

¿Máquinas inteligentes?

Enrique Soto

Instituto de Fisiología
Universidad Autónoma de Puebla

A Julio y Antonella

Una pregunta que se hacen los estudiosos de las neurociencias y de la computación es sobre la posibilidad de desarrollar sistemas inteligentes. Esta pregunta es relevante no sólo desde el punto de vista tecnológico, que es el más evidente luego de treinta años de robots televisivos, sino en sus aspectos teóricos, ya que una respuesta afirmativa implica adherirse a la idea de que no existe en nuestros cerebros nada semejante a “un soplo divino”, un alma o espíritu. Una respuesta negativa, si bien no necesariamente representa que se adopte una postura idealista, si implica que la estructura de la entidad que da origen a nuestra conciencia es tal que, en principio, es irreproducible y, con toda probabilidad, incognoscible.

El problema no es trivial y tiene que ver también con la posibilidad de conocernos a nosotros mismos, de entender cómo es que pensamos y tenemos conciencia, cómo se origina la imaginación, nuestra memoria, gustos y pasiones. En última instancia, las neurociencias pretenden entender la inteligencia y la conciencia humanas.

Es importante distinguir entre lo que pudiera ser una “máquina pensante” y lo que son las computadoras. Al hablar de máquinas inteligentes, la mayoría de nosotros tendemos a pensar en nuestra computadora personal y su increíble necesidad. De ninguna manera es a éstas que me refiero. Las computadoras con las que usualmente tenemos relación son capaces de realizar ciertas operaciones básicas a muy alta velocidad y nada más. No es válido hacer trivial el asunto preguntando por ejemplo: ¿si acaricio una computadora le gusta y le excita o no? Claro que no

se trata de saber si una microcomputadora responde a mis caricias, o si las computadoras con Windows son más sabias que aquellas que usan MS-DOS. Se trata de imaginar y entender cuáles son los elementos básicos que constituyen eso que llamamos inteligencia, y discutir si es posible generar sistemas altamente organizados capaces de reproducir dichos procesos.

La inteligencia natural

Como punto de partida acordemos que la inteligencia se refiere a la capacidad para resolver problemas. Esta capacidad se desarrolla a lo largo de la evolución y ofrece a las distintas especies mejores niveles de adaptación a su medio ambiente. Ciertamente queda aún sin respuesta una pregunta básica: ¿qué ventajas precisas ofrece la aparición de la conciencia del yo a la especie humana?

Desde el punto de vista de las células que participan en los procesos mentales, podemos decir que el sistema nervioso está formado por neuronas y células gliales. Las neuronas son las células que forman circuitos y redes, y tienen propiedades que permiten explicar los principios de organización del cerebro.

Las neuronas pueden concebirse como unidades operacionales que modifican su actividad en función de ciertos cambios en el medio. Estas células reciben influencias de otras neuronas o de células sensoriales por medio de conexiones denominadas sinapsis. Las neuronas tienen propiedades de tipo no lineal que les permiten responder de forma compleja y, si bien está aún pendiente caracterizar exhaustivamente sus propiedades, el conocimiento que tenemos de ellas per-

mite construir modelos capaces de reproducir la actividad de algunos conjuntos neuronales.

El cerebro debe sus propiedades al conjunto de elementos que lo constituyen y a las relaciones estructurales y funcionales que entre ellos establecen. Tenemos así varios factores que caracterizan lo que podría denominarse como la complejidad del sistema nervioso: desde el punto de vista numérico está constituido por billones de unidades (neuronas); éstas poseen una dinámica no lineal que les permite responder en función de su historia y estado actual; existe un complejo lenguaje químico por el cual los elementos del sistema se comunican entre sí y con su medio externo; las neuronas están organizadas estructuralmente en forma de redes y conjuntos de redes que forman subsistemas funcionales; la organización del sistema es flexible y puede modificarse en función de sus interacciones con el medio, desde el nivel de las unidades constitutivas, hasta el nivel de las redes y la manera como éstas interactúan conformando grandes sistemas funcionales (véase Figura 1).

