia de imágenes de una grilla que representa el espacio del modelo. La grilla está formada por 400 (20 x 20) cuadrados que representan a cada uno de los agentes del modelo. Estos agentes tienen un color específico, dependiendo del estado del mismo. Un agente independiente es amarillo, un empresario es rojo, un trabajador es azul y un empresario efectivo es verde. Las imágenes muestran la evolución del mercado desde que la simulación empieza hasta que llega al primer peak, previo a la primera recesión. La creación de empresas es evidente. Sin duda la razón del crecimiento económico son las empresas, y en especial, las empresas efectivas.

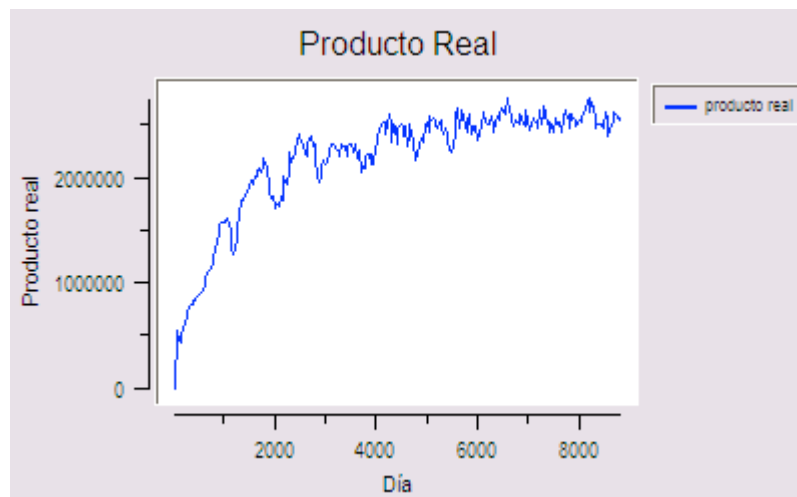


Figura 2.15

### 5.11.2. Ciclos

Es posible observar ciclos de producción como fenómenos emergentes del modelo MECHA. Al igual que en el mundo real, es difícil determinar qué es lo que produce un resultado específico dentro del modelo. La ventaja de usar un modelo de simulación es que se puede alterar parámetros para ver el efecto que tienen en los resultados emergentes y así establecer dependencias.

¿Por qué el modelo presenta ciclos? Dado que cada depresión va acompañada de un alza en los precios, pareciera ser que cada ciclo es producto de una quiebra a gran escala de empresas.

Efectivamente, vemos en la Figura 2.18 que el número de agentes que quiebra sigue una tendencia cíclica similar al producto de la economía.

Pareciera ser que los ciclos económicos son producto de la creación y quiebra de empresas. Vemos en las Figura 2.15 y 2.16 que durante los primeros meses se concentra la máxima cantidad de creación de empresas. El nivel de endeudamiento es muy alto. Algunas de estas empresas lograrán ser efectivas, otras no. En la Figura 2.17 vemos que los precios bajan producto del aumento de la oferta. Las empresas efectivas podrán pagar su deuda y seguir creciendo, mientras que las empresas que no logren la efectividad seguirán luchando hasta que su deuda llegue al nivel de bancarrota. Unas empresas llegarán antes que otras a este nivel, dependiendo de la efectividad de cada cual. Cuando quiebra la primera empresa, todos los acreedores de esa empresa deben absorber la deuda que ésta dejó, por lo que el sistema se ve afectado considerablemente. Por esta razón, la quiebra de una empresa trae consigo la quiebra de otras empresas y así se logra un efecto dominó que termina cuando quiebran todas las empresas que no son lo suficientemente eficaces como para sobrevivir a la crisis. Existe una retroalimentación negativa en el sistema. Los precios suben producto de la disminución en la oferta. Durante el proceso de quiebras el producto de la economía cae, formando el primer ciclo del modelo. Una vez que las empresas que no son lo suficientemente eficaces han desaparecido, se reinicia la creación de empresas. Aumenta

nuevamente el nivel de crédito ya que nuevas empresas entran en deuda.<sup>18</sup> Los precios vuelven a bajar producto de la nueva formación de empresas. Una vez que la deuda de las empresas que no son lo suficientemente eficaces es mayor al punto de quiebra, se produce el segundo efecto dominó que traerá consigo a un segundo ciclo económico.

Esta evolución cíclica de empresas se puede ver más claramente en la secuencia de fotos de la Figura 2.16. En un inicio solo hay agentes independientes en el modelo, a medida que este está en ejecución, unos agentes crean empresas y otros se emplean. Cada vez la densidad de empresas es mayor, y eventualmente aparece las primeras empresas efectivas. Continúan apareciendo empresas efectivas, hasta el punto en que quiebran las primeras empresas, dando inicio a la primera recesión. Este punto es el peak del primer ciclo.

