
Apuntes sobre una historia de la mente mecánica

Guillermo Cecchi

¿Sabes lo que pienso?”, dice ella. “Que los recuerdos de la gente son tal vez el combustible que queman para permanecer con vida. Si esos recuerdos tienen importancia real o no, no es relevante para el mantenimiento de la vida. El fuego no está pensando, ‘Ah, esto es Kant’, o ‘Ah, ésta es la edición vespertina del Yomuri’, o ‘Lindos pechos’, mientras quema. Para el fuego, no son más que desechos. Es exactamente lo mismo. Recuerdos importantes, recuerdos no tan importantes, recuerdos totalmente inútiles: no hay distinción –son todos combustible.

Haruki Murakami, *After Dark*

El concepto de mente, y en particular las concepciones del funcionamiento mental basadas en analogías mecánicas, tienen una historia tan larga, quizá, como la de la palabra escrita. Está más allá de los límites del presente artículo el proveer un análisis exhaustivo de las ideas sobre la mente, habida cuenta su complejidad, variedad y de enorme interconexión con prácticamente todos los aspectos de la sociedad. Nos limitaremos, por tanto, a examinar solamente algunos hitos en la historia mecánica de la mente que, vistos con la perspectiva de principios del siglo XXI, parecen especialmente relevantes. En particular, trataremos de enfocarnos en cómo el estado de la técnica y las teorías científicas dominantes en distintos períodos han co-evolucionado con el concepto de mente, para llegar a cómo la explosión de nuevas tecnologías puede afectar cómo la forma en que pensamos sobre la mente en cuanto máquina.

MENTE, MÁQUINA Y MATERIA

Si cada instrumento pudiese, en virtud de una orden recibida o, si se quiere, adivinada, trabajar por sí mismo, como las estatuas de Dédalo o los trípodes de Vulcano “que se iban solos a las reuniones de los dioses”; si las lanzaderas tejiesen por sí mismas; si el arco tocase solo la cítara, los empresarios prescindirían de los operarios, y los señores de los esclavos.

Aristóteles, *Política*

Tratándose de pensar en la mente, todos somos hijos putativos de Platón y Aristóteles, queramos o no. Si bien, y con razón (tal fue la intención del señor griego), el pasaje citado de la *Política* es normalmente utilizado para discutir el concepto aristotélico de esclavitud en relación a las necesidades económicas de la sociedad, su contenido nos parece relevante como *prefigurador* de nuestras ideas sobre los autómatas. En primer lugar, la “lanzadera” (objeto que se usaba para enrollar el hilo) que camina por sí misma, y el pasaje en su totalidad, no parecen ser una metáfora o un análisis mecanístico de la mente. Sin embargo, es interesante notar que luego de la obvia implicación entre automatización y esclavitud, el próximo ejemplo es el de una cítara capaz de ejecutar música por sí misma, una actividad sin valor económico evidente, y que implica valores puramente humanísticos. De la infinitud de ejemplos que se podrían haber utilizado, ¿por qué este en particular? Nos arriesgamos a sugerir, atrevidamente, que Aristóteles prefigura la idea del autómata como reflejo o copia imperfecta del ser humano, en lugar de como, simplemente, una herramienta compleja. Esto se hace evidente también por el hecho de que en este ejercicio retórico la lanzadera se propone como reemplazo del esclavo, a quién Aristóteles no priva de su humanidad, sino de quien considera que su esclavitud es consecuencia de la necesidad y de la tendencia natural de los hombres a ser amos o esclavos. De hecho, el texto hace referencia explícita al esclavo como instrumento vivo en comparación al instrumento inanimado, de modo que la lanzadera, en virtud de su autonomía, podría considerarse como un esclavo.

Podría pensarse que Aristóteles utiliza este recurso retórico como argumento para demostrar la necesidad absoluta de la esclavitud, dado que sería imposible en la Grecia del siglo IV a.C. imaginar una lanzadera automática. Sin embargo, no parece ser tal la intención. Esta imposibilidad no

se presenta de manera explícita, es decir, no se dice “la lanzadera jamás podrá caminar”, o que sea imposible reemplazar el trabajo del esclavo. Tal vez nuestro viejo sabio ya había visto suficientes herramientas movidas por el viento o el agua sin intervención humana como para descartar de plano esta posibilidad. Pero hay una parte del pasaje, su mismo inicio, que resulta a mi entender extremadamente reveladora: “en virtud de una orden recibida o, si se quiere, adivinada”. Aquí Aristóteles implícitamente nos sugiere cómo debería ser el autómeta que reemplace al esclavo, y en su *gedanken experiment* prefigura lo que será el ideal de autómeta como instrumento, como reflejo, copia o imitación del hombre: la lanzadera que camina por sí misma y es capaz de adivinar los deseos de su amo.

Aquí entonces marcamos, de manera arbitraria, y usando con descaro el viejo truco del argumento *ad verecundiam* (si lo dijo Aristóteles, será cierto) el primer hito: aún cuan indirecta, probablemente la primera mención al concepto de autómeta (al menos en la tradición occidental) se hace en referencia a la posibilidad de: (i) reemplazar una herramienta conocida, (ii) reemplazar al esclavo que la opera, y (iii) dotarlo al autómeta de facultades cognitivas.

Fast forward. Desde el punto de vista de la tecnología, del concepto de autómeta y de lo que hoy podríamos reconocer como teorías de la mente, no vamos a recalar en un largo período de la historia; volveremos por él más tarde. Sencillamente, el próximo puerto es Descartes. Nuestro René, como es bien sabido, vivió en un período de extraordinaria efervescencia científica y tecnológica (1596 a 1650), y junto a Leibniz (1646 a 1716) fundó las bases del dualismo: en las palabras más modernas de Leibniz, “los espíritus obran de acuerdo a las leyes de las causas finales a través de sus deseos, fines y medios. Los cuerpos obran de acuerdo a las leyes de las causas eficientes del movimiento. Los dos reinos, el de las causas eficientes y el de las causas finales, están en armonía entre sí. De acuerdo a este sistema los cuerpos obran como si (para suponer lo imposible) no hubiesen espíritus, y los espíritus como si no hubiesen cuerpos, y aún así tanto cuerpo como espíritu obran como si el uno influenciara al otro”. (Este pasaje proviene de: Leibniz, *Monadología*, §79 y §81; de hecho, Leibniz se refiere a Descartes explícitamente en §80.)

