Instituto de Investigaciones Gino Germani
VI Jornadas de Jóvenes Investigadores
10, 11 y 12 de noviembre de 2011
Mg. Santiago Koval
UBA / UADE
santiagokoval@gmail.com
Eje 7: Políticas del cuerpo.
Título: El Algoritmo de Dios

Abstract

Ciencia *real* y ciencia *ficción* han construido durante los últimos años mundos posibles presentados como técnicamente probables, que giran alrededor de la noción de que las fronteras que separan al hombre de sus productos técnicos se irán perdiendo, hasta desaparecer, en el futuro cercano. En el punto extremo de las parábolas del discurso científico y ficcional, la mente humana, máxima expresión de la capacidad organizativa de la naturaleza, se iguala al cerebro artificial, máximo estandarte de la capacidad creativa de la cultura. La construcción de tal dispositivo supone, desde un punto de vista teórico, el descubrimiento del *Algoritmo de Dios* (aquel conjunto finito de reglas que, con reminiscencias bíblicas, Dios usó en la noche de los tiempos para crear la mente humana). En el origen de los discursos que componen la plataforma imaginaria de nuestra era, se hace evidente, así, la presencia extendida del *dualismo cartesiano*, esquema de interpretación mecanicista y reduccionista que conduce a suponer, de un lado, que la mente humana puede ser reducida a fórmulas matemáticas y, de otro, que puede existir sin el cuerpo biológico en que se desarrolló durante millones de años.

1. Introducción

La llegada de las Tecnologías de la Información y la Comunicación (TIC) a partir de la década de 1970 produjo cambios de orden cualitativo en lo *tecnológico realizable* (lo que es técnicamente factible) y en lo *tecnológico concebible* (lo que es teóricamente pensable). Estas transformaciones se han expresado respectivamente en cambios de registro material y conceptual en nuestro espíritu de época, una nueva era en la historia de la humanidad signada por el uso cotidiano y natural de las tecnologías digitales. Tanto en el dominio de la *ciencia*

real como en el de la *ciencia ficción*, los discursos relativos a la *integración hombre-máquina*, que discurren acerca de la fusión entre biología y tecnología, apuntan en conjunto a la idea de que las TIC darán lugar, en un punto cercano, a una *singularidad tecnológica*, punto histórico de inflexión y cambio trascendental expresado en la aparición de *androides* y *poshumanos*, figuras artificiales ontológicamente idénticas, e incluso superiores, a los seres originales en que se inspiran.

La revolución tecnológica iniciada con la llegada del microprocesador se ha consolidado en los años noventa a partir de la convergencia material y conceptual entre biología, ingeniería genética, electrónica e informática. La aplicación de regiones tecnológicas diversas al desarrollo de máquinas humanas y humanos mecánicos ha dado lugar a un régimen de nuevos desarrollos técnicos, expresado en el surgimiento de seres artificiales híbridos, a mitad de camino entre biología *natural* y tecnología *cultural*. Este conjunto de nuevas posibilidades ha ido transformando el horizonte narrativo de los discursos asociados a esta rama de desarrollo, alimentando conforme a ello las fantasías y aspiraciones de realidad en los centros de investigación más influyentes del planeta. La mayoría de estos centros se ubican en universidades e institutos especializados en robótica, cibernética, nanotecnología, ingeniería genética, biotecnología, informática, etc., y sus autores y principales promotores son mayormente inventores y especialistas en tecnologías de punta que han participado desde hace años en el desarrollo de las técnicas sobre las que ahora reflexionan.

El peso teórico de sus argumentos ha ido influyendo en el imaginario social, mezclándose y alineándose lentamente con otro tipo de discursos, como el cinematográfico y el literario. Estos discursos de la ficción científica, comprometidos con la realidad tecnológica que los rodea, absorben desde una posición privilegiada –producto, en gran parte, de su carácter mitológico- el capital de nociones vinculado a estas nuevas posibilidades, proyectando, de acuerdo con su propia lógica discursiva, mundos posibles (relatos, historias, alegorías, etc.) poblados de *robots*, *androides*, *ciborgs* y *poshumanos*. Desde la perspectiva teórica del discurso tecnocientífico, el imaginario de la ciencia ficción es atravesado, en particular a partir de la década del setenta, por la idea fundamental, antes inconcebible, de que las TIC llevan las posibilidades de fusión entre hombres y máquinas a un nivel cualitativamente nuevo.

El objetivo de este trabajo consiste en revisar los argumentos construidos por las *representaciones de la tecnología* de nuestra era, expresadas en el contenido de los discursos teóricos y ficcionales que discurren acerca del futuro cercano de una humanidad abandonada a

la voluntad de la técnica. En particular, nos interesa puntualizar en los siguientes aspectos específicos, a saber: primero, el carácter mixto y sincrético de los discursos que componen la plataforma imaginaria, que nace como producto de la unión, en un nivel posible de análisis, entre ciencia real y ciencia ficción; segundo, la tendencia, ubicua en esta plataforma de discursos, a privilegiar lo mental y abstracto sobre lo corporal y concreto; por último, la propensión a aceptar que será posible, dado el avance tecnológico suficiente, replicar y descargar el contenido mental de un ser humano en un dispositivo técnico, planteo que hunde sus raíces filosóficas en el *dualismo cartesiano*, un esquema de pensamiento instalado en el mundo occidental hace casi 400 años.

Este trabajo podrá dar lugar a análisis posteriores orientados a ubicar a las representaciones de nuestra era en un contexto más amplio de valoraciones culturales y de regulaciones sociales hegemónicas, que se orientan, en conjunto y por separado, a clasificar el cuerpo, el cerebro y la mente humanas como metáforas nacidas de nuestra compresión de la técnica. En definitiva, nuestra ontología es deudora de preceptos tomados de la tecnología. Y así como el corazón se definió durante siglos como una bomba o un simple mecanismo de relojería, ahora nuestro cerebro y nuestra mente se entienden como máquinas electrónicas con poder de computación.

