

Entrevistas y artículos de interés

“Computador, cerebro y cognición: cambios revolucionarios en nuestra concepción de lo humano”

Marcelo Díaz, filósofo y académico de la Universidad de Santiago de Chile.

- ¿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, maldíceme 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: “Frankenstein o el moderno Prometeo” – Volumen II. Capítulo 2. Madrid, Valdemar, Gótica, 1994.

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, le 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. Hoy también como el científico Frankenstein creamos artefactos, nos beneficiamos de ellos y observamos su inmersión en un mundo (lamentablemente) todavía concebido antropocéntricamente a pesar de las revoluciones copernicana y darwiniana, cambios conceptuales que nos han quitado una tajada importante de esa torta de la que todos comemos, la torta de nuestra infinita falta de modestia, de creer aún que el mundo, la naturaleza, está ahí para saciarnos de ella. Actualmente estamos – o debiéramos estar - como el Doctor Frankenstein con sus primeros remordimientos, los monstruos que hemos creado son los artefactos tecnológicos. No son letales como el personaje de la novela, no poseen intrínsecas malas intenciones 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 Mary Shelley.

Frankenstein en la revolución informática y el computador

Hablar del computador es hablar de un mecanismo, de una máquina. Actualmente el impacto social de 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 de 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 la máquina (digital) que nosotros conocemos, el computador, con todas las limitaciones que pueda tener (de capacidad, de memoria) 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.

En 1950 escribía en "Computing Machinery and Intelligency" (*Mind*, 59, pp. 433-460): "Creo que, a fines de siglo, el uso de las palabras y la opinión general culta habrán cambiado tanto que uno podrá hablar del pensamiento de las máquinas sin esperar que se le contradiga". La investigación en inteligencia artificial, paso a paso, parece corroborar experimentalmente en los centros de computación de diversas universidades y centros tecnológicos del mundo las arriesgadas tesis de Turing. 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 (A. Rivière: *Objetos con mente*, 1991) Este desafío lo planteó 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 de que un observador externo no fuera capaz de determinar o precisar si la respuesta que ha recibido a la pregunta que él ha hecho 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.

Las nociones que son propias de las ciencias cognitivas se encuentran en la modernidad o mucho antes en la antigüedad, pero no es menos cierto que el momento en que se consolida lo que llamamos "paradigma cognitivo", como consecuencia de la "revolución informática", coincide con la opinión que el estudio de la mente y la cognición había estado demasiado entregado a filósofos y psicólogos (especialmente conductistas) y que era el momento para desarrollar una ciencia de la mente y la cognición que pudiera expresar los procesos subyacentes a la actividad mental con la precisión de las ciencias empíricas y formales, por ejemplo, a través de mecanismos explícitos y formalismos lógico-matemáticos. Las figuras protagonistas de este gran cambio fueron Warren McCulloch (matemático y neurólogo); Walter Pitts (lógico); Kart 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 reflejo de las épocas, así en la Antigüedad y 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 fuerza. En 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-autómata 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 (la reminiscencia del monstruo de Mary Shelley es notable). A mediados del siglo XX surge la noción de máquina como computador cuya característica es que procesa información y es programable y se investiga en los ámbitos de la teoría de la información, cibernética y electrónica haciendo la comparación con el cerebro en tanto este sería un computador, una máquina electrónica, una máquina con pensamientos. Y, finalmente, en nuestra época tenemos las células con sus características de adaptación y reproducción tratadas al interior de la bioquímica y 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 son los artefactos inteligentes, la inteligencia artificial (Ursúa 1993:265).

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 peor aun una máquina de vapor, sino que tienen en común la manera de

funcionar y esto quiere decir que todas las máquinas funcionan algorítmicamente. Las máquinas son realizaciones de algoritmos. Las ciencias cognitivas en su origen 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) 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). Filosóficamente esta concepción es ontológicamente neutral en sus inicios, sus epígonos no quisieron pronunciarse sobre la naturaleza del sistema que procesa la información: ¿es un cerebro con sus neuronas y conexiones sinápticas 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 tendamos a creer peligrosamente que información y conocimiento es lo mismo.

