

# COMPORTAMIENTO ADAPTABLE

K. VELA VELUPILLAI (\*)

"Lo que comúnmente se denomina adaptación, al traducirse a términos de mecánica es equilibrio directo"

Herbert Spencer, 1866 (1996), pág. 234

## I INTRODUCCION

"Se puede preguntar si existen soluciones o trayectorias para un modelo [adaptable, evolutivo, económico] y cuál es el carácter dinámico y económico de las mismas. En particular, podemos preguntar si cualquiera de esas soluciones tendría el carácter de un equilibrio general e intertemporal, y en qué condiciones la economía evolutiva adaptable convergería hacia uno".

Day, 1993, pág. 21

Fritz Machlup describió una vez la palabra *estructura*, según se emplea en economía, como una palabra equívoca<sup>1</sup> — "una palabra que hace una declaración evasiva o engañosa"<sup>2</sup>. Algo similar puede ser sugerido por el creciente uso que se está haciendo de los términos *adaptable*, *evolutivo*, *cuasi racional*, *restringidamente racional* y muchos términos evocativos análogos aprovechados por los teóricos imaginativos para describir el comportamiento económico que parece ser más realista en sentido descriptivo y en sentido predictivo. Como Holland y Miller, ellos mismos contribuyentes influyentes de un fascinante programa de investigación so-

---

(\*) Departamento de Economía Queen's University - Belfast Belfast BT7 1NN Reino Unido y  
Departamento de Economía Universidad de las Américas - Puebla Apartado Postal 332 Sta.  
Catarina Mártir 72820 Puebla México 17 de noviembre de 1997

bre modelación de entornos económicos como *complejos sistemas adaptables* poblados por *agentes artificialmente adaptables* (AAA):

"Habitualmente hay sólo una forma de ser totalmente racional, pero existen muchas formas de ser menos racional".

Holland y Miller (1991, pág. 367)

Esto sugiere que las diversas formas en que teorizamos acerca de la racionalidad menos que total — como agentes adaptables, restringidamente racionales, evolutivos, etc. — deben considerarse *aproximaciones*, de una u otra clase para la "única forma de ser totalmente racional". Esto está ejemplificado, por ejemplo, en ... la definición de la racionalidad aproximada a como restringidamente racional:

Pero ¿podemos no seguir el otro camino? ¿Puede no tratarse de "racionalidad total" en la aproximación a algo más "realista" en sentido descriptivo y en sentido predictivo? ¿Realmente incluso en un sentido normativo? Más concretamente, ¿no hay más que una forma de "ser totalmente racional"? Volveré a tratar algunos de estos temas más adelante en este capítulo y en otros capítulos pero, por el momento, puede ser útil recordar que las denominadas teorías de la "racionalidad total" sólo son buenas en cuanto a ser aproximadamente correctas en sentido descriptivo y en sentido predictivo. Entonces, ¿por qué un modelo "totalmente racional" de comportamiento racional debería ser un punto de referencia — si sus aptitudes predictivas y descriptivas son sólo aproximadamente correctas? Seguramente, es al revés que en las ciencias naturales y puras. John Harsanyi, uno de los grandes defensores de las teorías normativas del comportamiento humano, del modelo totalmente racional, planteó algunos de estos temas clave de la siguiente forma:

"Los seres humanos a menudo no alcanzan la racionalidad perfecta postulada por las teorías normativas del comportamiento racional ... Uno de los motivos es la *limitada capacidad de procesamiento de información* del sistema nervioso central humano, destacada por la teoría de Simon de la "racionalidad limitada".

... En los hechos reales estas limitaciones obvias de la racionalidad humana por lo general no impiden que las teorías del comportamiento racional arrojen *predicciones aproximadas razonablemente*

*buenas* acerca del comportamiento social a través de una amplia gama de situaciones sociales. Por supuesto, será necesaria gran cantidad de experimentación ulterior con *varios tipos de teorías del comportamiento racional* antes de poder afirmar qué tan lejos nos pueden llevar estas teorías al explicar y predecir el comportamiento social empírico en cada campo en particular".

Harsanyi, 1977, pág. 17; segundo juego de cursiva agregado.

Luego, ¿existe más de "una forma de ser totalmente racional", como parece sugerir Harsanyi? Esta es mi opinión y trataré de exponer mis motivos para apoyar un modelo alternativo totalmente racional en el presente capítulo y los siguientes. En mis obras anteriores he identificado un agente totalmente racional con una Máquina Turing (aceptando la tesis de Church-Turing).

En la próxima sección I trato de examinar la caracterización teórica de recursión de una definición de comportamiento adaptable recientemente sugerida por Lucas (1986). En la sección III continúo examinando la conexión entre el denominado comportamiento totalmente racional y el comportamiento adaptable, pero esta vez de una forma más realista a nivel descriptivo. Finalmente, en la sección IV que contiene las conclusiones, trato de deducir algunos indicios, sobre la base de la analítica en la sección II y la sección III, en el camino hacia definiciones de formas alternativas de ser totalmente racional.

## II COMPORTAMIENTO RACIONAL DINAMICO

"Se podría preguntar si [los procesos que involucran más o menos seguimiento mecánico de una *regla de determinación*] implican mucha opción real al utilizar el sentido diario de dicha expresión"

Nelson y Winter, 1982, pág. 68; cursiva agregada.

Lucas (op. cit) analiza una caracterización dinámica del comportamiento racional al sugerir que "todo comportamiento es adaptable" (ibid, pág. S408; cursiva agregada) "ya que las funciones de demanda tradicionales son reglas de determinación" (ibid, pág. S410) y a su vez éstas son "constantes de cierto proceso de adaptación" (pág. S402, cursiva agrega-

da) "que han llegado luego de un proceso de deliberada experimentación y evaluación de resultados" (pág. S410). La evaluación de resultados se realiza por medio del tradicional formalismo de preferencias. De este modo, Lucas parece estar haciendo explícitos los formalismos subyacentes de la *dinámica implícita* del comportamiento racional. De hecho, Lucas manifiesta explícitamente:

"El comportamiento supuesto por estas reglas de determinación es "racional" en el sentido en que los economistas emplean dicho término. Pero no sólo es compatible con el comportamiento adaptable, de ensayo y error, sino que el experimento concebido para descubrir esta racionalidad presupone que, de existir, constituye el resultado de cierto proceso (indeterminado) de adaptación".

(ibid, pág. S4; cursiva agregada)

Para examinar las caracterizaciones e implicaciones teóricas de recursión trabajaré con las tres equivalencias siguientes:

- (a). Se *supondrá* que los procesos de adaptación son sistemas dinámicos (en el sentido técnico estricto de la expresión);
- (b). A partir de Máquinas Turing pueden construirse sistemas dinámicos capaces de universalidad computacional;
- (c). El comportamiento racional por parte de un agente económico es equivalente al comportamiento computacional de una Máquina Turing;

Teniendo a mano estas equivalencias será posible formular preguntas teóricas de recursión acerca de las "constantes de ... procesos de adaptación", es decir, constantes de sistemas dinámicos o la configuración con pausa de las Máquinas Turing que, por (c), representa la opción racional apropiada. Ahora, la equivalencia que figura en (c) ha sido demostrada en Velupillai (1997). Las equivalencias de (b) serán indicadas en esta sección, mientras (a), como ya se dijo, simplemente será supuesta. La hipótesis de la equivalencia en (c) es bastante corriente y, en realidad, está implícita en las inspiradas descripciones de Lucas.

