

Rodolfo J. Rodríguez Rodríguez

## Modelos cognoscitivos para la filosofía contemporánea de la mente

**Summary:** *This article examines models of the mind that have been developed from the Cognitive Science research. The first set are the information processing models such as the symbolic, the subsymbolic (connectionism) and the modularity. The second set are the cognitive neurobiology models such as the F. Crick and P. S. Churchland reductionist models of mental processes, but specially the Neural Darwinism of the Nobel Prize: G. Edelman. Finally the author suggests other kind of mental interpretations such as: the Quantum Thought of R. Penrose, the subjective ontology of J. Searle and the interpretation of the conscious mind of D. Dennett.*

**Resumen:** *Este artículo examina modelos de la mente que han sido desarrollados a partir de las investigaciones en Ciencias Cognoscitivas. El primer conjunto son los modelos de procesamiento de la información como el simbólico, el subsimbólico (conexionista) y la modularidad. El segundo conjunto son los de la neurobiología cognitiva tales como los modelos reduccionistas de los procesos mentales de F. Crick y P. S. Churchland, pero especialmente el Darwinismo Neural del Premio Nobel: G. Edelman. Finalmente el autor sugiere otro tipo de interpretaciones mentales tales como el pensamiento cuántico de R. Penrose, la ontología subjetiva J. Searle y la interpretación de la mente consciente de D. Dennett*

### Introducción

En el presente siglo una de las áreas de investigación más revolucionarias dentro de la computación electrónica ha sido la inteligencia artificial. Esta disciplina tiene sus raíces en los desarrollos de la lógica formal, generada a finales de siglo pasado e inicios del presente con figuras como G. Boole, G. Frege, G. Peano, B. Russell, A.N. Whitehead, quienes establecen estructuras sólidas de carácter veritativo funcional para la mecanización simbólica de los razonamientos. Como prehistoria de esta disciplina se encuentran los trabajos de C. Shannon (1938)<sup>1</sup>, McCulloch, W. y W. Pitts (1943)<sup>2</sup> y de N. Wiener (1948).<sup>3</sup> El origen propiamente dicho de la I.A. se puede establecer a partir del artículo de A.M. Turing publicado en *Mind*: "Computing Machinery and Intelligence" (1950).<sup>4</sup>

A partir de Turing se tiene pues un núcleo teórico básico sobre el que se fundamentará todo un modelo estructural de investigación en Inteligencia Artificial. Dicho modelo fue desarrollado por expertos en distintas áreas como J. McCarthy, A. Newell, H. A. Simon y M. Minsky como los más conocidos.<sup>5</sup>

El objetivo inicial fue desarrollar máquinas (hardware) que puedan correr programas (software) capaces de simular y que inclusive generen el razonamiento humano de manera automática (acudiendo a cálculos formales como el proposicional

en un primer momento). Los primeros desarrollos en esta área fueron los probadores de teoremas matemáticos y de lógica formal, como también los primeros jugadores artificiales de ajedrez. Todo ello representó el punto de partida de nuevas áreas de investigación sobre la inteligencia en particular y sobre la cognición en general. Las limitaciones de estos primeros enfoques plantearon la necesidad de investigar otras disciplinas relacionadas con la cognición, tales como la lingüística, la psicología, la neurociencia, la antropología; surgen entonces así las Ciencias Cognoscitivas.<sup>6</sup>

### 1. Modelos de procesamiento de la información

El primer modelo cognoscitivo generado a raíz del desarrollo de la I.A., surge en el contexto de las investigaciones sobre el procesamiento electrónico de la información. Tuvo enorme significación la obra de Claude Shannon, matemático del MIT, que con su tesis de maestría: "*Análisis simbólico de los circuitos de relé y conmutación*" (1938),<sup>7</sup> demuestra que los circuitos, como los que aparecen en un aparato electrónico, podían expresarse mediante ecuaciones semejantes a las de Boole; es decir bivalentemente, en la oposición: verdadero-falso, equivalente a la oposición abierto y cerrado (o conectado o desconectado) en un circuito. Shannon determina que cualquier operación puede describirse mediante esos relés de "conmutación". La obra de Shannon sentó las bases para la fabricación de máquinas capaces de ejecutar tales operaciones veritativo funcionales, y además sugirió nuevas formas de diseñar circuitos. En el plano teórico dio la pauta -al establecer un conjunto de instrucciones codificadas que debían seguirse de forma minuciosa- para lo que sería la programación de computadoras, determinando que ésta debía concebirse como un problema de lógica formal y no como un problema aritmético, idea que proviene de los trabajos de Boole. Shannon introduce de esta manera el tema lógico-cognoscitivo en el mundo de la incipiente computación electrónica de su tiempo.

Alan M. Turing, ya en 1936 ("*On Computable Numbers with an Application to the Entscheidungsproblem*")<sup>8</sup>, había definido una clase de autómatas (conocida como la máquina de Turing) con la que demostró que cualquier miembro de esta clase podía computar cualquier función entre un conjunto de clases. Este tipo de máquina formal de Turing,

es la base del funcionamiento de la mayoría de las computadoras digitales.