Es decir, en la visión cartesiana el cuerpo actúa como un autómeta regido por las leyes estrictas de la materia, mientras que la mente lo hace por sus propias reglas, como el deseo. Cuerpo y mente no interactúan, no pue-

den hacerlo, porque ello implicaría que la mente también está sujeta a las leyes de la materia, lo cual es inaceptable. Al menos esta es la interpretación cartesiana de Leibniz, que se ha consolidado como la canónica. Se han llenado bibliotecas enteras acerca del por qué de este rechazo a la materialidad de la mente. Sin duda, una de las más fuertemente utilizadas es la razón teológica de preservar la inmaterialidad e inmortalidad del espíritu: si la mente está atada al cuerpo como esposa y no como invitada, entonces está sujeta a las mismas leyes de decaimiento y muerte del cuerpo. Un segundo argumento, también de origen teológico pero con múltiples ramificaciones, requiere que rompamos la linealidad histórica de nuestro análisis.

Rewind, siglo IV d.C. La Iglesia, entre tantos disensos teológicos, debate una cuestión fundamental: si Dios es todopoderoso, ¿por qué permite la existencia del Mal? La respuesta, que en esencia fue adoptada por la Iglesia hasta nuestros días, la proveyó San Agustín: Dios le otorga al hombre el libre albedrío, la libertad para elegir hacer el Bien usando su capacidad intelectual. Esta libertad, aunque influenciada por las pasiones, es finalmente irreductible. “No puede negarse que tenemos libre albedrío”, dice San Agustín en *Sobre el libre albedrío*. Si el hombre no tiene libertad para elegir, es decir, que está determinado por la materialidad de su naturaleza, es entonces el mismo Dios quien ha determinado que exista el Mal, y los hombres son inimputables. Esta conclusión es inaceptable para el orden establecido en la sociedad, aún en el siglo XXI.

Fast forward, siglo XVI. La revolución científica que se había iniciado con Copérnico y Galileo estaba tomando la forma que se cristalizaría con Descartes, Newton y Leibniz: el mundo material es un gran mecanismo de relojería, en donde los objetos se mueven debido a la fuerza que otros objetos ejercen sobre ellos, con leyes que pueden describirse matemáticamente y que están dominadas por lo que puede llamarse determinismo clásico. Pierre-Simon Laplace, ya en el siglo XVIII, lo expresaría célebramente diciendo que si uno conociera la posición y velocidad iniciales de todas las partículas del universo, sería posible determinar unívocamente el futuro del mismo. El enorme éxito de la teoría newtoniana en explicar la mecánica celeste se refleja en la confianza de Laplace en este determinismo absoluto: si se conocen las leyes, se conoce el futuro. El detalle de las condiciones iniciales parece menor, un mero tecnicismo.

Es comprensible entonces que Descartes propusiera su doctrina dualista, aún antes de Newton. La tecnología, aunque en expansión, sólo podía ofrecer como ejemplo más refinado el mecanismo de relojería. Curiosamente, si bien René podía disponer de relojes de cuerda, no fue sino hasta seis años después de su muerte que Huygens introdujo el mucho más preciso reloj de péndulo (1656). El reloj, más que cualquier otra herramienta, estaba diseñado para que fuera exacto, repetible, sujeto a las leyes inmutables del movimiento y la gravedad. ¿Cómo pensar que aún el más complejo de los relojes pudiera siquiera compararse con la vida mental? Sin embargo, sin embargo ... Don René, tal vez por motivos similares a los de Aristóteles, se esfuerza en desarrollar un argumento contra la posibilidad de que aún autómatas altamente sofisticados posean una mente comparable a la del hombre:

“Y aquí me extendí particularmente, haciendo ver que si hubiese máquinas tales que tuviesen los órganos y figura exterior de un mono o de otro cualquier animal, desprovisto de razón, no habría medio alguno que nos permitiera conocer que no son en todo de igual naturaleza que esos animales; mientras que si las hubiera que semejasen a nuestros cuerpos e imitasen nuestras acciones, cuanto fuere moralmente posible, siempre tendríamos dos medios muy ciertos para reconocer que no por eso son hombres verdaderos; y es el primero [...] no se concibe que ordene en varios modos las palabras para contestar al sentido de todo lo que en su presencia se diga, como pueden hacerlo aun los más estúpidos de entre los hombres; y es el segundo que, aun cuando hicieran varias cosas tan bien y acaso mejor que ninguno de nosotros, no dejarían de fallar en otras, por donde se descubriría que no obran por conocimiento, sino sólo por la disposición de sus órganos, pues mientras que la razón es un instrumento universal, que puede servir en todas las coyunturas, esos órganos, en cambio, necesitan una particular disposición para cada acción particular; por donde sucede que es moralmente imposible que haya tantas y tan varias disposiciones en una máquina...”. (Descartes, *Discurso del método*).

Comencemos por el segundo argumento cartesiano. Claramente, lo que se está expresando es la idea del determinismo clásico, inspirado por las leyes de la mecánica y por los engranajes, poleas y pesas de la maquinaria contemporánea. Podríamos fácilmente reconocer en esta descripción lo que

hoy llamaríamos Máquina de Estados Finitos Determinista: un sistema definido por un número finito de estados internos, y por transiciones entre ellos unívocamente determinadas por el estado interno presente, al que puede incluirse un número también finito de estados de entrada. Es una máquina perfectamente laplaciana: si conocemos los estados iniciales, conocemos su futuro.

Sin sorpresas. Como el reloj de Huygens.

Fast forward. En la segunda mitad del siglo XIX el desarrollo de la termodinámica y la mecánica estadística, gracias al trabajo de Clausius y Boltzmann, comienza a modificar la concepción de la materia. En lugar del universo cósmico regido por las leyes exactas de la mecánica celeste newtoniana, el universo molecular parece estar dominado por la irreversibilidad de la segunda ley de la termodinámica, y por el elemento errático, impredecible que introduce el azar de la temperatura. Primer golpe al determinismo laplaciano, y al segundo argumento cartesiano: las máquinas no tienen por qué operar según “la disposición de sus órganos” como mecanismos de reloj: si son lo suficientemente pequeñas, serán inevitablemente sujeto de los caprichos del azar termodinámico.