Para mostrar una construcción particularmente simple de un sistema dinámico que incluye las actividades computacionales de una Máquina Turing (y, por lo tanto, el comportamiento racional de un agente económico) describiré, en forma concisa, una ingeniosa construcción reciente de Moore (1995). Sin embargo, antes de proceder a la descripción de la construcción de Moore, permítaseme señalar que, desde que von Neumann (1966) construyó un autómata celular de 29 estados (CA), se han logrado varios éxitos en la construcción de sistemas dinámicos simples capaces de soportar la universalidad computacional — es decir, al simular el comportamiento de una Máquina Turing (universal). Más recientemente, por ejemplo, Berlekamp y otros (1982) han mostrado la forma en que puede implementarse la universalidad computacional en la configuración inicial del juego LIFE.

Ahora, el punto de partida de Moore era una construcción sugerida por Minsky en su obra seminal (Minsky, 1967, capítulos 11 y 14). Empleando lo que él denominó equivalencia de Máquina de Programa de la Máquina Turing — llamada Máquina de Registro en la bibliografía más reciente. Minsky mostró, en los teoremas 14.1.1 y 14.2.1 (ibid) cómo construir una Máquina de Programa con sólo 1 o 2 registros. El resultado de 2 registros es:

"Para cualquier Máquina Turing  $T$  existe una Máquina de Programa  $M_T$  con sólo dos registros que se comporta igual que  $T$  cuando se hace arrancar con cero en un registro y  $2^l 3^m 5^n$  en el otro registro."

ibid, pág. 257

Moore, luego, combina los estados finitos de la Máquina Turing planteando:

$$I = 2^l 3^m 5^n \quad (1)$$

y dejando:

$$x = \eta I + s \quad (2)$$

Donde:

$\eta$  : número (finito) de estados en la Máquina Turing dada

$s$ : 0, 1, ..., n-1

A continuación, Moore incluye esto en un mapa analítico de los reales. Luego, para verificar si la máquina se encuentra en el estado  $s$  o no, definir:

$$h_s(x) = \left( \frac{\sin \pi x}{n \sin \frac{\pi(x-s)}{n}} \right)^2 \quad (3)$$

Se puede apreciar que:

$$h_s(x) = 1 \text{ en valores enteros si } x \bmod n = s$$

y, de lo contrario,

$$h_s(x) = 0$$

A continuación, para verificar si el registro  $k_s$ , para  $k_s = 2, 3$  o  $5$ , es igual a cero, calcular:

$$g_s(I) = \left( \frac{\sin \pi I}{k_s \sin \frac{\pi I}{k_s}} \right)^2 \quad (4)$$

Luego:  $k_s = 0 \Leftrightarrow g_s(I) = 1$  si  $I \bmod k_s = 0$

y:  $k_s = 1 \Leftrightarrow g_s(I) = 0$  si  $I \bmod k_s \neq 0$

Luego, como en una Máquina Turing estándar, dependiendo del símbolo leído y del estado actual, existe una transición; o sea, dependiendo de  $s$  y de a cuál de los registros se está accediendo,  $I$  debe ser modificado convenientemente y la Máquina Turing debe ir a un nuevo estado, digamos  $t$ . Llamémosles  $\alpha_s$ ,  $\beta_t$ ,  $\gamma_s$  y  $\delta_t$  respectivamente. Luego la transición puede ser afectada calculando:

$$f_s(x) = g_s \left( \left( \frac{x-s}{n} \right) (x-s) \gamma_s + \delta_t \right) + \left( 1 - g_s \left( \frac{x-s}{n} \right) \right) ((x-s) \alpha_s + \beta_t) \quad (5)$$

Si definimos:

$$f(x) = \sum_{s=1}^{n-1} h_s(x) f_s(x) \quad (6)$$

Este es el mapa que contiene la Máquina de Programa de Minsky que, a su vez, simula la Máquina Turing original que es una representación del comportamiento de opciones del agente económico racional. Además este mapa, como un sistema dinámico y una Máquina Turing, tiene incluida en él la indeterminación del problema de la pausa. Las indeterminaciones de los conjuntos de límites de su dinámica, por ejemplo, no serán debidas a una dependencia sensible de las condiciones iniciales, es decir, aunque se conozcan exactamente las condiciones iniciales, existirán indeterminaciones de propiedades particulares de los conjuntos de límites de la dinámica de los mapas.

A un nivel más general, una vez que se efectúa la construcción de Minsky, el camino de (1) a (6) puede ser repetido fácilmente para sistemas dinámicos hechos a la medida de prescripciones específicas. Además, recordemos que, en un sentido preciso, toda Máquina Turing es también una Máquina Turing Universal, es decir, que dicha máquina puede ser programada para computar cualquier cómputo teóricamente posible e imitar la acción de cualquier Máquina Turing arbitraria. A los efectos de facilitar la referencia en los resultados a ser planteados y demostrados como teoremas necesito dos definiciones simples y sintéticas.

### Definición 1

Se dice que un sistema dinámico — discreto o continuo — es *capaz de universalidad computacional* cuando, utilizando sus condiciones iniciales, puede ser programado para simular la acción de cualquier Máquina Turing arbitraria, en particular, la acción de una Máquina Turing Universal.

### Definición 2

Un sistema dinámico capaz de universalidad computacional se denominará *Sistema Dinámico Universal*.

### Observación 1

Obviamente hubiera sido suficiente para mis fines utilizar simplemente una de las Cas — digamos la configuración LIFE pertinente — que ha demostrado ser capaz de universalidad computacional como el sistema dinámico universal. Pero la construcción de Moore tiene las virtudes de ser: (a) simple, (b) repetible, y (c) de baja dimensión (en realidad, de 1 dimensión).

Ahora puedo volver a la conexión sugerida por Lucas entre el comportamiento adaptable y la racionalidad, y establecer dos resultados precisos.

### Teorema 1

Sólo los procesos de adaptación capaces de universalidad computacional son compatibles con la racionalidad "en el sentido en que los economistas emplean dicho término".

### Demostración

Recordemos, de resultados anteriores, que el agente racional puede ser modelado como una Máquina Turing y que el comportamiento racional del agente económico es equivalente al comportamiento computacional apropiado de una Máquina Turing (equivalencia (c) que antecede). Por la construcción anterior se mantiene la equivalencia (b) y (a) es una suposición. Luego, por el resultado de Kramer sobre autómatas *finitos* (proposición 3.4) sabemos que tales Máquinas Turing, restringidas, son incompatibles con la racionalidad "en el sentido en que los economistas emplean dicho término". Por lo tanto, se deduce que cualquier sistema dinámico modelado por medio de autómatas no *capaces de universalidad computacional* es análogamente incompatible, y de ahí que sólo los procesos de adaptación *capaces de universalidad computacional* son compatibles con la racionalidad "en el sentido en que los economistas utilizan dicho término".

L.Q.Q.D.

## Teorema 2

No existe ningún procedimiento eficaz para saber si determinadas clases de reglas de determinación son "constantes de (cierto) proceso de adaptación".

