

**LAS ONTOLOGÍAS DEL PARADIGMA COGNITIVO:
COMPUTACIONALISMO, CONEXIONISMO Y
MATERIALISMO EMERGENTISTA¹**

Ontologies of the cognitive paradigm: computationalism, connectionism and
emergent materialism

Marcelo Díaz*
Alex Espinoza**

¿Cómo podré conmoveros? ¿No conseguirán mis súplicas que os apiadéis de vuestra criatura, que suplica vuestra compasión y bondad? Creedme, Frankenstein: yo era bueno; mi espíritu estaba lleno de amor y humanidad, pero estoy solo, horriblemente solo. Vos, mi creador, me odiáis ¿Qué puedo esperar de aquellos que no me deben nada? Me odian y me rechazan (...). Escucha mi relato: y cuando lo hayas oído, máldiceme o apiádate de mí, según lo que creas que merezco. Pero, escúchame. Las leyes humanas permiten que los culpables, por malvados que sean, hablen en defensa propia antes de ser condenados. Escúchame, Frankenstein. Me acusas de asesinato; y sin embargo destruirías, con la conciencia tranquila, a tu propia criatura. ¡Loada sea la eterna justicia del hombre! Pero no pido que me perdones; escúchame y luego, si puedes, y si quieres, destruye la obra que creaste con tus propias manos.

Mary Shelley

Resumen

¿Qué ontología subyace en los principales modelos de la cognición? ¿Qué trajo el paradigma cognitivo? ¿Qué compromiso filosófico adquirimos cuando, conscientes o no, teñimos nuestra concepción de lo real, especialmente de los sistemas biológicos inteligentes, con la explicación que damos del funcionamiento de un sistema informático? La hipótesis que defendemos es que con respecto a la cognición y dentro de un marco conceptual naturalista y emergentista, debemos (y podemos) reconocer el límite entre lo que es, en principio, explicable y lo que no lo es.

Palabras clave: Computacionalismo, conexionismo, materialismo emergentista.

Abstract

Which ontology underlies the main models of cognition? What did the cognitive paradigm bring about? Which philosophical commitment we assume when,

¹ Este artículo forma parte del Proyecto de Investigación Mayor N° 3730-07 de la Universidad de Tarapacá: “Estudio epistemológico de la relación entre lo observable y lo inobservable”. (2007-2009).

conscious or not, we tinge our conception of what is real, especially of intelligent biological systems, with the explanation we give of the functioning of an information system? The hypothesis we maintain is that, with regards to cognition and the naturalistic and emergentist conceptual frame, we must (and can) recognize the limits of what is explainable and what is not.

Key words: Computationalism, connectionism, *emergent* materialism.

INTRODUCCIÓN

La famosa novela gótica y romántica de Mary Shelley (1797-1851), publicada en 1818, ha llegado a ser un clásico y como tal permite una lectura compleja en la que podemos reconocer la culminación de la filosofía natural de la época moderna que, tomando saberes y prácticas provenientes de la química, la física, la biología y la medicina permite a su autora concebir un ser de ficción creado de partes informes unidas y estimuladas eléctricamente. Se trata de un monstruo que alcanza a tener conciencia de sí, de su soledad y de un trágico destino. Es una tragedia la que vive el monstruo pero, también, es el sino que parece acompañar a los momentos estelares de la ciencia y la tecnología.

Actualmente, estamos (o debiéramos estar) como el Doctor Frankenstein con sus primeros remordimientos, los monstruos que hemos creado son los artefactos tecnológicos. Estos no son letales como el personaje de la novela, no poseen intenciones intrínsecamente malas, sino que el monstruo está en nosotros mismos, en sus artífices, únicos responsables de que nuestros artefactos puedan irse por los caminos del personaje terrorífico y, a la vez, humano de M. Shelley.

Así como la novelista imaginó un monstruo que escapó al control de su creador, podemos preguntarnos si tenemos o no tenemos un control de la máquina, del computador, que encarna, sin duda maravillosamente, *Dizyplinas* científicas tan decisivas en la cultura de las últimas décadas, sean ellos la lógica, la teoría de la información, la física y otras. También podemos preguntarnos: ¿Cómo encaja este artefacto en una concepción global de lo real, esto es, qué lugar ocupa en la ontología? ¿Qué compromiso filosófico adquirimos cuando, conscientes o no, teñimos nuestra concepción de lo real, especialmente de los sistemas biológicos inteligentes, con la explicación que damos respecto al funcionamiento de un sistema informático?

Nuestra hipótesis es que debemos (y podemos) reconocer tanto el ámbito de lo mental que, en principio, puede ser explicable con nuestros conceptos actuales, como el ámbito que nuestro aparato conceptual no está en condiciones de recoger. Para los efectos de nuestra hipótesis, se expondrán y

se evaluarán parte de las filosofías que subyacen a los modelos cognitivos computacionalista y conexionista.² Esta tarea se hará teniendo como base el punto de vista emergentista de una filosofía materialista.