2. ¿Realidad científica o metáfora ficcional?

La alineación conceptual entre el discurso de la academia científica y el discurso de la ciencia ficción indica la existencia de una tendencia común dentro del imaginario vinculado a la integración hombre-máquina, caracterizada por una concepción de las nuevas tecnologías como factores cualitativos de transformación de lo real. De alguna manera, acertadas o no, las narraciones heterogéneas que discurren acerca de las consecuencias cercanas de la fusión entre humanos y entidades mecánicas se estrechan la mano alrededor de un núcleo común de ideas claramente discernibles. En términos de los discursos que las promueven, las TIC darán lugar a una nueva etapa en la historia de la humanidad marcada por el advenimiento de seres artificiales que realizan los sueños milenarios en que se fundan.

Ahora bien, tomando en consideración los discursos generados por la academia científica, lo cierto es que la distancia existente entre sus proyecciones y fantasías por un lado, y las realidades tangibles y efectivamente existentes en el mundo material sobre el que teorizan del otro, sigue siendo, pese a todo, manifiestamente notoria. Los sueños poshumanistas trascendentales de transbiomorfosis o descarga del contenido de la mente en

un sustrato digital y las fantasías singularistas de máquinas emocionales y robots universales dotados de inteligencia artificial que superan a la humanidad en todos sus aspectos, lejos de concretarse en realidades inmediatas, continúan siendo, en su calidad de discursos, propuestas ficcionales acerca de un futuro probable. A raíz de ello, cabe pensar que las parábolas y profecías construidas por los científicos y teóricos de nuestra era forman parte, no tanto de lo tecnológico realizable, sino más bien de lo tecnológico concebible. Su formulación teórico-científica, sus argumentos y metodologías rigurosas, no pueden ocultar el hecho de que, en esencia, estos discursos no difieren mucho de las propuestas de las ficciones científicas. Lejos de poder confirmar en los hechos las metáforas construidas con palabras, los discursos teóricos del mundo académico se subsumen así, irremediablemente, en el universo narrativo de la ciencia ficción y pasan a compartir, conforme a ello, la plataforma conceptual con los discursos imaginarios de la ciencia ficción.

Alineados, pues, respecto de los mundos posibles que proyectan y estrechamente ligados a un universo narrativo común, lo imaginario científico y lo imaginario ficcional constituyen, en conjunto, una instancia de discurso de nivel superior, fundamentalmente imaginaria, que reflexiona acerca de los mundo futuros que derivarán del desarrollo tecnológico reciente. Los argumentos teóricos provenientes de los centros de investigación más importantes del mundo por un lado y las parábolas narrativas procedentes de los relatos literarios y cinematográficos por el otro, conforman un ideario cultural de enorme significación conceptual, y el valor imaginario de sus aportes al desarrollo tecnocientífico debe tomarse seriamente en consideración a la hora de ensayar un diagnóstico contemporáneo acerca del estado del arte de la fusión entre humanos y tecnología.

Las proyecciones elaboradas desde esta plataforma imaginaria de carácter mixto giran en torno a los posibles efectos culturales y sociales que traerá a la humanidad la inminente llegada de una singularidad tecnológica en los próximos años. Dentro del enorme abanico de miedos y fantasías asociados a la emergencia de figuras como los robots, los androides, los ciborgs y los poshumanos, los discursos de nuestra era insisten en determinadas parábolas fundamentales que no hacen otra cosa que repetir y reformular, una y otra vez, temores ancestrales arraigados desde los orígenes de la humanidad misma: el terror a la rebelión de la criatura, sea un monstruo de carne o un artificio de metal, sea un autómata aislado o una especie de máquinas organizada; el temor al castigo divino como respuesta al desafío del hombre a los límites impuestos en su naturaleza por los dioses que lo crearon; el sueño de la fuerza de trabajo dócil, barata e infatigable, un esclavo mecánico incondicional sin pretensiones morales, que dejaría a los hombres en un paraíso terrenal; el anhelo de superar

las limitaciones físicas del cuerpo por medio de una elevación mística del espíritu y la mente, que llevaría a la inmortalidad; el miedo al desequilibrio en el mercado laboral por la introducción de trabajadores automáticos, que dejaría a gran parte de la población desempleada; la fascinación por la creación de vida y conciencia artificiales, que cuestionan de raíz la identidad y unicidad existencial del ser humano y ponen en jaque la moralidad de sus aspiraciones demiúrgicas; y así siguiendo.

Es de notar que los miedos y fantasías construidos por la ciencia ficción se consideran normalmente como mundos probables y no como escenarios imposibles, como sí lo son algunos mundos construidos por el género fantástico. La distinción trazada por Philip K. Dick entre fantasía y ficción científica se relaciona precisamente con esta noción: «[l]a *fantasía* trata de lo que la opinión general considera imposible; la *ciencia ficción* trata de lo que la opinión general considera posible en las circunstancias apropiadas»[1]. Los discursos de la ciencia ficción cuentan, de esta forma, con el poder persuasivo de aquello que puede llegar a ocurrir si se cumplen ciertas condiciones tecnológicas, que en el momento de su enunciación son teóricamente posibles.

Discursos de ficción científica que construyen mundos teóricamente posibles en condiciones técnicamente probables, las *representaciones de la tecnología* de nuestra era centran sus esfuerzos narrativos en la idea central de singularidad tecnológica, un cambio trascendental en la evolución humana producto de la llegada de nuevas posibilidades técnicas. La cuestión central, pues, que se encuentra en la base de estos planteos gira en torno a la pregunta fundamental acerca de si la singularidad tendrá efectivamente lugar o no en el futuro cercano. ¿Pasará a ser, en lo sucesivo, la metáfora ficcional de hoy realidad científica del mañana? Así como el sueño del viaje a la Luna, proyectado discursivamente en 1865 por Julio Verne, se consolidó a la postre, un siglo más tarde, como avance tecnológico en el dominio científico, así también, cabe preguntarse si las proyecciones narrativas de la ciencia ficción contemporánea acerca de la llegada de una singularidad tecnológica se convertirán ellas también, en un futuro inmediato, en presencia material de lo tecnológico realizable.