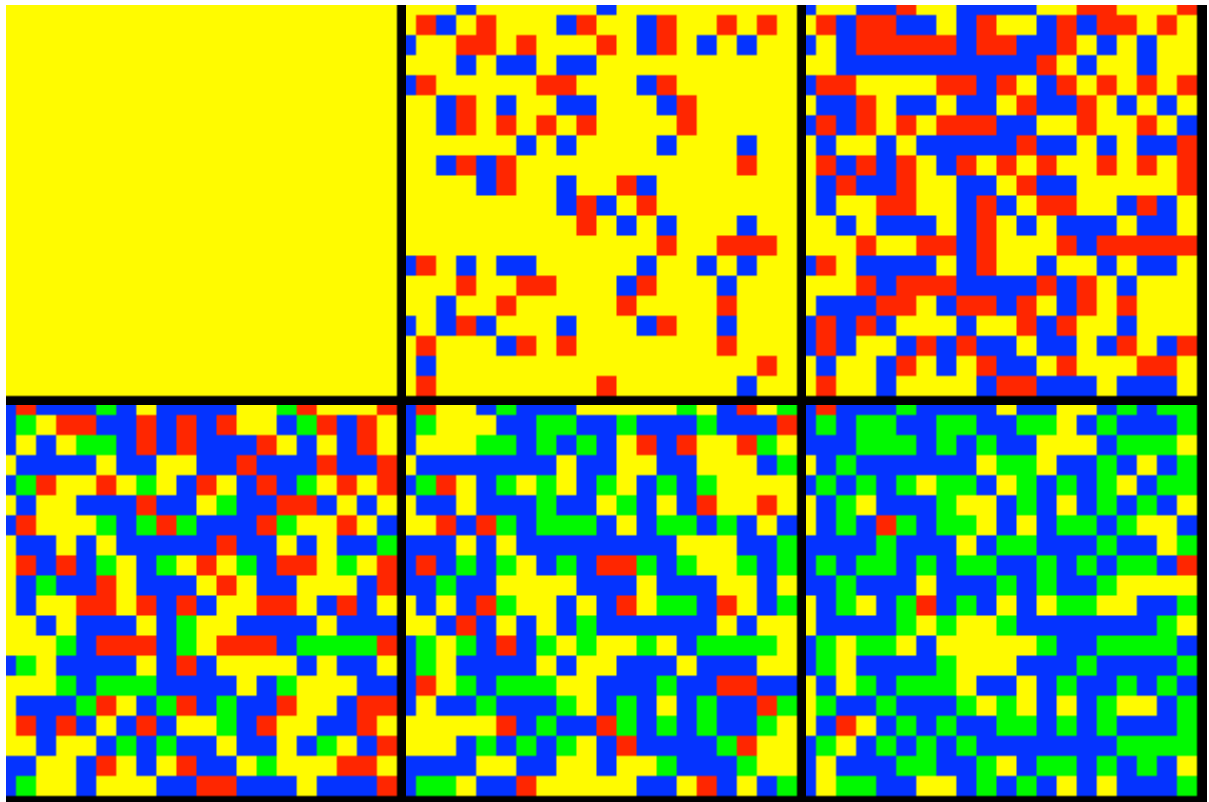


Figura 2.16

<sup>18</sup> Se observa que el nivel de crédito nunca logra igualar el nivel previo al primer ciclo, ya que es durante el inicio del modelo cuando se crea la mayor cantidad de empresas.



Es interesante destacar que, al igual que en el mundo real, el modelo simula un mercado competitivo en el cual sobreviven los mejores, logrando un proceso evolutivo que causa ciclos en la economía.

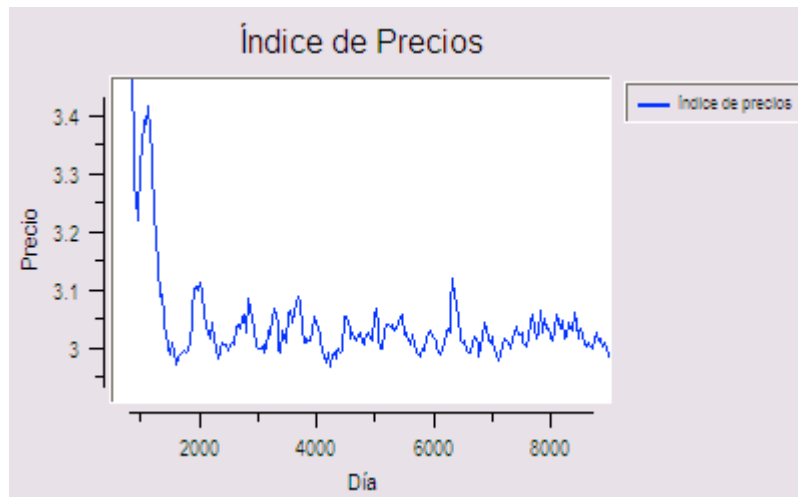


Figura 2.17

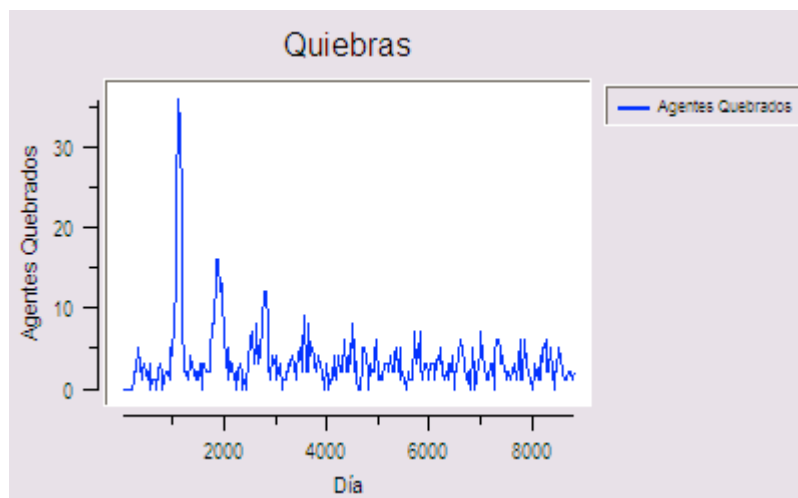


Figura 2.18

Para estudiar si los fenómenos emergentes ya mencionados son propios del modelo, se observó los resultados del mismo usando cinco semillas distintas y aleatorias en el generador de números aleatorios de la simulación, ver Figura 2.19. Los resultados generados por el modelo usando cada una de las cinco semillas son similares. Además, todos poseen crecimiento y ciclos económicos.