Los argumentos originales de Alan Turing se complementan con los de Alonzo Church, al postular las capacidades universales de resolución de problemas de las computadoras, y sugiriendo que el cerebro debe comprenderse como si fuera una computadora. Particularmente Church sugiere que si existe un método consistente y finito para resolver un problema dado, luego existe un método que puede correr en una máquina de Turing y dar exactamente los mismos resultados. Por ello, para aquellos problemas que pueden ser resueltos consistentemente en tiempos finitos y especificados, una máquina de Turing es tan poderosa como cualquier otra entidad que pueda resolver el problema, incluido el cerebro.<sup>9</sup>

Los argumentos de Shannon, Turing y Church estructuran la base de los modelos cognoscitivos de procesamiento de la información, y éste logra eventualmente un gran éxito computacional. En este sentido lo primero es elegir un problema significativo que todo el mundo esté de acuerdo que requiere para su solución el uso de la inteligencia, luego se debe identificar los elementos de información que se necesitan para lograr la solución al problema, se debe determinar cómo podría representarse esta información en una computadora, encontrar un algoritmo que pueda manipular esta información para solucionar el problema; escribir el código computacional que implemente ese algoritmo y finalmente se debe poner éste a prueba contra instancias muestradas (usualmente simples) del problema.<sup>10</sup>

#### 1.1. Modelo simbólico

El modelo simbólico de la I.A. es el primer núcleo sólido en I.A., enmarcado dentro del contexto del "macromodelo" de Procesamiento de Información y alcanza su máxima expresión con la "*Hipótesis del sistema de símbolos físicos*" de A. Newell y H. A. Simon.<sup>11</sup> Para ellos, un sistema de símbolos físicos es una máquina que produce a lo largo del tiempo una colección evolutiva de estructuras simbólicas. El sistema de símbolos tiene como referente un mundo de objetos más amplio de tan solo las *expresiones simbólicas*. El término "físico" hace referencia a que tales sistemas obedecen las leyes de la física, es decir, son realizables por sistemas de ingeniería o computación electrónica.

La relación entre el referente y los símbolos se dará por medio de designación e interpretación. En dicho modelo se considera que un sistema de símbolos físicos tiene los medios necesarios y suficientes para una acción inteligente. En este caso el concepto de “necesario” se entiende en el sentido de cualquier sistema que se considere inteligente de manera general, se puede entender como un sistema de símbolos físicos. Lo de “suficiente” hace referencia a que cualquier sistema de símbolos físicos de suficiente tamaño puede ser organizado hasta llegar a mostrar acción inteligente general.<sup>12</sup>

En el corazón de la teoría de los Sistemas de símbolos físicos, se encuentran las arquitecturas simbólicas las cuáles sirven de substrato para la computación electrónica, que es posible por la manipulación de señales físicas por medio de reglas explícitas que también están compuestas por otras señales. La manipulación está basada solamente en las propiedades físicas de las señales, no por sus significados. Todo procesamiento implica combinaciones basadas en reglas de señales simbólicas; esto es, hileras de señales. El sistema entero es semánticamente interpretable; esto es, todos sus componentes se refieren a objetos o a funciones de estados.<sup>13</sup>

Este sistema de símbolos físicos es un claro ejemplo de máquina general de las ideadas por Turing, consolidándose así el modelo estructural propuesto originalmente por este último y que se ha denominado como: **MODELO SIMBÓLICO** y que representa el primer modelo consolidado dentro de la inteligencia artificial para poder explicar los procesos mentales, aún cuando con el tiempo han salido a flote muchas de sus limitaciones.<sup>14</sup>

## 1.2. Modelo subsimbólico

El primer modelo cognoscitivo a partir del surgimiento de la inteligencia artificial por parte de la filosofía contemporánea de la mente es el simbólico, anteriormente descrito. Pero no es el único modelo para la filosofía de la mente, ni siquiera es el único en inteligencia artificial, ni en procesamiento de la información.

Como alternativa al modelo simbólico se buscó desarrollar una tecnología que se acercase lo más posible al cerebro, como fuente de la inteligencia. Esto pues el modelo simbólico partía de un supuesto que se consideró no del todo correcto, al postular que la inteligencia era posible producirla

en computadoras seriales de tipo convencional y el cerebro no funciona de esta manera sino en forma masivamente paralela, donde cientos de miles o millones de neuronas desarrollan sutiles procesamientos de la información en el cerebro y a mucho más altas velocidades. Los perceptos coherentes se forman en el tiempo que exceden a los tiempos elementales de reacción de las neuronas singulares por un factor de poco más de diez. Especialmente en lo que concierne a los procesos perceptuales básicos, como la visión, esta observación excluye las formas de procesamiento de información interactivas que tendrían que barrer los datos que ingresan serialmente, o pasarlos a través de muchas etapas de procesamiento intermediarias.<sup>15</sup>

Asimismo, dentro del cerebro, el conocimiento no se almacena en ninguna forma que se parezca a un programa de computadora convencional, sino que se almacena estructuralmente, en forma de patrones distribuidos de pesos sinápticos excitatorios o inhibitorios, cuyas magnitudes relativas determinan el flujo de las respuestas neuronales que las que constituyen la percepción y el pensamiento.<sup>16</sup>