Pero esto es sólo el comienzo, porque el determinismo laplaciano recibirá su golpe de gracia en las primeras décadas del siglo XX, por medio del doble ataque de la mecánica cuántica y la teoría de sistemas dinámicos. La mecánica cuántica establece que, al menos en el mundo atómico, es imposible conocer con certeza toda la información acerca de una partícula (por ejemplo, su posición y su velocidad); más aún, los sistemas cuánticos se comportan de forma tal que parecen ocupar varios estados distintos al mismo tiempo, y sólo la intervención del experimentalista hace que el sistema “elijan”, de forma azarosa, un estado en particular. Esta interpretación de la cuántica llevó al mismo Einstein a renegar de ella aduciendo que “Dios no juega a los dados”. Curiosamente, una de sus contribuciones seminales durante el Año Maravilloso de 1905 fue en la teoría del Movimiento Browniano, la descripción estadística de partículas moviéndose bajo la influencia de la temperatura. Aún una de las mentes más brillantes en la historia de la ciencia no pudo superar una estructura epistemológica (o metafísica, si se quiere) tan profundamente arraigada en la cultura. Más aún, otra de sus contribuciones fundamentales (y que fuera la principal

justificación para otorgarle el premio Nobel) fue a la mismísima fundación de la mecánica cuántica, con su teoría de la cuantización de la luz en fotones para explicar el efecto foto-voltáico.

Sin embargo, podría aducirse, al menos al principio del siglo xx y antes del desarrollo de la electro-fisiología moderna, que los procesos atómicos y moleculares, sujetos a las caprichosas leyes de la cuántica y la termodinámica, no se aplican al cerebro, y menos aún a la mente. Después de todo, ciertamente no se aplican a los mecanismos de relojería que inspiraran al determinismo clásico. Pero el asedio no tiene fin: el trabajo de Aleksandr Lyapunov y Henri Poincarè sobre sistemas dinámicos introdujo un nuevo palo en la rueda. Si bien las ideas tomaron forma contemporáneamente en paralelo a la teoría cuántica, la disponibilidad de computadoras en la segunda mitad del siglo xx las hizo populares en la ciencia y fuera de ella, en donde suele llamarse “teoría del caos”. El concepto básico, atribuible a Lyapunov (1857-1918), es que aún cuando estemos en presencia de un sistema completamente libre de azar (ya sea cuántico o termodinámico), en tanto sea un sistema “legal”, es decir regido por leyes que pueden expresarse matemáticamente (como con ecuaciones diferenciales), para determinar con certeza la evolución de dicho sistema es necesario conocer con precisión sus condiciones iniciales. Hasta aquí, Laplace; pero, ¿qué sucede si el sistema es “divergente”, es decir, si condiciones iniciales similares nos llevan a destinos muy distintos? En tal caso, la posibilidad de predecir el futuro depende de la capacidad de determinar las condiciones iniciales. En casos extremos, en el “caos”, el sistema “olvida” infinita, rápidamente su condición inicial, de manera tal que para todo propósito práctico se vuelve impredecible: sabemos sus leyes, que no incluyen el azar, pero no sabemos dónde estará en el futuro inmediato. En otras palabras, un sistema caótico está determinado “por la disposición de sus órganos”, pero nadie puede preveer lo que tal disposición conlleva. Este no es un concepto puramente académico, dado que una enorme cantidad de sistemas, desde la dinámica de poblaciones hasta la de las condiciones atmosféricas, son susceptibles de comportamiento caótico. Cuanto más complejo, más probable es que el sistema sea divergente o inestable, lo cual sabemos hoy día se aplica especialmente al tejido nervioso, cuyos elementos primarios, las neuronas, son en sí mismas altamente complejas y divergentes.

Recapitemos entonces: el segundo argumento cartesiano para negar la equivalencia entre la mente humana y la mente mecánica se basa en una concepción limitada del determinismo material, que fue destronada abruptamente en el plazo de unas pocas décadas por nuevas formas de describir el comportamiento del mundo material: la termodinámica, la cuántica y la teoría de sistemas dinámicos. Subyacente al argumento antimecánico se encuentra la idea del libre albedrío, que sufre de manera similar una erosión a manos de las nuevas teorías: visto desde fuera, un sistema que es irreduciblemente o prácticamente impredecible es indistinguible de uno que actúa como si no tuviera ataduras materiales, como si fuera radicalmente libre.

Ahora bien, Don René, viejo zorro, nos tendió otra trampa: su primer argumento, el que de modo arbitrario decidimos discutir en segundo lugar. Con él trata de convencernos de que los autómatas jamás podrán establecer una conversación, tal y como aún los hombres más simples pueden hacer. Aunque va más allá de la intención del presente artículo, cabe señalar que este argumento se relaciona con la idea de la radicalidad subjetiva del cartesianismo, es decir, que es últimamente imposible comprender, experimentar o compartir la subjetividad del otro, dado que para el uno el otro siempre será objeto. Únicamente el sujeto puede experimentar su subjetividad, la objetividad está fuera del sujeto y pertenece a la esfera pública, cuya forma más concreta o mesurable es el lenguaje. En nuestra interpretación, Descartes nos dice que el sujeto sólo puede conocer al otro (el autómata) como objeto por medio de la comunicación. La asignación de inter-subjetividad ya es de por sí tentativa cuando se trata de humanos, es decir, sólo puedo inferir que el otro es un sujeto tal cual lo soy yo, porque la subjetividad es intransferible e imposible de medir; ¿qué decir entonces de los autómatas, que ni siquiera pueden proferir palabras? Un argumento muy difícil de rebatir, particularmente en tiempos de engranajes y poleas.