3. Singularidad tecnológica

Más allá del centro de un *agujero negro*, las leyes de la física y la matemática no se aplican, y todos los criterios universales dados desde siempre por sentado deben descartarse. El

[1] Dick (1981: 10).

perímetro del agujero negro, el *event horizon* (horizonte de eventos), marca el punto crítico de inflexión, a partir del cual la realidad misma (las dimensiones, la materia, el espacio y el tiempo) gira sobre sí misma y deviene otra cosa. En su centro, ocurre la singularidad, allí donde "el espacio-tiempo tiene una curvatura infinita y la materia tiende a una densidad infinita bajo la presión de la gravedad infinita. En una singularidad, el tiempo y el espacio dejan de existir tal como los conocemos, y con ellos, todas las leyes de la física"[2].

Nacida originalmente para el terreno de la física, la noción de *singularidad* o *punto singular*, aplicada a otros terrenos científicos, refiere un punto de cambio cualitativo y trascendental más allá del cual todos los modelos científicos dejan de tener validez predictiva y explicativa, y deben, por lo tanto, ser reemplazados por un nuevo paradigma de conocimiento. Utilizada incluso más ampliamente para todos los terrenos del conocimiento, la singularidad es, siguiendo al físico y matemático ruso Alexander Friedmann, un punto en el cual la teoría en sí misma se rompe.

Valiéndose del concepto de la física para el campo de la tecnología, y retomando los conceptos de John von Neumann, Raymond Kurzweil (1999) define la singularidad tecnológica como un período futuro en el que el ritmo de cambio tecnológico será tan rápido, y su impacto tan profundo, que la vida humana se transformará de manera irreversible. En este mismo orden, el matemático y autor de ciencia ficción Vernor Vinge (1993) constata que, tomando en cuenta el desarrollo tecnológico creciente de los últimos años, sobrevendrá inevitablemente entre 2020 y 2050 una singularidad tecnológica marcada, de un lado, por la aparición de entidades artificiales con una inteligencia superior a la humana (ordenadores sensibles superinteligentes, supercerebros diseñados genéticamente, redes electrónicas que adquieren conciencia, etc.) y, de otro, por la superación tecnológica de las capacidades humanas por medio de interfaces naturales hombre-máquina, que dará lugar a la emergencia de seres poshumanos superinteligentes.

Entendida como una pérdida fundamental de los puntos de referencia, esta noción se ha convertido en el eje conceptual por excelencia de los argumentos teóricos de nuestro tiempo, y a su alrededor se han concentrado todas las posturas discursivas, ficcionales o no, que ensayan acerca de los escenarios posibles que sobrevendrán en un futuro inminente. Promovida por las Tecnologías de la Comunicación y la Información, la singularidad tecnológica se propone como el resultado de la posibilidad real de alcanzar un nuevo estadio en el linaje evolutivo de los hombres y las máquinas.

---

[2] Cfr. sitio Web oficial de la NASA, disponible en http://imagine.gsfc.nasa.gov/docs/ask_astro/answers/010912a.html (consultado en agosto de 2011).

Del lado de la *integración exógena* (la tendencia de nuestra cultura a reproducir al hombre en un dispositivo técnico, es decir, a humanizar la máquina), la singularidad se expresa en la emergencia del *androide*, ser artificial de naturaleza idéntica al ser humano, dotado en su máxima expresión de inteligencia y conciencia artificiales. Del lado de la *integración endógena* (la propensión a potenciar lo humano por medios tecnológicos, esto es, a mecanizar al hombre), en la aparición del *poshumano*, ser humano de naturaleza idéntica a una máquina, descargado en su máxima expresión a un sustrato digital. De uno y de otro lado, la singularidad, entendida como un punto extremo en el largo proceso histórico de réplica y potenciación tecnológicas, alcanza en ambas tendencias la cúspide de sus posibilidades casi exclusivamente en el dominio de lo mental. Allí donde los promotores de la integración exógena encuentran el mayor obstáculo y el más grande anhelo en la búsqueda de la imitación de lo humano, allí también los defensores de la vertiente endógena descubren la más fuerte inspiración en su afán de expansión biológica: la *mente humana*.

Así pues, la tendencia general en la producción de seres artificiales pareciera orientarse, en los últimos años, hacia el dominio exclusivo de lo mental. Conforme a las posibilidades que brindan las TIC de abordar el organismo humano en el más mínimo y microscópico detalle, cualidad de nuestra era que Paul Virilio dio en llamar «endocolonización»[3], el inexplorado terreno de lo mental comenzó a ser lentamente invadido por procesos y elementos de la tecnología que permiten extremar las aspiraciones de réplica e incremento artificiales. No contentos con los avances técnicos en la invasión e imitación del cuerpo, sus más entusiastas promotores se volcaron hacia el misterioso universo de la mente.[4]

Para las posturas exógenas, concentradas alrededor de la *inteligencia artificial dura*, la singularidad ocurrirá si de la tecnología computacional emergen la inteligencia y la conciencia artificiales; para las posiciones endógenas, agrupadas en torno al *poshumanismo trascendental*, la singularidad sobrevendrá si el contenido mental de un cerebro humano puede ser extraído y descargado enteramente en un entorno artificial. Descarga mental poshumana y emergencia artificial de la conciencia, se trata, en última instancia, de las caras opuestas de un mismo problema filosófico: si es factible replicar la mente en un cerebro artificial, sin dudas será posible descargar su contenido en un sustrato digital; esto es, si la mente no es más que un complejo sistema de información reducible a fórmulas matemáticas

[3] Virilio (1989).