## Demostración

De las discusiones, suposiciones y resultados que anteceden, observemos en primer lugar que las "constantes de procesos de adaptación" son "constantes de sistemas dinámicos". Luego, dada una clase de reglas de determinación, preguntar si corresponden a constantes de sistemas dinámicos es equivalente a preguntar si las mismas pertenecen a la cuenca de atracciones de (cierto) sistema dinámico tal. Sin embargo, por el teorema anterior, el sistema dinámico en cuestión debe ser capaz de universalidad computacional, o sea, un sistema dinámico universal. Ya que sino, las constantes violarían alguna forma del postulado de racionalidad.

A continuación, afirmo que las cuencas de atracción de los sistemas dinámicos universales son recursivamente enumerables pero no recursivas. El ser recursivamente enumerables se deduce del simple hecho de que las trayectorias que les pertenecen pueden ser efectivamente listadas probando sistemáticamente las condiciones iniciales.

Por otro lado, supongamos que son recursivas. Entonces el problema de la *pausa* tendría solución. Puesto que, dada la cuenca de atracción y puntos iniciales arbitrarios, esto significará que la dinámica efectivamente contestará a las preguntas de los miembros. Ya que se trata de un sistema dinámico universal significa que puede contestarse una pregunta acerca de si una Máquina Turing Universal se detendrá o no, por entradas arbitrarias, en una configuración en particular. Esto contradice la irresolubilidad del problema de la *pausa*. Por lo tanto, por el teorema de Rice, no existe ningún procedimiento eficaz para determinar si cualquier clase dada de reglas de determinación es la constante de (cierto) proceso de adaptación (o no).

L.Q.Q.D.

Estos resultados, en pocas palabras, expresan que no existe ningún procedimiento sistemático, disponible *a priori*, para examinar y caracterizar las propiedades a largo plazo de los sistemas dinámicos interesantes.

Por "sistemas dinámicos interesantes" quiero decir los que pueden encapsular el comportamiento totalmente racional. Cualquier dinámica económica que lleve a conjuntos de límites que son sólo puntos límite (nodos, focos, puntos de concavidad, incluso ciclos límite) no puede representar el comportamiento totalmente racional en un sentido muy preciso. Es un aspecto de este sentido preciso lo que he intentado captar en los dos teoremas que anteceden.

Para decirlo de otra forma, la dinámica económica no capaz de universalidad computacional será incompatible con la hipótesis del no arbitraje: existirán redituables oportunidades inexplotadas en el límite. Esta es obviamente una de las razones clave de la necesidad de complementar la dinámica simple con *sensacionalismos ad hoc*<sup>3</sup> en las teorías del crecimiento y el ciclo, por ejemplo.

Pero, como siempre, estos resultados aparentemente negativos tienen aspectos positivos. Después de todo, estos resultados niegan solamente la viabilidad de los *procedimientos efectivos* utilizables, para contestar preguntas *sistemáticas* acerca de la opción racional (a nivel dinámico). Ellos no afirman que no existen *procedimientos con fines especiales*, o que no pueden encontrarse los mismos, para analizar los *procesos de la opción racional*.

Por ejemplo, aun cuando Lucas sugiere que "todo comportamiento es adaptable" y que las reglas de determinación, tales como las funciones de demanda, son "constantes de cierto proceso de adaptación", parece haber tenido siempre presente un estudio de la dinámica con fines especiales.

"En términos generales, consideramos o modelamos un individuo como un grupo de reglas de determinación (normas que dictan la acción a emprender en determinadas situaciones) y un conjunto de preferencias utilizadas para evaluar los resultados que surgen de combinaciones de situaciones y acciones en particular. Estas reglas de determinación están siendo constantemente revisadas y modificadas, se prueban y comprueban nuevas reglas de determinación frente a la experiencia, y las normas que dan resultados deseables sustituyen a las que no los dan. Empleo el término "adaptable" para referirme a este proceso de ensayo y error por medio del cual se determinan nuestras modalidades de comportamiento". *ibid*,pág.S-401

Contra el telón de fondo establecido por el formalismo de la presente sección, y los resultados obtenidos, ¿cómo podemos abordar el estudio sistemático de los "procesos de ensayo y error que demostraron funcionar en una cantidad de situaciones", desde un punto de vista teórico de recursión? Algunas respuestas sugeridas componen el tema de la sección siguiente.

### **III UNA BASE COMPUTABLE PARA EL ESTUDIO DE "PROCESOS DE ENSAYO Y ERROR" EN LA OPCION RACIONAL**

"Es totalmente válido considerar la marcha de una economía como el proceso de computar respuestas a los problemas planteados a la misma".

Goodwin, 1951, pág. 1

Richard Goodwin hizo la interesante sugerencia anterior, la que, según tengo entendido, fue la primera declaración explícita con un contenido económico computable. Lo que Goodwin sugería, tantos años atrás, era que cualquier sistema dinámico que encapsulara "la marcha de una economía" podría interpretarse como una adecuada máquina computadora; si, además de eso, "la marcha de una economía" está orientada por los procesos de la opción racional o representa los mismos, entonces la máquina computadora debe ser una Máquina Turing y el sistema dinámico debe ser un sistema universal. Junto con el concepto de Goodwin mis observaciones y resultados de las secciones que anteceden al admirable resumen de Ian Stewart de los recientes resultados de la interconexión entre la teoría de los sistemas dinámicos y la computabilidad:

"Una enorme gama de interrogantes ... ha demostrado ser indeterminable. La indeterminabilidad alcanza a muchas interrogantes en la teoría de los sistemas dinámicos. Entre éstas están si la dinámica es caótica, si la trayectoria a partir de un punto inicial dado eventualmente pasa por cierta región específica del espacio de fase ... En realidad, virtualmente cualquier interrogante "interesante" acerca de los sistemas dinámicos es — por lo general — indeterminable".

Stewart, 1991, pág. 664

Los resultados de la sección anterior podrían considerarse como un trampolín hacia un intento de dar contenido positivo a estas fértiles ideas. En esta sección doy otros pasos descriptivamente conjeturales en la misma dirección.

A partir del formalismo y los resultados del capítulo III es fácil ver que la función proposicional que debe evaluar el agente económico racional, o la función que debe computar la Máquina Turing, incluye — genéricamente — la **O EXCLUYENTE**; es decir, dadas las proposiciones  $X$  e  $Y$ , el agente racional debe evaluar el valor de verdad de  $X$  O SINO de  $Y$ . Esto es verdadero si, y sólo si, exactamente una de las proposiciones es verdadera.

Lo que es aún más importante, e independientemente de estas observaciones que anteceden, cada proposición (en el cálculo predicado) puede expresarse como una cadena de:

$$\sim [(X \wedge \sim Y) \vee (Y \wedge \sim X)] \quad (7)$$

(para la proposición dada  $X$  e  $Y$ ).