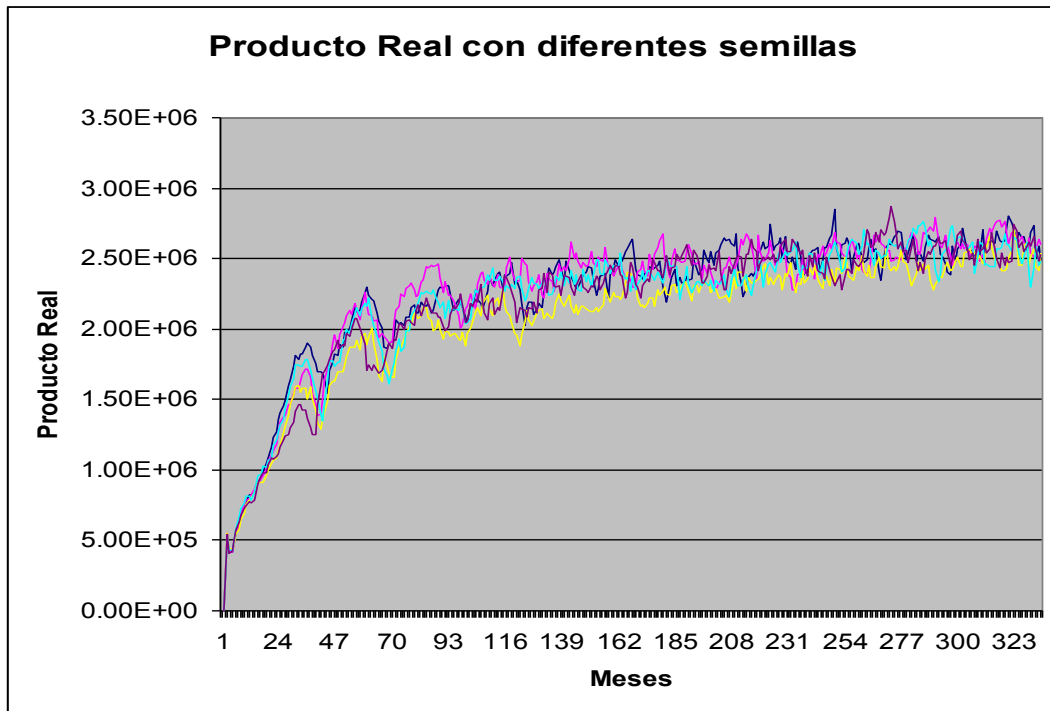


Figura 2.19

### 5.11.3. Equidad

Uno de los principales objetivos del modelo MECHA es mostrar relaciones entre el crecimiento y la distribución de ese crecimiento dentro de la economía. Como índice de distribución de recursos, el modelo muestra el coeficiente Gini para la distribución de la riqueza.<sup>19</sup>

En la serie de tiempo del coeficiente Gini emergen ciclos al mismo tiempo que la serie de crecimiento económico, ver Figura 2.20. Además la tendencia de la serie Gini es decreciente, es decir que el modelo presenta una distribución más igualitaria de los recursos a medida que la economía crece. Sin embargo, en el corto plazo a medida que la economía crece, la distribución de la riqueza es más desigual. En estricto rigor, cada vez que la economía está creciendo, el coeficiente Gini aumenta, ya que son las empresas eficaces las

<sup>19</sup> Para mayor información sobre el coeficiente Gini, ver Anexo.

que lucran con este crecimiento. Una vez que las empresas empiezan a quebrar y se forma el primer ciclo, la distribución de los recursos se hace más igualitaria, la quiebra es absorbida por gran parte de las empresas eficaces. Por esta razón, el coeficiente Gini cae junto con el crecimiento económico. Este ciclo se repite al igual que el ciclo de crecimiento. Es interesante destacar que cada vez que el coeficiente Gini cae, desciende más que lo que aumentó en el ciclo anterior, lo que significa que cada ciclo de crecimiento tiene como resultado una mejor distribución de riqueza. Esto tiene sentido, ya que los ciclos se deshacen de las empresas altamente inefectivas, que son las que poseen menos riqueza dentro del modelo.



Figura 2.20

Finalmente, podemos ver una comparación gráfica y directa de las principales series en la Figura 2.21, donde la correlación entre los ciclos es clara. Las series graficadas en esta Figura son Producto Real  $\times 10^{-6}$ , Gini  $\times 10$  e Índice de Precios. El gráfico muestra una evidente relación entre los ciclos de las series.

