

Inteligencias artificiales ¿Son posibles y deseables?

Manuel Carabantes

Introducción

«Me llamo Robinette Broadhead, pese a lo cual soy varón. A mi analista (a quien doy el nombre de Sigfrid von Schrink, aunque no se llama así; carece de nombre por ser una máquina) le hace mucha gracia este hecho. [...] –Rob, hoy no estás cooperando mucho –dice Sigfrid a través del pequeño altavoz que hay en el extremo superior de la alfombra. A veces utiliza un muñeco de aspecto muy real, que está sentado en un sillón, da golpecitos con un lápiz y me dedica una rápida sonrisa de vez en cuando. [...] A veces intento esto con él, diciendo alguna verdad dolorosa con el tono de quien pide otro ponche de ron al camarero de una fiesta. Lo hago cuando quiero esquivar su ataque. No creo que surja efecto. Sigfrid tiene muchos circuitos Heechee en su interior. Es mucho mejor que las máquinas del instituto al que me enviaron durante mi episodio. Observa continuamente todos mis parámetros físicos: conductividad cutánea, pulso, actitud de ondas beta, en fin, de todo. Obtiene indicaciones de las correas que me sujetan sobre la alfombra, acerca de la violencia con que me retuerzo. Mide el volumen de mi voz y lee sus matices en el espectro. Y también conoce el significado de las palabras. [...] Pero Sigfrid no es real. Es una máquina. No puede sentir el dolor. Así pues, ¿adónde *van* a todo ese dolor y ese cieno? Trato de explicarle todo esto, y acabo diciendo: –¿No lo entiendes, Sigfrid? Yo te traspaso mis problemas y tú los traspasas a alguien más, así que tienen que desembocar en algún sitio. No me parece real que desemboquen en forma de burbujas magnéticas en una pieza de cuarzo que nadie *sienta* jamás»¹.

El escritor de ciencia ficción Frederik Pohl fantasea en estos fragmentos de su novela *Pórtico* con la posibilidad de que la ciencia nos brinde en el futuro máquinas nuevas. Pero no como las que conocemos hoy, que, a lo sumo, son capaces de poner tornillos en una cadena de montaje, sino verdaderamente inteligentes. Poner tornillos es una tarea sencilla para cualquier ser humano normal –e incluso con cierta discapacidad psíquica–. Sin embargo, Sigfrid va mucho más allá, porque es capaz de realizar el trabajo del universitario con más años de estudios: un médico. En particular un médico psiquiatra. Sigfrid, por tanto, no es una máquina destinada a reemplazar la fuerza humana –como hace el brazo robótico que pone tornillos–, sino a reemplazar la mente humana. Es lo que denominamos una IA (inteligencia artificial).

Dentro de las investigaciones en IA podemos distinguir dos corrientes. Por un lado, la IA fuerte, que pretende crear mentes artificiales como Sigfrid. Y por otro, la IA débil, cuyo propósito es más modesto, puesto que se conforma con crear modelos informáticos que

¹ Pohl, Frederik, *Pórtico* (Barcelona, Planeta, 2006), varios capítulos.

resulten útiles para el estudio de la mente. En la actualidad ambas coexisten proporcionándose beneficios mutuos. Así, por ejemplo, los avances en el conocimiento de la mente humana obtenidos por el programa débil de la IA, sirven a los científicos que trabajan en el programa fuerte para intentar construir sus mentes artificiales.

En el presente ensayo nos ocuparemos de describir y discutir sólo el programa fuerte de la IA, en tanto que nos parece el más problemático y atractivo de los dos. Lo primero será preguntarnos si es posible crear mentes artificiales. ¿Puede la ciencia en nuestros días ofrecernos tan rico fruto de su árbol, o es sólo una veleidad que deseamos inductivamente tras haber visto lo rápido que han crecido sus ramas durante el último siglo? Una vez tengamos una respuesta a este interrogante, lo que procederá es reflexionar sobre la conveniencia de tales inventos. ¿Qué ventajas y peligros se despliegan ante nosotros si la ciencia, efectivamente, consigue crear IAs en sentido fuerte?

Para estas tareas utilizaremos diversas fuentes, desde un matemático como Alan Turing, hasta ingenieros informáticos como Joseph Weizenbaum y Jeff Hawkins, o un psicólogo como Howard Gardner, y filósofos como John Searle y el profesor Javier Bustamante. Las investigaciones en IA son multidisciplinarias, y ese hecho ha de quedar reflejado en nuestra bibliografía.

¿Qué es un ordenador?

Alan Turing, un matemático inglés de la primera mitad del siglo XX, es considerado unánimemente como el precursor de la informática moderna. Debe su fama a haber inventado la máquina que lleva su nombre, la «máquina de Turing». Ésta consiste en una cinta magnética infinita –es decir, tan larga como sea necesario– y un cabezal que se desplaza a lo largo de la cinta leyendo y escribiendo valores (1 y 0) en función de los valores leídos anteriormente. La denominada «máquina de Turing universal» es un tipo especial de máquina de Turing en todo idéntica a la descrita, excepto en que la cinta magnética está en blanco y, por tanto, la máquina puede ejecutar cualquier programa que sea escrito en su cinta. Joseph Weizenbaum la describe así:

«Existe una máquina de Turing U (realmente una clase completa de máquinas) cuyo alfabeto consta de dos símbolos “0” y “1”, de tal modo que, dados cualquier procedimiento escrito en cualquier lenguaje preciso y no ambiguo y una máquina de Turing L , que formalice las reglas de transformación de ese lenguaje, la máquina de Turing U puede imitar a la máquina de Turing L en la ejecución de L de ese procedimiento»².

² Weizenbaum, Joseph, *La frontera entre el ordenador y la mente* (Madrid, Pirámide, 1976), p. 60.

Y cuando Turing dice que su máquina universal puede ejecutar *cualquier* programa, también está pensando que se puede escribir un programa para ejecutar *cualquier* procedimiento que seamos capaces de definir mediante un conjunto de reglas completas y consistentes³. Por ejemplo, poner tornillos. Es fácil pensar en todas las acciones que un programa debe ordenarle a un brazo mecánico para colocar tornillos en una cadena de montaje con el mismo grado de perfección –o incluso superior– con que lo haría un obrero. Como todos sabemos, esos brazos mecánicos enormes que hay en las fábricas de automóviles llevan a cabo las acciones necesarias para poner tornillos sin requerir la participación externa de un operario humano. Este tipo de procedimiento, en el cual las reglas determinan el siguiente estado sin participación externa, se denomina *procedimiento efectivo* o *algoritmo*⁴.