Algunos investigadores en I.A. buscan relacionarse con la neurociencia experimental de manera similar a como se relacionan la física teórica y la física experimental, de tal manera que pueda unificar los desarrollos y las conjeturas teóricas con la experimentación. Surge así otro modelo en inteligencia artificial que se ha dado en llamar como: “computación con redes neurales”. Pero la aterradora complejidad del cerebro opone grandes obstáculos a la materialización de los objetivos de dicho modelo, pues el cerebro humano consiste en aproximadamente 100 mil millones de neuronas, que posiblemente sean diez veces más. Las neuronas se comunican transmitiendo paquetes eléctricos (potenciales de acción) a una población de neuronas vecinas. De acuerdo con los estudios neurofisiológicos, la amplitud precisa y la forma de esos paquetes y el tiempo preciso para su llegada es un intervalo de dos milisegundos aproximadamente. De ello se puede modelizar cada paquete como un “bit” singular portador de información en la corriente de salida de una neurona y decir que una neurona produce información a una velocidad aproximada de cien bits por segundo (100 BPS). El estimado total del cerebro sería de 10 billones de bits por segundo, tomado un factor de cien, para el “ancho de banda” interno del cerebro.<sup>17</sup>

Una neurona transmite información a las neuronas vecinas en uniones neuronales llamadas “si-

napsis". Una sola neurona puede tener tanto como diez mil entradas sinápticas, aunque en algunos casos muchas entradas menos, para otros casos un aproximado de 100 mil entradas convergen en una sola neurona. El número total de sinapsis en el cerebro se puede estimar en 1000 billones.<sup>18</sup>

El punto de partida para los nuevos enfoques en I.A. deberá ser entonces, modelar el cerebro como un computador compuesto de unos 100000 millones de computadoras individuales, cada una de los cuales es una computadora analógica o digital e "híbrida", de suma complejidad, interconectando en formas intrincadas con muchas computadoras similares.<sup>19</sup>

La computación por redes neurales o conexionismo busca generar este modelo del cerebro de manera artificial, es decir por medio de una modelación tecnológica del mismo, acudiendo a la física estadística y a la ingeniería para poder lograrlo (no a la neurobiología, como podría suponerse). La I.A., queda parcelada en un número elevado de procesadores simples, con un crecimiento potencial enorme de la velocidad de computación.

El "paralelismo" ha demostrado ser difícil de aplicar en I.A., aunque ya ha surgido en la computación numérica como la única forma indefinidamente extensible de superar el "cuello de botella" de von Neumann (los límites fundamentales impuestos a la velocidad de los procesadores individuales por las leyes de la física, por ejemplo, dado que las señales no pueden viajar de una parte a otra de la computadora más rápido que la velocidad de la luz). Dado que estos sistemas son imitaciones de sistemas estadísticos, las nuevas estrategias también proporcionan métodos de aproximación estadística a problemas de optimización que se han revelado a todo intento de computación exacta.<sup>20</sup>

Este tipo de modelo de procesamiento de información mediante estados de procesadores simples y las conexiones entre ellos es llamado: CONEXIONISTA O SUBSIMBOLICO, sostenido por figuras como David. E. Rumelhart, James McClelland (1986),<sup>21</sup> Paul Smolensky (1988)<sup>22</sup> y el grupo de Investigación Procesamiento Distribuido en Paralelo (PDP).

El grupo de investigación PDP es el que le ha dado forma al Conexionismo (Rumelhart & McClelland, 1986) caracterizándolo de la siguiente manera: un conjunto de unidades de procesamiento (entendidas como neuronas artificiales) que están interconectadas por "*pesos de conexión*" o

"*cargas de conexión*" o "valores de conexión" (connection weights).<sup>23</sup>

Se generaron así sistemas redes neurales artificiales. Si se tienen entradas a alguno de estos sistemas se da una activación de la red que es modulada por los pesos de conexión. Las entradas se dan en una unidad o neurona artificial (pasando a ser activada desde su estado pasivo) por medio de una regla de activación. Los pesos de conexión son manipulados sobre la base de "*reglas de aprendizaje*", las cuales son funciones matemáticas. La manipulación está puramente basada en los valores de activación y pesos de conexión (no por sus significados). Las entradas en el sistema son convertidas en salidas por medio de las reglas de activación, las cuales consisten en la combinación y recombinación de la *activación de valores*.<sup>24</sup>

Los sistemas conexionistas son semánticamente representables, lo que hace que se mantenga dentro de la clase de los modelos de procesamiento de información. La diferencia de este modelo es que no acude al procesamiento serial, sino al procesamiento distribuido en paralelo, por medio de redes neurales artificiales, similares al procesamiento de las redes neuronales naturales. En estos casos el procesamiento no es simbólico, sino que dichas redes se interpretan subsimbólicamente es decir es un procesamiento de la información de manera pluridireccional (es en este sentido que busca asemejarse a las conexiones sinápticas neuronales).<sup>25</sup>

La diferencia característica de los sistemas conexionistas de los sistemas computacionales simbólicos, es que los primeros no son programables en el sentido de los segundos, es decir mediante el almacenamiento en la memoria del texto de un algoritmo cuyos pasos debe seguir la computadora de manera secuencial. A una red neural "se le enseña" por medio de pruebas con casos especialmente seleccionados. Es decir se le puede "enseñar" una letra o una palabra para que posteriormente pueda reconocerlas. Es decir, más que ser programada una red neural, esta "aprende". Esto es lo que se ha dado en llamar como "aprendizaje mecánico" que se realiza modificando en la red neural, los valores de conexiones o pesos de conexión (emulando las sinapsis en las redes neuronales) entre sus distintas celdas. De esto se deriva su nombre de conexionismo.<sup>26</sup>