REVOLUCIÓN INFORMÁTICA Y COMPUTADOR ¿ACASO UN NUEVO MONSTRUO?

El computador es una máquina que actualmente goza de un destacado impacto social. La nueva tecnología de los computadores es de tal magnitud que ha sido incorporada en nuestro lenguaje cotidiano (“chip”, “bit”, “software”, etc.). Pero, el mecanismo concebido por Turing, el famoso teórico de la informática, es abstracto, universal, en el sentido que se puede computar, combinar, cualquier algoritmo que esté bien definido al menos en principio. Turing planteaba esta posibilidad en el nivel teórico, de modo que el computador —la máquina (digital) que nosotros conocemos— con todas las limitaciones que pueda tener (de capacidad, de memoria, etc.) viene a ser en parte la concreción de aquella máquina de Turing. Entonces, una máquina de Turing que realiza un algoritmo, constituye el fundamento teórico de la existencia del computador digital. En su fundamento lógico, todo computador digital incorpora en sí dispositivos conceptuales ideados por este genial matemático.

Turing planteó la posibilidad de alterar la distinción de sentido común entre los objetos de los que predicamos funciones cognitivas, esto es, los seres vivos (OFC) y los objetos de los que negamos tales funciones cognitivas, los objetos inanimados (O~FC), con la inclusión de las máquinas entre los OFC.³ Este desafío lo planteó Turing en 1950 en el famoso test que lleva su nombre y que reza así: si una máquina puede imitar el comportamiento inteligente de una persona, al punto que un observador externo no fuera capaz de precisar si la respuesta que ha recibido a la pregunta que él ha realizado la ha dado una persona o una máquina, entonces, dice Turing, eso debe tomarse como una evidencia de que la máquina es capaz de pensar (1984 en su edición en castellano).

Las nociones propias de las ciencias cognitivas se encuentran en la modernidad. Sin embargo, el momento cuando se consolida lo que llamamos

² El modelo cognitivo se conoce, también, como modelo cognitivo clásico o de representaciones simbólicas; el modelo conexionista, a su vez, se suele llamar también modelo cognitivo no-clásico, modelo de redes neurales o modelo de procesamiento distribuido en paralelo. La semántica de ambos términos (“computacionalismo” y “conexionismo”) se despliega más adelante en la sección Modelos Cognitivos.

³ Angel Rivière (1991) distingue entre objetos de los que predicamos funciones mentales y objetos de los que las negamos. Hemos optado por la expresión “funciones cognitivas” (y no “funciones mentales”) para, desde el comienzo, situar nuestro trabajo en el contexto de la filosofía de las ciencias cognitivas.

“paradigma cognitivo”, como consecuencia de la “revolución informática” (aproximadamente a mediados de los 70), coincide con la opinión de que el estudio de la mente y de la cognición había estado demasiado entregado a filósofos y psicólogos. Así, se empieza a desarrollar una ciencia de la mente y la cognición teniendo como objetivo estudiar los procesos subyacentes a la actividad mental con la precisión de las ciencias empíricas y formales, por ejemplo, a través de mecanismos explícitos respetando los formalismos lógico-matemáticos. Las figuras protagonistas de este gran cambio fueron Warren McCulloch (matemático y neurólogo); Walter Pitts (lógico); Karl Lashley, George Miller, Jerome Brunner (psicólogos); John von Neumann, Norbert Wiener, John McCarthy, Alan Turing y Marvin Minsky (matemáticos); Claude Shannon (ingeniero); Noam Chomsky (lingüista); Herbert Simon, Allen Newell, P.N. Johnson Laird (especialistas en computación); Gregory Bateson (antropólogo), y varios otros teóricos e investigadores.

El concepto de máquina es complejo y es un reflejo de las épocas. En la Antigüedad y en la Edad Media el prototipo de máquina es la palanca, el aparejo; la disciplina que la estudia es la estática y su característica es la transformación de la fuerza. Entre los siglos XVII y XVIII, su prototipo es la rueda/engranaje, la bomba o el reloj; su estudio es parte de la mecánica con hidráulica. Sus características son el determinismo, automatismo e independencia y se le compara con los animales (hipótesis del animal-automata de Descartes), con el hombre-máquina (La Mettrie) y con el cosmos-reloj (Laplace). A comienzos del siglo XIX, el prototipo de máquina es el regulador de fuerza centrífuga que se caracteriza por su retroalimentación y autorregulación. A mediados de ese siglo surge la máquina de vapor que es tratada por la termodinámica y cuya característica es la transformación de energía. Y ya en los albores del siglo XX tenemos el motor eléctrico y el telégrafo que son del campo de estudio de la electricidad y que permiten hablar del cerebro como una red eléctrica (¡obsérvese la notable actualidad del monstruo de Mary Shelley!). A mediados del siglo XX surge la noción de máquina como computador, cuya característica es que procesa información y es programable. Se investiga el computador en los ámbitos de la teoría de la información, cibernética y electrónica haciendo la comparación con el cerebro, en tanto, éste sería un computador, una máquina electrónica, una máquina con pensamientos. En nuestra época tenemos las células con sus características de adaptación y reproducción, tratadas al interior de la bioquímica donde el cerebro se concibe como una máquina bioquímica; asimismo, tenemos al cerebro caracterizado por su creatividad en el contexto de la neurociencia donde la máquina es el artefacto inteligente, la inteligencia artificial (Cfr. N. Ursúa, 1991).