Así pues, estamos de acuerdo en que los ordenadores modernos, que a grandes rasgos no son más que versiones evolucionadas de la máquina de Turing universal⁵, pueden efectuar cualquier procedimiento efectivo. La tosca faena de poner tornillos es un procedimiento efectivo, pero ¿también lo es la íntima y delicada tarea de un psiquiatra? Los defensores del programa fuerte de la IA, entre los que se incluye Turing, están convencidos de que la respuesta es afirmativa. El fundamento de su convicción puede ser ilustrado con un juego.

El juego consiste en invitar a un detective experto en psiquiatría a ser tratado por dos psicoterapeutas distintos, A y B. Uno es un psiquiatra humano, y el otro es Sigfrid. Nuestro detective se comunicará con ellos de forma que no pueda apreciar las diferencias físicas obvias entre ambos, puesto que aquí lo relevante es el ámbito de lo mental. ¿Será capaz el detective de distinguir quién es la máquina? Turing propone llamar a esto el «juego de imitación»⁶, aunque nosotros lo conocemos hoy en día como el «test de Turing». Su formulación, en palabras de John Searle, es la siguiente: «Si un ordenador puede actuar de modo tal que un experto sea incapaz de distinguir la actuación del ordenador de la de un humano provisto de cierta facultad cognitiva –practicar la psicoterapia en el caso que estamos proponiendo–, entonces el ordenador posee también esa facultad»⁷. Es decir, que, según Turing, si una máquina *parece* pensante, entonces *es* pensante. Una transición del mundo de las apariencias al mundo del ser que es totalmente insostenible, pues ¿acaso no *parecen*

³ *Ibíd.*, p. 46.

⁴ *Ibíd.*, p. 48.

⁵ Omitimos la consideración de las denominadas redes PDP (Procesamiento Distribuido en Paralelo) por dos motivos. El primero, porque son máquinas todavía experimentales. Y el segundo, porque cualquier tarea que ellas puedan realizar es realizable por una computadora de procesamiento serial, es decir, una computadora corriente.

⁶ Turing, Alan, *¿Puede pensar una máquina?* (Valencia, Teorema, 1974), p. 12.

⁷ Searle, John, *Mentes, cerebros y ciencia* (Madrid, Cátedra, 1990).

pardos todos los gatos por la noche pero por el día, con la luz del sol, descubrimos que *son* de diversos colores? Asimismo, ¿no podría suceder que una computadora, como Sigfrid, dé la apariencia de estar pensando pero en realidad no sea pensante? Ante esta objeción, Turing responde:

«Este argumento parece ser la negación de la validez de nuestro test. Según la formulación más extrema de tal punto de vista, la única manera con la cual uno podría estar seguro de que una máquina piensa, consistiría en *ser* la máquina y sentirse pensar uno mismo. [...] Éste es, de hecho, el punto de vista solipsista. Puede que sea el punto de vista más lógico de mantener, pero hace difícil la comunicación de las ideas. [...] En lugar de debatir continuamente sobre este punto, suele mantenerse el gentil convenio de que todo el mundo piensa»⁸.

La opinión de Turing obedece al principio metodológico conductista según el cual sólo lo empíricamente observable puede ser objeto de la ciencia. Puesto que en el juego de imitación lo único que podemos observar de Sigfrid y del psiquiatra humano son sus conductas, entonces la conducta sería el único factor que decidiera científicamente quién es humano.

Obviamente, tal y como nos dice el sentido común, el test de Turing es un criterio falso. Su refutación más contundente es obra del filósofo americano John Searle. Éste señala que los programas informáticos son pura *sintaxis*, es decir, que consisten en conjuntos de reglas para manipular símbolos (1 y 0) detrás de los cuales no hay ningún contenido *semántico*. Supongamos que a la pregunta del detective «¿Usted cree que mi madre me odia?» Sigfrid respondiera «No, no creo que le odie. Las madres aman a sus hijos». Esta conducta podría perfectamente ser la misma que diera el psiquiatra humano, y por tanto Sigfrid pasaría con éxito el test de Turing. Sin embargo, el ordenador, dice Searle, no puede pensar en nada cuando enuncia su respuesta. Porque el ordenador opera con palabras que para él no tienen significado, pues en el fondo se reducen a series de 1 y 0. Mientras que para el ser humano las palabras como «amor» sí refieren a contenidos semánticos significativos (el escalofrío de aquel primer beso), a pesar de que sus neuronas funcionen por impulsos eléctricos semejables a 1 y 0. Podríamos introducir en la base de datos de Sigfrid cientos de películas románticas, o las obras completas de Stendhal, que nada cambiaría. Frases como «En amor no se goza sino de la ilusión que uno mismo se forja»⁹ no serían más que sucesiones de símbolos sin significado para él. Ahora bien, Searle admite que distinto sería el caso si la computadora en cuestión, en vez de tener la arquitectura similar a la máquina de Turing que tienen los ordenadores actuales, fuese una réplica del organismo humano que le permitiera adquirir contenidos semánticos mediante la acumulación de experiencia a lo largo de una vida inmersa

⁸ Turing, Alan, *¿Puede pensar una máquina?*, loc. cit., pp. 39 y 40.

⁹ Stendhal, *Del amor* (Madrid, Alianza, 2003), p. 110.

en un contexto cultural, tal y como hacemos los hombres. Pero esta empresa, la de un robot androide, no parece que vaya a estar al alcance de la neurofisiología en los próximos siglos, dice Weizenbaum¹⁰.

En conclusión, hay actividades superiores de la mente humana como la psicoterapia que *quizás* sean expresables en forma de procedimientos efectivos. Pero para su *buen* ejercicio requieren algo más que no está al alcance de los ordenadores: contenidos semánticos. Ciertamente, desde el punto de vista de la razón instrumental el único criterio para decidir quién es un *buen* psicoterapeuta es el resultado; un buen psicoterapeuta es aquel que cura con éxito a sus pacientes, aunque no comprenda lo que éstos quieren decirle cuando le hablan del amor o de cualquier otro asunto. En la segunda parte abordaremos las monstruosas consecuencias que vive nuestra sociedad por culpa de la aplicación hegemónica de la racionalidad instrumental. Por ahora, hemos alcanzado la conclusión de que una computadora no puede realizar cualquier tarea *igual* que lo haría un humano.