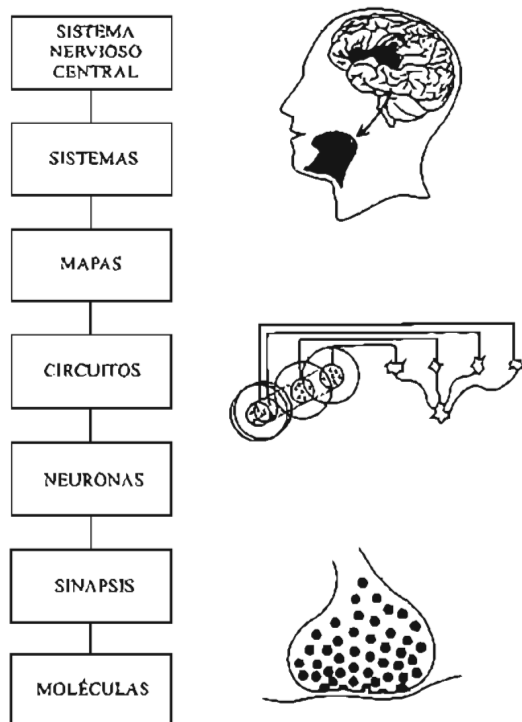


Figura 1. Para comprender el funcionamiento del sistema nervioso su estudio debe abordarse a varios niveles; desde el molecular hasta el sistémico que implica la comprensión global del cerebro.

¿De dónde surge esta complejidad estructural y funcional, cómo fue diseñada? La respuesta la concibo en el marco de referencia de la teoría de la evolución; la selección natural en el transcurso de millones de años ha llevado a la organización celular desde un simple conjunto con algunas interacciones elementales, al desarrollo del cerebro. Su sorprendente adaptación al medio ambiente es explicable ya que es precisamente el medio el escultor de esta obra.

El sistema nervioso, si bien está formado por unidades sencillas, tiene propiedades que son completamente diferentes a las de sus elementos constitutivos. Todavía más, difícilmente podemos, a partir del conocimiento de las unidades que forman el cerebro, entender los procesos psíquicos, ya que éstos no son reductibles a aquéllas. Se requiere de un acercamiento multidisciplinario, y en varios niveles, para poder desentrañar los procesos funcionales y las interacciones que determinan en el cerebro la emergencia del pensamiento y la conciencia.

La unidad de la mente y el pensamiento

Una idea muy difundida cuando se piensa en la actividad mental es la de concebir a la mente como un todo indivisible. Esta concepción es natural, sobre todo si hacemos un poco de introspección; en la imaginación nos aparece nuestro yo unitario y podemos pensarnos ahí, plácidamente, asoleándonos en una playa, pensándonos. En ningún momento requerimos de esfuerzo alguno para unir elementos que nos conformen; nos concebimos como un ser único e indivisible. Sin embargo, las funciones psíquicas superiores están organizadas de una forma a la que podríamos denominar modular, esto es, que hay regiones del cerebro que procesan únicamente cierto tipo de información, y que áreas diferentes de procesamiento pueden funcionar en forma independiente del resto. Existen reportes en la literatura que apuntan en esta dirección. Por ejemplo, sujetos que luego de una lesión cerebral pierden la capacidad de reconocer alguna parte de su cuerpo o de su entorno (véase Figura 2).

Aún más relevantes para el estudio de la inteligencia y la conciencia son los individuos que a

pesar de estar completamente ciegos, cosa que puede demostrarse por la carencia de todo tipo de reflejos visuales y por las lesiones que poseen, niegan estar ciegos; es decir, de alguna forma han perdido la conciencia de que existe un mundo susceptible de ser observado. Igualmente hay sujetos en los que puede demostrarse una cierta capacidad visual, por ejemplo cuando van a coger un objeto desconocido ponen la mano en posición adecuada antes de tocarlo y afirman, sin embargo, que no pueden ver.