El paradigma cognitivo, se ha consolidado, se ha institucionalizado, se ha propagado y, hoy tenemos legiones de especialistas en biología, lingüística, psicología, 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 Kuhn (*La estructura de las revoluciones científicas* 1992) propuso para explicar el desarrollo de las ciencias tendríamos que reconocer que este paradigma se encuentra en su etapa de ciencia normal, donde encontramos ámbitos donde el paradigma funciona, es decir permite explicar, controlar y crear; pero también tenemos ámbitos donde el paradigma no funciona o presenta evidentes fallas - las llamadas anomalías como, por ejemplo, el problema que plantea a los filósofos y neurobiólogos la explicación de la naturaleza de la subjetividad de la conciencia en tanto se resiste a ser incorporado en modelos cognitivos computacionalistas o conexionistas -. 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.

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, los blogs, facebook y otras tecnologías informáticas que parecen prolongar *ad infinitum* las posibilidades de acceso a información en red, nos preocupan o nos debieran preocupar por sus consecuencias en la salud corporal y psíquica, en lo social, económico y moral (recuérdese el artículo de Roberto Ampuero que menciona a Bezzos, una especie de superbiblioteca virtual en la forma de un aparato personal de no más tamaño que un notebook). No se trata de ser un retrogrado al estilo de vociferar que “muera la inteligencia” (en este caso la inteligencia informatizada) como aquél falangista furibundo que denostaba al filósofo Unamuno, sino de recordar que las tecnologías informáticas son justamente eso: “tecnologías” y justamente por ello tienen un impacto ambivalente.

Nadie niega que han traído el aumento exponencial en el acceso a la información (relevante o no), que sean una herramienta formidable en las nuevas metodologías de enseñanza, que permitan cirugías enormemente precisas donde el ojo o el pulso humano podría fallar, que permitan al solitario deprimido conectarse con potenciales nuevos confidentes haciendo algunos movimientos con sus dedos sobre el teclado del computador, que faciliten la transacción rápida y precisa de bienes, etcétera. Pero también está el lado oscuro del acceso informático. La informatización de crecientes áreas de la industria genera cesantía o discriminación con quienes se han quedado abajo del tren de la informatización, sea por avanzada edad o por ser simplemente analfabetos informáticos; 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 inversiones enormes para combatirlos, ocupando dineros que inicialmente pudieron ser orientados a la solución de carencias sociales inmediatas como la cobertura gratuita a la salud, etcétera. En fin, hoy al igual que en el pasado las tecnologías nos dejan con el sabor que necesitan un control social consensuado democráticamente. ¿Cómo? Eso se lo dejamos a los responsables, a las entidades gubernamentales, legislativas, a los medios de comunicación.

Modelos cognitivos

Pero hay algo que no podemos dejar en manos de otros, la formulación de los fundamentos de una filosofía, como diría Ortega y Gasset, “acorde con los tiempos”. Esto, a nuestro juicio, lleva a la cuestión de como la revolución informática afecta a nuestra concepción de lo humano (cabe advertir que no estamos haciendo mención de cómo el humanismo, secular o religioso, se pronuncia sobre las consecuencias nocivas o beneficiosas que tienen las innovaciones tecnológicas en nosotros). 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. 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, 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 (el joven Freud); (b) como un ser dotado de propiedades psíquicas epifenoménicas o que supervienen a la materia cerebral (Broad, Ayer, Kim); (c) como un ser física 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) o con propiedades disposicionales actualizables en el comportamiento (Hempel, Ryle, Wisdom).

Como veremos más adelante a estas concepciones de lo humano hay que agregar otras que se siguen directamente de la revolución informática, o 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, del computador, nos atrevimos 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 seguimos 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 predefinidas. Fue el momento del funcionalismo con los filósofos Putnam y luego Fodor y Lewis como sus más importante defensores proponiendo que la cognición fuera abordada mediante definiciones justamente 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 las 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 del sistema respecto de la conducta (outputs). Es el momento también en que el impacto social de la tecnología informática – ahora crecientemente al alcance de más gente en los países industrializados – da origen a la industria del conocimiento y empieza a maravillarse por sus logros inmediatos.