¿Qué es la inteligencia?

Ya sabemos que la inteligencia no es la conducta inteligente. Pues, cuando leemos un libro, ¿quién puede saber desde afuera si lo estamos entendiendo? El psicólogo Howard Gardner, exponente mundial en la materia, define la inteligencia como

«Un potencial psicobiológico para resolver problemas y para generar resultados que sean apreciados al menos en un determinado contexto cultural»¹¹.

Los seres humanos afrontamos diversos tipos de problemas. En base a este hecho y al estudio de sujetos intelectualmente raros –como los afectados por el *síndrome de savant*–, Gardner distingue en la última revisión de su teoría de las inteligencias múltiples un total de nueve inteligencias:

«1, Lingüística: Dominio y amor por el lenguaje y las palabras, junto al deseo de explorarlos. 2, Lógico-matemática: Confrontación y valoración de objetos, abstrayendo sus relaciones y principios subyacentes. 3, Musical: Capacidad no sólo de componer e interpretar piezas con tono, ritmo y timbre, sino también de escuchar y de juzgar. Puede estar relacionada con otras inteligencias, como la lingüística, la espacial o la corporal-cinética. 4, Espacial: Habilidad para percibir el mundo visual con precisión, para transformar y modificar lo percibido y para recrear experiencias visuales incluso en ausencia de estímulos físicos. 5, Corporal-cinética: Dominio y orquestación de los movimientos del cuerpo. Manipulación hábil de objetos. 6 y 7, Inteligencias personales: Determinar con precisión el estado de humor, los sentimientos y otros estados mentales de uno mismo (inteligencia intrapersonal) y de los otros (interpersonal), utilizando esta información como guía de conducta. 8, Naturalista: Identificación y caracterización de objetos naturales.

¹⁰ Weizenbaum, Joseph, *La frontera entre el ordenador y la mente*, loc. cit., p. 177.

¹¹ Gardner, Howard, *Inteligencias múltiples. La teoría en la práctica* (Barcelona, Paidós, 1999).

9, Existencial: Captación y reflexión sobre cuestiones fundamentales de la existencia»¹².

Los ordenadores actuales destacan por su inteligencia lógico-matemática, e incluso por la musical, y su inteligencia espacial (4 y 8) va en aumento gracias al desarrollo de dispositivos como células fotoeléctricas, infrarrojos y GPS. Por desgracia, son poco o nada competentes en las demás. Su comprensión del lenguaje natural es imposible (1, 7, 8 y 9), como señala Searle; y subir unas simples escaleras es una gesta épica para un robot (5).

A principios del siglo XX los psicólogos pensaban que sólo había una inteligencia única, encargada de resolver todos los problemas, ya fueran matemáticos o existenciales. En tal contexto una computadora tendría que ser *multi-purpose* –en el sentido más amplio– para poder ser calificada como inteligente. Es decir, que tendría que ser capaz de resolver de forma mínimamente competente cualquier tipo de problema, ya fuera matemático o existencial. Obviamente, eso es imposible, como ya hemos demostrado. En cambio, dentro del vigente marco de las inteligencias múltiples, elaborado por Howard Gardner y otros como L. Thurstone y J. Guilford, sí es posible conceder el calificativo de inteligente a una computadora, pues sólo tiene que demostrar su competencia en alguno de los nueve campos arriba distinguidos. Por tanto, la posibilidad de crear IAs depende no sólo de la arquitectura de los ordenadores y de la pericia de los programadores para aprehender las operaciones de la mente humana en forma de procedimientos efectivos, sino también de la noción de inteligencia que se considere verdadera.

¿Cómo sería una IA en sentido fuerte?

El ingeniero informático Jeff Hawkins se adscribe a la teoría de las inteligencias múltiples. Su propósito, dice, no es

«construir humanos. Quiero entender la inteligencia y construir máquinas inteligentes. Ser humano y ser inteligente son asuntos separados. [...] Si se desea construir máquinas inteligentes que se comporten como humanos –es decir, que pasen el test de Turing en todos sus aspectos– es probable que se tenga que recrear buena parte de la restante composición que hace a los humanos como son. Pero [...] para construir máquinas que sean inteligentes de verdad, pero no exactas a los humanos, podemos centrarnos en la parte del cerebro estrictamente relacionada con la inteligencia»¹³.

Esa parte del cerebro es la corteza. Allí reside casi toda la inteligencia. Comprenderla es el camino para fabricar cortezas artificiales que funcionen como auténticas IAs.

La corteza se organiza de modo jerárquico y bidireccional, dice Hawkins. Jerárquico porque se divide en regiones conectadas piramidalmente. Abajo las regiones encargadas de

¹² *Ibíd.*

¹³ Hawkins, Jeff, *Sobre la inteligencia* (Madrid, Espasa, 2005), p. 56.

procesar la información sensorial, que es la más básica. A medida que ascendemos, las regiones se ocupan de tareas cada vez más abstractas, hasta llegar a las áreas de asociación, que integran la información proveniente de los cinco sentidos para formar eso que, heideggerianamente, llamamos la sensación de «estar ahí». Esta jerarquía es bidireccional porque la información no sólo tiene una dirección ascendente, sino también descendente. La información ascendente es la que nos llega del mundo exterior, y sirve para formar memorias, a todos los niveles, de lo que experimentamos. La descendente sirve para elaborar predicciones del mundo exterior sobre la base de las memorias adquiridas. Hawkins denomina a este modelo suyo de la inteligencia el «modelo de memoria-predicción»¹⁴. Un flujo continuo de información está llegando constantemente a nuestro cerebro, hagamos lo que hagamos –y aún cuando no hacemos nada–. Pongamos por caso, el acto de subir unas escaleras. Lo hacemos sin pensar, de manera automática. Esto quiere decir, que lo hacemos sin participación de las regiones corticales superiores. Las inferiores pueden encargarse de una tarea tan simple. La cosa cambia cuando sucede algo inesperado, como por ejemplo un tropiezo. En ese caso, el patrón inesperado

«continuará propagándose hacia arriba de la jerarquía cortical hasta que alguna región superior pueda interpretarlo como parte de su secuencia de hecho normal. Cuanto más necesite ascender el patrón inesperado, más regiones de la corteza cerebral participan en la resolución de la entrada inesperada»¹⁵.