Son muchas las variantes de alteración parcial de la conciencia por lesiones cerebrales. Un fenómeno notable es el del "miembro fantasma". Esto sucede en algunas personas que han sufrido una amputación y sienten dolor en el miembro inexistente, a pesar de tener una conciencia explícita de su ausencia. Respecto a las funciones ideativas más abstractas hay también ejemplos que indican su modularidad; destacan, entre otros, los casos de sujetos con retraso cognitivo profundo que, sin embargo, poseen habilidades sorprendentes para realizar operaciones matemáticas, o una extraordinaria capacidad de interpretación musical, o aquellos en que lesiones de las regiones occipitales les producen una incapacidad para reconocer visualmente caras conocidas.

Todos estos casos indican que eso que llamamos inteligencia no es indivisible, sino que tenemos inteligencias: diferentes capacidades para resolver problemas y que, muy probablemente, la actividad mental se realiza en módulos que operan de forma paralela y semi independiente. Finalmente, de forma que aún nos es desconocida, toda esta actividad se integra dando su unidad a la actividad mental.

Quizá el ejemplo más sobresaliente de unidad es el que se deriva de la percepción visual, que nos permite apreciar claramente los objetos como un todo y que, sin embargo, se genera a partir de un mundo constituido de pequeños puntos coloreados. Es más, las imágenes visuales se proyectan en las regiones occipitales del cerebro divididas en dos mitades, de forma que los extremos están representados en las regiones centrales. A pesar de ello, el mundo visual aparece mentalmente como un todo, y podemos reconocer claramente los objetos, su continuidad, textura y

color; sin embargo, la actividad sensorial que produce esta percepción es discontinua estructural y temporalmente. Más aún, nos resulta invisible un área del campo visual en la que carecemos de receptores sensoriales (conos y bastones). Todos poseemos esta región denominada punto ciego; sin embargo, no experimentamos su presencia. Entonces, la "impresión" de unidad que tenemos del campo visual se debe a las propiedades de nuestra actividad mental, pero puede claramente ser disecada en elementos constitutivos independientes.

Otro aspecto relacionado es el que tiene que ver con el surgimiento de la inteligencia a lo largo de la evolución. Existen corrientes que ven en el pensamiento abstracto y en la inteligencia algo exclusivo del hombre; nuevamente aparece la inteligencia y la conciencia como un todo, que se posee o no.

Si atendemos a los procesos fisiológicos y características estructurales de diversos mamíferos, así como a las evidencias derivadas del análisis de comportamiento, habremos de aceptar que es muy probable que, en animales no humanos, existan diversos grados de conciencia e ideación abstracta que les permiten resolver problemas, y en ciertas circunstancias, ejecutar acciones que no pueden explicarse por simples instintos de supervivencia o similares.

En síntesis, lo anterior indica que la conciencia y la actividad mental superior no son un todo unitario, que se posee o no, sino que se producen como resultado de la actividad concertada de sistemas neuronales.

La inteligencia artificial

Obviamente la ficción lleva la delantera y abundan en la literatura y el cine los robots y las máquinas inteligentes. No puede decirse que en la actualidad exista una sola máquina que piensa; sin embargo, existen sistemas que presentan un procesamiento rudimentario, que manejan infor-



Figura 2. Dibujo de reloj hecho por un niño disléxico, de siete años once meses de edad, con negligencia del campo izquierdo Stein y Walsh, 1997.

mación compleja y que poseen algo similar a la intuición, elementos que sugieren que es factible realizar, a largo plazo, la síntesis del pensamiento (véase Figura 3).

Como ejemplo cabe mencionar el tipo de máquina que han desarrollado en Cycorp en Austin, Texas, basados en el análisis crítico de experimentos previos de inteligencia artificial, que lamentablemente terminan por lo general en sistemas que mucho se parecen a un sabio idiota: máquinas que traducen fuera de contexto, o que en el diagnóstico médico aceptan como factible que un automóvil pueda tener rubeola. El objetivo ha sido entonces crear máquinas que traten el contexto en que se dan las cosas; por ejemplo, sabiendo que tanto el aire como un libro están ambos "dentro de" un cuarto, sea claro para la máquina que el libro puede sacarse de ahí y el aire no. Para crear esta contextualización, en una etapa inicial, la máquina aprende miles de con-

ceptos, algunos de ellos no simples definiciones, sino ideas operativas y en ocasiones contradictorias. Aún más, el diseño de aprendizaje del sistema permite introducir contextos de ficción, lo que le confiere la posibilidad de entender metáforas.