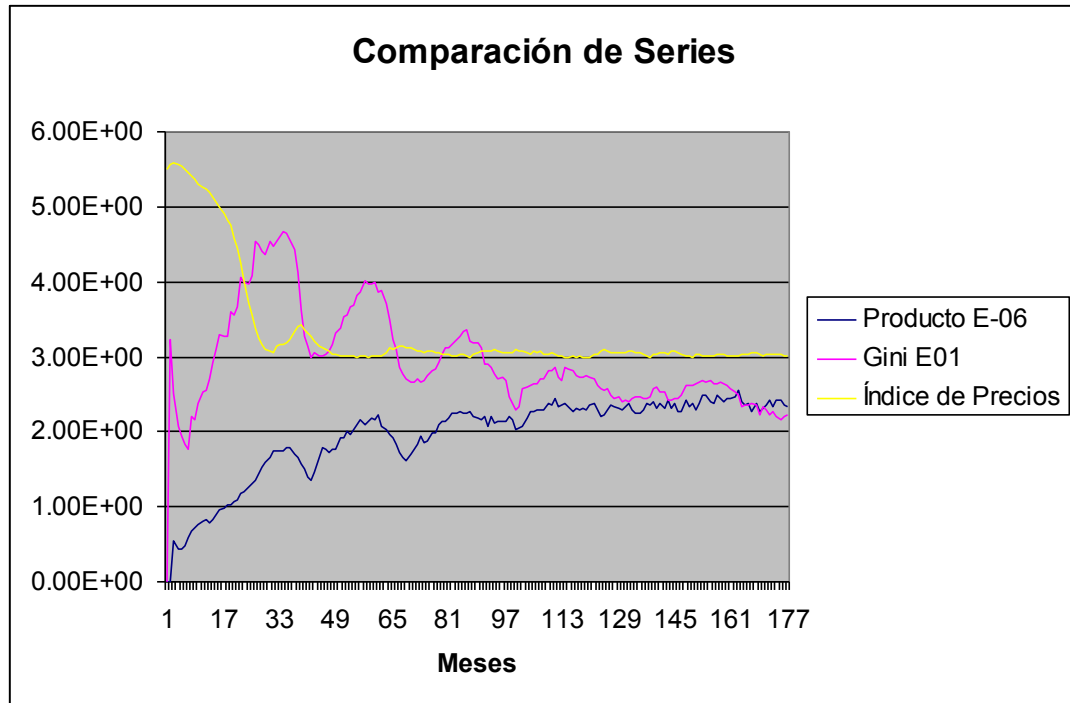


Figura 2.21

## 5.12. Comentarios sobre la Validación Operacional del Modelo

Ya se vio que del modelo emergen ciclos económicos. Es interesante destacar que los ciclos económicos existen en economías de libre mercado, pero no en economías centralizadas. Esto se debe a que los recursos de una economía centralizada están bajo control directo y cualquier recurso no utilizado sería puesto en uso. En un sistema de libre mercado, las empresas pueden despedir trabajadores para mantener utilidades durante un periodo de baja demanda.<sup>20</sup>

<sup>20</sup> Más adelante, se realiza un experimento de redistribución de la riqueza mediante el cobro de un impuesto. Ahí podemos ver como desaparecen los ciclos económicos a medida que centralizamos la economía mediante aumentos en el impuesto.

Joseph Schumpeter (1939) estudió los ciclos económicos. Según Schumpeter la alteración creada por empresarios que entran al mercado e introducen nuevos productos generan un crecimiento económico cíclico. Este proceso fue llamado “destrucción creativa” ya que en el sistema económico se destruyen empresas menos eficientes para que puedan entrar nuevas empresas. El modelo de esta tesis muestra claramente el proceso de destrucción creativa.

Karl Marx (1885) sostuvo que las recesiones económicas eran inevitables en un sistema capitalista.

Numerosos economistas han estudiado los ciclos económicos a través de la historia, entre ellos están Joseph Kitchin (1923), Clement Juglar (1916) y Simon Kuznets (1930) .

Estos estudios respaldan los resultados entregados por el modelo y le dan validez.

## 5.13. Experimento: Redistribución de la Riqueza

Habiendo ya visto los principales fenómenos emergentes del modelo MECHA (crecimiento, desarrollo y ciclos económicos) y sus respectivas correlaciones, se estudia la posibilidad de manipular exógenamente los recursos de la economía y ver así si existe algún beneficio en términos de crecimiento y distribución de recursos. Queremos estudiar el efecto en la economía del sistema, de la redistribución de la riqueza, mediante la recaudación de un impuesto y su posterior distribución en partes iguales a todos y cada uno de los integrantes de la economía.

Este efecto se estudió con cuatro tasas de impuesto distintas, 25%, 50%, 75% y 100%, además de incluir en el análisis la opción inicial de cero impuestos. En definitiva lo que tenemos son cinco modelos económicos que van desde un sistema de libre mercado a uno totalmente centralizado.

El resultado de este experimento es muy interesante. En la Figura 2.22 tenemos el producto real para cada modelo económico. Vemos que con un impuesto de 25% el producto de la economía aumenta considerablemente en relación al modelo capitalista inicial, éste sin impuesto. La economía llega a un producto estacionario de 2.781.851 Unidades Monetarias. Esto significa 274.329 Unidades Monetarias más que el producto al que llega el modelo del sistema sin impuestos, o bien un aumento del 10,9% en el producto. Además, como era de esperar, la distribución de la riqueza es bastante más pareja, ver Figura 2.23. Sin embargo, al aumentar los impuestos a 50% y 75% el producto disminuye a 2.317.825 y 2.344.326 Unidades Monetarias respectivamente, lo que significa una disminución absoluta del producto de 189.697 y 163196 Unidades Monetarias y una disminución relativa del 7,5% y 6,5%, respectivamente. La distribución de los recursos para el modelo de 50% impuesto es aún más pareja que en el caso del impuesto 25%, y para el modelo 75% esta distribución es más pareja que la del modelo 50% impuesto. El modelo con 100% impuesto muestra una baja importante en el producto a 1.987.950 Unidades Monetarias, es decir, 519.572 Unidades Monetarias menos que el producto producido en un sistema libre de impuestos, lo que significa una baja del 20%. Como es de esperar, este modelo presenta la distribución de la riqueza más pareja de todos. Se deduce entonces que la relación entre el producto y los impuestos no es directa, y pareciera ser que un impuesto sobre la riqueza de alrededor del 25% es el indicado para impulsar el crecimiento económico global del sistema.