Cabe precisar que cuando se habla de las similitudes entre el cerebro y las nuevas máquinas no se está diciendo, por ejemplo, que el cerebro sea una máquina de relojería (o, mucho peor, una máquina de vapor). Lo que tienen en común es la manera de funcionar, esto quiere decir que todas las máquinas funcionan algorítmicamente. Las máquinas son realizaciones de algoritmos.

En su origen, las ciencias cognitivas proponen un modelo de lo humano, un modelo tecnológico, informático o, si se quiere, maquinista. El hombre (varón o mujer) sería un complejo procesador de información y el cerebro o la mente inmaterial (da lo mismo, como veremos más adelante) sería el medio interno de este sistema cognitivo, en el que hay recepción de información que proviene del medio ambiente (*inputs*), procesamiento de la misma (que constituye propiamente el *software* o programa) y salida conductual (*outputs*). Ontológicamente, esta concepción fue neutral en sus inicios; sus epígonos no quisieron pronunciarse sobre la naturaleza del sistema que procesa la información: ¿Es acaso un cerebro con sus neuronas, complejas conexiones sinápticas, etc., o es una mente inmaterial con sus pensamientos? Ni dualismo psicofísico, ni monismo materialista son opciones que estén necesariamente conectadas a este modelo. Simplemente, si concebimos cualquier sistema cognitivo —sea orgánico, como los humanos o los monos antropoides; o inorgánicos como los computadores— en los términos de entrada, procesamiento y salida de información estamos hablando de sistemas propiamente inteligentes. De ahí que tendemos a creer, peligrosamente, que información y conocimiento es lo mismo.

El paradigma cognitivo se ha consolidado, institucionalizado, propagado. Hoy tenemos legiones de especialistas en biología, lingüística, psicología, informática, etc., que diseñan, investigan, enseñan, publican y piensan imbuidos de modelos cognitivos que se han ido desarrollando al alero de tal paradigma. Si siguiéramos, en parte, el esquema que Thomas Kuhn propuso para explicar el desarrollo de las ciencias en *The Structure of Scientific Revolutions* (1962) tendríamos que reconocer que este paradigma se encuentra en su etapa de ciencia normal, donde encontramos ámbitos en los que el paradigma funciona (ejemplos paradigmáticos), es decir, permite explicar, controlar y simular procesos tan complejos como los cognitivos-representacionales-intencionales que, en el mundo animal, están ligados con la percepción y la acción.

En contra de la opinión de autores como Husserl, que identifican la intencionalidad con la propiedad de un acto de representación consciente, en las ciencias cognitivas el concepto de intencionalidad es considerado independientemente de toda toma de conciencia de un contenido de pensamiento. Pero, justamente, la conciencia o el carácter subjetivo de la experiencia consciente (conciencia fenomenológica), presenta un formidable desafío a las

ciencias cognitivas actuales. En términos kuhnianos, nos encontramos con una anomalía, con un ámbito donde el paradigma no funciona. Se trata de una brecha explicativa, en el sentido de que no disponemos de una teoría satisfactoria acerca de nuestra naturaleza física o funcional que explique nuestra experiencia subjetiva. Sin embargo, no podríamos decir que tal traspás sea, por el momento, indicador de una crisis del paradigma, al menos en el sentido que los cognitivistas estén atrapados en un escepticismo sistemático y terminal.

En la actualidad estamos en un escenario muy distinto al que tuvo Turing y los demás forjadores del paradigma. La Internet y los recursos que ella ha permitido como el chat, blogs, facebook y otros, parecen prolongar *ad infinitum* las posibilidades de acceso a la información en red. Ello nos preocupa, o nos debiera preocupar, por sus consecuencias en la salud corporal y psíquica, en lo social, económico y moral. No se trata de ser un retrógrado al estilo de vociferar que “muera la inteligencia” (en este caso la inteligencia informatizada) como aquel falangista furibundo que denostaba al filósofo Unamuno (citado por Erich Fromm, 2008) sino de recordar que las tecnologías informáticas son justamente eso: “tecnologías” y, justamente, por ello tienen un impacto ambivalente.