Para 1994 este sistema era capaz, ante una pregunta abstracta tal como "muestra una persona feliz", de responder presentando una imagen con la leyenda "un hombre viendo a su hija que aprende a caminar". Éste es un ejemplo de lo que hemos alcanzado y parece que las máquinas, o más bien, los sistemas de máquinas capaces de aprender de su experiencia, desarrollar conceptos, entender contextos y generar decisiones intuitivas, no parecen tan lejanos.

Un punto importante al desarrollar sistemas de inteligencia artificial es que, si se aprovechan únicamente la velocidad de procesamiento y la capacidad de "computar" de la máquina, rápidamente se tiene el problema de la explosión combinatoria. ¿En qué consiste este problema? Si se pretende que una máquina tenga en cuenta todas las posibles soluciones al enfrentar una cierta circunstancia, encontraremos que existen más situaciones posibles de las que la máquina puede manejar. Igualmente sucede con nuestro cerebro; no podemos manejar todos los escenarios de un problema; de ahí la necesidad de la intuición. Las máquinas actualmente desarrollan procesos de intuición por la vía de la lógica de conjuntos difusos y de lo que Edelman ha denominado "degeneración de la definición de grupo".

Otra área paralela de la inteligencia artificial es el trabajo que se ha desarrollado con pequeños robots y diferentes artilugios mecánicos que interactúan de forma variada con el mundo que les rodea. Hay grupos de investigadores que trabajan con robots que aprenden a caminar con razonable éxito en el mundo real. Hay robots artificiales, otros que trabajan en el mundo virtual, que roban latas de Coca Cola, otros más que han logrado desarrollar sistemas de control motor distribuido. Algunos de estos artefactos imitan ya la actividad locomotora de los insectos, y la multiplicidad de sus diseños sugiere en cierta forma que estamos en una etapa análoga a la del cámbrico, cuando la naturaleza desarrolló millones de variantes en el diseño de los seres vivos. El

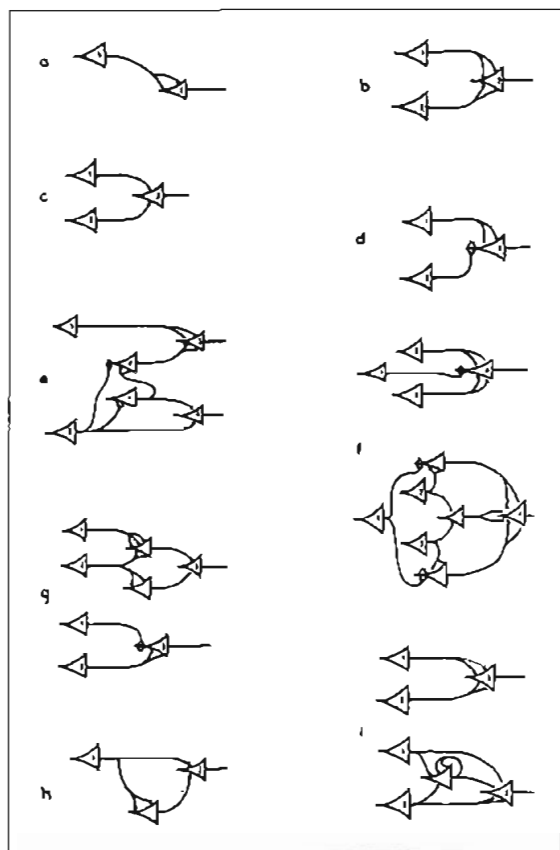


Figura 3. Neuronas ideales de McCulloch y Pitts (1943). Estos autores fueron de los primeros en pensar las neuronas como pequeños circuitos eléctricos. Su objetivo era identificar analogías entre los procesos biológicos y los electrónicos.