Esto, es decir, (7), es la negación de la O EXCLUYENTE:

$$\begin{aligned} X \text{ O SINO } Y &= (X \vee Y) \wedge (\sim (X \wedge Y)) & (8) \\ &= (X \wedge \sim Y) \vee (Y \wedge \sim X) & (8a) \end{aligned}$$

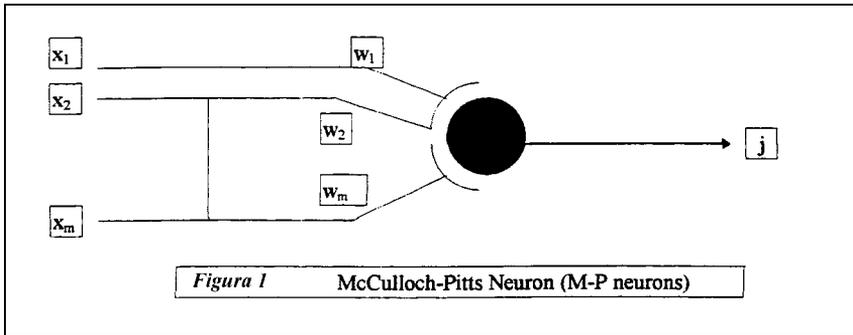
La negación de (8a) da (7), que, dada su propiedad de ser capaz de expresar cada proposición (en el cálculo predicado) como cadenas de sí misma es llamada una **proposición universal**. En otras palabras (7) es una **proposición universal** en el mismo sentido en que definimos **sistemas dinámicos universales**: *es capaz de universalidad computacional*.

Recordemos ahora que las razones convencionales dadas para invocar comportamiento restringidamente racional por parte de los agentes económicos son las fisiológicas, las físicas y otras limitaciones naturales a la capacidad de procesamiento de la información de los seres humanos. Para citar a Harsanyi nuevamente: "Los seres humanos ... no alcanzan la racionalidad perfecta [debido a la] *limitada capacidad de procesamiento de información del sistema nervioso central humano* ..." En nuestro caso se requiere que el "sistema nervioso central humano" procese cadenas de la proposición universal. ¿Cuál es la arquitectura de la unidad mínimamente

compleja del "sistema nervioso central humano" capaz de procesar la proposición universal? Utilizando las unidades neuronales estándar McCulloch-Pitts podemos analizar algunas respuestas a esta pregunta.

(a) Definición de Neuronas McCulloch-Pitts (neuronas M-P)

*Notas Elementales sobre Redes Neurales*



Las neuronas M-P se caracterizan por

- (i)  $m + 1$  números: entradas  $\rightarrow x_1, x_2, \dots, x_m$   
salida  $\rightarrow y$ .
- (ii) Número umbral:  $\theta$
- (iii) Ponderaciones "sinápticas":

$w_1, w_2, w_3, \dots, w_m$ , (en donde  $w_i$  está asociada con  $x_i$  e  $y_i$ ). El período refractario se toma como la unidad de tiempo.

*Supuesto:*  $y_j(n+1) = 1$  si  $\sum w_{ij} x_i(n) \geq \theta$ , es decir, la neurona dispara un impulso a lo largo de su axón en el momento  $n + 1$  si la suma ponderada de sus entradas en el momento  $n$  excede el umbral de la neurona.

$w_{ij} > 0$  - sinapsis estimuladora  
 $w_{ij} \leq 0$  - sinapsis inhibitoria

(b) Definición de Red Neural

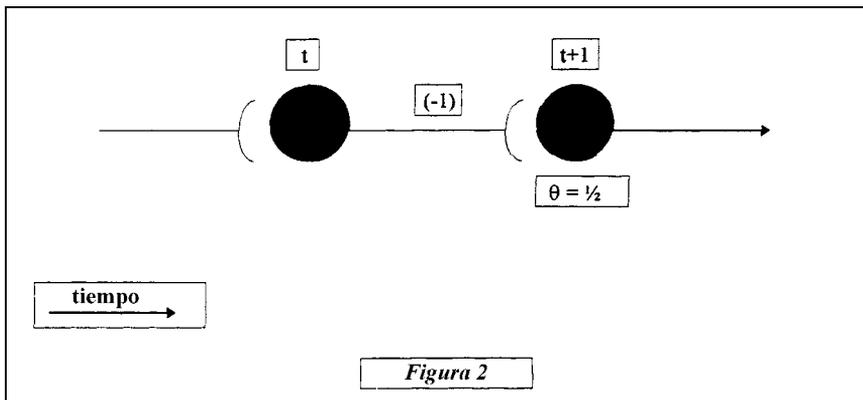
Una *red neural* es un conjunto de neuronas M-P, cada una con el mismo período refractario, interconectadas dividiendo la salida de cualquier neurona ("axón a las dendritas") en un número de líneas y conectando algunas o todas ellas a las entradas de otras neuronas.

*Proposición* - Toda red neural es un autómata finito.

(c) Ejemplos Simples de Redes Neurales para Evaluar Proposiciones

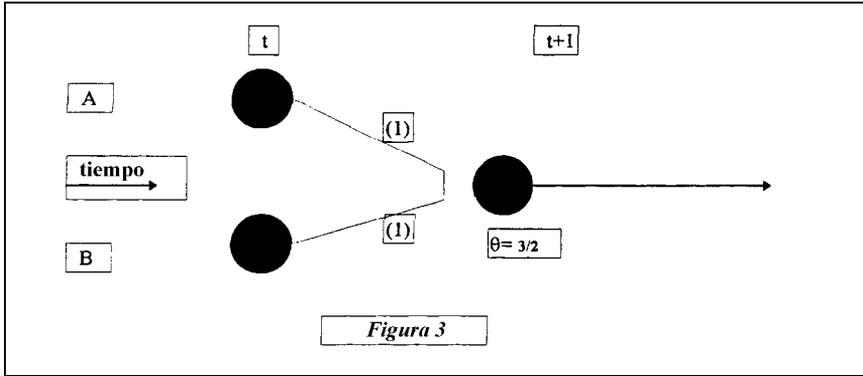
Tres ejemplos simples de cómo las redes neurales evalúan las proposiciones Booleanas prepara el fundamento para el ejemplo principal, universal, del texto.

## a. Negación



- (i) La neurona izquierda emite una "oleada" en el momento  $t$  si cierta proposición  $A$  es verdadera.
- (ii) Si el canal de transmisión tiene ponderación,  $w, (-1)$ , y la neurona derecha tiene el umbral  $\theta = -1/2$ , entonces, en el momento  $t + 1$ , la neurona derecha no emitirá una "oleada", es decir,  $A \rightarrow \sim A$ .
- (iii) Contrariamente, cuando  $A$  es falsa:  $\sim A \rightarrow A$ .