Fast forward, nuevamente. Es 1950, la guerra ha terminado y nuestro héroe, Alan Turing, ha contribuido no sólo a refundar la matemática, a fundar la ciencia de la computación y a redireccionar la biología matemática, sino también a descifrar el código de guerra nazi (¡en sólo 42 años de vida!). Su influencia en la ciencia y la tecnología es inconmensurable, pero popularmente es tal vez más conocido por un *gedanken experiment* que propuso en su artículo “Maquinaria computacional e inteligencia”, conocido como el

Test de Turing. Lo que hace Turing aquí es la más simple de las maniobras de judo: tomar la fuerza del enemigo y usarla contra él. El enemigo no es otro sino Descartes, y su fuerza es el primer argumento cartesiano contra la mente mecánica. Según vimos, Don René nos dice que los autómatas no son capaces de establecer una conversación, de forma que no pueden siquiera intentar convencernos de su subjetividad. Ante esto, Turing primero se hace la pregunta: “¿Pueden pensar las máquinas?”. De esto se concluye que debe definirse qué es “máquina” y qué es “pensar”, y propone un método no ambiguo para contestar la pregunta: el “juego de la imitación”. En este juego, un interrogador humano (C) interroga por escrito, y sin verlos, a un hombre (A) y una mujer (B); el objetivo de C es determinar el sexo de A y B, mientras que el de A es confundir a C. Luego de presentar las reglas del juego, Turing dice: “Ahora hacemos la pregunta, ‘¿Qué pasaría si una máquina tomara el lugar de A en el juego?’ ¿Decidirá equivocadamente el interrogador tan a menudo cuando el juego es jugado así como cuando es jugado entre un hombre y una mujer? Estas preguntas reemplazan nuestra original ‘¿Pueden pensar las máquinas?’”.

De modo brillante, Turing pone a Descartes de cabeza: si es difícil asignar subjetividad a otro ser humano, entonces es igualmente difícil determinar que un autómata no tiene subjetividad, y que en consecuencia su mente mecánica no puede compararse a la mente humana. Luego de presentar su juego, Turing desmorona una serie de objeciones que suelen utilizarse en contra de la mente mecánica, de manera más o menos jocosa. En un ataque directo al primer argumento cartesiano, propone la objeción de “las incapacidades varias”, entre las que enlista que las máquinas no pueden ser amables, hermosas, enamorarse o enamorar, entre otras. Y explícitamente, que no pueden usar las palabras de forma correcta. Acusa luego a quienes proponen esta objeción de inductivistas: si las máquinas nunca hicieron X (hablar apropiadamente, algo que puede objetivarse por medio del juego de la imitación) entonces nunca lo harán. Proviendo de quien construyera las primeras computadoras, es una acusación devastadora: unos pocos años antes, la electricidad sólo se usaba para mover motores y encender bombillas. De modo similar desarticula la objeción de la incapacidad de predecir el comportamiento humano aduciendo que, sencillamente, uno siempre puede incluir elementos estocásticos en el autómata. Visto a la

distancia, este argumento parece casi obvio, pero no lo era en su momento. De hecho, transcurrieron varias décadas hasta que el azar como tal fuera incorporado en el seno de la ciencia de la computación como herramienta, por ejemplo, en la optimización estocástica. Creemos, por el contrario, que el argumento es en parte reflejo de la nueva forma de pensar a la materia, en la que el azar y la inconsistencia, por así decirlo, tienen un rol central.

Podríamos decir que, dada la simplicidad de la propuesta de su Test, el edificio cartesiano era ya un castillo de naipes y que Turing sólo tuvo que gritar, como Nietzsche, que Dios había muerto. La potencia intelectual del Test de Turing y sus infinitas variaciones no ha disminuido con el paso del tiempo. Sería fútil y oneroso enlistar todas las formas de la prueba que se han diseñado, implementado o discutido, pero arbitrariamente podemos mencionar tres eventos. El primero es el Premio Loebner, posiblemente el más cercano a la idea original, instituido desde 1991. Si bien el premio mayor nunca fue otorgado, la enorme disponibilidad de recursos computacionales en años recientes ha generado creciente interés y expectativa; existen varias competencias similares. Un aspecto menos académico del legado de Turing es lo que se conoce como *captchas*, algoritmos intensamente utilizados en la web como “Test de Turing invertido”, con los que se intenta reconocer la presencia de un humano detrás del teclado, para separarlos de algoritmos robóticos en busca de datos (*malware*). Finalmente, la compañía de computación IBM, siguiendo la senda marcada por su éxito en la competencia de ajedrez contra Gary Kasparov en los 1990, participará en el show televisivo *Jeopardy!* en 2011. En este programa de concursos los participantes deben responder proveyendo la pregunta cuya respuesta es presentada, y que requiere conocimientos de cultura popular, arte, ciencia, deportes y una variedad de información sin estructura precisa. Si bien tanto la compañía como los organizadores de *Jeopardy!* no pretenden presentar la competencia como un Test de Turing, las comparaciones son inevitables. Una de las razones para esto es que, al igual que ocurriera con Deep Blue (el competidor de Kasparov), la arquitectura y programación del nuevo autómatas están basadas en consideraciones estrictamente ingenieriles, sin pretensión de asemejarse a estructuras neuronales o mentales. Sin embargo, una buena participación del sistema en una competencia típicamente “humana” nos confrontará con la misma

pregunta que hiciera Turing: “¿Pueden pensar las máquinas?” (En pruebas piloto, el sistema parece competir a un nivel comparable al humano).

Cerraremos aquí un arco que lleva de Descartes a Turing, que fuera tensado por Aristóteles y San Agustín. Proponemos que al estar ligada (aunque no “determinada”) la historia de la mente mecánica al estado de la tecnología y a las teorías de la materia, el paradigma cartesiano reinante durante varios siglos se comienza a desarticular en dos frentes más o menos simultáneos. Por una parte, las teorías físicas representan al mundo material como caprichoso, estocástico e impredecible, haciéndolo incompatible con el determinismo mecánico que inspirara la tecnología mecánica y la física newtoniana: los relojes y el sistema solar. Por otra parte, impulsado por el empuje incontenible de la digitalización y la computación electrónica, Turing cuestiona el subjetivismo como último bastión de exclusividad de la mente humana, y en el proceso libera a los proponentes de la mente mecánica, en su sentido más amplio.

IRRUPCIÓN DE LA BIOLOGÍA

Nuestro recorrido serpenteante nos lleva ahora, nuevamente, a los comienzos del siglo XX: a partir de ahí vamos a seguir un nuevo sendero que arribará a nuestra época. Tratándose de una historia de la mente mecánica, resultará curioso que ignoremos casi por completo, como se verá, una rama de la ciencia dedicada precisamente a replicar computacionalmente capacidades mentales: la Inteligencia Artificial (AI, por sus siglas en inglés). Nos referimos a ella al menos en su forma más tradicional, basada en la formalización de procesos con acceso introspectivo, es decir, deducción, razonamiento y planeamiento. La razón será breve: en su forma simbólica (dominante en sus inicios), a la AI parece ser sólo una extensión del paradigma cartesiano, y como tal no ofrece nuevos retos filosóficos. La imitación del procesamiento mental consciente en base a funciones lógicas no difiere mucho de la operación de engranajes y poleas, y de hecho ya estaba disponible en tiempos de Aristóteles.