Hoy en día somos testigos del aumento exponencial en el acceso a la información y de todas las consecuencias que ello implica para el desarrollo humano. Sin duda que las nuevas tecnologías de la información son una herramienta formidable en las nuevas metodologías de enseñanza; permiten cirugías enormemente precisas donde el ojo o el pulso humano podrían fallar; facilitan, también, que el solitario y deprimido se conecte con potenciales nuevos confidentes; ayudan a la transacción rápida y precisa de bienes, etc.

Pero, también, está el lado oscuro del acceso informático: la informatización de crecientes áreas de la industria genera cesantía; los jóvenes alumnos discriminan, a veces *a priori*, al viejo y sabio profesor renuente a incorporarse a la autopista de la información; surgen redes delictuales en red que obligan a los gobiernos a hacer enormes inversiones para combatirlos. En fin, hoy —al igual que en el pasado— las tecnologías nos dejan con el sabor de que ellas necesitan un control social consensuado democráticamente.

MODELOS COGNITIVOS

Pero hay algo que no podemos dejar en manos de otros, como es la formulación de los fundamentos de una filosofía, como diría Ortega y Gasset en *El tema de nuestro tiempo* (1923), “a la altura de los tiempos”. Esto, a nuestro juicio, lleva a la cuestión de cómo la revolución informática afecta a nuestra concepción de lo humano.

Traeremos a colación que el ser humano individual —al menos en las filosofías dominantes en la última centuria, más preciso hasta los 70— es una entidad concebida en el contexto de avances notables, sea en la psicología (conductismo), la psicofisiología, la física y la biología. El ser humano individual fue concebido como: (a) un ser dual, dotado de un espíritu inmaterial y un cuerpo material que o bien interactúan en tanto sustancias cualitativamente distintas (Bergson, Popper, Sherrington, Adrian, Eccles, Penfield) o bien, concebidos el espíritu y la materia como series de acontecimientos psíquicos y series de acontecimientos físicos, que discurren en paralelo o sincronizados (Freud); (b) como un ser dotado de propiedades psíquicas epifenoménicas o que supervienen a la materia cerebral (Broad, Ayer, Kim); (c) como un organismo con propiedades disposicionales actualizables en el comportamiento (Hempel, Ryle, Wisdom); (d) como un ser físico y totalmente material cuyo actividad y propiedades psíquicas se identifican con las propiedades fisico-químicas del cerebro (Carnap, Feigl, Armstrong, Smart, Place, Mountcastle, Pribam, Changeux). Veremos más adelante que a estas concepciones de lo humano hay que agregar otras *post* años 70 que se siguen directamente de la revolución informática, o que están condicionadas por ésta en tanto sus propuestas son discutidas sobre el trasfondo de esta revolución y el paradigma a que dio lugar.

Fue a fines de la década de los 70 cuando tomando prestada la imagen de la máquina procesadora de información, esto es del computador, los teóricos de la cognición se atrevieron a pensar que nuestra propia capacidad de razonar —aquello que tanto maravillaba a los griegos al punto de considerarlo el aspecto divino de nuestra condición— encarnaba una máquina de Turing.

Primero, se siguió un modelo estrictamente computacionalista (cognitivo clásico o de representación de símbolos), en el que la cognición es una combinación de representaciones (símbolos) regulado por reglas claramente definidas. Ese fue el momento del funcionalismo filosófico con autores como Putnam, luego Fodor y Lewis como sus representantes de mayor importancia, proponiendo que la cognición fuera abordada mediante definiciones estrictamente funcionales, es decir, que se hiciera caso omiso de la materia que subyace al sistema cognitivo (sea neuronas en un cerebro, pensamientos en una mente inmaterial o microchips en un computador) para centrarse en la función y, todo ello, en el marco de las relaciones causales del sistema con el ambiente que le provee de los *inputs* necesarios, de las relaciones causales al interior del sistema y de las relaciones entre el sistema y la conducta (*outputs*). Es el momento, también, cuando el impacto social de la tecnología informática —crecientemente al alcance de más gente en los países del primer

mundo— da origen a la industria del conocimiento y empieza a maravillarse por sus logros inmediatos masificando “la sociedad del conocimiento”.

Pero, ¿Cuál es la naturaleza de las representaciones en el computacionalismo? Es decir, si la revolución informática trajo aparejada la concepción que conocer es procesar información mediante la combinación de símbolos o representaciones, tenemos el derecho a preguntarnos por éstas, especialmente si recordamos la antigua tesis —reconocible en la lingüística y en la filosofía— según la cual nosotros reflejamos con nuestros símbolos el mundo (concepción especular del lenguaje).

Téngase presente que las representaciones son configuraciones de símbolos discretos que se manipulan según reglas secuenciales (un problema se resuelve después de otro). Los símbolos tienen 3 niveles:

- a) Un nivel físico, por ejemplo, un estado de carga eléctrica, un estado neurofisiológico.
- b) Un nivel sintáctico o formal en virtud del cual la máquina o cerebro manipula los símbolos con independencia de su sentido, este es el nivel más importante para los computacionalistas.
- c) Un nivel semántico en el que los símbolos representan el mundo exterior. En este nivel se plantea la cuestión de si una representación, o un símbolo, es adecuado o no, es decir, si es o no es una buena representación de un estado de cosas.