Este enfoque ha sido utilizado para explicar la conducta a un nivel de descripción funcional bajo la presuposición que esto mismo ocurre al nivel psicológico.<sup>27</sup> Sin embargo, adolece de una explica-

ción funcional de la inteligencia, pues no es claro cómo a partir de las partes (neuronas y redes neurales) emergería la misma.<sup>28</sup>

En computación el modelo conexionista es una elegante solución al problema de *altas velocidades*. Así si por otra parte el conexionismo sugiere que podría entenderse los procesos neurofisiológicos de la percepción como "vectores" que codifican las distintas entradas sensoriales, y también lo hace para varias de las salidas motoras.<sup>29</sup>

Bajo el modelo conexionista en computación se pueden llegar a lograr sistemas masivamente paralelos que representen y manipulen transformaciones vectoriales como las dadas en los procesos neurofisiológicos del cerebro. El procesamiento computacional se puede hacer más rápido que el procesamiento cerebral. Este último procesa en milésimas de segundo, el computacional lo logra en millonésimas de segundo.<sup>30</sup>

Asimismo las redes neurales artificiales aparte de ser computacionalmente más poderosas, son resistentes contra daños (o su equivalente de lesiones cerebrales), rápidas y modificables.

Las estrategias formales del conexionismo se han fundamentado en la física estadística y comparten con el modelo simbólico la noción implícita de que los objetos y sucesos, las categorías y la lógica están dados, y que la naturaleza del trabajo cerebral es procesar información sobre el mundo con algoritmos que conduzcan a conclusiones que lleven a la conducta.

Un problema de base que ha presentado por los modelos conexionistas es su dificultad para realizar operaciones computacionales recursivas, que son representativas de la computación humana. Esto pues las facultades cognitivas humanas para usar información contextual semántica y pragmática que facilita la comprensión sobrepasa en gran medida a la de cualquier computadora de las hasta ahora construidas.<sup>31</sup>

Se puede decir finalmente sobre el conexionismo que más que resolver el problema natural de la inteligencia se ha convertido más bien en una herramienta poderosa de la computación electrónica. Es revelador advertir que todos los proyectos relevantes para diseñar y construir grandes maquinarias paralelas hacen uso de estructuras altamente artificiales para la comunicación y el procesamiento. Esto se aplica a la Máquina de Conexión de la Thinking Machines Corporation (comunicación en hipercubo y en matriz rectangular), al Procesador Masivamente Paralelo de la NASA, al

Procesador Digital de Matrices del ICL (matriz rectangular), al Hipercubo de Intel Corporation (comunicación en hipercubo), al RP3 de IBM y a la Ultracomputadora de New York University (comunicación en red omega). Al final de cuentas, el conexionismo se ha dirigido más a la eficiencia computacional que a la explicación de las constricciones y alcances del mismo en la emulación de las redes neuronales naturales.<sup>32</sup>

### 1.3. Modularidad

A este modelo se adscriben figuras como J. Fodor (1983),<sup>33</sup> H. Gardner (1985),<sup>34</sup> y Z. Pylyshyn (1981),<sup>35</sup> y en cierta medida: R. Jackendoff (1987) con su dicotomía de mente computacional y mente fenomenológica.<sup>36</sup>

La modularidad parte del supuesto que la cognición se inicia con un sistema de entrada (Input) independiente (como las responsables de la percepción de objetos) y que procesos cognitivos de mayor generalidad serán dados en un sistema central. Los sistemas de entrada son el dominio específico en el sentido de que cada tipo de información recibida de procesamiento es diferente. Los sistemas de entrada se asemejan a los niveles de procesamiento de información en términos de influencias selectivas por variables.

La más importante propiedad es la *encapsulación*, que es proceso la información solo influye a partir de un dominio de módulo de entrada. El caso -por ejemplo- para el habla, es influenciado sólo por entradas de habla, no por contextos situacionales y lingüísticos. Los módulos de entrada son cognitivamente impenetrables, es decir no son objeto de control volitivo. Los sistemas de entrada se direccionan hacia sus salidas, así por ejemplo, el módulo del lenguaje para la salida de acceso léxico se da solución significados de palabras. Los sistemas de entrada operan independientemente uno de los otros y no se comunica. Finalmente, un módulo de entrada es asociado con la estructura neuronal específica.<sup>37</sup>

En contraste con los sistemas de entradas, los sistemas centrales están influidos por muchas diferentes variables. Los sistemas centrales tienen acceso a las salidas de todos los sistemas de entrada y a todo el conocimiento en la memoria. En esta dicotomía, los sistemas de entrada se consideran como sistemas computacionales, mientras que el sistema central corresponde a lo que el organis-

mo "cree". Los sistemas de entrada pueden ser estudiados como sistemas computacionales, mientras que los centrales no pueden ser considerados de tal manera. Esto porque son muchos los factores que influyen en sus procesos.<sup>38</sup>

En breve, el enfoque modular plantea una tricotomía de niveles, *sistemas de salida, de entrada y centrales*. Este enfoque comparte muchas de las premisas originales del modelo de procesamiento de información, particularmente el supuesto de sistemas separados de la percepción y de la acción. Esta explicación funcional de diferenciar sistemas de entrada y salida y entre sistema de entrada y central no sustenta substratos sólidos desde el punto de vista de la psicología de la percepción y de la neurofisiología, pues en el cerebro y el sistema nervioso en general no hay muestras de tales divisiones entre sistemas, por lo que no parece ser muy seguro a la hora de explicar la cognición.