Hecho el descargo, volvamos a nuestra estación finisecular. Mientras el mundo de la física y de la matemática transitaban por una revolución fundacional que reformulaba el concepto de materia, un cambio igualmente significativo, aunque menos estruendoso, tenía lugar en las ciencias de la

vida. Luego del espectacular impacto de Darwin en la ciencia y en la sociedad en general, dos desarrollos concernientes a la mente en cuanto manifestación del cerebro tenían lugar casi en paralelo. Por un lado, el trabajo de Ramón y Cajal y de Golgi comenzaba a revelar las bases fisiológicas del cerebro. Si bien Cajal y Golgi recibieron conjuntamente el Nobel en 1906, disputaban intransigentemente la validez de sus teorías: Cajal sostenía que la neurona es la unidad esencial e irreductible del cerebro, mientras que Golgi sostenía la teoría reticular del funcionamiento “global” o distribuido del mismo. A un punto tal que Golgi dedicó su discurso de aceptación del Nobel enteramente a criticar a Cajal. Pero más allá de esta disputa, es evidente que ambos entendían que el tejido nervioso es el que provee la base fisiológica del cerebro, y de la actividad mental que produce. Es un tejido extremadamente complejo, tan distante del reloj como Cajal de Descartes, y mucho más cercano al mundo estocástico de Boltzmann.

En sentido contemporáneo, uno de los discípulos del gran neurólogo Jean-Martin Charcot, Sigmund Freud, comenzaba a transformar a la psicología en una disciplina tan poderosa e impactante que la llevaría desde los claustros universitarios hasta los salones burgueses, y la mesa de los bares. Para lo que nos atañe, sin embargo, el surgimiento de Freud (y en similar medida de Jung, aunque con menos popularidad) es relevante porque impone en el discurso público la idea de que la mente opera en dos niveles que interactúan entre sí mientras mantienen su separación: la actividad consciente, sujeta a la introspección cartesiana, y la inconsciente, en la que operan pulsiones animales difíciles de controlar, como Eros, la pulsión de la vida, y Tánatos, la de la muerte. Si bien la idea del inconsciente no era del todo novedosa, si lo era en el marco de una teoría científica y de una práctica clínica. Freud, al tanto de los adelantos en la neurología y neurofisiología, hace explícita la conexión con los avances representados por Cajal y Golgi: en su *Proyecto para una psicología científica* (1895), propone que “la intención es proveer una psicología que sea una ciencia natural”. Incluso el texto contiene diagramas rudimentarios de redes neuronales, lo que enfatiza la noción de que, eventualmente, Freud quiere basar su teoría en el sustrato biológico. De modo similar, Jung, con su teoría de los arquetipos y del inconsciente colectivo, resalta el valor de lo *a priori* en el funcionamiento de la mente en el marco de la teoría evolutiva darwiniana. En palabras de un jungiano contemporáneo:

“Los arquetipos son concebidos como unidades neuropsíquicas que evolucionaron por medio de la selección natural y que son responsables de determinar las características del comportamiento tanto como de las experiencias afectivas y cognitivas de los seres humanos” (A. Stevens, *Psiquiatría evolutiva*). La lucha por la supervivencia, entonces, está unida irrevocablemente a la vida mental, y requiere o impone condicionamientos a su libertad.

La irrupción de la biología y de la teoría evolutiva en la psicología conforman un nuevo ataque a la distinción cartesiana entre mente humana y mente mecánica: no solamente tiene la mente estructuras *a priori* que la determinan, como ya lo propusiera Kant, sino que tienen un origen animal: la necesidad de sobrevivir y reproducirse. La “disposición de sus órganos” es parte de la mente; más aún, en particular para la posición jungiana, es la que le da sentido, impulso vital. Sólo que ahora el condicionamiento y la dependencia mutua entre la mente y la materia, a través del prisma de la tecnología, tiene otros actores. Por un lado, la nueva física y la nueva biología, átomos y neuronas. Por el otro, las computadoras electrónicas, que operan con la nueva física, de forma más estridente en el caso del transistor, basado enteramente en un fenómeno cuántico.

Ahora bien, ¿de qué forma este *a priori* evolutivo impacta a las teorías mecanísticas de la mente, en específico? Vamos a tomar dos ejemplos de pensadores para quienes, de formas distintas, la biología tiene una influencia directa y medible en el funcionamiento de la mente, aún para las operaciones cognitivas más abstractas. El primer caso es el de Noam Chomsky, quizás el intelectual más influyente de la segunda mitad del siglo xx. Chomsky se propuso estudiar el lenguaje humano con las herramientas derivadas de la teoría de la computación, para entender cuál es el grado de complejidad de las estructuras gramaticales (es decir, sin referencia explícita al contenido semántico) del lenguaje natural. Para ello propuso una jerarquía de lenguajes incrementalmente más complejos, los que pueden ser decodificados por máquinas de complejidad comparable. En el fondo de la jerarquía, las máquinas son autómatas de estados finitos (ya presentados anteriormente), mientras que en la cima son Máquinas de Turing Universales: básicamente, autómatas finitos pero con acceso a una memoria de capacidad de potencial infinito. Una de sus líneas de investigación ha sido entender cuán alto en esta jerarquía hay que subir para poder

decodificar distintos lenguajes y aspectos del lenguaje (como morfología y sintaxis). Su trabajo ha sido enormemente influyente en la lingüística, e incluso en la ciencia de la computación.

Justamente sus estudios sobre la complejidad de las estructuras del lenguaje lo llevaron a concluir que no es posible para los adultos aprenderlo con la facilidad y la rapidez con la que lo hacen los niños, sin la presencia de una maquinaria preconcebida y dedicada. Esta maquinaria, innata y específica al lenguaje, posee los principios de la Gramática Universal, de la que los distintos lenguajes particulares se derivan cuando se fijan ciertos parámetros (¿tienen género los sustantivos, hay pasado participio?). Una de las razones que usa para justificar la existencia de la Gramática Universal es la de la “pobreza del estímulo”: es verdaderamente muy poca información la que los niños reciben de sus padres y del resto de su entorno para construir, desde una tabla rasa, la maquinaria requerida para el lenguaje. Otras habilidades con carga cognitiva similar o aún menor, como por ejemplo la aritmética, son adquiridas de forma mucho más lenta y requieren un alto grado de entrenamiento y tutelaje. ¿Qué hace especial al lenguaje? La respuesta es simple: como herramienta de comunicación, tiene un valor superlativo para la supervivencia, y en términos darwinianos ha permitido a la especie humana encontrar un nicho evolutivo único, al constituir la base de la coordinación grupal y el conocimiento acumulado.