[4] De alguna manera, la preferencia por el dominio abstracto de lo mental se vincula con el concepto propuesto por Tomás Maldonado acerca del *proceso de desmaterialización* propio de nuestra sociedad tecnológica, caracterizado por «un progresivo rebajamiento de la materialidad del mundo». Véase Maldonado (1994: 13).

computables, entonces tanto su descarga como su réplica en un soporte físico que emule la estructura de un cerebro humano deberían ser posibles.[5]

La cuestión teórica que está detrás de esta controversia se conoce en filosofía de la mente como el *problema mente-cuerpo*, y no son pocos los filósofos que han reflexionado en los últimos siglos sobre el tema, algunos de los cuales comienzan ahora a aplicarlo a la noción de singularidad. El advenimiento de una singularidad tecnológica, la emergencia del androide y del poshumano como posibilidades técnicas, está sujeto necesariamente a la cuestión fundamental acerca de si los fenómenos tecnológicos creados por el hombre son o no análogos a los fenómenos biológicos labrados por la naturaleza. En particular, a la pregunta acerca de si la mente humana, máxima expresión de la capacidad creativa de lo natural, puede ser reducida en su totalidad a fenómenos puramente físicos y mecánicos, que podrían por consiguiente recrearse, llegado el nivel de sofisticación apropiado, por medio de elementos computacionales tomados de la tecnología cultural.

4. Dualismo cartesiano y el problema mente-cuerpo

Como materia del pensamiento moderno, el problema mente-cuerpo hunde sus raíces en el *dualismo cartesiano*, que definió el marco conceptual predominante en el paradigma occidental desde el siglo XVII hasta nuestros días. El dualismo responde a un esquema binario, esto es, a un sistema de ideas o de pensamientos con dos valores, como la lógica, en la que los teoremas son válidos o inválidos; la epistemología, en la que las proposiciones son verdaderas o falsas; y la ética, en la que los individuos son buenos o malos y sus acciones, correctas o incorrectas. Asimismo, la doctrina de las dos verdades, la sagrada y la profana o la religiosa y la secular, es una respuesta dualista al conflicto entre religión y ciencia.[6]

El *dualismo cartesiano de la sustancia*, instalado en la tradición filosófica occidental por René Descartes en 1641, es la idea según la cual la realidad del ser humano consiste en dos partes separadas, dos órdenes del ser divididos por una brecha insorteable. El mundo material es una serie indefinida de variaciones en la forma, tamaño y movimiento de una materia homogénea, única y simple llamada *res extensa*. Se incluyen aquí a todos los eventos físicos y biológicos, incluso el complejo comportamiento animal, que Descartes consideraba como el resultado de procesos puramente mecánicos. En este sentido, el cuerpo

[5] Es de notar que en la base de esta comparación entre los procesos biológicos y los tecnológicos, se encuentra una analogía metafórica trazada entre la oposición mente/cerebro y el contraste *software*/*hardware*.
[6] Véase Ryle (1949).

humano es una sustancia extensa, está en el espacio, sujeto a leyes físicas y mecánicas, y sus procesos y estados pueden ser controlados por observadores externos. Sin embargo, detrás de la materia, se erige la *res cogitans*, el yo pensante, núcleo irreductible más allá de toda duda metódica:

> [D]el hecho mismo de que yo sé que existo, y de que advierto que ninguna otra cosa en absoluto atañe a mi naturaleza o a mi esencia, excepto el ser una cosa que piensa, concluyo con certeza que mi existencia radica únicamente en ser una cosa que piensa. Y aunque quizás [...] tengo un cuerpo que me está unido estrechamente, puesto que de una parte poseo un clara y distinta idea de mí mismo, en tanto que soy solo una cosa que piensa, e inextensa, y de otra parte una idea precisa de cuerpo, en tanto que es tan solo una cosa extensa y que no piensa, es manifiesto que yo soy distinto en realidad de mi cuerpo, y que puedo existir sin él.[7]

El dualismo de la sustancia propone, así, que a la par de la *res extensa*, que compone o constituye el universo material, hay una *res cogitans* que es independiente de la materia. La distinción entre sustancia pensante y sustancia extensa es absolutamente clara porque una se define por la exclusión de la otra: lo pensante no es extenso; lo extenso, no piensa. A raíz de ello, se forman dos órdenes del ser, separados por una brecha insorteable. He aquí el dualismo.

Ahora bien, ¿cómo es posible que mente y cuerpo, dos sustancias tan distintas, guarden estrechísimas relaciones? ¿Cómo puede ser que los estados de una sustancia no mecánica ni espacial puedan interactuar causalmente con estados de una sustancia mecánica que está en el espacio? ¿Cuál es la verdadera naturaleza de la conciencia, la inteligencia, el dolor y los demás procesos y estados mentales que emergieron misteriosamente en determinado punto de la evolución de las especies? Las respuestas a estas y a otras preguntas definen precisamente el objeto de estudio de la filosofía de la mente, concentrado alrededor del problema mente-cuerpo, y son estas mismas cuestiones las que se encuentran en la base conceptual de las posturas optimistas y pesimistas acerca de la llegada de una singularidad tecnológica.

Producto directo de la historia conceptual de la que nacen, los argumentos sostenidos por los pensadores de nuestra era acerca de la llegada de una singularidad, encarnados en la emergencia de la conciencia artificial (*inteligencia artificial dura*) y en la descarga del contenido de la conciencia (*poshumanismo trascendental*), encuentran sus raíces en un debate filosófico de enorme envergadura teórica, que ha dividido las aguas de la filosofía de

[7] Descartes (1985 [1641]: 71).

la mente en los últimos siglos. La cuestión elemental que está detrás de toda la discusión es si la mente humana, máxima manifestación de la capacidad organizativa del mundo natural, forma parte del mismo conjunto de fenómenos de lo corpóreo material y si puede, a raíz de ello, ser reducida a fenómenos físicos y mecánicos definibles por fórmulas matemáticas computables. La cuestión no es menor: si efectivamente puede reducirse a un algoritmo matemático claramente definido, entonces será solo cuestión de tiempo para que, llegado el nivel de sofisticación técnica necesaria, sea posible descargar o reproducir totalmente su contenido en un dispositivo electrónico de estructura análoga a la del cerebro humano en el que se ha desarrollado durante millones de años.