Por otra parte es importante señalar que la noción similar de *encapsulamiento* ha sido muy fértil en la Programación orientada a objetos (POO) que ha representado en el fondo una nueva "filosofía" de la programación computacional. Dicha programación se centra alrededor de algunos conceptos de mayor generalidad como los tipo de datos abstractos y clases, los tipos de jerarquías (subclases), herencia (procesos de generalización y especialización) y polimorfismo. El dominio hacia los que se orienta dicha programación son los "objetos". Un objeto es una variable declarada que pertenece a una clase específica. Así, un objeto *encapsula* un estado, conteniendo una copia de todos los campos de datos que están definidos en una clase. La encapsulación es un principio usado cuando se desarrollan una estructura de un programa completa, tal que cada componente de un programa debe ser encapsulado en un diseño de decisión particular.<sup>39</sup>

Este concepto de encapsulación similar al enfoque de la modularidad ha sido exitoso en la POO, y resulta una alternativa significativa para los modelos de procesamiento de la información, pero de nuevo, resultan insatisfactorios a la hora de querer ser aplicados a la cognición humana, tal y como se ha señalado anteriormente.

## 2. Modelos de neurobiología cognoscitiva

Algunos de los enfoques más importantes surgidos en el contexto de las Ciencias Cognoscitivas y que por consecuencia intenta resolver el problema

del funcionamiento del la mente-cuerpo, han sido aquellos generados en el ámbito de la Neurociencia (neuroanatomía, neurofisiología, psicofisiología, neuropsicología, neurobiología, neurofarmacología), y pueden ser etiquetados como NEUROBIOLOGIA COGNOSCITIVA.

Un intento neurofilosófico por tratar de generar una teoría unificada de la mente-cerebro ha sido el dado Patricia Smith Churchland (1986).<sup>40</sup> Este enfoque propone un trabajo cooperativo entre los investigadores de la mente y los investigadores del sistema nervioso. Enfoques como el funcionalismo clásico (Newell, Simon, Minsky) no dan cuenta de los complejos y variados niveles con que trabaja un neurofisiólogo. Así un enfoque que solo da cuenta de las funciones psicológicas sin acudir a la neurociencia, es parcial. Para P. Smith C. resulta contraproducente para lograr una ciencia unificada de la mente-cerebro, postular que la psicología es irreducible a sus fundamentos neurofisiológicos y neuroanatómicos. Por el momento es conveniente postular que ambas ciencias coevolucionan y que seguirán haciéndolo por algún tiempo - señala P. Smith C., pero como se ha dado anteriormente en la historia de la ciencia, todo parece mostrar que van en camino de la reducción. Ejemplos claros de ello han sido las investigaciones en Bioquímica de la conducta, en Neurofarmacología, Neuroquímica y sobre las bases moleculares (biología molecular) del sistema nervioso. Consecuentemente desde esta postura, cualquier dualismo entre mente y cuerpo es claramente demostrado como erróneo. Otro enfoque reduccionista de los procesos mentales es el propuesto por el premio nobel (por el descubrimiento de la estructura de la doble hélice del ADN junto con J.D. Watson): Francis Crick (1994).<sup>41</sup> Su "hipótesis sorprendente" -como él la llama- consiste en que cualquier individuo, sus alegrías y aflicciones, sus recuerdos y ambiciones, su sentido de identidad personal, su libre albedrío, es de hecho el reflejo de la conducta de un vasto ensamblaje de células y de sus moléculas asociadas. Así por ejemplo, para que haya conciencia y memoria a corto plazo lo que se necesita es la actividad de circuitos reverberatorios, que se encargan de mantenerlas. Asimismo la conciencia requiere la actividad de varias áreas corticales así como del tálamo. Aún cuando su hipótesis ya no resulta tan sorprendente su aporte para la "búsqueda científica del alma" ha resultado muy influyente.

Pero el que pareciera más prometedor es el modelo propuesto por el también premio nobel (fisiología): G. Edelman y sus colaboradores en la Rockefeller University<sup>42</sup>, quienes se enfrentan a los enfoques de la inteligencia como procesamiento de la información y proponen una explicación neurobiológica de la inteligencia, planteando su teoría de la SELECCION DEL GRUPO NEURONAL (SGN), a partir de los modelos de conformación del desarrollo del sistema nervioso. Un enfoque similar y complementario al de Edelman ha sido del Jean Pierre Changeux, neurobiólogo del laboratorio molecular del Institut Pasteur en París. Changeux parte de los estudios de Edelman, añadiendo a la epigénesis neuronal, la impresión de la cultura en el desarrollo del cerebro después del nacimiento de los individuos, es decir añade aspectos antropogenéticos<sup>43</sup>.