El dilema con el que se encontró Chomsky es el siguiente: aún cuando podamos construir un autómata que sea capaz de comprender y generar lenguaje, y que fuera óptimo para la implementación, como una máquina de Turing, ¿cómo es que tal máquina está allí en primer lugar, y es a la vez tan eficiente? Chomsky no duda que el lenguaje sea automatizable; al contrario, lo que propone es que la mente es un autómata con un módulo altamente complejo y especializado pero *formalizable*. Un módulo que la misma biología dio a luz.

Sin necesidad del *Fast forward*, avancemos hacia las décadas finales del siglo xx para concentrarnos ahora en nuestro segundo pensador, Antonio Damasio. A diferencia de Chomsky, Damasio es un neurólogo que gracias a los grandes avances en las técnicas experimentales ha podido indagar más profundamente en las bases neuronales de los fenómenos mentales, en particular por medio del uso de la imágenes por resonancia magnética

(MRI), y en conjunción con una gran disponibilidad de pacientes con diversos tipos de lesiones cerebrales. Sus contribuciones a la neurobiología se centran en el estudio del procesamiento de las emociones; entre otros hallazgos, demostró que una persona con una lesión en el área del cerebro conocida como amígdala (evolutivamente antigua) es incapaz de reconocer emociones expresadas facialmente (alegría, horror, disgusto), mientras que su habilidad para reconocer figuras geométricas abstractas (sin valor emocional) permanece intacta. Esta interacción entre una operación estrictamente formal, reconocimiento de patrones geométricos, y una subjetiva, como el sentimiento de disgusto, llevó a Damasio a mirarse en el espejo cartesiano. Sí, nuevamente debemos hacer *rewind* y confrontarnos con nuestro viejo amigo... En su muy popular libro *El error de Descartes: emoción, razón, y el cerebro humano*, Damasio propone que el error fundamental de Don René fue haber ignorado por completo el papel que juegan las emociones en la totalidad de la vida mental, aún en aquellas actividades que tienen la apariencia de ser puramente cognitivas. Es decir, el error cartesiano es haber separado razón de emoción. Las consecuencias de esta idea se hacen evidentes en uno de sus más elegantes resultados experimentales, publicado en el artículo “Decidiendo ventajosamente antes de conocer la estrategia adecuada”. Damasio reporta lo siguiente: un grupo de sujetos son invitados a participar de un juego con ciertas reglas. Desconocido para los participantes, el juego está “cargado”, de manera que si siguen las reglas ya establecidas terminarán perdiendo; eventualmente, los participantes descubren la “trampa” e implementan la estrategia de juego adecuada. Los experimentalistas también determinan el momento preciso en el que cada participante descubre, conscientemente, la “trampa”. Hasta aquí, nada curioso. Sin embargo, mientras el juego transcurre, a los participantes se les mide el nivel de conductividad eléctrica de la piel (SCR), una variable que está determinada por el sistema simpático, que a su vez controla las glándulas de la transpiración, junto a varios otros órganos internos. La SCR también está asociada con estados emotivos tales como la excitación o el estrés; de alguna forma, es posible entender que el estado de los órganos internos sea reflejado como un estado emocional básico, tal como el malestar asociado con el hambre y el frío. Es de esperar, entonces, que el descubrimiento de la “trampa” conlleve un cambio emotivo, disparando un estado de exci-

tación o sorpresa, reflejado eventualmente en la SCR. Pero, de forma inesperada, el cambio en la SCR tiene lugar mucho antes que el descubrimiento consciente. Es decir, el sistema simpatético, origen de las emociones primarias, predice aquello que eventualmente el procesamiento consciente decidirá. ¿Cómo es esto posible? La respuesta de Damasio es que razón y emoción no pueden separarse con facilidad. Que el procesamiento inconsciente previo a la decisión consciente incluye un “marcador somático” que, siguiendo los mismos lineamientos de la señal del hambre, dispara una emoción primaria para que la razón consciente actúe sobre ella, o en función de ella. De la misma forma, las emociones pueden ser elaboradas cognitivamente para transformarse en sentimientos como el amor, el odio o la compasión, que pueden ser objeto de análisis introspectivo. En este sentido, Damasio no dista demasiado de Freud y Jung, pero va más allá cuando propone que estos impulsos vitales actúan directamente sobre la razón, como lo demuestra el experimento de la “trampa” bajo su interpretación: mientras la conciencia se ocupa de los aspectos inmediatos del juego, un sistema de procesos inconscientes acumula suficiente evidencia, posiblemente desestructurada, a cerca de que algo “huele mal”, algo que “se siente en las entrañas”, para luego sonar la alarma y forzar a la actividad consciente a intervenir. Es más, sin esta fuerza vital la razón consciente no tendría motivos para actuar, no se vería movida a hacerlo.

¿Cuán equivocado estaba Descartes, según Damasio? Tal vez nos hayamos ensañado demasiado con nuestro viejo amigo René, pues es fácil hacer leña del árbol caído, después de todo. Deberíamos hurgar un poco más profundamente en su pensamiento. En otro pasaje de las *Meditaciones*, nos dice que “La Naturaleza me enseña a través de las sensaciones de hambre y sed, etc., que no estoy presente en mi cuerpo meramente como un marinero está presente en un barco, sino que estoy unido muy cercanamente, si se quiere estoy entremezclado con él, de manera que yo y mi cuerpo formamos una unidad”, y más adelante que “Estas sensaciones de hambre, sed, dolor, etc., no son más que modos confusos de pensamiento que surgen de la unión y entremezcla, si se quiere, de la mente y el cuerpo”. ¿Entonces? Lo que Descartes trata de decirnos es que, si bien el cuerpo y la mente están en íntimo contacto, la mente, por medio de la reflexión introspectiva, el *cogito*, puede separarse del cuerpo y seguir sus propias reglas. En el expe-

rimento de Damasio, por el contrario, la razón parece más un sirviente de la emoción. De manera asombrosa, algunos de los participantes del juego que no reconocieron conscientemente que había una trampa, sin embargo adoptaron la estrategia adecuada luego de que su SCR sonara la alarma; ¿deberíamos considerarlos autómatas cartesianos por eso?