El modelo de memoria-predicción de Hawkins, como vemos, sería muy útil para mejorar inteligencias en las que los ordenadores todavía no son diestros, como la cinestésica. Y también la lingüística, pues en una conversación predecimos constantemente lo que va a decir nuestro interlocutor. Eso nos permite entenderle aún cuando un intenso ruido de fondo no nos deje escuchar cada palabra, pues las que faltan las deduce predictivamente nuestro cerebro. Weizenbaum está de acuerdo en que este fenómeno predictivo, «tanto a nivel sintáctico como a niveles contextuales más amplios»¹⁶, es una de las claves de la comprensión lingüística.

Las dificultades técnicas principales para realizar el proyecto de Hawkins de crear cortezas artificiales son tres. Por un lado, la plasticidad de la memoria. La corteza cerebral tiene 32 billones de sinapsis, donde se almacena la memoria. En cuanto a capacidad, el disco duro de un ordenador puede dar la talla. Pero el cerebro conserva sus recuerdos a pesar de la muerte diaria de miles de neuronas, mientras que para un ordenador la más mínima pérdida suele resultar fatal para su correcto funcionamiento. El segundo problema estriba en la conectividad. La materia blanca que conecta las regiones de la corteza cerebral contiene un

¹⁴ *Ibíd.*, p. 15.

¹⁵ *Ibíd.*, p. 186.

¹⁶ Weizenbaum, Joseph, *La frontera entre el ordenador y la mente*, loc. cit., p. 160.

astronómico número de conexiones imposible de imitar mediante silicio. La solución en este caso consistiría en implementar axones compartidos a la corteza artificial, tal y como compartimos un mismo hilo telefónico entre todos los vecinos de un bloque. Y el tercer problema, que es el más importante, es que la corteza cerebral humana no funciona por sí sola. Es el lugar *principal* donde reside la inteligencia, igual que el motor es la pieza principal para el movimiento de un coche. Pero sin unas bujías con las que arrancar el motor, el coche no se moverá ni un centímetro. Asimismo, el hipocampo ocupa un lugar imprescindible en la cima de la jerarquía cortical para la formación de nuevas memorias, aunque éstas se almacenen en las sinapsis de la corteza. Y el tálamo es la vía intermedia para la retroalimentación demorada entre regiones corticales. Construir una corteza cerebral en un laboratorio se antoja difícil, aunque posible en tanto que conocemos bien su estructura y funcionamiento. Sin embargo, recrear un tálamo o un hipocampo es pura ciencia ficción a día de hoy.

El proyecto de Hawkins ha de superar una última objeción. Pero no de carácter técnico, sino ético. Supongamos que fuera posible construir una corteza cerebral que diera lugar a una IA, sin tálamo ni hipocampo ni ninguna parte del cerebro viejo. Es posible que esta entidad, a diferencia de los ordenadores actuales –basados en la máquina de Turing–, tuviera representaciones semánticas, y por tanto fuera pensante. Hasta tendría conciencia¹⁷, según Hawkins. Pues bien, cuando esta entidad pensase en «amor» vendrían a su conciencia recuerdos relacionados con ese término: imágenes, sonidos y otras memorias provenientes de sentidos como, por ejemplo, el radar y la visión infrarroja, que nosotros ni siquiera podemos concebir porque no los tenemos. ¿Podría entonces esta entidad IA ejercer la psiquiatría sin ninguna objeción ética?

La respuesta es no. Tendrá contenidos semánticos. Pero sin tálamo, sin hipocampo, sin un rostro, sin manos de cinco dedos, y sin ser tratado como un hombre por otros hombres en sociedad; sin todo esto y mucho más, la respuesta es no. Nosotros no podemos concebir cómo sería «ver» con un radar. De forma análoga, la IA descrita no podría «ver» el mundo emocionalmente, porque carecería de sistema límbico –formado por el tálamo y el hipocampo, entre otros–, que es el *principal* artífice de nuestra vida afectiva. Sus recuerdos serían apáticos, inhumanos. Y aunque siguiéramos siendo generosos con el progreso de la ciencia, la fabricación de un sistema límbico en laboratorio tampoco sería suficiente, pues el sistema límbico es la parte *principal*, pero no la única necesaria para nuestra vida afectiva –así

¹⁷ «La conciencia es lo que se siente al tener corteza cerebral». Jeff Hawkins, *Sobre la inteligencia*, loc. cit., p. 226.

como señalamos antes que la corteza es la parte principal, pero no la única necesaria para la inteligencia—. De esta manera, tirando del hilo, pronto nos daríamos cuenta de que el ser humano es un todo. Como el todo es más que la suma de las partes¹⁸, ninguna de sus partes puede bastar para sustituirle en aquellas tareas que requieren de él en su integridad. Un brazo mecánico sin cuerpo es suficiente para poner tornillos a un automóvil. Pero para poner un tornillo a un hombre, hace falta un psiquiatra entero.

«Sostengo que el ser humano individual, como cualquier otro organismo, se define por los problemas que afronta. [...] Ningún otro organismo, y por supuesto ningún ordenador, está dotado para confrontar problemas auténticamente humanos en términos humanos. Y puesto que el dominio de la inteligencia del hombre está determinado (salvo para un pequeño conjunto de problemas formales) por la humanidad de éste, cualquier otra inteligencia, por grande que sea, será extraña al dominio del hombre»¹⁹.

¿Cómo es nuestro mundo?

Uno de los grandes hitos de la modernidad fue la creación del Estado: una nueva entidad política constituida sobre la identidad cultural de los habitantes que comparten un espacio. Nuestro mundo actual se caracteriza justamente por la liquidación de los Estados modernos. Vivimos en un mundo globalizado, y la globalización es, en palabras de Ulrich Beck,

«la ampliación de la política más allá de la vieja categoría de Estado nacional»²⁰.

O dicho de otro modo, la globalización es

«el conjunto de acciones mediante las cuales los Estados nacionales participan del punto de vista que admite que los espacios cerrados ya no existen»²¹.