Pero bien, la propuesta de Edelman y sus colaboradores se remonta al período de formación cerebro del embrión, cuando la selección entre células neuronales en competencia y sus procesos determinan la forma anatómica y los patrones de conectividad sináptica del sistema nervioso<sup>44</sup>. Esta selección para la conectividad se elabora mediante mecanismos evolutivos de adhesión y movimiento de células, crecimiento diferencial, división celular y muerte de células. Dada sus propiedades dinámicas, estos mecanismos selectivos introducen variación individual en las redes neuronales. Más tarde, durante la experiencia posnatal, la selección entre diversos grupos de células preexistentes, complementada por la modificación diferencial de fuerzas o eficacias sinápticas sin cambios en el patrón de conectividad, da forma al repertorio conductual del organismo de acuerdo con lo que posee por él valor adaptativo en su nicho<sup>45</sup>.

La SGN es una teoría que comienza a ponerse a prueba, y que de acuerdo con distintas experiencias ha comenzado a ser exitosa. Así, se han construido una serie de autómatas para la simulación computada para poner a prueba la consistencia de la SGN, así como para demostrar la habilidad de los sistemas de reconocimiento selectivos para realizar interesantes tareas de reconocimiento y categorización. Esos modelos pueden ser invaluable para ayudar a que los neurobiólogos se concentren en aspectos experimentales, que eventualmente ayuden a desarrollar computadoras capaces de desarrollar tareas de clasificación sensible, pues en la actualidad no se tienen<sup>46</sup>.

Para Edelman, en vista de la complejidad de los sistemas biológicos, es necesario comenzar a analizar esos sistemas en términos de las estructuras y funciones básicas necesarias y sus modos de origen, su desarrollo tanto como su evolución. La separación entre hardware y software implícita en la estrategia de la I.A. tradicional tiene que abandonarse, aun cuando haya servido de principio orientado en el desarrollo de las computadoras tipo von Neuman, que pueden ser máquinas lógicas -y en alguna medida máquinas culturales- pero no máquinas biológicas. Edelman considera que la I.A. solo se alcanzará en sistemas no-von Neuman en los que las variantes especializadas de hardware, basadas en el tema común de la selección y el pensamiento de la población, trábajarán sin programas para adaptarse a todos los ambientes particulares en los que se encuentren, tal como lo hacen los organismos biológicos. Los programas y la inteligencia basada en la comunicación podrán venir después<sup>47</sup>.

## Conclusiones

Al comprobarse que con tan sólo los aportes de las ingenierías eléctrica, electrónica, mecánica y computacional, no se tenían los instrumentos adecuados para simular o generar razonamiento en los modelos originales de la Inteligencia Artificial, fue necesario acudir a otras áreas como la Psicología y la Lingüística que pudieran dar criterios más precisos sobre la naturaleza del razonamiento en particular y de la Inteligencia y la Conciencia en general. Así como se acude a la Psicología y a la Lingüística en los años sesenta, hacia los setenta y ochenta se acude a otras disciplinas en busca de fuentes para un mejor conocimiento de los procesos de cognición humana, a disciplinas como la fisiología del cerebro y la neuroanatomía (neurociencia), a la antropología y la filosofía, generándose así un nuevo campo llamado CIENCIAS COGNOSCITIVAS, que comprende un total de seis áreas: Computación, Lingüística, Psicología Cognoscitiva, Neurociencia, Antropología y Filosofía. La Inteligencia Artificial ha hecho contacto con la Filosofía en muchos puntos, tales como la lógica, la epistemología, la filosofía del lenguaje, la filosofía de la ciencia y la filosofía de la mente. Esta última tiene por central el problema mente-cuerpo, que ha tenido una larga historia en Filosofía.

La filosofía contemporánea de la mente debe superar los idealismos y dualismos de su pasado histórico y más que partir de presupuestos metafísicos poco fiables o sin ninguna fundamentación, debe buscar un marco explicativo unificador en el contexto de las Ciencias Cognoscitivas, para los fenómenos neurofisiológicos y psicológicos y así aportar visiones de conjunto sobre los modelos de la cognición humana.

Los dos grandes grupos de modelos de la mente-cerebro aquí presentados, como lo han sido los del procesamiento de la información y como los de la neurobiología cognitiva, resultan fundamentales para los estudios de filosofía de la mente y no se podrá hacer ninguna aporte relevante a este campo si no son tomados en cuenta.

Ha faltado mencionar algunos modelos de la psico-cognoscitiva, como los de J.A. Anderson<sup>48</sup> y R. Sternberg<sup>49</sup>, así como los modelos de ontología subjetiva J. de Searle (1980,1992)<sup>50</sup>, los modelos hermenéutico-existenciales de T. Winograd y F. Flores (1987)<sup>51</sup>, y de F. Varela, E. Thomson y E. Rosch (1993)<sup>52</sup>, estos últimos autores buscan introducir dentro de las Ciencias Cognoscitivas la “experiencia humana” a partir de la tradición del budismo zen.

Así enfoques sumamente controversiales han sido el de H. Putnam (1989)<sup>53</sup>, que se retracta de su postura de los años sesenta, cuando postuló el enfoque computacional de la mente. Asimismo el físico R. Penrose (1989, 1994)<sup>54</sup> ha atacado fieramente los intentos de la Inteligencia Artificial, postulando el carácter cuántico del pensamiento y la necesidad de una nueva física para lograr los fines de la I.A. Este enfoque ha despertado fieras discusiones.