Pero, ¿cuál es la naturaleza de las representaciones en el computacionalismo?, es decir si la revolución informática trajo aparejado la concepción de que conocer es procesar información mediante la combinación de símbolos o representaciones, tenemos el derecho a preguntarnos por éstas, especialmente si recordamos que la tesis según la cual nosotros reflejamos con nuestros símbolos el mundo es muy antigua y de larga tradición en la filosofía y en la lingüística. Hay que tener 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) Nivel físico: un estado de carga eléctrica, un estado neurofisiológico; (b) Nivel sintáctico: en virtud de su estructura lógica o formal la máquina o cerebro manipula los símbolos con independencia de su sentido, este es el nivel más importante para los computacionalistas; (c) Nivel semántico: 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 la combinación de símbolos a nivel sintáctico constituye o no una buena representación. Será adecuada la representación cuando sirve para solucionar el problema planteado. Por ejemplo, si me lanzan una pelota al rostro, el cerebro realiza combinaciones de acuerdo a reglas de manera que finalmente esquivo la pelota.

Pero surgen problemas con la computación. ¿En términos físicos qué significa “computar”? Si hablamos de cintas de almacenamiento electrónico sobre la que se hacen marcas no hay problema, pero sí lo hay con los símbolos. Y es así porque el símbolo, por definición es, como mencionamos 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 la crítica que hizo Searle (1980) a una versión fuerte de la tesis que pensar es computar (= Inteligencia artificial fuerte o informativismo), 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 metáfora útil. El argumento o parábola es un experimento mental en el que Searle se imagina así mismo desempeñando el papel de un computador que está preparado para responder preguntas en un idioma desconocido para él, el idioma chino.

- En la habitación hay un montón de cajas con símbolos en chino (base de datos).
- Recibe Searle preguntas, estímulos o inputs a través de un ranura.
- Las preguntas se las envían chinos auténticos. (No se ven entre los chinos parlantes y Searle)
- Existe dentro de la habitación un libro de reglas que permite saber –al consultarlo- que si hay tales símbolos en chino se debe contestar con estos otros símbolos chinos. Es decir, al consultar el libro de reglas (software) puede reconocer los símbolos por su forma y ejecutar determinadas operaciones con los símbolos de acuerdo con las reglas del programa, puede, entonces, realizar los pasos del programa.
- Finalmente, puede devolver manojos de símbolos en chino a los chino parlantes que están fuera de la habitación (respuestas, salida de información o outputs)

Obsérvese que algo curioso está ocurriendo: 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 fuera de la habitación, 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 que está en la habitación. Sin embargo, este individuo no sabe chino. “En consecuencia –escribe Searle 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 entiende 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 él quiere establecer que una máquina (no biológica) no piensa. El test de Turing no sirve, no se puede decir que el computador piense. Reducido a su esqueleto el argumento queda así:

- 1º Los programas de computador son enteramente sintácticos.
- 2º las mentes tienen una semántica.
- 3º 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.

Hace dos décadas un 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” (Young 1992: 183). A poco andar efectivamente se impuso el modelo de procesamiento paralelo similar a los procesos del cerebro.

En filosofía los filósofos 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, aquella que al decir de Paul Churchland, existe hace 2000 años y que en lo esencial se ha mantenido inmodificada. La postura radical de los 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, etc. (Churchland 1981).

Si le damos crédito a Churchland tendríamos que imaginarnos como 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 según Brunner), 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 Nuñez 2008:8). 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 actitud intencional (Dennett), 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 como tampoco engañar y detectar el engaño.