Esta transformación era inevitable desde el momento en que surgió el capitalismo, pues, como dice Immanuel Wallerstein,

«el capitalismo es, dada su propia lógica interna, necesariamente global»²².

Un empresario textil, por ejemplo, comienza fabricando y vendiendo sus productos a nivel local. Luego, nacional. Y cuando el mercado nacional se le queda pequeño, da el salto al internacional. Surge entonces el problema de dirigir un negocio cuyas partes se encuentran separadas por grandes distancias. La ubicación de cada una se ha elegido buscando reducir los costes de producción al mínimo posible. Así, las plantaciones de algodón pueden estar en África, los telares en Paquistán, y los almacenes repartidos por el mundo entero. ¿Cómo *controlar* un negocio repartido por una docena de franjas horarias diferentes? Cuando el empresario comenzó con un pequeño taller, podía controlar el negocio personándose allí todas

¹⁸ Aristóteles, *Metafísica*, 1045 a10.

¹⁹ Weizenbaum, Joseph, *La frontera entre el ordenador y la mente*, loc. cit., p. 184.

²⁰ Beck, Ulrich, *¿Qué es la globalización?* (Barcelona, Paidós, 2008), p. 16.

²¹ *Ibid.*, p. 29.

²² *Ibid.*, p. 58.

las mañanas. La aparición de las grandes empresas multinacionales obligó a desarrollar una nueva forma de control: la burocracia, que literalmente significa «el poder de la oficina». Legiones de trabajadores fueron encerrados en cubículos para controlar el negocio haciendo «papeleo». Cada eslabón obedecía las órdenes de los eslabones superiores. Y, para evitar que algún obrero cambiase las órdenes por iniciativa propia, éstas debían ir firmadas en un papel.

Paralelamente, los Estados crecieron con las empresas. La población aumentaba a causa del creciente nivel material de vida. Y la amenaza del comunismo obligaba a controlar a esa población creciente por dos motivos. Uno, para neutralizar a los elementos subversivos. Y dos, para distribuir atenciones sociales que aumentarían la calidad de vida hasta el mínimo necesario para evitar que las contradicciones de clase alcanzaran el nivel crítico. La solución fue también la implantación de la burocracia.

A mediados del siglo XX, los Estados Unidos tenían 150 millones de habitantes. Habían doblado su población en menos de 50 años. El tamaño de esa sociedad era ya difícilmente controlable por la burocracia y sus oficinistas.

«Tareas de cálculo sin precedente aguardaban a la sociedad americana al final de la Segunda Guerra Mundial, y, el ordenador, casi milagrosamente, llegaría para hacerse cargo de ellos»²³.

Como señala Javier Bustamante, el computador surge a imagen de una sociedad burocrática y para potenciarla en su desarrollo²⁴.

Por tanto, vemos que la sociedad globalizada y burocratizada en la que vivimos hoy no es producto de la computadora, pero sí es cierto que la computadora evitó que reventara como una burbuja que es demasiado grande para seguir existiendo. La computadora «enjabonó» la burbuja. Sus cualidades, como la velocidad y su enorme capacidad de cálculo, la hacen imprescindible para ciertas tareas que se han implantado en nuestra sociedad. La alternativa al uso de las computadoras para actividades como, por ejemplo, la asistencia social habría sido la descentralización de los servicios.

«Pero el ordenador se utilizó para automatizar la administración de los servicios sociales y centralizarla a lo largo de líneas políticas establecidas. [...] El ordenador, pues, se utilizó para conservar las instituciones políticas y sociales de los Estados Unidos, afianzándolos e inmunizándolos (al menos, temporalmente) de las enormes presiones del cambio»²⁵.

Es frecuente oír hablar de la «revolución del ordenador». Pero si una revolución ha de medirse por la profundidad de las revisiones sociales que entraña, entonces realmente no ha existido la revolución del ordenador. Más bien, todo lo contrario: los ordenadores han sido un

²³ Weizenbaum, Joseph, *La frontera entre el ordenador y la mente*, loc. cit., p. 33.

²⁴ Bustamante, Javier, *Sociedad informatizada, ¿sociedad deshumanizada?* (Madrid, Gaia, 1993), p. 151.

²⁵ Weizenbaum, Joseph, *La frontera entre el ordenador y la mente*, loc. cit., pp. 35 y 36.

invento antirrevolucionario, en tanto que usado para conservar el sistema político y de producción capitalista. Los ordenadores cerraron la posibilidad de un mundo distinto. ¿Cómo habría cambiado el mundo si los ordenadores no se hubieran inventado? Imposible saberlo.

Lo que sí sabemos es cómo es el mundo hoy. Alineándonos con la Escuela de Frankfurt, consideramos que el concepto fundamental para comprender nuestro mundo de capitalismo avanzado es la «racionalidad instrumental»²⁶. La racionalidad instrumental es aquella que se ocupa de maximizar la relación entre medios y fines. Es decir, de hallar los medios más eficientes para la consecución de los fines. En este marco, la tecnología es una actividad de primer orden, pues tradicionalmente se define como el conjunto de habilidades (herramientas, maquinaria, instrumentos, materiales, procedimientos sistemáticos, etc.) que hacen posible la consecución de ciertos fines humanos²⁷. Y, dentro de las tecnologías, la informática destaca por haberse convertido en la gran condición necesaria para el espectacular desarrollo exponencial de todas ellas que venimos observando durante las últimas décadas.

¿Cómo nos afectan los ordenadores?

Para diseñar motores, fármacos y carreteras; o para predecir fenómenos climáticos, demográficos y económicos. Los ordenadores sirven a todas estas actividades tecnológicas gracias a dos características. La primera es su condición de máquinas universales, que ya explicamos en el primer capítulo. Y la segunda –derivada de la primera– es su capacidad para realizar simulaciones. Antiguamente, durante el proceso de diseño de un avión había que construir montones de maquetas a escala y probarlas en un túnel de viento para descubrir sus fallos. Era un proceso largo y costoso. Hoy la tendencia es que las simulaciones reemplacen a la experimentación. En vez de construir maquetas, se programa un ordenador que simule el comportamiento del avión en un túnel de viento virtual.