Finalmente uno de los enfoques más influyentes en Ciencias Cognoscitivas y que se presenta como la propuesta más original a la filosofía de la mente contemporánea ha sido la esbozada por el filósofo D. Dennett (1991, 1994).<sup>55</sup> Dennett presenta una teoría integradora para la explicación de la conciencia, tanto desde el punto de vista de su estructura neuronal, como desde el punto de vista psico-evolutivo.<sup>56</sup> Para Dennett las mentes conscientes humanas son máquinas virtuales más o menos seriales implantadas en un hardware eminentemente paralelo suministrado por la evolución.<sup>56</sup>

## Notas

1. Shannon,C.E. 1938 .“A Symbolic Analysis of Relay an Switching Circuits”, tesis del M.I.T.; publica-

da en *Transactions of American Institute of Electrical Engineers*. 57, 1-11.

2. McCulloch,W y Pitts, W.1943.“A Logical Calculus of the Ideas Immanent in Nervous Activity”, *Bulletin of Mathematical Biophysics*, 5, 115-33.

3. Wiener, N. 1948. *Cybernetics, or Control and Communication in the Animal and the Machine*, Cambridge, Mass., MIT Press. Traducción al español por Francisco Martín: “*Cibernética o el control y comunicación en animales y máquinas*. Barcelona: Tusquets editores, 1985.

4. Turing, A.M.1950.“Computing Machinery and Intelligence”, *Mind*, Vol. LIX, No. 236, reimpresión en: Hofstadter, D. & D.C. Dennet. 1981. *The Mind's I. Fantasies and Reflections on Self and Soul*. New York,-Batam Books, pp. 53-67. Traducción al español por Francisco Martín en: Anderson, A.R (comp) *Controversias Sobre Mentes y Máquinas*. Barcelona: Tusquets editores, 1984.

5. Gardner, H. (1988). *La Nueva Ciencia de la Mente. Historia de la Revolución Cognitiva*. Barcelona: Ediciones Paidós. 1ra. edición en español: 1988. Traductor L. Wolfson.

6. Rodríguez R., R.J. (1994). “Epistemología e Inteligencia Artificial. ¿Es posible una epistemología androide?” En: *Repertorio Científico*. Vol. 2. No. 1. Enero-Abril. Accesable vía internet en el Gopher: ARENA-L.UNED.AC.CR: Revistas: Repertorio Científico: Vol. 2-1.

7. Shannon. C., Idem.

8. Turing, A.M. 1936. “On Computable Numbers with Applications to the Entscheidungs-Problem”, en: *Proceedings of London Mathematical Society*, Serie 2, 42, 230-65.

9. Cfr.

Hofstadter, D.R. 1979. *Gödel, Escher, Bach: An Eternal Golden Braid*. New York: Basic Books. Existe traducción al español: *Goedel, Escher Bach. Un eterno y gracil bucle*. Barcelona: Tusquets Editores, 1987.

Putnam, H. 1960. “Minds and Machines” en S. Hook(comp). *Dimensions of Mind*, New York, New York University Press. Traducción al español por F. Martín en: Anderson, A.R (comp) *Controversias Sobre Mentes y Máquinas*. Barcelona: Tusquets editores, 1984.

10. Cfr.

Newell, A., J.C. Shaw, & H.A. Simon, 1963. “Empirical Explorations with the Logic Theory Machine a Case Study in Heuristics” Newell, A.& H.A. Simon. 1963. “GPS, A Program that simulates Human Thought” Ambos en: Feigenbaum, E.A. & J. Feldman. *Computers and Thought*. U.S.A.: McGraw-Hill, Inc. 1963.

Simon, H.A. 1979. “Information Processing Models of Cognition”. En: *Annual Review of Psychology*. 30:363-96.

Posner, M.I. & McLeod. 1982. “Information Processing Models-In Search Of Elementary Operations”. En: *Annual Review of Psychology*. 33: 477-514.