(b) Conjunción



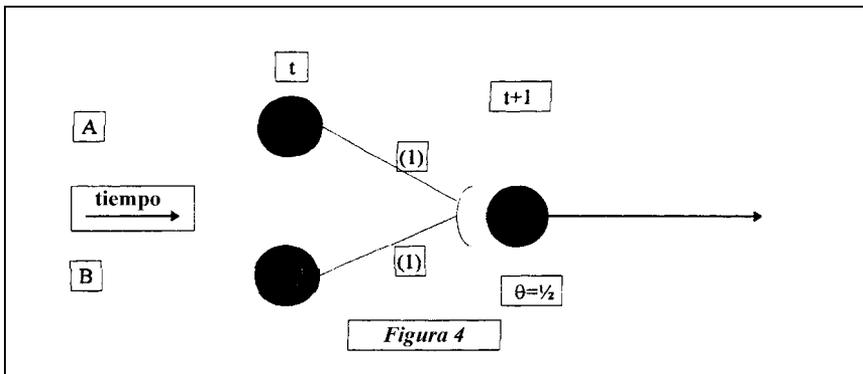
La neurona derecha emitirá la oleada en  $t + 1$  si ambas neuronas de la izquierda emitieron la oleada en el momento  $t$ , indicando que las proposiciones probadas por estas dos neuronas ambas han sido verdaderas. Por lo tanto:

$$Aw_1 + Bw_2 = 1 \times 1 + 1 \times 1 = 2 > \theta = 3/2$$

cuando  $A$  y  $B$  son verdaderos. (Nótese el peligro de formalismos diofantinos.) Si, por otra parte,  $A$  es verdadero y  $B$  es falso,

$$1 \times 1 + 0 \times 1 = 1 < \theta = 3/2$$

(c) La O Inclusiva



Si al menos una de las neuronas izquierdas emite la oleada en el momento  $t$ , la neurona derecha emitirá la oleada en  $t + 1$ . Por lo tanto, si  $A$  es verdadero (1), y  $B$  es falso (0), tenemos

$$1 \times 1 + 0 \times 1 = 1 > \theta = 1/2$$

Nada más que una colección de estas, e igualmente simples, unidades elementales son la base de una arquitectura de redes neurales para computar cualquier cosa que una Máquina Turing pueda computar.

(d) Una Red Neural para X O SINO Y

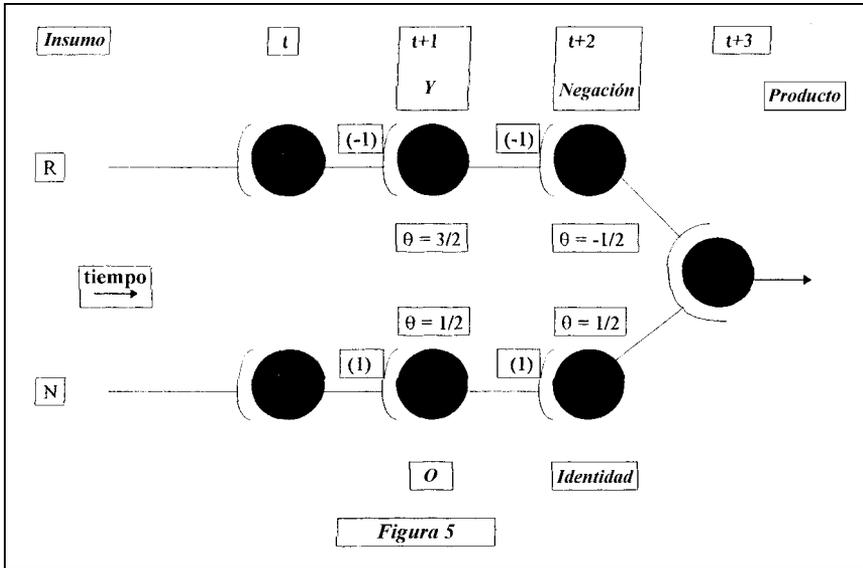


Figura 5

Observación 1: La función del operador de "identificación" es almacenar  $R \vee N$  mientras  $R$  estaba siendo evaluada — en este caso negada.

Observación 2: Ninguna opción de ponderaciones y umbrales permitiría a la red neuronal computar (evaluar)  $R$  O SINO  $N$  en un *solo* paso.

(e) La "Geometría" de (d)

*La Geometría*

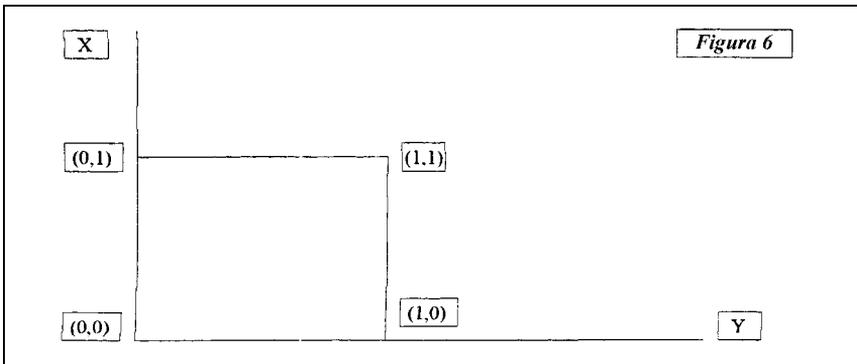
La señal (entrada) puede contener  $x$ ,  $y$ , ambas o ninguna. Por lo tanto, el agente tendría que identificar cuatro posibles entradas y asignar valores de verdad a las señales de salida adecuadamente.

En forma tabular:

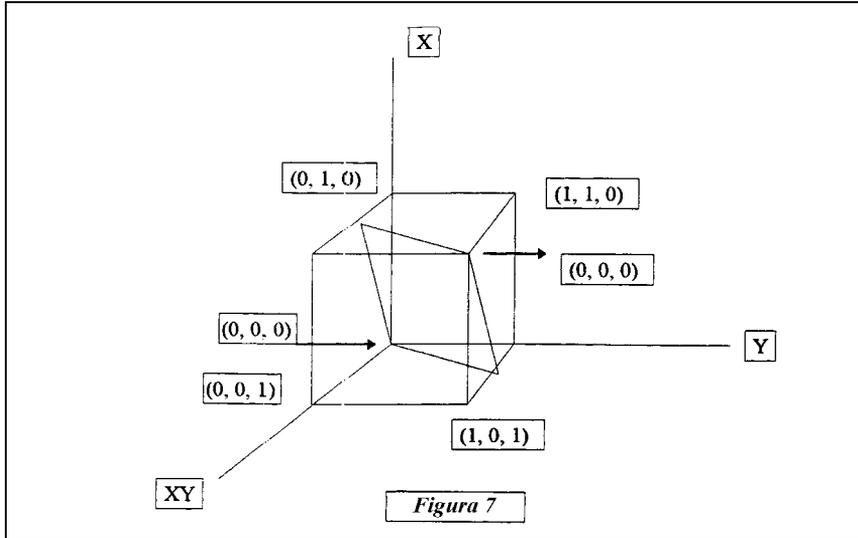
<i>Entrada</i>			<i>Salida</i>
$x$	$y$		
0	0	→	0
1	0	→	1
0	1	→	1
1	1	→	0

Donde:      0 = falso  
                  1 = verdadero

Es fácil ver en la Figura 6 que no puede trazarse ninguna "línea" de modo tal que  $\langle 10,01 \rangle$  esté de un lado y  $\langle 00,11 \rangle$  esté del "otro" lado, es decir, los dos conjuntos de señales de entrada no son separables (linealmente).



No obstante, si admitimos predicados de segundo orden podemos visualizar una posible separación como en la Figura 7. Introduciendo el predicado de segundo orden  $XY$  podemos separar  $X$  e  $Y$  en 3 espacios.