Pero surgen problemas con la computación. En términos físicos, ¿Qué significa “computar”? Si hablamos de cintas de almacenamiento electrónico (como aquellas concebidas por los primeros teóricos de la computación) sobre la que se hacen marcas, entonces no hay problema, pero sí lo hay con los símbolos. Y es así porque el símbolo, por definición es —como sugerimos anteriormente— una unidad discreta que no obedece a leyes intrínsecas como los objetos físicos. Es decir, la computación depende de algo externo: la programación.

A propósito, recordemos, sucintamente, la crítica que hizo Searle (1980) a una versión fuerte de la tesis que pensar es computar (Inteligencia artificial fuerte o informativismo). Tal es el argumento de la sala china: La IA fuerte sostiene que la mente es un programa de computador. Un computador es un mecanismo en el que se manipulan símbolos formales que codifican información mediante un código binario de 0 y 1, donde la información codificada se traduce en impulsos eléctricos y donde se procesa la información de acuerdo con reglas formales sintácticas que son propias del programa de computación. En cambio, la inteligencia artificial débil es simplemente la idea de que el computador es una herramienta útil para hacer simulaciones de la mente y, en general, de cualquier cosa que se pueda describir con precisión. En la versión débil de la inteligencia artificial, se

considera que la analogía que se hace entre la mente y el cerebro por un lado y el *software* y el *hardware* por el otro, es solamente una analogía útil.

El argumento o parábola de la sala china es un experimento mental en el que Searle se imagina a sí mismo desempeñando el papel de un computador que está preparado para responder preguntas en un idioma desconocido para él, como es el idioma chino. La situación es la siguiente:

En una habitación hay un montón de cajas con símbolos en chino (base de datos). Searle está dentro de la habitación y recibe preguntas (estímulos o *inputs*) a través de una ranura. Las preguntas se las envían chinos auténticos. Searle no se ve con los chinos parlantes.

Dentro de la habitación hay un libro de reglas en inglés (el idioma de Searle) que, al consultarlo, le permite saber a Searle (*software*) que si hay determinados símbolos en chino se debe contestar con otros símbolos chinos. Es decir, Searle puede reconocer los símbolos por su forma y ejecutar determinadas operaciones con ellos de acuerdo con las reglas del programa. Puede, entonces, realizar los pasos del programa.

Finalmente, Searle puede responder a través de otra ranura a los chino-parlantes que están fuera de la habitación (respuestas, salida de información, *outputs*).

Obsérvese que algo curioso está ocurriendo en esta sala china. Si tenemos a la vista el test de Turing (test considerado crucial por la IA fuerte para adjudicar pensamientos a una máquina), resulta que el individuo dentro de la habitación sí está pasando el test, pues, los que están afuera —los chinos auténticos— no serán capaces de discriminar entre las respuestas que les habría dado un chino-parlante auténtico y las respuestas que reciben del individuo (Searle) que está en la habitación. Sin embargo, este individuo no sabe chino. En consecuencia, escribe Searle rememorando la parábola en su libro *El misterio de la conciencia*, “si yo no entiendo chino cuando ejecuto un programa de computador para entender el chino, entonces tampoco entiendo chino ningún otro computador digital que opere exclusivamente sobre estas bases, pues ningún computador tiene nada que yo no tenga” (2000:24).

Lo interesante aquí es que Searle quiere establecer que una máquina (no biológica, es decir artificial) no piensa. El test de Turing no sirve, no se puede decir que el computador piense. Reducido a su esqueleto el argumento queda así:

1° Los programas de computador son enteramente sintácticos.

2° Las mentes tienen una semántica.

3° La sintaxis no es lo mismo que la semántica, o no es suficiente para que se dé la semántica. *Ergo* los programas no son mentes.⁴

⁴ Una presentación actualizada y breve del argumento de la sala china se encuentra en M. Díaz

Hace casi dos décadas, el científico J. Z. Young, consciente de las dificultades del computacionalismo y de la necesidad de seguir modelos de la cognición que fueran más cercanos a las operaciones cerebrales, a la biología de un sistema cognitivo, escribía

Es interesante que los investigadores en inteligencia artificial, que están interesados en comprender los procesos cerebrales de máxima complejidad, sientan la necesidad de considerar estas soluciones analógicas. Son conscientes, en gran medida, de la inoperancia de los intentos de establecer comparaciones entre las actividades cerebrales y las de un computador digital (1992:183).

A poco andar, efectivamente, se impuso el modelo de procesamiento en paralelo, similar a los procesos del cerebro.