La búsqueda del *Algoritmo de Dios*,[8] aquel conjunto finito claramente definible por fórmulas matemáticas que, con reminiscencias bíblicas, Dios usó en la noche de los tiempos para crear la mente humana, se presenta, así, como el más grande anhelo y el más fuerte desafío al momento de intentar revivir el secreto máximo de la existencia en un sustrato artificial.

5. Inteligencia artificial dura

La *Prueba de Turing*, propuesta en un famoso artículo publicado en 1950 por el matemático y filósofo Alan Turing[9], uno de los padres de la ciencia de la computación y la informática moderna, consiste en un experimento destinado a comprobar si una computadora (una *máquina universal de Turing*[10]) puede pensar tal como lo hace un ser humano. La prueba consiste en un interrogatorio llevado a cabo por un humano que presenta una serie de preguntas a dos individuos ubicados detrás de sendas puertas, uno de los cuales es una máquina. Si el interrogador no puede distinguir cuál de los dos sujetos es el ser humano y cuál la máquina, entonces ésta ha pasado la prueba, y podemos afirmar de ella que tiene una capacidad de pensamiento análoga a la de un ser humano. Turing especulaba que para el año 2000 una computadora electrónica podría pasar aproximadamente un 30 por ciento de éxitos frente a un interrogador promedio durante un interrogatorio que durara cinco minutos.

[8] Originalmente, el término se utiliza para referirse a la mejor solución posible en juegos o acertijos complejos, como el cubo de Rubik. Nosotros hemos aplicado esta noción al problema mente-cuerpo.
[9] Véase Turing (1950).
[10] La *máquina universal de Turing*, ideada por Alan Turing en 1947, es un tipo especial de máquina que puede realizar el trabajo de otras máquinas, y se la considera como el prototipo del actual sistema operativo en un ordenador moderno.

El argumento más conocido contra la Prueba de Turing es el propuesto por el filósofo de la mente y del lenguaje John Searle[11], y promovido por el físico-matemático y también filósofo de la mente Roger Penrose[12], conocido como el *Experimento de la Habitación China*. Éste consiste en imaginar una habitación completamente aislada del exterior, excepto por una ranura por la que pueden entrar y salir textos escritos en chino. Un sujeto humano, que no conoce una sola palabra de chino, está dentro de la habitación y cuenta con una serie de instructivos y diccionarios que le definen las reglas sintácticas del alfabeto completo de dicho idioma (reglas del tipo *si entran tales caracteres, escribe tales otros*). El sujeto recibe preguntas en chino del exterior de la habitación y para responderlas usa los manuales que tiene a disposición, devolviéndole respuestas satisfactorias en dicho idioma, a pesar de no conocerlo. Pues bien, el argumento de Searle es que si un ser humano puede responder correctamente preguntas en chino solo conociendo las normas sintácticas de dicho idioma, entonces una máquina universal de Turing que disponga de tales reglas de *sintaxis* (el nivel de las reglas de combinatoria) podrá asimismo responderlas, sin que por ello debamos concluir que conoce efectivamente la *semántica* (el nivel del significado). Los postulados de la Prueba de Turing son, por lo tanto, falsos: la máquina universal no entiende chino, solo simula entenderlo.

Los experimentos de Turing y Searle representan, respectivamente, los dos polos de la inteligencia artificial, dividida de acuerdo al nivel de réplica de la mente que cada uno de ellos considera que podrán llegar a alcanzar las computadoras en los próximos años. Según la tesis de la *inteligencia artificial dura*, derivada de la Prueba de Turing, es teóricamente posible recrear, por medio de una máquina electrónica, una mente humana. Todas las cualidades de la mente (la inteligencia, el dolor, el placer, la conciencia, el libre albedrío, etc.) emergerán de modo natural en forma de *software* cuando el comportamiento algorítmico de la computación alcance determinado nivel de sofisticación en el dominio del *hardware*. Llegará el momento, sostienen sus defensores, en que a causa de la creciente complejidad de sistemas computacionales emerja la inteligencia; es todo cuestión de tiempo.

Antes bien: la contraparte moderada de esta tesis robusta se conoce como i*nteligencia artificial débil*, y deriva conceptualmente del Experimento de la Habitación China de Searle y Penrose. Esta tesis sostiene que, a pesar de que la tecnología computacional alcance un nivel de sofisticación sumamente elevado, las máquinas no podrán reproducir determinados comportamientos no mecánicos ni algorítmicos propios del

[11] Véase Searle (1985).
[12] Véase Penrose (1996).

cerebro y, por lo tanto, se limitarán a *simular* un comportamiento inteligente análogo al observable en los seres humanos, sin que por ello debamos adscribirles una conducta inteligente o consciente.

Para Penrose[13], acérrimo defensor de esta última posición, es en principio posible construir por medio de una computadora un modelo artificial que simule la acción del complejo sistema neuronal que tiene lugar en el cerebro. Sin embargo, constata, la tesis de la inteligencia artificial dura, que afirma que la simple aplicación de un algoritmo matemático puede provocar conocimiento consciente, es teóricamente insostenible. La idea de que toda actividad física, incluyendo la conciencia como fenómeno emergente de la operación cerebral, no es otra cosa que la activación de un enorme y complejo cómputo, y de que, por lo tanto, las ciencias computacionales podrán llegar a elaborar, llegado cierto punto de sofisticación, un algoritmo matemático complejo y de estructura probabilística cuya ejecución dé como resultado una conciencia artificial, es para Penrose una noción errada que no contempla el hecho central de que la conciencia no es, como suele creerse, un fenómeno producido *accidentalmente* por un cómputo complicado. Más bien, la impronta de la conciencia es una formación de juicios, por definición, no algorítmica. Siguiendo el Principio Antrópico –por el cual la naturaleza del universo en que estamos inmersos está fuertemente condicionada por la exigencia de que deben estar presentes seres sensibles como nosotros para observarla–, la conciencia, lejos de ser el resultado computable de un algoritmo matemático, se presenta como el fenómeno por excelencia en el que se hace conocida la misma existencia del cosmos. De ahí que, concluye Penrose, pensar a la conciencia como un proceso reducible a fórmulas precisas es una metáfora pobre que limita significativamente su carácter probabilístico no algorítmico ni computable.