«La simulación consiste en la representación de un sistema complejo real mediante un modelo matemático, es decir, un conjunto de datos y parámetros organizados de forma que constituyen una representación teórica de dicha realidad»²⁸.

Naturalmente, las simulaciones no siempre reflejan de forma exacta la parcela de la realidad que pretenden emular. Por eso, las predicciones meteorológicas a veces fallan. Pero de los errores se aprende, y los ingenieros trabajan duro para mejorar sus simuladores. Tan importantes son las simulaciones informáticas, que hay algunas ciencias, como la sociología, que no pudieron probar sus modelos hasta la aparición de los ordenadores, pues la aeronáutica

²⁶ Bustamante, Javier, *Sociedad informatizada, ¿sociedad deshumanizada?*, loc. cit., p. 43.

²⁷ *Ibid.*, p. 60.

²⁸ *Ibid.*, p. 159.

puede experimentar construyendo aviones a escala y estrellándolos, pero la sociología no puede –o al menos, no debe– infligir hambre a un grupo de población para estudiar sus reacciones.

Así contemplada, la tecnología en general –y la informática en particular– consigue parecer una actividad neutral a ojos de la opinión pública, en tanto que se ocupa de optimizar los medios sin inmiscuirse en la proposición de fines. Se oye comentar que los ordenadores «no son ni buenos ni malos», sino que son meras herramientas que nos afectan para bien o para mal en función de la finalidad que se les dé. «Si se utilizan para programar redes sociales que nos ponen en contacto con nuestros amigos, son buenos. Si se utilizan para distribuir pornografía infantil, son malos». Tal es la forma de pensar del hombre de a pie –que padece miopía intelectual–. Es el mismo argumento que esgrimen en Estados Unidos aquellos que defienden el derecho de la población civil a poseer armas de fuego. Dicen que una pistola puede servir para proteger a tu familia del asalto de unos ladrones. Pero lo cierto es que, con una pistola en la mesilla de noche, el mundo se ve de otra manera. Aunque no la utilices nunca, el mero hecho de saber que la tienes puede convertirte en una persona distinta. Si antes dejabas que tu vecino hiciera ruido por las noches, ahora te atreverás a tocar su puerta para exigirle que guarde silencio. Si antes dejabas que alguien se te colase en el turno de la carnicería, ahora levantarás la voz para decir que tú ibas antes. Los instrumentos y las máquinas, sean pistolas o computadoras,

«simbolizan las actividades a que dan lugar (su propio uso). Un remo es un instrumento que sirve para remar y representa la capacidad de remar en toda su complejidad. Aquel que no haya remado no puede ver en él verdaderamente un remo»²⁹.

En este sentido, los instrumentos trascienden la mera condición de medio práctico para convertirse en transmisores de una visión del mundo. Por eso son *pedagógicos*. Ni siquiera es necesario que un individuo use directamente un instrumento para que éste transforme su visión del mundo. La visión del mundo de la señora que se colaba en la carnicería ha cambiado.

Ya hemos visto cómo los ordenadores resultaron decisivos tras la Segunda Guerra Mundial para mantener un tipo de sociedad. Este hecho por sí solo refuta la tesis de neutralidad de los ordenadores. Pero hay más hechos que se suman a dicha refutación. Distinguiremos dos: la crisis de la identidad humana, y la reducción de la capacidad humana de control en distintas actividades sociales.

²⁹ Weizenbaum, Joseph, *La frontera entre el ordenador y la mente*, loc. cit., p. 25.

Crisis de la identidad humana

El empresario textil que antes pusimos de ejemplo quiere obtener el mayor beneficio posible de su negocio –como todos los empresarios–. La informatización es un *medio* que le ayuda a conseguir esa *finalidad*. Sin embargo, los ordenadores que ha comprado no son capaces de gestionar por sí solos la totalidad del negocio. Quizás en el futuro las IAs podrán hacerlo. Pero los ordenadores actuales requieren ser manipulados por trabajadores. Así se constituye el binomio «trabajador humano–computadora». Veamos lo que aporta cada uno. La computadora por su lado ofrece principalmente las cualidades de la exactitud y la velocidad; realiza cualquier tarea con rigor matemático en cuestión de segundos. El trabajador, al no poder ser tan preciso y veloz, se convierte en el «cuello de botella» del sistema informático³⁰. De cara al empresario el trabajador se ha transformado así en un obstáculo para la maximización de la productividad. Y los obstáculos deben ser eliminados, naturalmente, como dicta la racionalidad instrumental. El trabajador humano ya no tiene ningún valor añadido que le sirva para compensar al empresario por su falibilidad y lentitud, pues la creatividad, que antes era su principal virtud, es rechazada de plano por la computadora, en tanto que ésta sólo permite trabajar dentro del programa preestablecido. Sencillamente, las ideas nuevas no son computables, y por tanto no pueden entrar en el sistema. Ni siquiera en una reunión entre personas para cambiar el programa preestablecido puede el trabajador librarse de la máquina, porque las ideas que allí proponga serán sometidas a una simulación guiada por la finalidad única de incrementar los beneficios económicos. La solución óptima a un problema es, por tanto, sólo una; quedando los caminos alternativos como rechazados, y sus posibles frutos, descartados. Las computadoras son rodillos de pensamiento único.

«En una sociedad entendida según el modelo del computador, donde la sincronía y funcionalidad de todos y cada uno de los componentes son factores esenciales para su correcto funcionamiento, queda cada vez menos espacio para el ser humano y sus características esenciales: la pasión, la esperanza, la falibilidad, el dolor»³¹.

La única vía de supervivencia para el trabajador humano pasa por replantearse sus propias características esenciales. Para comer ha de trabajar, y para trabajar ha de mimetizarse con la computadora. Dicho y hecho. Se comienza purgando al concepto de inteligencia de todas aquellas operaciones que no puede realizar el computador. Muchas de las inteligencias múltiples –que vimos más arriba– distinguidas por Howard Gardner son suprimidas. Ahora «ser inteligente» no significa «ser creativo», sino «ser calculador». La nueva tabla de valores que el trabajador ha de interiorizar para sobrevivir en este mundo está compuesta por los

³⁰ Bustamante, Javier, *Sociedad informatizada, ¿sociedad deshumanizada?*, loc. cit., p. 81.