En filosofía, los pensadores que adhirieron y proyectaron esta manera de concebir la cognición y la mente humana fueron, principalmente, los eliminativistas Paul y Patricia Churchland. Estos criticaron, ácidamente, al computacionalismo o modelo de procesamientos de símbolos, haciendo notar —entre otras cosas— que dicho modelo reivindica la psicología popular (*folk psychology*), aquella que, al decir de Paul Churchland, existe hace 2.000 años y que en lo esencial se ha mantenido inmodificada. La postura radical del matrimonio Churchland era clara: dicho marco conceptual, el de la psicología popular, debe ser eliminado de nuestra ontología porque no ha servido para solucionar problemas tan candentes como la naturaleza de las imágenes mentales, de la memoria, el fundamento de la diferencia de inteligencia entre los individuos, etc. (Churchland, 1981).

Si damos crédito a Paul Churchland, tendríamos que imaginarnos cómo sería nuestra vida sin la psicología popular, aquella psicología originada en la tendencia animista a explicar la naturaleza en términos de creencias y deseos. Esos conceptos, las creencias y los deseos, son los cimientos principales de los que se sirve la “*mêtis*” (inteligencia narrativa), esa inteligencia heraclíteica y dinámica, “ligada a los contextos interpersonales y prácticos, relacionadas con la sagacidad y las capacidades que se ponen en juego en las complejas y veloces interacciones entre personas” (Rivière y Núñez, 2001:8). Esta “*mêtis*” sería diferente de la inteligencia parmenídea o *nôus*, que consiste en organizar un mundo abstracto, estático e impersonal de relaciones invariables. Pero renunciar a la psicología popular es renunciar, ni más ni menos, que a la comprensión de nuestra vida de relación, a nuestra existencia-con-los-otros, porque sin psicología popular estamos cercenados de la mirada mental o de la

“La teoría de la mente de John Searle. Apuntes de clases”, *Alcances*, Vol. II N° 2 (2007):113-127.

actitud intencional,⁵ esa mirada en la que leemos la mente de los otros y la nuestra en las acciones, expresiones, palabras y hechos. Sin ese mentalismo hábil no podremos comprender racionalmente la conducta como producto de los deseos y creencias; no podremos ni comprender ni predecir la conducta de otros y tampoco podremos realizar actividades comunicativas complejas; menos engañar y detectar el engaño. En otras palabras, al renunciar a la “*mêtis*” estamos descuidando las capacidades cognitivas que están en la base de las competencias sociales.

El conexionismo es un paso adelante en el paradigma cognitivo y, a su vez, es una reactualización del programa cibernético que quedó detenido a fines de los 60 “cuando Marvin Minsky y Seymour Papera demostraron que el perceptrón (primer sistema conexionista de reconocimiento de formas) sólo puede resolver una clase restringida de problemas simples” (Cfr. P. Bourguine, 2003:94).

El conexionismo, entonces, ya no es un modelo de procesamiento de símbolos discretos sino que es el procesamiento de una red de elementos o redes formales, que se suponen son la idealización de las conexiones neuronales humanas. Este modelo está motivado en un intento de imitar el modo de operar del sistema nervioso, puesto que diferentes partes del cerebro parecen llevar a cabo funciones de cálculo separadas e independientes, por ejemplo, el proceso de información visual en el córtex visual.

En el sistema conexionista, la capacidad de reconocimiento de formas o el aprendizaje se caracteriza porque los cambios de las unidades que, a su vez, producen cambios en sus pesos sinápticos, lleva a una continuidad en el sistema hasta encontrar una solución adecuada al problema. Este tipo de sistemas es muy diferente a los sistemas computacionales clásicos de la representación. En un sistema conexionista, todo el conocimiento que tiene la red se refleja en sus conexiones. Por lo tanto, un sistema conexionista aprende en cuanto ha variado el peso sináptico en sus conexiones.

Como hemos visto, en los sistemas clásicos computacionales la representación se expresa a través de una sintaxis (esto es, de la manipulación de símbolos discretos según su sintaxis), vale decir, se expresa a través de la lógica formal; en cambio, en los sistemas conexionistas esto no se da. En los sistemas cognitivistas clásicos los cambios son secuenciales, simbólicos; en cambio, en los sistemas conexionistas los cambios son concurrentes, simultáneos y numéricos. Son cambios (o procesos) continuos (y no discretos).

⁵ Daniel Dennett (1987) sostiene que la intencionalidad de los estados mentales, más que una propiedad intrínseca a los sistemas, es una estrategia o postura intencional desde la que contemplamos a la gente en la vida diaria, esto es, atribuyéndoles creencias y deseos. Dennett piensa que esta postura puede ampliarse a sistemas no naturales.

En los sistemas conexionistas no se da paralelismo entre sintaxis y semántica, como ocurre en los sistemas clásicos cognitivistas (computacionalistas o de representaciones simbólicas). En estos sistemas clásicos la sintaxis que rige el programa refleja, supuestamente, la estructura del mundo real, el contenido, la semántica. Eso es lo que se espera, pero ya sabemos que esto presentó graves dificultades a la teoría cognitiva de Jerry Fodor, uno de los máximos filósofos defensores del modelo cognitivo clásico.