paralelismo entre la explosiva variedad en los insectos y la diversidad de diseños de robots, define claramente el hecho de que en este campo nos encontramos en una etapa inicial de evolución y tanteando por ensayo y error. Lamentablemente es eso: el diseño actual de máquinas autónomas no ha resultado más que en insectos, animáculos destinados a realizar una danza repetitiva. La pregunta es: ¿serán estos artilugios mecánicos los ancestros de magníficos robots del tipo de Daneel o Giskard?

Algunos aspectos que habremos de indagar para decidir si es factible crear sistemas inteligentes son, entre otros, si el cerebro es un sistema que, con base en una entrada de información, realiza ciertas operaciones (descarga neuronal, activación sináptica, activación de una red), o si existe en nuestro cerebro algo como un analizador de significados. En el primer caso estaríamos entonces en posibilidad de crear sistemas que realizan operaciones que pudieran ser análogas a las del cerebro y cuya actividad sería independiente de los elementos constitutivos; es decir, no se requieren neuronas para generar la actividad mental. En el segundo caso, los elementos constitutivos tienen propiedades que les confieren la capacidad de análisis simbólico, sus propiedades son las del sistema en conjunto y son, por tanto, probablemente irreproducibles. Habremos entonces de definir si la actividad mental es independiente del sistema en que se encuentra materializada.

Comprender el significado, interpretar un sistema de señales o elementos como simbólico, es comprender el sentido y especificar lo que significan esos elementos (véase Figura 4). La comprensión se realiza en el marco de una coherencia que en principio parece especificable bajo ciertas reglas que, en última instancia, tienen que ver con la realidad. Como sea, si verdaderamente nuestra capacidad de pensar se origina en la capacidad de nuestro cerebro para manipular símbolos de manera "automática", entonces es factible reproducir el pensar por medio de un sistema formal como lo son las computadoras; los formalistas tienen un lema que resume en cierta forma estas ideas: "si cuidas la sintaxis, la semántica se cuidará sola".

Otro aspecto fundamental que diferencia a los



Figura 4. Los programas de computadora son formales, cuidan la sintaxis; la mente humana tiene contenido simbólico, cuida la semántica. Tomado de Searle, 1990.

hombres de las computadoras es el de la individualidad y el libre albedrío. Evidentemente, el que un sistema aprenda de su experiencia le confiere ya cierta "individualidad". Por otra parte pienso que el desarrollo de la teoría de los sistemas dinámicos y de la teoría del caos, provee, por primera vez, maneras racionales de encarar esta cuestión. En la teoría matemática de los sistemas dinámicos se conoce que ciertos sistemas que se denominan caóticos evolucionan de manera impredecible, ello a pesar de tener una causalidad perfectamente definida. La manera como estos sistemas evolucionan, su trayectoria, depende especialmente de sus condiciones iniciales. Es decir, en un sistema caótico, pequeñas variaciones en las condiciones iniciales determinan cambios notables en su devenir. De hecho, perturbaciones mínimas pueden también alterar de manera notable la trayectoria del sistema.

Un ejemplo de estos sistemas lo constituye el clima. Podemos sin duda especificar los procesos que intervienen y determinan el clima, sin embargo, por tratarse de un fenómeno con una dinámica caótica, resulta impredecible a largo plazo (véase Figura 5).

De lo anterior se desprende que fenómenos como nuestra individualidad son susceptibles de análisis y muy probablemente se generen como resultado de una dinámica no lineal, caótica, en nuestro sistema nervioso. De ser así, es factible

introducir no linealidades en el diseño de sistemas, lo que determinaría que adquirieran una dinámica caótica. Entonces, variaciones iniciales, o pequeñas perturbaciones generarán un sistema impredecible a largo plazo, o sea, una cierta individualidad caprichosa.