La irrupción de la biología, de la mano de las nuevas teorías de la materia, parece indicarnos que la respuesta es, en parte, afirmativa: somos autómatas, pero con reglas que pueden ser caprichosas e impredecibles, y en las que emoción y razón funcionan como dos caras de la misma moneda, la una mirando hacia adentro, y la otra exterior y pública, aún cuando sus elucubraciones se hagan en silencio.

CONVERGENCIA

Es hora, finalmente, de acercarnos al siglo XXI, dominado no tanto por nuevas teorías físicas sino por el continuo avance de la experimentación biológica y, fundamentalmente, por las nuevas tecnologías de la información. Para ello vamos a pasar primero por una estación ineludible: la década de 1960, con toda su carga de revolución, exceso e imaginación. En particular, queremos hablar con el profeta de los medios, el visionario de las nuevas formas de comunicación: Marshall McLuhan. Este pensador es famoso por su estilo propagandístico, acuñador de frases “virales” antes que existieran las redes sociales electrónicas: “el medio es el mensaje” y “la aldea global” son las más conocidas. Para lo que nos concierne, la mente mecánica, McLuhan propone una idea que es a la vez radical y una nueva vuelta de tuerca a la tabla rasa aristotélica: que la mente humana y la cultura en general están determinadas por la tecnología en uso, y muy especialmente por las formas de comunicación que éstas permiten. En *La galaxia Gutenberg* hace especial énfasis en cómo las diferentes transiciones de la palabra hablada a la manuscrita, de ésta a la mecánicamente impresa y finalmente a los medios audiovisuales, fueron concomitantes con enormes, revolucionarios cambios culturales. McLuhan explica esto basado en la idea de que el modo sensorial dominante en cada forma de comunicación determina la estructura de la mente, al tener que adaptarse a sus requerimientos físicos: “Cuando la tecnología extiende uno de nuestros sentidos, una nueva conversión de

la cultura ocurre tan rápido como la nueva tecnología es interiorizada. Una teoría del cambio cultural es imposible sin el conocimiento de los cambios en las relaciones sensoriales efectuados por distintas externalizaciones de nuestros sentidos” (Marshall McLuhan, *La galaxia Gutenberg*).

En particular, explica que la palabra escrita impone una linealidad al pensamiento, y por ende a la cultura, mucho más estricta que en la tradición oral, y que se ve acentuada por la introducción de la imprenta y del tipo móvil por Gutenberg. Propone así que el surgimiento de los medios audiovisuales, al liberar la comunicación humana de esta linealidad, deberán tener un efecto revolucionario: “Le imprenta creó al público. La tecnología eléctrica creó a las masas. El público consiste de individuos separados caminando con puntos de vista separados, fijos. La nueva tecnología demanda que abandonemos el lujo de esta posición, esta perspectiva fragmentaria”. Uno de sus seguidores, Leonard Shlain, va más allá y argumenta en *El alfabeto versus la diosa* que la transición de la oralidad a la palabra escrita fue responsable del abandono del culto a la diosa y la femineidad de la cultura, y que igualmente el dominio de la comunicación visual, no lineal y espacialmente integrativa, dará paso a una nueva era de igualdad entre los sexos.

McLuhan lleva la idea de la herramienta como extensión de nuestras facultades físicas y mentales a su conclusión más obvia: la tecnología electrónica no es sino una extensión del sistema nervioso central. Si la rueda es extensión del pie, y el martillo de la mano, el teléfono lo es de la voz y el oído. Nada nuevo. A menos que lo acoplemos con su idea de que cada externalización trae aparejada una nueva internalización igualmente física y mental. Si el uso del automóvil nos hace sedentarios, ¿por qué no habría de afectar nuestro pensamiento el uso de los medios de comunicación?

De hecho, la pregunta correcta no es si lo afecta, sino de qué forma. Todos conocemos por experiencia propia el efecto de la tecnología electrónica en nuestra vida cotidiana: cómo el uso de teléfono celulares, televisores, radios, del omnipresente correo electrónico nos distraen, nos enfocan, nos multiplican y nos dividen. También sabemos que están afectando a la sociedad en general, a la política, economía y cultura en formas que aún no podemos comprender.

Pero queremos indagar específicamente en el impacto de las nuevas tecnologías en nuestro concepto de mente mecánica, o tal vez debamos llamarla, por

el momento, electrónica. Se nos antoja hacer esta vez nuestro propio *gedanken experiment*, usando el *fast forward* hacia un punto indeterminado del futuro.

La experimentación en neurobiología en este futuro hipotético ha avanzado enormemente desde los primeros trabajos de, entre otros, el grupo de Miguel Nicolelis en la Universidad Duke de Carolina del Norte, en Estados Unidos. Sus estudios de neuroprostética demostraron resultados impactantes en su momento. Utilizando electrodos implantados en el cerebro de un mono, fue capaz de controlar remotamente, transfiriendo datos a través de Internet, un brazo robótico situado a cientos de kilómetros, en el Massachusetts Institute of Technology, que repetía los movimientos del brazo del mono. En un experimento relacionado, la señal del cerebro del mono controlaba a un robot más complejo, capaz de caminar en una cinta mecánica. El mono, caminando en su propia cinta, no sólo tenía electrodos en una región que produce comandos motores, sino que también recibía estimulación eléctrica retroalimentada a partir de la posición del robot, como si proviniera de su propio cuerpo. Luego de establecer el circuito cerrado entre el cerebro del mono y el movimiento del robot, el mono, en cierto momento, aprendió que no tenía por qué mover sus piernas, porque el robot las movía por él. Es decir: el cerebro del mono interiorizó, en palabras de McLuhan, la externalización de su nueva herramienta, de su nueva extensión. El robot, vale decir, estaba ubicado en Japón.