Inteligencia artificial dura e inteligencia artificial débil, la Prueba de Turing y el Experimento de la Habitación China, se trata en suma, de uno u otro lado, de la cuestión acerca de si una fórmula algorítmica ejecutada por un programa apropiado en un sustrato digital dotado de suficiente poder de computación puede o no dar como resultado emergente la compleja naturaleza de la mente humana. Para ello, necesariamente, debe antes definirse si la mente es o no el producto de un proceso puramente material y mecánico, completamente explicable por leyes físicas y matemáticas.

[13] Véase Penrose (1996: 334-397).

6. Ascensión poshumana

Para los *poshumanistas trascendentales* como Hans Moravec (1995), Raymond Kurzweil (1999), Vernor Vinge (1993) y otros tantos *extropianos*, los seres humanos son objetos puramente físicos y mecánicos, y la actividad consciente no es otra cosa que el resultado de procesos completamente materiales que se pueden, o se podrán en un futuro cercano, reproducir por medios tecnológicos. La abstracción absoluta de la materia orgánica a través de una descarga o *transbiomorfosis* se logrará, consecuentemente, por medio de una traducción tecnológica de las redes neuronales de nuestra mente a la memoria de un ordenador concebido especialmente para tal fin, que será una réplica exacta, neurona por neurona, de la compleja estructura del cerebro natural.

Esta versión extrema del transhumanismo, basada en una premisa mecanicista de los procesos mentales, sostiene, así, que será posible extraer la *res cogitans* de la *res extensa*, transfiriendo a un sistema informático su contenido abstracto y trascendental, y desechando en ello el continente material inservible en que ha evolucionado. El ser líquido-fluido posbiológico resultante, pura conciencia emancipada de la sustancia material que imprimía limitaciones a su potencialidad, será tras este proceso autoprogramable, autoconfigurable, ilimitado y potencialmente inmortal, un *Übermensch* en estado puro.

Para el ya citado Raymond Kurzweil, la nanotecnología permitirá el diseño de *nanobots*, robots diseñados a un nivel molecular medido en nanómetros, como por ejemplo, los *respirocitos* (glóbulos rojos artificiales mecánicos), que invadirán el organismo y permitirán el desarrollo de modelos detallados del cerebro, acortando, de este modo, cada vez más la distancia entre humanos y dispositivos informáticos.[14] Así lo cree David Ross, acérrimo defensor de la noción de descarga poshumanista, quien propone invadir el cuerpo con nanomáquinas que ataquen a las neuronas y las reemplacen por programas (*software*) que cumplan las mismas funciones. El *infomorfo*, réplica informática del sujeto descargado, habitaría en un ciberespacio impecablemente fiel al sustrato físico del cuerpo, de forma tal que el universo sutil y complejo de todas las sensaciones y percepciones físicas tendrían su correlato indistinguible en el nuevo entorno virtual.

Ahora bien, formuladas las metáforas y las proyecciones, la pregunta surge necesariamente: ¿qué pasará con una mente descargada a un sustrato digital por medio de un proceso tecnológico técnicamente correcto? En la novela *El hombre de silicio*, de 1991, el

[14] Véase Kurzweil (2005).

autor de ciencia ficción Charles Platt, imagina precisamente este problema: «[l]a trepanación y la digitalización fueron correctas, su inteligencia está intacta. Solo que… no vuelve a la vida. El problema es que no sabemos aún qué es la conciencia»[15].

Para los autores que se oponen a la noción de ascensión poshumana, el sueño de la mente sin cuerpo hunde sus raíces en un desconocimiento de la naturaleza de la conciencia y del organismo humano como un todo. Las premisas del poshumanismo, sostienen sus opositores, parten de una confusión cartesiana profunda: un reduccionismo materialista y científico que, en aras de exaltar lo abstracto en detrimento de lo concreto, supone erróneamente que la mente puede vivir sin el cuerpo en el que se ha desarrollado a lo largo de un extenso y lento proceso de evolución natural ocurrido durante millones de años.

Los cerebros humanos, constata William Calvin, filósofo y profesor de Neurociencias y Biología evolutiva de la Universidad de Washington, son «las configuraciones de materia más elegantemente organizadas de todo el universo»[16]. La ciencia humana, aun a principios del siglo XXI, solo ha podido indagar superficialmente su compleja naturaleza gracias a la llegada reciente de tecnologías y dispositivos que permiten por primera vez invadirlo y *endocolonizarlo*. De ahí que el entendimiento alcanzado luego de menos de un siglo de investigación acerca de la naturaleza de sus procesos emergentes sea necesariamente débil y sesgado, y que las proyecciones construidas por los poshumanistas se apoyen en una comprensión pobre e insuficiente de su verdadera naturaleza.

Así lo cree también Erich Harth, profesor de Física de la Universidad de Siracusa, quien sostiene que la división entre mente y cuerpo es un sinsentido. La neurobiología y la conciencia están inextricablemente unidas, constata Harth, por lo cual la noción de *descarga* es teóricamente imposible:

> La información que intentamos transmitir es específica al cerebro en el cual se desarrolló [...]. [N]ecesitaríamos un sistema que, a diferencia de un ordenador no especializado, se adecuase a la información almacenada. Un equivalente del cerebro que no solo fuese genéticamente idéntico al cerebro original, sino que también contuviese la miríada de modificaciones aleatorias en su circuito que tienen lugar entre concepción y madurez. La información necesaria para especificar un sistema así es astronómica. Que incluso una pequeña porción de esa información pudiese extraerse de un cerebro vivo sin destruirlo, es dudoso.[17]

[15] Platt (1991: 232).
[16] Citado en Dery (1998).
[17] Ibídem, 348.