El conexionismo, entonces, ya no es un modelo de procesamiento de símbolos discretos sino 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. El sistema conexionista funciona de manera parecida al sistema autoorganizativo que menciona Varela. El sistema conexionista es capaz, por ejemplo, de reconocer la voz y transformarlo en escritura a máquina. Es decir, hay aprendizaje supervisado en el sentido que al haber error se le pide que modifique su peso sináptico y establezca nuevas relaciones hasta encontrar la solución. La capacidad de reconocimiento de formas o aprendizaje se caracteriza porque los cambios de las unidades, que a su vez producen cambios en sus pesos sinápticos, llevan a una continuidad en el sistema hasta encontrar una solución adecuada al problema. Este tipo de sistemas son muy diferentes de 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 (es decir, de la manipulación de símbolos discretos según su sintaxis). Se expresa a través de la lógica formal; en cambio, en los sistemas conexionistas esto no se da. En los sistemas clásicos cognitivistas 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). En los sistemas conexionistas no se da paralelismo entre sintaxis y semántica como ocurre en los sistemas

clásicos cognitivistas. Varela escribe al respecto: “Esta separación entre forma y sentido fue el golpe maestro que dio origen al enfoque cognitivista, y en verdad a la lógica moderna. Pero esta maniobra también implica una flaqueza cuando se encaran los fenómenos cognitivos más profundos: ¿cómo adquieren los símbolos su sentido?” (Varela... :78. En estos sistemas clásicos cognitivistas la sintaxis que rige el programa refleja la estructura del mundo real (el contenido, la semántica). Bueno, a menos eso es lo que se espera, pero ya sabemos que esto presentó graves dificultades a la teoría cognitiva de uno de los máximos filósofos defensores del modelo cognitivo clásico: Jerry Fodor.

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 decir es que ningún rasgo de, por ejemplo, letras escritas a mano, está codificado en una unidad. 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 solo 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 coche se representara 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 a partir de esta sintaxis limpia la semántica. Teniendo como tenemos un conocimiento matemático preciso de los cambios, no podemos, sin embargo, tener un conocimiento preciso de la reconstrucción del objeto. Las redes conexionistas pueden en el curso del aprendizaje 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 que el modelo cognitivo conexionista se afianzó nos volvimos a fijar en el cerebro vivo y con ello se desarrolló la neurociencia cognitiva. Esto nos 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 que ciencia y filosofía forman un continuo racional (los filosofía está en los fundamentos de la ciencia y también en las consecuencias de la ciencia aplicada).

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 comprometedor: el naturalismo. De modo que ser naturalista resulta, en apariencia, menos comprometedor que declararse materialista, solo hay que vérselas con entidades naturales sin excluir necesariamente las entidades que habitualmente consideramos inmateriales pero que no son sobre-naturales 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 sobrenaturalismo (la creencia en la existencia e influencia de fantasmas, deidades, etc.).

Si bien el naturalismo y el materialismo coinciden en los objetos que estudian la física, la química y la biología, el primero resulta ser muy restrictivo respecto a lo social y tecnológico. Los seres humanos somos animales y somos capaces de producir cosas artificiales (que por definición no son naturales) como los utensilios que condicionan nuestra vida. Entonces “materialismo” y “naturalismo” coinciden parcialmente. 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 naturalistas. Pero claro esto obliga a concebir una forma de materialismo de tipo emergentista, es decir, que no sea reduccionista o (metodológicamente) fiscalista en el sentido que toda explicación provenga de la ciencia física. Una combinación, como la propuesta por Mario Bunge, Ferrater Mora, Vollmer, etc., entre un pluralismo de propiedades y niveles de realidad (físico, químico, biológico, social, tecnológico y semiótico) con un monismo de sustancia (la materia como única sustancia) parece ser la vía para evitar una forma de materialismo fiscalista o nivelador donde todo se resuelve en el nivel físico de la realidad.

Pero distinguir entre el materialismo fiscalista y el no fiscalista no es fácil, tendemos a asociar materialismo y fiscalismo por el peso que las explicaciones fiscalistas 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. También creían 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, de figura, volumen, número, etc. En otras palabras, lo cualitativo, lo observable, se explica por lo cualitativo, que es inobservable. Es innegable la elegancia de esta explicación y el influjo que ha ejercido sobre muchas mentes brillantes. En la modernidad y aun en nuestro todavía naciente siglo XXI el materialismo fiscalista, heredado de La Mettrie, Hobbes y de los científicos y filósofos del siglo XX como Feigl, Carnap, Armstrong, Smart o, más recientemente, Penrose, 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 justos, el emergentismo que defendimos nosotros 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.