En los sistemas conexionistas los procesos son distribuidos, es decir, la representación está distribuida y no localizada en un punto concreto. Supongamos que el sistema conexionista ya ha sido entrenado (reconoce). Si decimos que es distribuido lo que queremos afirmar es que ningún rasgo de, por ejemplo, letras escritas a mano, está codificado en una unidad, sino que cada rasgo se codifica en varias unidades y cada unidad codifica varios rasgos, o sea, hay distribución (en un sistema conexionista puede extraerse una unidad y el sistema sigue funcionando tal como ocurre con el cerebro a pesar de la muerte continua de neuronas).

En un sistema conexionista la codificación es a escala microcósmica y sólo la composición de esos micro-rasgos da la representación como tal del objeto, por ejemplo, la codificación de una letra se da a nivel microcósmico, es como si en lugar de representar un vehículo se representaran las moléculas que lo constituyen. La representación en un sistema conexionista se da en un contexto, es decir, que cada micro-rasgo viene definido por la colección de los otros micro-rasgos. Ejemplo: el lomo del libro al representarse va a depender del ángulo en que esté el libro. Al variar la representación de todos los otros rasgos del libro se seguirá que la representación del lomo del libro varíe. Incluso la forma de entrenamiento (reconocimiento) que ha tenido el sistema hará que varíe la forma en que se va a codificar el objeto.

Como en el sistema conexionista hay variación continua en los pesos de las conexiones (peso sináptico), podemos tener una conexión precisa —una sintaxis limpia— pero no podemos reconstruir la semántica a partir de esta sintaxis limpia. Aunque tenemos un conocimiento matemático preciso de los cambios, no tenemos, un conocimiento preciso de la reconstrucción del objeto.

En el curso del aprendizaje, las redes conexionistas pueden dar la posibilidad de elegir las conexiones que se van a realizar, esto es lo que se llama “capacidad semántica” o “autonomía semántica”. Lo que Searle criticó con su famoso argumento de la sala china a los sistemas clásicos cognitivistas puede evitarse aquí porque los sistemas conexionistas pueden cambiar su sistema de representación a través del aprendizaje.

En la medida en que el modelo cognitivo conexionista se afianzó, nos volvimos a fijar en el cerebro vivo y con ello se desarrolló la neurociencia

cognitiva. Esto permite referirnos al órgano del conocimiento, al biosistema que es el cerebro. De ahí, al naturalismo —o materialismo, con el cual simpatizamos— no hay más que un paso: el que va de la biología (de la ciencia natural) a la filosofía, asumiendo la tesis, por cierto filosófica, que ciencia y filosofía forman un continuo racional.

En la actualidad uno de los mayores desafíos que tienen pensadores y científicos es dar una explicación de la cognición y de la mente que incluya, de manera clara y sin caer en reduccionismos metodológicos ni epistemológicos, los avances en las ciencias cognitivas en el marco de una concepción unitaria del mundo. Tal esfuerzo demanda esclarecer la compleja red de conceptos involucrados en la visión materialista y naturalista. Los tiempos están para un materialismo sofisticado y no-fisicalista, según bosquejaremos a continuación.

FISICALISMO Y MATERIALISMO EMERGENTISTA

Nuestro punto de vista es, claramente, naturalista y materialista. Pero exige algunas precisiones conceptuales, porque hoy ser (ontológicamente) materialista no es una etiqueta deseable para la mayoría de los filósofos. Reconocerse como materialista quiere decir para muchos: “ser moralmente materialista” (o “centrado en la búsqueda de placer”) o “marxista no renovado”. Los científicos, por su parte, tampoco simpatizan con la tarea de poner sobre el tapete sus presuposiciones filosóficas. El fantasma del reduccionismo fisicalista, o vulgar, baila sobre sus cabezas y los lleva a preferir, al igual que a los filósofos —o a ellos mismos cuando filosofan— una etiqueta menos comprometedora: el naturalismo.

De modo que ser naturalista resulta, en apariencias, menos comprometedor que declararse materialista. Sólo hay que vérselas con entidades naturales, sin excluir, necesariamente, las entidades que habitualmente consideramos inmateriales pero que no son sobrenaturales, como es el caso de las ideas o representaciones, el espíritu, la razón, etc. Así como lo opuesto al materialismo es el inmaterialismo (o idealismo), lo opuesto al naturalismo es el sobre-naturalismo (la creencia en la existencia e influencia de fantasmas, deidades, seres mitológicos, etc.).