Pero volvamos al futuro. Impulsada por la virtualmente infinita disponibilidad de almacenamiento y ciclos de computación, y la nano-bio-tecnología, la neuroprostética ha logrado sortear los problemas que le impedían en tiempos de Nicolelis implantar más allá de algunos cientos de electrodos, y sólo en animales experimentales. Ahora es posible propalar la actividad de millones de neuronas en el cerebro humano, e igualmente enviarle otras tantas señales codificando todo tipo de información, sin peligro para la salud ni necesidad de estar acoplado físicamente por medio de cables. Es posible también grabar en forma continua horas y días de actividad cerebral, para luego editarla a nuestro antojo, compartirla o volver a recrearla por medio de nuestro portal de *upload*, tal como lo prefiguraran innumerables novelas y películas de ciencia-ficción.

Proliferan compañías que prometen todo tipo de productos que utilizan esta corriente virtual, el *stream*, para nuestro beneficio (¡avísenme cuando

esté bajando la guardia con mi ex!) o el de otros (¡hice viral el *stream* embarazoso de mi ex!). El portal LiveStream permite que compartamos nuestras experiencias más íntimas en tiempo real con miles de amigos y desconocidos, transformando la idea de aldea global en la de dormitorio común —la familia de William Gibson ha iniciado juicio por derechos de autor—.

El significado de “Lo siento” y “Estoy contigo” ha cambiado de forma radical.

Mientras el mundo virtual fluye y refluye, los filósofos se preguntan qué ha quedado de nuestra subjetividad individual irreductible. ¿Quién puede decir que es imposible ‘saber lo que es ser un murciélago’, como afirmara Thomas Nagel? Hay quienes sostienen que saberlo es sólo una cuestión de grado, que podemos estar tan cerca de ello como queramos gracias a la tecnología del *stream*; otros siguen afirmando que la única forma de realmente entender lo que significa ser un murciélago es ... ser uno. Como siempre, hay otros que niegan y reniegan de los dos anteriores, y que sostienen que el ser humano siempre puede mantener su subjetividad eligiendo no participar en el juego del flujo virtual, optando por retener su subjetividad. Finalmente, los miembros del grupo Herederos de Turing critican a quienes atesoran estas nociones. Argumentan que su posición es escapista, que temen confrontarse con la verdad. Que la convergencia entre el sistema nervioso central y el sistema electrónico exógeno ha borrado las viejas distinciones, y que las subjetividades de los individuos se han expandido de manera tal hacia el exterior, superponiéndose entre sí y llenando todo el espacio perceptivo disponible, que no hay lugar para discriminaciones. Que esto ha ocurrido como consecuencia de una necesidad, una pulsión irrefrenable de la misma condición humana, y que la internalización de las nuevas tecnologías ha sido tan súbita y natural que sólo cabe concluir que, de una manera u otra, siempre estuvo allí esperando ser revelada. Que el hombre entrega su subjetividad de buena gana porque siempre ha sido un espejismo, producto de una fase histórica de la relación entre hombre y naturaleza. Que la mente mecánica del autómatas es ahora la mente bio-electrónica del híbrido humano-digital.

DIVERGENCIA

No queremos hacer futurología, y no deberíamos hacerla si fuéramos fieles a las ideas sobre *impredictibilidad* de los sistemas complejos que discutimos previamente. El propósito de nuestro pequeño experimento es enfatizar que, habida cuenta la historia de los efectos de la tecnología y las teorías de la materia sobre nuestra concepción de la mente como mecanismo, sólo podemos esperar que el prodigioso avance de las tecnologías digitales y biológicas produzcan nuevas concepciones. Creemos, arriesgándonos, que estas concepciones implicarán un cambio radical en la noción de autómeta y de subjetividad, que se moverán en la dirección de la convergencia para la primera, y divergencia para la segunda. Es decir, convergencia entre ser humano y tecnología de la que ya tenemos atisbos, y divergencia de la individualidad en una multi-subjetividad que solo podemos imaginar. ❧

REFERENCIAS

Libros

- Aristóteles, *Politics*. The MIT Internet Classics Archive, <http://classics.mit.edu>.
- Noam Chomsky, *Language and Mind*. Cambridge: Cambridge University Press, 2006.
- Antonio Damasio, *Descartes' Error: Emotion, Reason and the Human Brain*. Nueva York: Putnam, 1994.
- René Descartes, *Discourse on the Method of Rightly Conducting One's Reason and of Seeking Truth*. The Gutenberg Project, <http://www.gutenberg.org>.
- Sigmund Freud, *Project for a Scientific Psychology* (1895), en *The Standard Edition of the complete psychological works of Sigmund Freud*, J. Strachey (Ed.), vol. I, pp. 283-387. Londres: Hogarth Press, 1950.
- Marshall McLuhan, *The Gutenberg Galaxy: the Making of Typographic Man*. Toronto: University of Toronto Press, 1962.
- Haruki Murakami, *After Dark*. Nueva York: A.A. Knopf, Nueva York, 2007.
- Gottfried Leibniz, *Monadology*. Philosophy Department Internet Archive, University of Leeds, <http://www.philosophy.leeds.ac.uk>.
- San Agustín, *On Free Choice of the Will*. Indianápolis: Hackett, 1993.

Leonard Shlain, *The Alphabet Versus the Goddess: the Conflict Between Word and Image*. Nueva York: Penguin, 1998.

Anthony Stevens y John Price, *Evolutionary Psychiatry*. Londres: Routledge, 2000.

Steven H. Strogatz, *Nonlinear Dynamics And Chaos: With Applications To Physics, Biology, Chemistry, And Engineering*. Cambridge: Westview Press, 2000.

Artículos

Antoine Bechara, *et al.*, “Deciding advantageously before knowing the advantageous strategy”, *Science*, vol. 275, 1997, pp. 1293-1295.

Murray Campbell, *et al.*, “Deep Blue”, *Artificial Intelligence*, vol. 134, 2002, pp. 57-83.

Carmelo Golgi, “Nobel Lecture”, 1906, <http://nobelprize.org>.

Mitsuo Kawato, “Brain controlled robots”, *HFSP Journal*, vol. 2, 2008, pp. 136-142.

Thomas Nagel, “What is it like to be a bat?”, *The Philosophical Review*, vol. 83, 1974, pp. 435-450.

Alan M. Turing, “Computer Machinery and Intelligence”, *Mind*, vol. 59, núm. 236, 1950, pp. 433-460.

Jonah Wessberg, *et al.*, “Real-time prediction of hand trajectory by ensembles of cortical neurons in primates”, *Nature*, vol. 408, 2000, pp. 361-365.