³¹ *Ibid.*, p. 118.

valores de la tecnología: eficiencia, funcionalidad, sincronización, etc. La tecnología, por tanto, no es un simple medio para realizar cualquier fin determinado externamente a ella, sino que ella, internamente, dicta un conjunto de fines axiológicos, o valores. El buen trabajador es el que trabaja «como una máquina». ¿Cuántas veces no habremos escuchado esa frase hecha de «ser un máquina» utilizada como un elogio? Los comportamientos que sean fieles a dicha pauta serán premiados, y los demás, castigados.

Observemos la moda de los videojuegos para medir y ejercitar la inteligencia, los llamados *brain trainers*. Yo –si se me permite hablar en primera persona– soy redactor de *Micromanía*, la revista de videojuegos más vendida de España, gracias a lo cual tengo la ocasión de probar todos los videojuegos del mercado. Los *brain trainers* consisten básicamente en realizar ejercicios mentales en un tiempo limitado. A menos tiempo empleado en la tarea, mayor es la puntuación otorgada por la máquina. Es decir, más inteligente eres. ¿Qué otros parámetros aparte del tiempo diacrónico puede si no medir una computadora? –Según este baremo, la inteligencia espacial de Miguel Ángel sería paupérrima porque tardó trece años en pintar la Capilla Sixtina–. Sin embargo, la gente anda en el autobús enganchada a esos *brain trainers* con el afán de mejorar su inteligencia. En esos programas sólo hay desafíos como resolver rompecabezas, encontrar palabras de cierto número de sílabas y efectuar cálculos matemáticos. Operaciones todas ellas resolubles por la computadora. Antaño la capacidad de pintar o escribir obras bellas eran parámetros que definían la inteligencia humana, pero han sido aniquilados porque no son aprehensibles en forma de algoritmo, es decir, no son computables.

A estas alturas del ensayo ya estamos en disposición de entender por qué Sigfrid sería un buen psiquiatra en nuestra sociedad. Lo sería precisamente «porque» es una máquina, y no «a pesar de que» es una máquina. Los psiquiatras humanos tienen necesidad de descansar unos minutos entre paciente y paciente. Sigfrid no, y por tanto es más productivo. Los psiquiatras humanos pueden hacer huelga. Sigfrid no, y por tanto es más productivo. Los psiquiatras humanos pueden negarse a aplicar terapias efectivas pero inmorales. Sigfrid no, y por tanto es más productivo. Productividad, productividad y productividad. ¿Qué más da que Sigfrid no entienda a sus pacientes? En un mundo dominado hegemónicamente por la racionalidad instrumental, el valor supremo de un trabajador es su productividad. Sólo importa el rendimiento que se le puede exprimir.

Crisis de la capacidad humana de control

Windows Vista, uno de los sistemas operativos más modernos, tiene 50 millones de líneas

de código. Si imprimiéramos ese código en un rollo de papel continuo, éste mediría 250 kilómetros. Cubriría la distancia entre Madrid y Valencia. Para la creación de tan gigantesco programa fue necesaria la participación de 2.000 desarrolladores. Entre ellos había jefes y subordinados, como en todas las empresas, pero no había ningún jefe en la cúspide que fuera capaz de entender en su totalidad el funcionamiento de Windows Vista. La consecuencia es que el programa a veces no funciona como debería, y nadie sabe con certeza dónde está la línea de código que está produciendo el error³².

Muy lejos de las oficinas de Microsoft se encuentra un trabajador ante su computadora que se ha bloqueado. Para seguir realizando su tarea necesita que el ordenador vuelva a funcionar, pero eso no está al alcance de su mano. Indefenso, ha perdido el control de la situación, pues depende de que la empresa desarrolladora, Microsoft en este caso, lance un parche que solucione el problema.

Esta situación de pérdida de control es frecuente. Todos la hemos vivido, y no tiene especial trascendencia más allá de contribuir a la sensación de dependencia de la computadora. Más grave sería que el programa informático que controla los misiles nucleares del ejército de una nación se accionase por error. Para tranquilizarnos, tendemos a creer que eso es imposible, que los ingenieros militares habrán tomado la prudente precaución de condicionar el lanzamiento de misiles nucleares al mandato humano. Pero la realidad es justo al revés. Stanley Kubrick explica por qué en su película *Teléfono rojo, ¿volamos hacia Moscú?* (*Dr. Strangelove*). La sinopsis de la cinta la tomamos de Juan Antonio Rivera³³:

»Estados Unidos dispone, ya en la década de 1960, de flotillas de bombarderos B-52; cada aparato puede dejar caer una carga nuclear de 50 megatonnes, «equivalente a sesenta veces la potencia explosiva de las bombas y obuses utilizados por todos los ejército beligerantes en la Segunda Guerra Mundial». Estos B-52 vuelan por turnos las veinticuatro horas del día y están todos ellos a dos horas de sus objetivos en el interior de Rusia.

»Una flotilla de estos aparatos recibe la orden cifrada «Ataque de escuadrilla según el plan R». La orden ha sido emitida desde una base aérea estadounidense por el general de brigada Jack D. Ripper (Sterling Hayden), un anticomunista delirante, que ve conspiraciones bolcheviques por todos lados y que ha decidido emprender por su cuenta y riesgo una guerra nuclear contra Rusia. [...]

»A partir de este punto, nadie, [...] ni el mismísimo presidente de Estados Unidos [...] puede hacer nada para detener la descabellada operación.

»La idea, sigue diciendo éste (un general norteamericano) cada vez más entusiasmado por el plan, es que la amenaza de una represalia no interrumpible disuadiera a los rusos de cualquier ataque a Estados Unidos.

³² *Ibid.*, p. 169.

³³ Rivera, Juan Antonio, *Lo que Sócrates diría a Woody Allen* (Madrid, Espasa, 2003), pp. 179 a 185.

«El embajador ruso, ante las dimensiones que empieza a tomar el asunto, [...] comunica a los militares del Pentágono allí reunidos, y para consternación general, que su país tiene prevista la contrarréplica adecuada al plan R: la Máquina del Apocalipsis, un ingenio «que destruirá toda la vida humana y animal de la Tierra».

Entonces entra en escena el Dr. Strangelove, para hablar con el presidente:

«—Pero ¿cómo es posible? —pregunta el presidente de los Estados Unidos— que ese ingenio [se refiere a la Máquina del Apocalipsis] se pueda activar automáticamente y que después sea imposible desactivarlo?