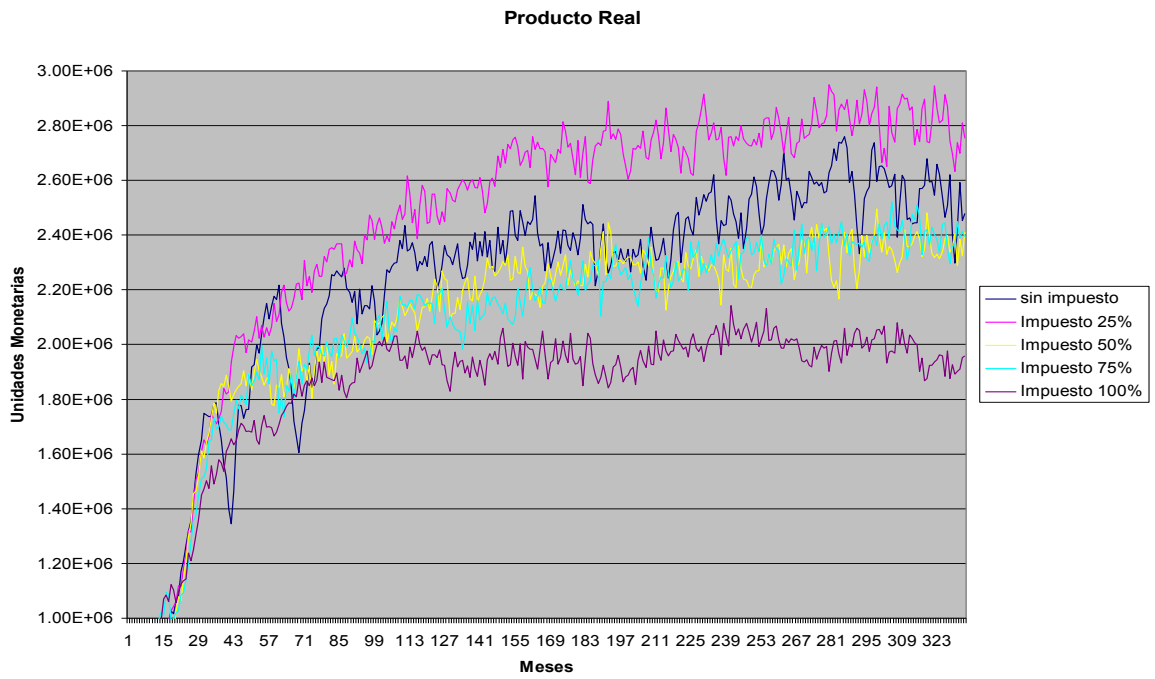


Figura 2.22

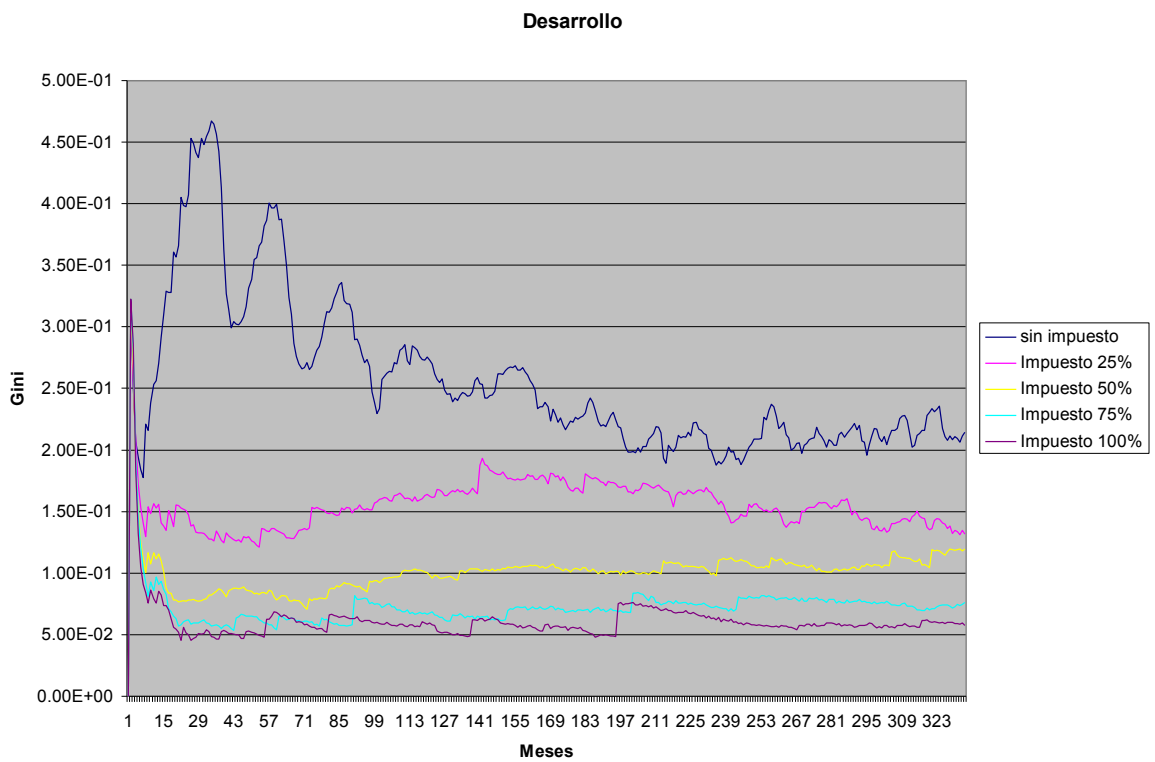


Figura 2.23

Los cinco modelos muestran precios similares. Sin embargo, a mayor impuesto se ve una serie menos cíclica en los precios, ver Figura 2.24.