Si bien el naturalismo y el materialismo coinciden en los objetos que estudian la física, la química y la biología, el primero (el naturalismo) resulta ser muy restrictivo respecto a lo social y tecnológico. Los seres humanos somos animales, seres naturales, y somos capaces de producir cosas artificiales (que, por definición, no son naturales) como los utensilios que condicionan nuestra vida. Entonces, es el materialismo (y no el naturalismo) el que tiene que explicar los artefactos (del lat. *arte factus*, hecho con arte) de

tipo conceptual como la lógica y las matemáticas que se resisten a ser explicados en términos estrictamente naturalistas (asunto que, por cierto, aquí no será tratado). Pero esto obliga a concebir una forma de materialismo de tipo emergentista, es decir, que no sea reduccionista o (metodológicamente) fisicalista, en el sentido que toda explicación provenga de la ciencia física.

Una combinación, como la propuesta por Bunge (2002), Ferrater Mora (1983) y otros autores entre un pluralismo de propiedades y niveles de realidad (físico, químico, biológico, social, tecnológico y semiótico) con un monismo de substancia (la materia como única substancia), parece ser la vía para evitar una forma de materialismo fisicalista, o nivelador, donde todo se resuelve en el nivel físico de la realidad.

Pero distinguir entre el materialismo fisicalista y el no-fisicalista no es fácil, pues tendemos a asociar materialismo y fisicalismo por el peso que las explicaciones fisicalistas han tenido desde los atomistas griegos. Éstos proponían que la materia son los átomos, entidades infinitamente pequeñas, no observables, que al unirse millones de ellas dan lugar al espectáculo que tenemos frente a nuestra vista y a los demás sentidos. La intuición atomista nos hacía suponer que en la realidad, además de los átomos, está el vacío, que es el medio en el cual se desplazan los átomos, entidades infinitamente pequeñas con propiedades geométricas, figura, volumen, número, etc. En otras palabras, lo cualitativo, lo observable, se explicaba por lo cuantitativo, que es inobservable. Es innegable la elegancia de esta explicación y el influjo que ha ejercido sobre muchas mentes brillantes.

En la modernidad, y todavía en nuestro naciente siglo XXI, el materialismo fisicalista heredado de La Mettrie, Hobbes y de los científicos y filósofos del siglo XX como Feigl, Carnap, Armstrong, Smart o, más recientemente, Penrose (una suerte de neofisicalista) hacen difícil la explicación de la novedad o de la emergencia de propiedades que no habríamos logrado prever, incluso, si hubiésemos dispuesto de todas y las mejores teorías con fuerte componente predictivo. Sin embargo, seamos autocríticos, el emergentismo que defendimos es, por el momento, no una alternativa explicativa sino solamente descriptiva, que describe los estadios emergentes pero no aborda la causalidad que lleva de un estadio a otro superior. En pocas palabras: estamos solamente recogiendo un hecho bruto de la naturaleza.

*Universidad de Santiago de Chile**
Departamento de Filosofía, Facultad de Humanidades
Alameda Bernardo O'Higgins 3363. Estación Central, Santiago (Chile)
mdiazsoto@vtr.net

*Universidad de Tarapacá***
Departamento de Psicología y Filosofía
Avda. 18 de Septiembre 2222, Arica (Chile)
aespenoza@uta.cl

BIBLIOGRAFÍA

- BOURGINE, P. ; HOUDÉ, O. *et al. Diccionario de Ciencias Cognitivas.* Buenos Aires: Amorrortu, 2003.
- BUNGE, Mario. *Crisis y reconstrucción de la filosofía.* Barcelona: Gedisa, 2002.
- CHURCHLAND, Paul. *Matter and Consciousness.* Massachusetts: MIT Press, 1981.
- DENNETT, Daniel. *The Intentional Stance.* Cambridge: MA:MIT Press, 1987.
- DÍAZ, Marcelo. "La teoría de la mente de John Searle. Apuntes de clases", en *Alcances*, Vol. II, N° 2 (2007):113-127.
- FERRATER MORA, José. *De la materia a la razón.* Madrid: Alianza, 1983.
- FROMM, Erich. *El corazón del hombre.* México: F. C. E., 2008.
- KUHN, Thomas. *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press, 1962.
- ORTEGA Y GASSET, José. *El tema de nuestro tiempo.* Madrid: Revista de Occidente, 1923.
- RIVIÈRE, Angel. *Objetos con mente.* Madrid: Alianza, 1991.
- RIVIÈRE, A.; M. NUÑEZ. *La mirada mental.* Buenos Aires: Aique, 2001.
- SEARLE, John. "Minds, brains and programs", en *Behavioral and Brain Sciences* 3. (1980):417-424.
- SEARLE, John. *El misterio de la conciencia.* Barcelona: Paidós, 2001.
- TURING, Alan. "Maquinaria computadora e inteligencia", en A. Ross. *Controversia sobre mentes y máquinas.* Barcelona: Tusquets, 1984.
- URSÚA, Nicanor. *Cerebro y conocimiento. Un enfoque evolucionista.* Barcelona: Anthropos, 1991.
- YOUNG, J. Z. *Filosofía y cerebro.* Barcelona: Sirmio, 1992.