«—Señor presidente, no es únicamente posible; es *esencial*, es la idea en que se basa esta Máquina —aclara el doctor Strangelove, muy obviamente enardecido por poder comentar los aspectos teóricos de la cuestión—: la disuasión es el arte de producir en la mente del enemigo el *miedo* al ataque. Por lo tanto, como el proceso decisorio es automático e irrevocable y funciona fuera del control humano, la Máquina del Apocalipsis es terrible. Es fácil de entender y absolutamente creíble y convincente.

Nótese que es *esencial*, como señala el doctor Strangelove (interpretado magistralmente por el inconmensurable Peter Sellers), que el ataque nuclear sea automático, irrevocable y fuera del control humano. En un mundo como el nuestro, cada vez más obsesionado por la seguridad, es seguro que las grandes potencias tienen artefactos similares a la Máquina del Apocalipsis, manejados por programas informáticos que constan de tantos millones de líneas de código, que escapan al control de sus propios programadores.

Conclusión

Las IAs en sentido fuerte son el siguiente paso de la ingeniería informática. Muchos sueñan que en el futuro los robots harán nuestro trabajo mientras nosotros descansamos. Sería el final del castigo bíblico a ganarnos el pan con el sudor de nuestra frente. Pero esto ni debe suceder, ni sucederá.

Que no debe suceder es algo que ya hemos explicado al analizar la problemática de Sigfrid —el psicoterapeuta robótico—, y lo hemos ampliado en el apartado anterior. Hay tareas propiamente humanas que no deben ser delegadas en las máquinas por motivos éticos —Sigfrid— y de seguridad —la Máquina del Apocalipsis—.

Que no sucederá es una predicción basada en la experiencia histórica. La revolución tecnológica de las IAs no cambiará nada a nivel social, como nada cambiaron las revoluciones tecnológicas en el pasado. La máquina de vapor no redundó en una mejora de la calidad de vida de los obreros, sino en un incremento de los beneficios de los dueños de los medios de producción. Análogamente, las IAs serán propiedad de los ricos, y para ellos trabajarán. Pensemos en las consecuencias de la invención de una IA capaz de entender el lenguaje natural.

«¿Cuál es verdaderamente su utilidad? Pienso que no existe ningún problema humano agobiante que

podría ser resuelto con mayor facilidad mediante la intervención de dicha máquina. Sin embargo, estas máquinas, caso de poder construirse, facilitarían considerablemente el control de la comunicación verbal. Acaso la única razón de que haya tan poca vigilancia oficial en las conversaciones telefónicas en muchos países del mundo es que exige mucha mano de obra. Cada conversación a través de un teléfono intervenido debe ser eventualmente escuchada por un agente humano. Pero las máquinas de reconocimiento del discurso podrán suprimir todas las conversaciones irrelevantes y ofrecer transcripciones de las restantes»³⁴.

Toda nueva tecnología será propiedad de los ricos, y por tanto utilizada por éstos para aumentar su dominación sobre los pobres. Que nadie piense que Internet es la gran excepción a esta regla. Es sabido que Google colabora con los servicios de inteligencia de los Estados Unidos proporcionándoles información acerca de las búsquedas realizadas por los usuarios de su buscador. Lo mismo hace Microsoft, incorporando programas espía entre los millones de líneas de código de sus sistemas operativos.

La tecnología, en tanto que no es un instrumento neutro sino creador de valores acordes a la racionalidad instrumental, no es un aliado, sino un obstáculo para la realización de una revolución social que, guiada por una racionalidad axiológica, termine con la brutal explotación capitalista. Así lo denuncia Marcuse:

«Nuestra sociedad se caracteriza antes por la conquista de las fuerzas sociales centrífugas por la tecnología que por el terror, sobre la doble base de una abrumadora eficacia y un nivel de vida más alto»³⁵.

Bibliografía:

- Aristóteles, *Metafísica* (Madrid, Gredos, 2003).
- Beck, Ulrich, *¿Qué es la globalización?* (Barcelona, Paidós, 2008).
- Bustamante, Javier, *Sociedad informatizada, ¿sociedad deshumanizada?* (Madrid, Gaia, 1993).
- Copeland, Jack, *Inteligencia artificial* (Madrid, Alianza, 1996).
- Franklin, Stan, *Artificial minds* (Cambridge, Massachusetts, The MIT Press, 1995).
- Gardner, Howard, *Inteligencias múltiples. La teoría en la práctica* (Barcelona, Paidós, 1999).
- Hawkins, Jeff, *Sobre la inteligencia* (Madrid, Espasa, 2005).
- Horkheimer, Max, *Teoría crítica* (Madrid, Amorrortu, 1974).
- Jackson, Peter, *Introduction to expert systems* (Wokingham, Addison-Wesley, 1986).
- Marcuse, Herbert, *El hombre unidimensional* (Barcelona, Orbis, 1984).

³⁴ Weizenbaum, Joseph, *La frontera entre el ordenador y la mente*, loc. cit., p. 223.

³⁵ Marcuse, Herbert, *El hombre unidimensional* (Barcelona, Orbis, 1984), p. 20.

- Pardo, José Luis, *El alma de las máquinas*, Sibila nº 7, octubre 2001, pp. 28-31.
- Pohl, Frederik, *Pórtico* (Barcelona, Planeta, 2006).
- Queraltó, Ramón, *Tecnología y valores en la sociedad global: El caballo de Troya al revés* (Madrid, Tecnos, 2003).
- Rivera, Juan Antonio, *Lo que Sócrates diría a Woody Allen* (Madrid, Espasa, 2003).
- Rumelhart, David, *Parallel distributed processing: Explorations in the microstructure of cognition, vol. I: Foundations* (Cambridge, Massachusetts, The MIT Press, 1989)
- Scheich, Henning, *Control del pensamiento*, en *Mente y cerebro*, 04/2003.
- Searle, John, *Mentes, cerebros y ciencia* (Madrid, Cátedra, 1990).
- Stendhal, *Del amor* (Madrid, Alianza, 2003).
- Turing, Alan, *¿Puede pensar una máquina?* (Valencia, Teorema, 1974).
- Weizenbaum, Joseph, *La frontera entre el ordenador y la mente* (Madrid, Pirámide, 1976).