(f) Algebra Simple de (d)

$$w_1 X + w_2 Y > \theta$$

en donde:  $\theta \geq 0$

$$w_1 > \theta$$

$$w_2 > \theta$$

y

$$w_1 + w_2 \leq \theta$$

lo que es imposible. En el lenguaje de las anotaciones que anteceden,  $X$  O SINO  $Y$  (la o excluyente) no pueden ser evaluadas (computadas) mediante simples unidades McCulloch-Pitts. Debemos no plantear linealmente el problema (una pequeña novedad en una profesión que plantea todo linealmente) de la siguiente forma:

$$w_1 X + w_2 Y + w_{12} XY > \theta$$

en donde:  $\theta \geq 0$

$$w_1 > \theta$$

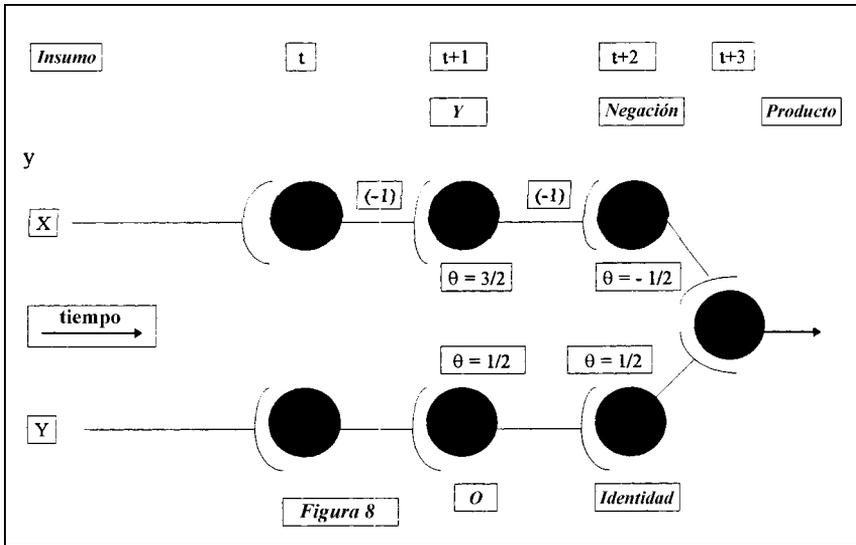
$$w_2 > \theta$$

y

$$w_1 + w_2 + w_{12} \leq \theta$$

Una posible solución es dada por  $\theta = 0, w_1 = w_2 = 1$ , y  $w_{12} = -2$ . La separabilidad y la no linealidad van juntas.

Una Red Neural para X O SINO Y:



Observación 1: La función del operador de "identificación" es almacenar  $X \vee Y$  mientras X estaba siendo evaluada — en este caso negada.

Observación 2: Ninguna opción de ponderaciones y umbrales permitiría a la red neuronal computar (evaluar) X O SINO Y en un *solo* paso.

(g) Discusión

Por la geometría y el álgebra del problema de procesar una proposición Universal, resulta evidente que es esencial cierta forma de *no linealidad*. Por la arquitectura de la red neuronal, notamos también que los *intervalos de*

*tiempo* son esenciales para procesar una proposición universal; esto es, por supuesto, porque ninguna opción de ponderaciones y umbrales permitiría a la red neuronal computar (evaluar)  $X \text{ O SINO } Y$  en un *solo paso*.

Tenemos ahora elementos suficientes para dar cierta sustancia concreta a los resultados formales de la sección II. Nótese en primer lugar que todas las unidades M-P, y las colecciones de unidades M-P, entre aquellas que han sido identificadas inicialmente como grupos de unidades de entrada y salida, son llamadas las **unidades ocultas**. Los procesos adaptables, por aproximaciones al tanteo, a los que se hace referencia en la sección anterior son, entonces, aquellos **procedimientos** que asignan conjuntos de ponderaciones a las unidades ocultas de modo tal que, para determinadas "señales" de entrada, pueden obtenerse las salidas resultantes.

Permítaseme plantear el problema de forma algo diferente. El problema general que enfrenta el agente económico racional es el de evaluar proposiciones arbitrarias para valores de verdad. Dado que una determinada proposición puede construirse a partir de cadenas de la proposición universal, el problema general será asignar ponderaciones adecuadas ( $w_{ij}$ ) a las unidades ocultas. La pregunta clave, entonces, es:

Dada una proposición arbitraria, ¿existe un procedimiento efectivo, determinable *a priori*, para asignar ponderaciones adecuadas (para, digamos, construir el necesario "hiperplano separador"), para evaluarla?

Los distintos teoremas de la sección anterior formalizan y establecen la respuesta negativa al tipo de pregunta que antecede. Podemos identificar las razones de las respuestas negativas universales: la necesidad de no linealidad e intervalos de tiempo<sup>4</sup>. O, si escribimos la expresión algebraica para el comportamiento de entrada-salida de la red neural como:

$$y = f(w,y) \tag{9}$$

Donde:

- $f$ : mapeado no lineal;
- $w$ : vector de ponderaciones de conexión forzada de número entero o número racional;
- $y$ : variables de vector de entrada, salida y ocultas;

Entonces, encontrar un procedimiento efectivo para asignar ponderaciones para evaluar una proposición dada es equivalente a encontrar un procedimiento efectivo para resolver la ecuación diofantina (9). Sabemos, a partir de la celebrada solución negativa al 10º Problema de Hilbert que no existen procedimientos generales eficaces, disponibles *a priori*, para resolver ecuaciones diofantinas dadas, arbitrarias.

Además, podemos considerar (9) de un modo dinámico, siendo las soluciones de (9) los atractores para un sistema dinámico:

$$\frac{dy}{dt} = y - f(w, y) \quad (10)$$

Sabemos a partir de la "geometría" del problema de evaluación por la o excluyente que el mapeado  $f$  es tanto no lineal como de dimensión  $\geq 3$ . Los atractores de dicho sistema dinámico, como soluciones a (10) serán, en general, "extraños" en el sentido en el que esto está definido en la teoría formal del sistema dinámico. Los sistemas dinámicos con atractores extraños son mínimamente requeridos si la universalidad computacional va a ser factible para ellos. Las cláusulas finales del presente informe tienen la intención de unir estos cabos dispares en un modo que puede sugerir una síntesis micro-macro aunque en una forma muy especulativa.

Ahora podemos resumir y unir esos cabos sueltos y dispares. El problema racional del agente económico de evaluar el valor de verdad de las proposiciones dadas, arbitrarias puede identificarse con:

- (i). Encontrar procedimientos para asignar ponderaciones "apropiadamente" a unidades ocultas en una red neural;
- ó (ii). Encontrar algoritmos para resolver ecuaciones diofantinas;
- ó (iii). Métodos para caracterizar atractores extraños;

No existen procedimientos eficaces conocidos para resolver ninguno de estos problemas de un modo universal *a priori*, y en el caso de (ii), se sabe que en general no existen dichos procedimientos eficaces (soluciones negativas al 10º Problema de Hilbert). Lo mejor que puede hacerse es encontrar un procedimiento uniforme para clases especiales de proposiciones. Encontrar dichos procedimientos uniformes, clasificar clases de pro-

posiciones, etc es lo que se considera como procedimientos "al tanteo". Esta clase de actividades es la que hace "todo comportamiento adaptable" — excepto para la clase de actividades altamente limitada donde las proposiciones pueden ser evaluadas por sus valores de verdad sin ninguna unidad oculta. Cuando se han clasificado una clase de proposiciones y se han encontrado procedimientos uniformes para evaluarlas entonces existe un sentido en el cual puede definirse una "constante" para ellos. Además, el modelado de "un individuo como un conjunto de reglas de determinación" puede ser interpretado de un modo similar: conjunto de procedimientos uniformes que se utilizan para evaluar clases de proposiciones.