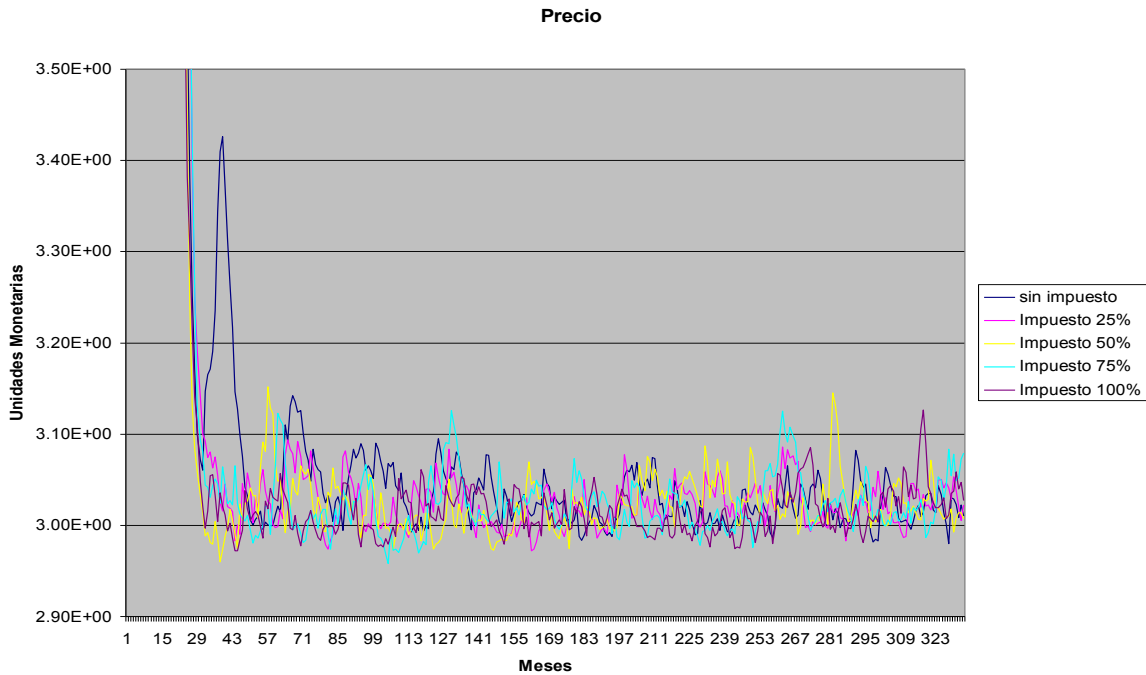


Figura 2.24

¿Por qué la economía crece con impuestos del 25% mientras se contrae con impuestos del 50%, 75% y 100%? Este impuesto tiene el efecto de redistribuir artificialmente la riqueza. Así, se benefician las empresas que se encuentran en peor situación económica a costa de las empresas que se encuentran en mejor situación económica. Dentro del grupo de empresas en mala situación, unas son mejores y otras son peores (más o menos efectivas, en términos del modelo), y no todas (necesariamente) se verán beneficiadas por la redistribución de riqueza. Esto depende del nivel de la tasa de impuesto. Un impuesto bajo logrará ayudar a las mejores empresas pertenecientes al grupo de empresas en mala situación económica del sistema, ya que éstas solo necesitan un pequeño empuje para lograr ser competitivas en el mercado. Un impuesto alto, logrará ayudar a gran cantidad de empresas dentro del grupo de las empresas en mala situación económica del sistema. La redistribución de riqueza sería tal, que las empresas tendrán un ingreso similar, independiente de su eficacia. Esto hace que desaparezcan los ciclos, ya que son muy pocas



las empresas que quiebran, y la evolución de empresas desaparece, limitando el crecimiento económico. Pareciera ser que es necesaria la “limpieza” de empresas altamente inefectivas cada cierto tiempo para el crecimiento económico.

Teniendo como objetivo lograr el máximo crecimiento económico, la idea es lograr un equilibrio en el cual se ayude únicamente a las empresas que realmente necesitan esta ayuda para salir al mercado. Dada la alta volatilidad del producto, sólo podemos afirmar que en el modelo MECHA se observa que ese equilibrio se logra con un impuesto sobre la riqueza que se encuentra entre 23% y 25%.

## Capítulo 6:

# TRABAJO FUTURO

El modelo MECHA es bastante flexible y útil. Dado que está escrito en Objective-C es posible expandir sus funciones sin mayor problema. Además es posible realizar una gran cantidad de experimentos, de los cuales solo se realizó uno.

Sin embargo, la mayor limitante del modelo es su función de demanda marginalmente decreciente, ver Figura 2.12. Esta función fue utilizada en este modelo sobre la base que la riqueza agregada del mismo es cero, es decir unos agentes deben y otros agentes prestan y todos se mantienen dentro de los márgenes de la función de demanda. Sin embargo, si quisiéramos abrir la economía, o modelar una entidad que inyecte capital sin crear deudas (como sería CODELCO para el caso chileno) los agentes sobrepasarían los límites de la función de demanda, logrando así demandas similares para todos los agentes independientemente de su riqueza.

Como desafío futuro, se espera perfeccionar el mecanismo por el cual opera la demanda en el modelo, y así ampliar su capacidad, en un estudio de Doctorado.

## Capítulo 7:

# CONCLUSIONES

El estudio de los fenómenos macroeconómicos usando modelos basados en agentes requiere de estructuras altamente complejas. El modelo entregado en esta tesis posee un mercado de consumo de bienes conforme a los conceptos de Hayek, un mercado del trabajo y un mecanismo que limpia las deudas sobre cierto límite (quiebra).

Dentro del mercado de bienes, emerge un equilibrio espontáneo, tal como lo señaló Hayek en su teoría de consistencia de planes en el equilibrio económico.

El modelo produce un desarrollo económico que tiene una tasa de crecimiento marginalmente decreciente. Esto era de esperar, ya que la función de consumo tiene un límite que impide crecer no sólo a la demanda sino que también a la capacidad de aprendizaje de las empresas, ya que las empresas aprenden en función de la demanda. Una vez que un cierto nivel ha sido alcanzado, la productividad no puede crecer.

Más sorprendente aún es la existencia de ciclos regulares. Debido a las reglas de quiebras, los agentes deudores desaparecen uno tras otro, dando inicio al ciclo. Una vez que se limpió el espacio de agentes altamente deudores, se crea nuevas empresas, dando término al ciclo. Dado que estos ciclos no pueden ser previstos mediante un estudio del modelo, son un fenómeno emergente del mismo.

Otro resultado importante es la tendencia decreciente del coeficiente Gini a largo plazo. A medida que la economía crece, la riqueza es distribuida en forma más pareja, aunque en el corto plazo el crecimiento y la desigualdad se mueven juntos. Un alto volumen de las deudas limita el crecimiento y crea una tendencia a una distribución más desigual de la riqueza. Este fenómeno es observable en el corto plazo del modelo. Para que la economía pueda crecer, es necesario que disminuyan las deudas y se revierta la tendencia a la

distribución desigual. Mediante la quiebra de empresas, la economía absorbe a los deudores, empareja su distribución de riqueza y puede seguir creciendo.