La urea y la mente

Cabe destacar el paralelismo que se da entre lo que fueron las ideas vitalistas y lo que son actualmente las explicaciones dualistas-idealistas de la mente. En su momento, el vitalismo fue un recurso teórico que pretendió dar cuenta del fenómeno de la vida. Actualmente, los dualistas pretenden explicar la mente y el pensamiento. Si se examina la historia del vitalismo se encuentran diversas posiciones, desde las que consideran que el "principio vital" tiene su origen en la divinidad, hasta las que consideran que este principio vital se extingue con la muerte. De forma similar, existen dualistas que asignan el origen de la mente a la divinidad, y que en su expresión más elemental, pero dominante, consideran que esta mente, el ánima, sale del cuerpo en el momento de la muerte para reunirse con la divinidad. Por otra parte tenemos dualistas mucho más evolucionados que no atribuyen un origen divino a la mente, pero que la conciben como una entidad que se relaciona con el mundo material a través de interacciones —desconocidas— que ocurren en ciertas áreas del cerebro del hombre.

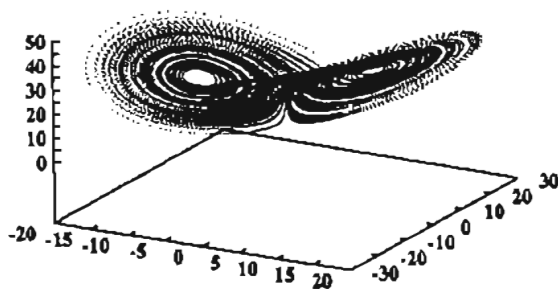


Figura 5. Atractor de Lorenz. Variaciones minúsculas de las condiciones iniciales generan trayectorias diferentes, pero siempre con formas análogas.

Uno de los principales postulados del vitalismo fue que las moléculas orgánicas no podían ser sintetizadas a partir de componentes elementales. Se pensaba que su síntesis era posible sólo en los seres vivos, lo cual les confería esta cualidad —el ser vivo. Cuando Whöler logró en 1828 la síntesis de la urea en el laboratorio, gran parte del edificio teórico del vitalismo se vino abajo. Ello a pesar de que la urea es una de las moléculas orgánicas más sencillas que se producen en el organismo. Whöler ofreció al mundo de la ciencia un dato empírico elemental que contradecía una de las ideas fundamentales del vitalismo; que la vida era irreductible y estaba presente en todas las moléculas que forman los seres vivos. Se inició así la decadencia del pensamiento vitalista, quedando demostrado que era posible sintetizar los elementos constitutivos de lo vivo; se puso en tela de juicio la existencia de propiedades exclusivas "inmanentes" a los seres vivos. Así como la Tierra, y luego el Sol, perdieron su lugar en el centro del universo, la vida perdió su carácter único y especial, pasando del dominio de la metafísica al de la química.

Un caso análogo ocurre con las máquinas que aprenden conceptos, o utilizan algoritmos genéticos, o perceptrones, o deciden intuitivamente, o juegan ajedrez, o reconocen caras, o leen la letra manuscrita y clasifican el correo. Alan Turing (uno de los creadores de la computación) decía que una máquina será inteligente cuando un explorador humano hable con ella y no pueda discernir si habla con una persona o una máquina; actualmente hay máquinas que logran sostener una conversación sin hacerse notar por espacio de más de media hora. En mi opinión, esto sugiere que las máquinas que juegan ajedrez y sus congéneres parlanchines, constituyen la prueba de que es posible realizar algo como la síntesis de la inteligencia. Ciertamente el juego del ajedrez o el reconocimiento de la letra manuscrita no es la inteligencia, tal como la urea no es la vida. Sin embargo, ambas son formas de expresión elemental de un fenómeno cuya síntesis es incompatible, uno con el vitalismo, el otro con el dualismo.