#### IV CONCLUSION

"Ahora si un organismo se enfrenta con el problema de comportarse aproximadamente en forma racional o *adaptable*, en un entorno particular, las clases de simplificación que son adecuadas pueden depender no sólo de las características — sensorial, neural y otras — del organismo, sino igualmente de la estructura del entorno. Por lo tanto, podemos esperar descubrir, mediante un examen cuidadoso de algunas de las características estructurales fundamentales del entorno, algunos indicios más en cuanto a la naturaleza de los mecanismos de aproximación utilizados en la toma de decisiones."

Simon, 1956 (1979), pág. 21; cursiva agregada

En las partes anteriores de este informe, investigué algunas de las "características del organismo" y las "características estructurales del entorno" que arrojan dudas sobre la utilidad de trabajar con la total parafernalia del comportamiento racional estándar. Ahora, si el dominio de las "características estructurales del entorno" pudiera también incluir los *problemas* que enfrenta un *solucionador de problemas*, entonces podremos investigar la "naturaleza de los mecanismos de aproximación" en uso — implícita y explícitamente — al hacerlos *solucionables*. Los mecanismos de aproximación han sido ideados por el agente racional — como un agente adaptable, como un agente restringidamente racional, como un autómata finito; ellos también han sido sugeridos para el "entorno": soluciones satisfactorias, dominios de definiciones elegidos adecuadamente como convenientes espacios topológicos, medidas teóricas y otros. Quisiera sugerir que existe una forma alternativa de abordar esta cuestión de las "aproximaciones" en

un contexto de resolución de problemas que conduce a una definición más natural del comportamiento restringidamente racional en el cual el agente se vuelve adaptable casi por necesidad.

Una forma de especificar el problema económico estándar (SEP), a ser resuelto por "el hombre racional de la economía" es formularlo como un problema de programación matemática (o un problema de control óptimo, juego teórico, etc.) Una simple formulación (de libro de texto)

### SEP

Minimizar  $f(x)$

a condición de que:  $g_i(x) \geq 0$ ,  $i = 1, \dots, m$ ;

y:  $h_j(x) = 0$ ,  $j = 1, \dots, p$ ;

con suposiciones estándar sobre  $f$ ,  $g$  y  $h$ . A menudo, las suposiciones garantizan la *existencia*<sup>6</sup> (por ejemplo) de un optimismo global, como corresponde al que "hace lo máximo, quien no se conformará con nada menos que *lo mejor*"

Ahora consideremos la siguiente variación del SEP:

### SEP<sup>α</sup>

Un problema de optimización es un par  $\{F, c\}$ , donde:

$F$ : cualquier conjunto (el dominio de las alternativas posibles)

$c$ :  $F \rightarrow \Sigma$  (p.ej.: función de criterio)

Problema: hallar  $f \in F$  s.t  $c(f) \leq c(g)$ ,  $\forall g \in F$

Obviamente, cierta estructura significativa, dictada por consideraciones económicas, debe imponerse sobre  $F$ ,  $c$ , etc.

Luego, consideremos otra variante más del SEP (y del SEP<sup>α</sup>):

### SEP<sup>†</sup>

(a). Dado un elemento combinatorio (o sea, un elemento especificado en forma numérico-teórica)  $f$  y un conjunto de parámetros  $S$ ,

*determinar* si  $f \in F$  (donde  $F$  está caracterizado por  $S$ ). Supongamos que este *procedimiento de determinación* es ejecutado por medio del algoritmo  $T_f$  (representando la Máquina Turing indexada por  $f$  que ha sido *efectivamente codificada*);

(b). Después de la determinación implementada por medio de  $T_f$ , utilicemos otro procedimiento de determinación (algoritmo) para calcular el valor  $c(f)$ , donde  $c$  está caracterizado por un conjunto de parámetros  $Q$ . Llamemos a este procedimiento de determinación  $T_c$ ;

(c). Observemos que  $S$  y  $Q$  están representados en forma numérico-teórica, por ejemplo, numerados por Gödel;

He aquí un ejemplo de **SEP**<sup>†</sup>

El problema de programación lineal estándar (entero) (SILP):

Minimizar  $c'x$   
 a condición de que:  $Ax = b$   
 y:  $x \geq 0$  (entero)  
 (vectorial, de dimensiones apropiadas)

De acuerdo con la interpretación del **SEP**<sup>†</sup>, esto significa que:

- (i). Los parámetros  $S$ , para el procedimiento de determinación  $T_f$ , están dados por:  $A, b$ ;
- (ii). Dado cualquier entero (vector)  $x$ ,  $T_f$  determina si  $Ax = b$  y  $x \geq 0$  se cumplen simultáneamente;
- (iii). "Luego"  $T_c$  es implementado el cual tiene  $c$  para  $Q$  para evaluar  $c'x$  para cada  $x$  determinada por  $T_f$ .

Observemos, sin embargo, que el "luego" del punto (iii) que antecede no implica necesariamente acciones secuenciales por parte de las Máquinas Turing. Las tareas de determinación más complejas, que encapsulan variedades más abundantes del SEA implicarán, por supuesto, un conjunto de Máquinas Turing operando en forma paralela y secuencial (sincrónica y asincrónicamente). Por lo tanto, desde este punto, todo agente y todo problema, **en general**, es un elemento combinatorio y un conjunto de procedi-

mientos de determinación; algunos de los procedimientos de determinación operan secuencialmente, siendo la salida de un subconjunto las entradas de otros subconjuntos; algunos operan (en forma sincrónica o asincrónica) en paralelo. El SEP<sup>†</sup> no viola el arreglo en el cual "el hombre racional de la economía" se mantiene en sus pautas. Pero sí da una perspectiva completamente diferente sobre la complejidad y la naturaleza dinámica de los problemas de determinación y la ejecución por parte de dicho agente racional.

Ahora, por la forma en que he reformulado el problema de la opción racional en Velupillai (op. cit) y en las secciones 2 y 3 que anteceden, el que importa es el SEP<sup>†</sup>. En general, se interpreta que el agente resuelve problemas de determinación de los cuales el SEA se considerará ahora como un caso especial. En dicho caso las "unidades ocultas" son, o se comportan *como si fueran*, variables ficticias. Toda la atención se centra solamente en las variables de entrada y salida; o sea, el dominio de la *racionalidad sustantiva*. Siempre que las "unidades ocultas" rompan los lazos que unen a las variables de entrada y salida, el SEP (y el SEP<sup>α</sup>) deben generalizarse al SEP<sup>†</sup> y debemos concentrarnos en la *racionalidad de procedimiento*, es decir, la *racionalidad del proceso de ponderar* las "unidades ocultas".