De hecho un filósofo fuertemente fiscalista como Armstrong escribe a fines de la década de los 60 que “resulta crecientemente verosímil que todo suceso químico y biológico sea explicable en principio como una aplicación particular de las leyes de la física que rigen los fenómenos no químicos y no biológicos” Esto equivale a decir que “el mundo entero explicado por la ciencia no contiene sino cosas físicas que operan conforme a las leyes de la físicas” (1968: 47-49). En consecuencia, frente a la pregunta antropológica esencial sobre la naturaleza de la especie humana, Armstrong propondrá de una manera profética que “si el progreso científico corrobora este punto de vista, parece que el hombre no es sino un objeto material y no tiene sino propiedades físicas” (1968:..). Obsérvese que este reduccionismo fiscalista no deja espacio para introducir la cuña del emergentismo, esto es, para la concepción de que emergen propiedades que aportan novedad cualitativa, las que resultan de los elementos debidamente estructurados que pertenecen a niveles ontológicos inmediatamente precedentes. El reduccionismo no permite explicar, por ejemplo, el fenómeno de la vida o el de la mente como propiedades que son irreductibles, respectivamente, a sus constituyentes físicos o neuronales.

En efecto Armstrong rechaza abiertamente la existencia de diferencias cualitativas entre lo biológico y lo físico. Según Armstrong de aceptarse propiedades novedosas en el cerebro tendría que aceptarse también la extrapolación de tales propiedades a otros sistemas físicos, como las máquinas, que dispongan de un alto grado de complejidad estructural. Al

hacerlo el partidario del emergentismo se refutaría a si mismo porque estaría borrando la línea divisoria que él ha querido trazar entre la física y la biología. Es una ironía histórica que en la época en que Armstrong negaba la posibilidad de máquinas pensantes porque ello implicaría borrar la diferencia entre las máquinas y los organismos, los herederos más radicales de Turing - como Putnam en su primer momento – replanteaban radicalmente la posibilidad de que una máquina de Turing pudiese pensar, dejando de lado, por irrelevante la cuestión de si el material del que está compuesto el sistema pensante son las neuronas cerebrales, los pensamientos inmateriales de un ser incorpóreo o los microchips de un computador. El estado de la cuestión vino a quedar no en la negación de que las máquinas piensen sino en como funciona un sistema pensante.

Finalmente, una breve reflexión de orden social: si como consecuencia de la revolución informática hemos llegado a creer erróneamente que información y conocimiento es lo mismo, entonces no debería extrañarnos que como consecuencia del vendaval de información al que tenemos acceso no sabemos como enfrentar las consecuencias negativas de tales tecnologías. No sabemos como contener lo que ha salido de la caja de Pandora que hemos abierto. La reacción correcta no debería ser una negación de la tecnología pues esa alternativa lleva al declive social. La vía correcta parece ser el control social de sus consecuencias imprevistas y desagradables, tanto en lo que atañe a nuestra existencia individual como colectiva. Tecnología y humanidad no se excluyen, en su justa combinación quizá es la clave para el progreso individual y social (Mario Bunge ha defendido un punto de vista similar al nuestro en *Crisis y reconstrucción de la filosofía*).

Referencias:

- Armstrong, David (1968) *A Materialist Theory of Mind*. London, Routledge and Kegan Paul.
- Churchland, Paul (1984) *Matter and Consciousness*, Massachussets, MIT Press, 1981.
- Kuhn, Th. (1992) *La estructura de las revoluciones científicas*. Fondo de Cultura Económica.
- Rivière, A. (1991) *Objetos con mente*. Alianza, Madrid.
- Rivière, A. y M. Nuñez (2001) *La mirada mental*. Aique, Buenos Aires.
- Searle, John (1980) "Minds, brains and programs" *Behavioral and Brain Sciences* 3, 417-424.
- Searle, John (2001) *El misterio de la conciencia*. Paidós, Barcelona.
- Turing, Alan (1950) "Computing Machinery and Intelligence", *Mind*, 59: 433-460.
- Ursúa, Nicanor (1991) *Cerebro y conocimiento. Un enfoque evolucionista*. Anthropos, Barcelona.
- Young, J. Z. (1992) *Filosofía y cerebro*. Sirmio, Barcelona.