En su conjunto existen distintas máquinas especializadas en realizar ciertas funciones; lo inte-

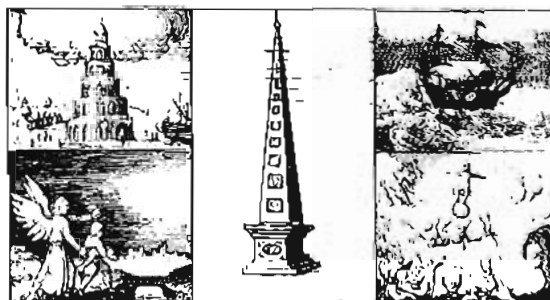


Figura 6. Dibujo publicado en 1619 por el místico paracelsiano Robert Fludd, quien otorgó una importancia fundamental a la capacidad imaginativa.

resante será cuando puedan incorporarse como elementos operativos de un sistema único. En tal caso, si construimos un sistema que incorpore las cualidades de estas máquinas, ¿tendrá propiedades que son la suma de sus elementos o emergerán nuevas propiedades? ¿Podremos incorporar en estos sistemas una individualidad autónoma y capacidad de aprender de su experiencia?

Concluyendo, aceptar la posibilidad de que existan máquinas que piensan, implica que la mente y el cerebro son entidades relacionadas funcionalmente. El desarrollo de las máquinas que piensan y en general de la inteligencia artificial tiene consecuencias prácticas evidentes, así como implicaciones teóricas fundamentales. Es muy probable que una respuesta clara a estas cuestiones deba esperar años. Estamos ante la desacralización del pensamiento; de entidad insustancial, sople divino y atributo único del hombre, el pensamiento y la conciencia podrían pasar a ser procesos que compartimos con entidades no vitales altamente organizadas. ¿Estaremos ya en la antecámara de los robots inteligentes?

El desarrollo de las neurociencias coincide con el gran avance en los terrenos de la computación y la microelectrónica. Se conjugan así preguntas, métodos y explicaciones que tienen el potencial para realizar la síntesis de la inteligencia. Ciertamente parece que nos toca vivir una época en la que se han de revolucionar nuestras ideas sobre la inteligencia, el hombre y la mente.

La posibilidad cada vez más evidente de que se puedan construir máquinas pensantes constituye un hecho doble, inquietante y fascinante. Inquietante porque pondrá al hombre y al humanismo un signo de interrogación que posiblemente signifique el final de la idea del hombre como ser único, singular e irrepetible; inquietante, también, por sus posibles implicaciones en la vida de las generaciones futuras; fascinante por el potencial explicativo que tiene el lograr que un conjunto de elementos, sean éstos neuronas o transistores, ordenados de manera compleja, sean capaces de desarrollar procesos psíquicos.

Lecturas recomendadas

- Asimov, I., *Los robots del amanecer*, Plaza y Janés, 1983.
- Crick, F., *La búsqueda científica del alma*, Editorial Debate, Madrid, 1994.
- Churchland, P.M., *Materia y conciencia*, Gedisa, 1992.
- Churchland, P.M. y Churchland, P.S., "Could a machine think", *Scientific American*, Vol. 262(1), 1990, pp. 26-31.
- Searle, J.R., "Is the brain's mind a computer program?" *Scientific American*, Vol. 262(1), 1990, 20-25.
- Edelman, G.M., *Neural Darwinism*, Basic Books, 1987.
- Hall, T.S., *History of general physiology*, T.S. Hall, Chicago, 1969.
- Haugeland, J., *La inteligencia artificial*, Siglo XXI, México, 1988.
- Lenat, D.B., "Artificial intelligence", *Scientific American*, Vol. 273(3), 1995, pp. 62-64.
- Minsky, M., "Will robots inherit the earth?", *Scientific American*, Vol. 271(4), 1994, pp. 87-91.
- Popper, K.R. y Eccles, J.C., *El yo y su cerebro*, Labor Universitaria, 1982.
- Sacks, O., *Un antropólogo en Marte*, Anagrama, 1997.
- Stein, J. y Walsh, W., "To see but not to read; the magnocellular theory of dyslexia", *Trends in Neuroscience*, Vol. 20, 1997, pp. 147-152.