Entonces, ¿en qué sentido podemos hablar de la existencia de "una forma de ser totalmente racional? En casi ninguno. Hay muchas maneras de ser racionales en el sentido de que clasificar tipos de problemas, hallar procedimientos uniformes, eficaces, etc., ¿son más Linneanos que todo lo demás?<sup>7</sup> Creo que la inducción, y sus formalismos, proporcionan reglas disciplinarias para este mundo Linneano de adopción de decisiones.

Sin embargo, permítaseme terminar con una definición tentativa de los agentes restringidamente racionales y notas generales análogamente tentativas.

#### Definición

Un agente restringidamente racional resuelve problemas de determinación en el sentido del SEP.

#### Observación

Esta definición es totalmente compatible con la definición dada por Osborne y Rubinstein en el contexto de los juegos repetidos (Osborne y Rubinstein, op. cit., pág. 164)<sup>8</sup>

Esta definición aclara, espero, que los agentes restringidamente racionales *no son agentes casi totalmente racionales*. Son restringidamente racionales, de acuerdo con estas definiciones, debido a que la índole de los problemas que enfrentan, y la forma en que consideran soluciones para los mismos, ponen *restricciones a las clases de problemas que ellos pueden resolver*. El agente sustantivamente racional — o "totalmente racional" — de formalismos estándar es un caso especial porque un agente tal se enfrenta con un *subconjunto especial* de la clase de problemas de determinación que enfrenta el agente restringidamente racional.<sup>9</sup>

Esto me lleva a un punto ulterior, general. En los últimos años se han hecho muchos intentos de modelar un comportamiento emergente sobre la base de la interacción entre conjuntos. Agentes simples — incluso simplistas. Relacionamiento, paralelismo masivo, autómatas celulares, etc., son algunas de las "contraseñas" de esta vasta literatura, que nadan entre dos aguas de fascinantes mundos interdisciplinarios. Creo que esto es una visión fundamentalmente errónea.

El agente económico (o social, o político) es cualquier cosa menos un elemento simple o simplista. Pienso que es imperativo que modelemos el agente en su total complejidad, es decir, como Máquinas Turing. Es el conjunto el que es simplista y puede no ser capaz de universalidad computacional. ¿Cómo las interacciones de un conjunto de agentes sofisticados — Máquinas Turing — pueden resultar en un sistema incapaz de lograr lo que los elementos individuales pueden conseguir en forma individual e independiente? Creo que hay dos razones:

- (a). Ante todo, el *sistema* no trata de resolver problemas de determinación; ése es el problema de cada agente. El sistema actúa simplemente como el Libro Mayor, el clasificador en el algoritmo "de bomberos voluntarios", el tablero de anuncios. Registra los créditos y débitos. Es un contador. ¡En el caso extremo es el Demonio de Maxwell ...!
- (b). En segundo lugar, como resultado de ser el contador, no es la universalidad computacional lo que importa para el comportamiento del sistema, sino cierto principio de conservación. La mayoría de las ciencias naturales, en sus variaciones fenomenológicas, tienen principios de conservación; esto es verdadero incluso para muchas de las ciencias puras (simetría, modularidad, etc.) Las ciencias sociales

parecen no haber tenido éxito en la formulación de principios de conservación útiles.

Por lo tanto, la pregunta a plantear finalmente es: ¿cómo un conjunto de interesantes Máquinas Turing puede dar por resultado macrosistemas — instituciones, leyes, etc. — con principios de conservación útiles? La universalidad computacional está en poder de agentes individuales que se comportan racionalmente e implementan el principio de "no arbitraje"; las leyes de conservación están en poder de agregados. Entonces, ¿cuál es la conexión o relación entre universalidad computacional y principios de conservación? Una respuesta a esta pregunta debería resolver la línea divisoria entre micro y macro.

**BIBLIOGRAFIA**

**K. Steiglitz (1982):** Complexity Prentice-Hall Inc., Englewood Cliffs, New Jersey.

**Simon, H. A. (1956), 1979:** "*Rational Choice and the Structure of the Environment*" en Models of Thought, Vol. 1, Yale University Press, New Haven.

**Spencer, H. (1866), 1996:** "*The Co-operation of the Factors*" [Capítulo XIII, § 169 (Vol. 1, Pt. III) en The Principles of Biology] en: Adaptive Individuals in Evolving Populations: Models and Algorithms, ed. por R. K. Belew y M. Mitchell Addison-Wesley, Menlo Park, California.

**Stewart, I (1991):** "*Deciding the Undecidable*" Nature, Vol. 352, 22 de agosto, págs. 664 y 665.

**Velupillai, K. (1977):** Computable Economics Oxford University Press, Oxford (de próxima aparición).

**Von Neumann, J. (1966):** Theory of Self-Reproducing Automata, (Editado y completado por A. W. Burks) University of Illinois Press, Urbana.

## NOTAS

- 1 En la descripción característicamente cáustica de Machlup:  
*"Estructura ... es ... una palabra equívoca empleada para evitar comprometerse con un concepto definido y claro"*.  
Machlup (1958), pág. 280.
- 2 Wordsworth Concise English Dictionary.
- 3 Según la oportuna expresión de Richard Day.
- 4 Como destacó perceptivamente Richard Goodwin:  
"Combinando las dificultades de las ecuaciones de diferencia con aquellas de teoría no lineal, obtenemos un animal de carácter feroz y es prudente no depositar demasiada confianza en nuestras conclusiones respecto a comportamiento."  
Goodwin (1950), 1982, pág. 79.
- 5 Sigo la excelente notación y formulación de Papadimitriou y Steiglitz (1982), aunque, por supuesto, podría haber elegido cualquier libro de texto estándar sobre optimización, control óptimo o teoría del juego, para formulaciones y definiciones alternativas — pero formalmente equivalentes.
- 6 A riesgo de extenderme sobre un punto árido, no puedo resistirme a repetir que garantizar la existencia de un óptimo es una cosa, encontrarlo es otra cosa, y encontrarlo fácilmente es otra más.
- 7 O, según la peregrina afirmación de Rutherford: "la ciencia es en gran parte "marcar y recopilar" y ... "
- 8 En particular la sección 8.4 y el capítulo IX de este excelente texto:  
"... nos concentramos más de cerca en las estrategias de equilibrio estructurado que en el conjunto de resultados de equilibrio, utilizando la herramienta de una máquina [concebida como una abstracción del proceso mediante el cual un participante implementa una estrategia en un juego reiterado]. ...  
... Un modelo que incluye tales aspectos "de procedimiento" de adopción de decisiones se conoce como un modelo de "racionalidad restringida" "  
ibid, págs. 163 y 164; cursiva agregada

- 9 En una curiosa colección de ensayos, con el atractivo título de *Inteligencia Artificial y Análisis Económico* (ed. por Moss & Rae, 1992)

Huw Dixon manifiesta que:

"Una simple generalización del enfoque ortodoxo es ... suficiente para captar algunos elementos de la racionalidad restringida sin necesidad de modelar el razonamiento en forma explícita".

ibid, pág. 136; cursiva agregada

Esto es casi exactamente la antítesis de mi definición y se basa en igualar una solución restringidamente racional con una solución aproximadamente óptima en una situación por lo demás ortodoxa. Si todo lo que significaba esa racionalidad restringida era más restricciones en el modelo estándar, entonces, es el declarativo un superconjunto del imperativo.