Los precios del mercado mostraron una tendencia constante, con pequeños ciclos y algo de ruido. Los ciclos emergentes en los precios son producto del alza y la baja en la oferta debido a la quiebra cíclica de empresas.

Los ciclos de las series de Producto, Gini y de Precio presentan una alta correlación.

Un experimento con distintas tasas de impuesto mostró que existe un punto óptimo para el nivel de impuestos sobre la riqueza que se ubica entre el 23% y el 25%. Si bien esta tasa de impuesto es válida solo dentro del modelo, su existencia indica que en el sistema económico real existe una tasa de impuestos óptima. Menor o mayor impuesto que este punto producirían contracción económica. Como es de esperarse, la distribución de la riqueza es más pareja mientras mayor es el impuesto. Además, a mayor impuesto los precios son más constantes debido a que la oferta es más fija ya que las empresas quiebran menos.

# ANEXO

El coeficiente Gini, diseñado por el estadístico italiano Corrado Gini, es un número entre cero y uno que mide el grado de desigualdad en la distribución del ingreso en una sociedad determinada. El coeficiente registraría cero (0,0= desigualdad mínima) para una sociedad en la que cada miembro recibiera exactamente el mismo ingreso y registraría un coeficiente de uno (1,0= desigualdad máxima) si un miembro recibiera todo el ingreso y el resto no recibiera nada.

Este coeficiente se basa en la curva de Lorentz, una curva de frecuencia acumulada que compara la distribución de una variable específica con la distribución uniforme que representa equidad total. Esta distribución de equidad total es representada por una línea diagonal, y mientras más se aleja la curva de Lorentz de esta línea, mayor es la desigualdad y por lo tanto mayor es el coeficiente Gini.

En la Figura 3.1 vemos un ejemplo.

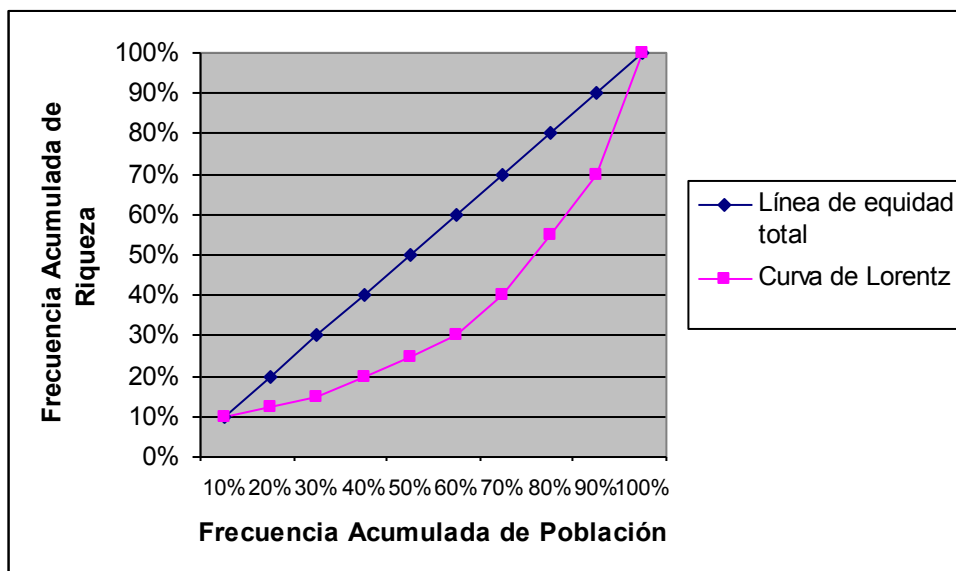


Figura 3.1

## ANEXO

La variable dependiente es la frecuencia acumulada de la población y la variable independiente es la frecuencia acumulada de la riqueza de forma que la curva de Lorentz nos dice que un 40% de la población posee un 20% de la riqueza y que un 90% de la población posee un 70% de la riqueza. Este ejemplo fue creado con fines explicativos, y no representa un sistema económico en particular.

El coeficiente Gini corresponde al doble del área existente entre la curva de Lorentz y la línea diagonal.