En este mismo orden, los biólogos y filósofos chilenos Francisco Varela y Humberto Maturana, que desarrollaron desde las ciencias cognitivas el concepto de *autopoiesis* como cualidad fundamental de los seres vivos, sostienen que la cognición y todos los procesos conscientes de la mente no pueden comprenderse si se los abstrae del cuerpo en que están encarnados. «[l]a mente no está en la cabeza», constata Varela para decir de modo simple que la cognición está *enactivamente encarnada:* «[l]a cognición no es la representación de un mundo predado por una mente predada, sino más bien la puesta en obra de un mundo y una mente a partir de una historia de la variedad de acciones que un ser realiza en el mundo»[18]. Ubicada en un espacio de «codeterminación entre lo interno y lo externo»[19], la mente ni existe ni no existe, no está en la cabeza ni separada de ella, ni separada del conjunto del cuerpo encarnado, es un resultado ontológicamente complejo que emerge enactivamente de estructuras encarnadas en la experiencia del ser con el mundo.

De alguna manera, este conjunto de reflexiones contrarias a la idea de ascensión poshumana es heredero de la fenomenología existencialista de Maurice Merleau-Ponty[20], quien formuló enormes aportes teóricos al estudio de la corporalidad y la percepción en un franco rechazo a los supuestos dualistas de la Modernidad. Para el filósofo francés, no hay conciencia a solas ni ser dicotómicamente escindido como propone el cartesianismo, sino que la conciencia es *conciencia encarnada*. El cuerpo es una tercera instancia en la relación entre el sujeto y el mundo: es la instancia de la mediación donde se experimentan y se perciben las cosas del mundo. La conciencia se hace presente a través de un cuerpo: es corporalizada y existencial, impregnada de *mundanidad*. Así lo confirma Jean François Lyotard, uno de los padres conceptuales de la Posmodernidad, para quien el pensamiento y la inteligencia están inextricablemente asociados a experiencias perceptivas y a las condiciones materiales de existencia propias del organismo físico, como el dolor y el sufrimiento.[21]

En suma, tanto si una mente puede despojarse efectivamente del organismo en que se originó para ser descargada a un sustrato virtual homólogo al cerebro humano, cuanto si por el contrario la naturaleza de la mente se enraíza indivisiblemente en el sujeto carnal del cual emerge, resultando por ello su escisión teóricamente imposible, de uno u otro modo, se trata en definitiva, lo mismo que en el campo de la inteligencia artificial, de la cuestión filosófica acerca de si la mente puede reducirse o no a un proceso físico y material que pueda

[18] Varela, Thomson y Rosch (1997, 157).
[19] Varela (2002: 242).
[20] Véase Merleau-Ponty (1970).
[21] Véase Lyotard (1998).

representarse por medio de un conjunto de reglas finitas. Si la conciencia es, en efecto, una conciencia encarnada, la supresión del cuerpo implicaría su muerte. Si por el contrario, la conciencia es un producto emergente de la complejidad estructural del cuerpo y ontológicamente distinta de éste, entonces será eventualmente factible su reproducción aislada del organismo en que se desarrolló.

7. Consideraciones finales

Toda especulación prospectiva toma, por lo general, una forma binaria que se divide entre utopías y distopías. En este sentido, la dicotomía trazada entre las posturas optimistas y las pesimistas construidas alrededor de la singularidad tecnológica remite, de algún modo, a la introducida por el semiólogo italiano Umberto Eco (1995) entre *apocalípticos* e *integrados*, más conocida como la oposición entre *tecnófobos* y *tecnófilos*.[22] Los primeros desconfían en mayor o menor grado de los avances tecnológicos, detectando en ellos *distopías apocalípticas* para el desarrollo de la humanidad. Los tecnófilos o integrados, en cambio, perciben el progreso de la tecnología como fundamentalmente beneficioso para el hombre y como el único medio que permitirá alcanzar un estadio de *utopía tecnológica*. Tanto unos como otros alientan al desarrollo de una singularidad, proyectando en el futuro inmediato, para bien o para mal, un cambio trascendental en todos los asuntos humanos.

¿Terminarán siendo, pues, las metáforas ficcionales de los tecnófilos y tecnófobos de hoy realidades científicas de mañana? Aunque así formulada la pregunta podría no tener aún una respuesta posible, con todo, las proyecciones de nuestra era acerca de la llegada de una singularidad no dejan de irrumpir en el seno de los discursos científicos, literarios y cinematográficos, emergiendo constantemente desde diversos soportes y medios de comunicación. En su conjunto, las posiciones, posturas y tonos discursivos heterogéneos se entroncan sobre una base común de pensamiento que gira en torno a la cuestión de si la realidad biológica desarrollada por el impulso de la naturaleza durante millones de años es o no análoga a la realidad tecnológica construida por obra del ser humano durante los últimos siglos.

En el punto extremo de la línea discursiva, las parábolas narrativas igualan a la mente humana, máxima demostración del esplendor creativo de lo natural, a un algoritmo matemático definido por reglas precisas, el *Algoritmo de Dios*, que será oportunamente

[22] Véase Eco (1995).

computable por un cerebro artificial, máxima expresión de la capacidad creativa del hombre. La materialidad de su organismo, la *res extensa*, despojada de la esencia de la vida alojada en la mente, la *res cogitans*, queda relegada a la de un desecho físico obsoleto, del cual la conciencia prescinde para subsistir. En el origen de los discursos de los defensores del poshumanismo trascendental y de la inteligencia artificial dura, se hace evidente así, a todas luces, la presencia extendida del dualismo cartesiano, que cubre la totalidad de las propuestas tecnoutópicas y tecnodistópicas a propósito de la llegada de una inminente singularidad tecnológica.