11. Cfr.
- Newell, A. 1980. "Physical symbol systems." en: *Cognitive Science*. 4:135-83. Gutiérrez, C. 1993a. Epistemología e Informática. Guía de Estudio. EUNED: San José, pp. 155-164.
12. Gutiérrez, C. 1993a, *Idem*.
13. Massaro, D.W. 1993. "Information Processing Models: Microscopes of the Mind" en: *Annual Review of Psychology*. 44: 383-425.
14. Cfr.
- Searle, J.R. 1980. "Mind, Brains and Programs", *The Behavioral and Brain Sciences*, 3,417-57. Reimpresión en: Hofstadter, D. & D.C. Dennet. 1981. *The Mind's I. Fantasies and Reflections on Self and Soul*. New York, Batam Books, pp. 353-373.
- Putnam, H. 1989. *Representation and Reality*. Cambridge, Mass. The MIT Press.
15. Schwartz, J.T. 1988. "El nuevo conexionismo: desarrollando relaciones entre la neurociencia y la inteligencia artificial", en: Graubard, S.R. *El Nuevo debate sobre la Inteligencia Artificial*. Gedisa Editorial: Barcelona, 1988.
16. *Idem*.
17. *Idem*.
18. *Idem*.
19. Thompson, R.F. 1975. *Introducción a la Psicología Fisilógica*. México: Harla. 1977. Tr. R. Naranjo V.
20. Reeke, G. y G. Edelman (1988). "Cerebros reales e inteligencia artificial" en Graubard, S.R. (1988) *El Nuevo debate sobre la Inteligencia Artificial*. Gedisa Editorial: Barcelona, 1988, pp. 167-21.
21. Rumelhart D.E., J. McClelland y PDP Research Group (ed). 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, V. 1, 2. Cambridge, Mass.: MIT Press.
22. Smolensky, P. 1988. "On the proper treatment of connectionism". *Behavioral Brain Sciences*. 11:1-73.
23. Massaro, D.W. 1993. p. 415.
24. *Idem*.
25. *Idem*.
26. Gutiérrez, C., 1993a, pp. 191-192.
27. Massaro, D.W., 1993, p. 416.
28. Gutiérrez, C. 1993a. *Idem*.
29. Churchland, P.M. 1984. *Matter and consciousness. A Contemporary Introduction to the Philosophy of Mind*. The MIT Press: Cambridge: Third Printing: 1990, p. 151.
30. Cfr
- Churchland, P.M. 1984, pp. 156-185 Churchland, P. M. 1989. *A Neurocomputational Perspective. The Nature of Mind and the Structure of Science*. The MIT Press: Cambridge, Massachusetts, Third Printing, 1993, pp. 197-230. Schwartz, J.T. 1988, p. 149.
31. Gutiérrez, C., 1993a, pp. 193-194.
32. Schwartz, J.T., 1988, p. 161.
33. Fodor, J.A. 1983. *Modularity of Mind*. Cambridge, Mass., Bradford.
34. Gardner, H. 1985. *Idem*.
35. Fodor, J.A. y Z.W. Pylyshyn. 1981. "How direct is the visual perception?: some reflections on Gibson's ecological approach". En: *Cognition*. 9:139-96.
36. Jackendoff, R. 1987. *Consciousness and the Computational Mind*. Cambridge, Mass.: The MIT Press, A Bradford Book.
37. Massaro, D.W., 1993, p. 418.
38. *Idem*.
39. Cfr.
- Wiener, R.S. & L.S. Pinson. 1988. *An Introduction to Object-Oriented Programming and C ++*. USA-Canada: Addison-Wesley Publishing Company. Cood, P. y E. Yourdon. 1991. *Object-Oriented Design*. U.S.A. Yordon Press. Pres. Pretince Building.
40. Churchland, P.S. (1986). *Neurophilosophy. Toward a Unified Science of the Mind/Brain*. The MIT Press: Cambridge: Seventh Printing, 1990.
41. Crick, F. 1994. *The Astonishing Hypothesis: The Scientific Search for the Soul*. New York: Charles Sribners's son.
42. Edelman. G. M. 1987. *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Basic Books.
43. Changeaux, J.P. 1985. *Neuronal Man. The Biology of Mind*. New York: Pantheon Books.
44. Reeke, G. y G. Edelman, 1988, p. 183.
45. *Ibid.*, pp. 184-185.
46. *Idem*.
47. *Ibid.* p. 197.
48. Anderson, A.R. 1983. *The Architecture of Cognition*. Cambridge, Mass. Harvard University Press.
49. Stemberg, R.J. 1983. "Components of Human intelligence", en *Cognition*, 15:1-48.
50. Searle, J.R. (1992). *The Rediscovery of Mind*. The MIT Press, Cambridge, Third Printing.
51. Winograd, T. & F. Flores (1986). *Understanding Computers and Cognition. A New Foundations for Design*. Ablex Publishing Corporation. Norwood, New Yersey, Third Printing 1987.
52. Varela, F. J., E. Thomson, & E. Rosch (1991). *The Embodied Mind. Cognitive Science and Human Experience*. The MIT Press, Cambridge, 1993.
53. Putnam, 1989, *Idem*.
54. Cfr.
- Penrose, R. 1989. *The Emperor's New Mind*. New York: Penguin Books: 1991.
- Penrose, R. 1994. *Shadows of the Mind: A Search for the Missing Science of Consciousness* New York: Oxford University Press.
55. Cfr.
- Dennett, D.C. 1991. *Consciousness Explained*. Little, Brown and Company, Boston Gutiérrez, C. 1993b.

“La explicación de la mente: A propósito de un libro de Daniel Dennet”. en *Revista de Filosofía de la U.C.R.* V. XXXI, No. 74, Julio 1993.

Beardsley, T. 1996. “Dennett’s Dangerous Idea”. en: *Scientific American*, February. 56. Gutiérrez, C. 1993a., p. 321.

Rodolfo J. Rodríguez R.  
Apto. 202-1100 Tibas  
E-Mail: rodolfo @ cariari.ucr.ac.cr  
rod ro @ ulativa. ac.cr  
Programa de posgrado en  
Ciencias Cognoscitivas  
Universidad de Costa Rica  
URL Home Page: <http://www.ucr.ac.cr/cogsci>  
cogsci.HTMLM.  
cogsi @ cariari.ucr.ac.cr.