# REFERENCIAS

- Auer, P., Burgsteiner, H. y Maass, W. (2001) *The p-delta Learning Rule for Parallel Perceptrons*, entregado para publicación, [http://www.igi.TUGraz.at/maass/p\\_delta\\_learning.pdf](http://www.igi.TUGraz.at/maass/p_delta_learning.pdf). 53
- Battiti, R. (1992) *First and Second Order Methods for Learning: Between Steepest Descent and Newtons Method*, *Neural Computation*, 4, Pags. 141-66. 15, 17
- Bruun, C. y Luna, F (1999) *Endogenous Growth with Cycles in a Swarm Economy: Fighting Time, Space, and Complexity*, *Economic Simulations in Swarm: Agent-Based Modelling and Object Oriented Programming*, Kluwer Academic Publishers. 57
- Carpenter, G. y Grossberg, S. (1986) *Neural Dynamics of Category Learning and Recognition: Attention, Memory Consolidation and Amnesia*, *Brain Structure, Learning and Memory*, (eds J. Davis, R. Newburgh and E. Wegman), AAAS Symposium Series. 3
- Carpenter, G. y Grossberg, S. (1987) *ART 2: Self-organization of Stable Category Recognition Codes for Analog Input Patterns*, *Applied Optics*, 26(23), pags. 4919-30. 3
- Caudill, M. y Butler, C. (1990) *Naturally Intelligent Systems*, MIT Press, Cambridge, MA. 4
- Clower, R.W. y Johnson, M.B. (1976) *Income, Wealth and the Theory of consumption*, pags. 76-129. 57
- Farren, D. (2004), *Experimento Económico Basado en Agentes*, Monografía para el curso Teoría Avanzada de Decisiones, Universidad Adolfo Ibañez. 1, 19
- Gini C. (1912), *Variabilità e Mutabilità*, Reimpreso en *Memorie di Metodologica Statistica*, Editorial Pizetti E., Salvemini T., Rome, 1955. 78
- Grossberg, S. (1982) *Studies of Mind and Brain, Boston Studies in the Philosophy of Science*, D. Riedel Publishing Company, Boston. 3
- Hayek, F. A. (1937) *Economics and Knowledge*, Individualism and Economic Order, University Of Chicago Press, Reissue edition, Chicago, IL, pags. 33, 37, 38
- Hayek, F. A. (1937) *Economics and Knowledge*, Individualism and Economic Order, University Of Chicago Press, Reissue edition, Chicago, IL, pags. 33-56. 32
- Hayek, F. A. (1949) *Individualism and Economic Order*, University Of Chicago Press, Reissue edition, Chicago, IL 36
- Hebb, D.O. (1949), *Organization of Behaviour*, Science Editions, New York, NY. 3
- Hecht-Nielsen, R. (1987) *Counterpropagation Networks, Proceedings of the IEEE First International Conference on Neural Networks* (eds M. Caudill and C. Butler), SOS Printing, San Diego, CA. 3
- Hornik, K., Stinchcombe, M. y White, H. (1989) *Multilayer Feed-Forward Networks are Universal Approximators*, *Neural Networks*, 2, pags. 359-66. 10, 18

## REFERENCIAS

- Hornik, K., Stinchcombe, M. y White, H. (1990) *Universal Approximation of an Unknown Mapping and its Derivatives using Multilayer Feed-forward Networks*, Neural Networks, 3, págs. 551-60. 18
- Jordan, M.I. (1986) *Attractor Dynamics and Parallelism in a Connectionist Sequential Machine*, Proceedings of the 8th Annual Conference of the Cognitive Science Society, Lawrence Erlbaum Associates, Hillsdale, NJ. 7
- Juglar, C. (1916) *A Brief History of Panics and Their Periodical Occurrence in the United States*, A. M. Kelley; 3rd edition, USA. 70
- Kauffman, S. (1995) *At Home in the Universe*, Oxford University Press, Oxford. 26
- Kitchin, J. (1923) *Cycles and Trends in Economic Factors*, Review of Economics and Statistics, Harvard Press, Vol, 5, págs 10-16 70
- Kohonen, T. (1984) *Self-organization and Associative Memory*, Series in Information Sciences, Springer-Verlag, Berlin. 3
- Kuznets, S. (1930) *Equilibrium Economics and Business Cycle Theory*, The Quarterly Journal of Economics, Harvard Press, Vol. 44, págs. 381-415, USA. 70
- Le Cun, Y. (1985) *Une Procédure d'Apprentissage pour Réseau à Seuil Asymétrique*, Proceedings of Cognitiva 85, Paris, págs. 559-604. 3
- Lee, T., White, H y Granger, C.W.J. (1989) *Testing for Neglected Nonlinearity in Time Series Models*, UCSD Discussion paper. 18
- Marx, K. (1885) *Capital: Critica de la Economía Política*, reedición de 1934, Ed. Cenit, Madrid, España. 70
- Minsky, M. y Papert, S. (1969) *Perceptrons - An Introduction to Computational Geometry*, MIT press, Cambridge, MA. 3
- Parker, D. (1985) *Learning Logic*, MIT Center for Computational Research in Economics and Management Science, T.R. 47. 3
- Rand, A. (1984) *Faith and Force: The Destroyers of the Modern World*, Philosophy, Who Needs It?, SIGNET Books, New York, NY, pág. 62. 30
- Rosenblatt, R. (1959) *Principles of Neuro-Dynamics*, Spartan Books, New York, NY. 3, 50
- Rumelhart, D.E. y McClelland, J.L. (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1: *Foundations*, MIT, Press, Cambridge, MA. 3, 10
- Sargent, R. (1988) *A tutorial on validation and verification of simulation models*, Winter Simulation Conference, Proceedings of the 20th conference on Winter simulation, San Diego, CA, págs. 33-39. 29, 31
- Schumpeter, J.A. (1939) *Business Cycles : a Theoretical, Historical and Statistical Analysis of the Capital Process*, Mc Graw-Hill. New York, US. 70
- Smith, A. (1776) *La Riqueza de las Naciones*, reedición de 1955, Editorial Bosh, Barcelona, España. 52



## REFERENCIAS

- Suazo, I. (2004) *Selección de Arquitecturas en Redes Neuronales Feed-forward*, Tesis de Magister en Ciencias, Universidad Adolfo Ibañez. 18
- Waldrop, M. (1992) *Complexity: The Emerging Science at the Edge of Chaos*. Simon & Schuster, New York, NY. 26
- Werbos, P. (1974) *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*, Disertación no publicada de Phd, Harvard University, Department of Applied Mathematics. 3
- White, H. (1988) *Multilayer Feed-forward Networks can Learn Arbitrary Mappings: Connectionist Nonparametric Regression with Automatic and Semi-Automatic Determination of Network Complexity*, UCSD Discussion Paper. 18
- White, H. (1991) *Learning in Artificial Neural Networks: A statistical Perspective*, Neural Computation, 1, pags. 425-64. 10, 15
- White, H. (1992) *Artificial Neural Networks: Approximation and Learning Theory*, Blackwell, Oxford. 3, 17
- Widrow, B. y Hoff, M. (1960) *Adaptive Switching Circuits*, 1960 IRE WESCON Convention Record, Institute of Radio Engineers, New York, NY. 3, 52, 53