En este sentido, el carácter dualista del discurso de la ficción científica, que opone el cuerpo a la mente y reduce uno a otro, deriva indefectiblemente en un empobrecimiento de la esencia del mundo natural, inherentemente complejo y variado en sus manifestaciones. De esta manera lo considera, entre otros, el paleontólogo y biólogo evolucionista Stephen Jay Gould, quien destaca la importancia de lo contingente, lo imprevisible, lo caótico y lo aleatorio en la conformación de todas las formas de vida que existen y existieron en el Universo durante millones de años de evolución. «[L]a ciencia», sostiene Gould señalando los sistemas de explicación totalizadores del discurso científico, «solo se refiere al dominio "superior" de la generalidad; la región "inferior" de la contingencia no solo es pequeña sino que se la aplasta, presionada por toda la grandiosidad que tiene arriba; es solamente el lugar de los detalles curiosos y sin importancia para el funcionamiento de la naturaleza»[23].

Así lo entiende también Fritjof Capra, doctor en Física teórica por la Universidad de Viena, quien propone una visión integral y holística de los organismos vivos, en oposición a la perspectiva cartesiana mecanicista y reduccionista de la ciencia clásica. De acuerdo con el nuevo paradigma, los sistemas vivientes son sistemas abiertos, plásticos, flexibles y adaptativos, capaces de autoorganización, autorrenovación, autoconservación y autotrascendencia, dotados de una naturaleza intrínsecamente dinámica caracterizada por un proceso de «interacción simultánea y recíprocamente dependiente de todos sus componentes múltiples»[24].

La analogía metafórica entre biología y tecnología, que facilita la elaboración de productos mecánicos y simplifica a un tiempo la comprensión de la naturaleza humana, conduce, pues, a una inevitable reducción del nivel de complejidad asociado a los sistemas vivientes. Basadas en esta metáfora reduccionista, las ficciones de nuestra era proyectan mundos poblados de androides y poshumanos, seres artificiales que anularían la distancia

[23] Citado en Sibilia (2005: 148).
[24] Capra (1992: 309).

fundamental entre naturaleza y artificio, y que arrastran irremediablemente desde el origen una concepción errada y limitada de las entidades que pretenden simular.

Más allá de todo, sin embargo, queda una sombra de duda acerca de si los productos culturales podrán abarcar en algún futuro probable la compleja constitución de la realidad biológica, y si será, en efecto, posible construir un dispositivo de tal condición y cualidad que sea capaz de comprender la enorme diversidad orgánica labrada por procesos naturales durante millones de años. La plataforma ficcional de nuestra era ofrece, a este respecto, mundos discursivos teóricamente posibles en condiciones técnicamente probables, que abren una y otra vez, en el entramado social, la pregunta constante por el destino de la humanidad abandonada al arbitrio de la tecnología.

8. Referencias

CAPRA, FRITJOF. (1992). *El punto crucial: ciencia, sociedad y cultura naciente.* Buenos Aires: Editorial Troquel.

DARWIN, CHARLES. (1998). *El origen de las especies.* Madrid: Editorial Debate.

DERY, MARK. (1998). *Velocidad de escape: la cibercultura en el final del siglo.* Madrid: Siruela.

DESCARTES, RENÉ. (1981 [1641]). *Meditaciones metafísicas:* l*as pasiones del alma.* Madrid: Ediciones Orbis.

DICK, PHILIP K. (1981). My definition of science fiction. En Lawrence Sutin, (Ed.). (1995). *The shifting realities of Philip K. Dick: Selected literary and philosophical writings*. Nueva York: Vintage Books.

ECO, UMBERTO. (1995). *Apocalípticos e integrados.* Barcelona: Tusquets Editores.

KURZWEIL, RAYMOND. (1999). *La era de las máquinas emocionales: cuando las computadoras superen la mente humana.* Barcelona: Planeta.

KURZWEIL, RAYMOND. (2005). *The singularity is near: When humans transcend biology*. Nueva York: Viking.

LYOTARD, JEAN-FRANCOIS. (1998). *Lo inhumano: charlas sobre el tiempo.* Editorial Manantial: Buenos Aires.

MATURANA, HUMBERTO y FRANCISCO VARELA. (1995). *De máquinas y seres vivos.* Santiago de Chile: Editorial Universitaria.

MERLEAU-PONTY, MAURICE. (1970). *Elogio de la filosofía*. Buenos Aires: Ediciones Nueva Visión.

MORAVEC, HANS. (1995). *Mind children: The future of robot and human intellingence.* Cambridge: Harvard University Press.

PLATT, CHARLES. (1991). *The Silicon Man*. Nueva York: Bantam Spectra.

PENROSE, ROGER. (1996). *La mente nueva del emperador: en torno a la cibernética, la mente y las leyes de la física.* México: Fondo de Cultura Económica.

RYLE, GILBERT. (1949). *The concept of mind*. Chicago: Chicago Press edition.

SEPÚLVEDA, LUZ MARÍA. (2007). *La utopía de los seres posthumanos.* México: Fondo de Cultura Económica.

SIBILIA, PAULA. (2005). *El hombre postorgánico: cuerpo, subjetividad y tecnologías digitales*. Buenos Aires: Fondo de Cultura Económica.

TURING, ALAN. (1950). Computing machinery and intelligence. Publicado en *Mind*, núm. 59, págs. 433-460.

VARELA, FRANCISCO. (2002). *El fenómeno de la vida.* Santiago de Chile: Dolmen Ediciones.

VARELA FRANCISCO, EVAN THOMPSON y ELEANOR ROSCH. (1997). *De cuerpo presente*. Barcelona: Gedisa.

VINGE, VERNOR. (1993). The coming technological singularity: How to survive in the post-Human era". Presentado en el Simposio VISION-21, patrocinado por el Centro de Investigación NASA Lewis y el Instituto Aeroespacial de Ohio, 30 y 31 de Marzo.

VIRILIO, PAUL. (1989). *La máquina de visión*. Madrid: Ediciones Cátedra.

WIENER, NORBERT. (1988). *Cibernética y sociedad.* Buenos Aires: Editorial Sudamericana.