

Inteligencia artificial

Ángel Gutiérrez

Introducción

Imagínese un escenario en el que el protagonista descubre una tierra desconocida donde todas las máquinas han sido eliminadas. Los habitantes humanos temen que puedan convertirse en una amenaza si llegan a hacerse inteligentes y adquirir una consciencia propia. Podría ser el argumento de una secuela de la serie "Terminator", ¿verdad? Sin embargo, acabo de describir someramente la trama de "Erewhon", un libro escrito por el británico Samuel Butler a finales del siglo diecinueve. Está claro que no es nueva la idea de que deberíamos tener miedo de a dónde pueden llegar las máquinas. Sobre todo si consiguen hacerse más inteligentes que nosotros...



Figura 1. Un Terminator, los sanguinarios esbirros del ordenador superinteligente Skynet. (Fuente: <http://1.bp.blogspot.com>)

||| Pero, al final, ¿a qué se le llama ser inteligente?

No es tan sencillo responder a esta cuestión. Podría pensarse, por ejemplo, que un buen jugador de ajedrez debe ser indudablemente inteligente. Sin embargo, la empresa IBM desarrolló durante los años noventa un ordenador, llamado *Deep Blue*, que fue capaz de derrotar al entonces campeón del mundo de ajedrez, Garry Kasparov. No me entienda mal. Con esto no quiero insinuar que los jugadores de ajedrez no sean inteligentes. Sí lo son, pero *Deep Blue* no lo era, y aun así consiguió vencer al mejor de todos ellos. Por tanto, con independencia de las acusaciones de fraude que Kasparov lanzó contra IBM, lo que *Deep Blue* dejó claro es que no hacía falta inteligencia para ser un jugador de ajedrez excepcional.



Figura 2. *Deep Blue*, ¿el mejor ajedrecista de la historia?
(Fuente: <http://archive.computerhistory.org>)

La inteligencia tampoco reside en la memoria, aunque ésta sea quizá un atributo fundamental de ella. Cualquier ordenador actual es capaz de almacenar más información que una persona corriente, y también de acceder a ella de un modo más rápido. Lo mismo puede decirse sobre la capacidad de cálculo; no ya de un superordenador, sino hasta de una simple calculadora que, estaremos de acuerdo, no tiene la menor pizca de inteligencia.

Por tanto, una vez más, ¿a qué se le llama ser inteligente? Hay un famoso comentario que hizo un miembro de la Corte Suprema de Estados Unidos, Potter Stewart, en una decisión que tomó sobre la definición de obscenidad y lo que se consideraba obsceno. El juez Stewart dijo “no puedo definirlo, pero sé qué es [obsceno] cuando lo veo”. Algo parecido podría decirse con respecto a la inteligencia:

“sabremos que algo es inteligente cuando lo veamos”. Pero está claro que la búsqueda de la inteligencia artificial necesita un criterio más concreto y, sobre todo, más científico. Es ahí donde entra el denominado test de Turing.

||| El test de Turing

Ya en los albores de la era informática, en los cincuenta, cuando ordenadores mucho menos potentes que los actuales ocupaban habitaciones enteras y eran un secreto de estado, hubo quien se preguntó qué prueba debería ser capaz de superar un ingenio artificial para poder ser considerado inteligente. La más exitosa de esas reflexiones fue, y sigue siendo, la del excéntrico y genial matemático y criptoanalista británico Alan Turing.

Turing planteó un escenario al que dio el nombre de “juego de imitación”. En él estaban envueltos una persona, un computador y un interlocutor humano. Este debía encontrarse aislado de los otros dos, que conoce sólo por X e Y; obviamente sin saber quién es quién, porque su misión, el objetivo del juego, es adivinar quién es la máquina y quién el humano. Para ello puede hacer a uno u otro las preguntas que considere oportunas, durante un tiempo prefijado. Las respuestas de la persona deben ayudar al interrogador a desenmascarar a la máquina, mientras que las de ella tienen que estar orientadas a engañarlo haciéndole creer que es la humana.

Este “juego” es la forma básica del test de Turing que, para muchos investigadores, representa la única herramienta válida para determinar si un sistema artificial es o no inteligente. El propio Turing, siempre tan

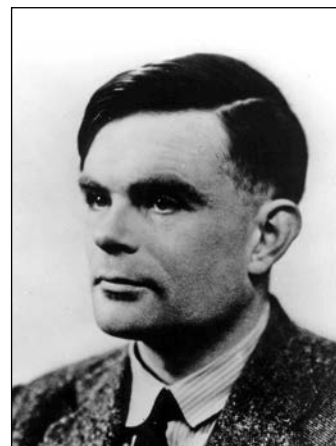


Figura 3. Alan Turing.
(Fuente: <http://www.bibleetnombres.online.fr>)

concienzudo, elaboró una lista de nueve objeciones al hecho de que una máquina digital pudiera llegar algún día a pensar por sí misma. Iban desde argumentos lógico-matemáticos a otros incluso teológicos. A todos ellos les dio una respuesta cabal, de modo que ninguna de esas objeciones parecía sostenerse.

Lo que ya no está tan claro es si el test de Turing es demasiado riguroso, en el sentido de que una máquina pensante podría no ser capaz, aun así, de superarlo. Hay un ejemplo curioso de un caso en que eso podría llegar a ocurrir. Imagínese que nuestra máquina es, además de pensante, honesta, y que bajo ninguna circunstancia se permite a sí misma mentir. En una situación como esa, al interrogador del test de Turing le resultaría muy fácil establecer quién es la máquina de los dos oponentes. Ésta no podría evitar irse de la lengua y “confesar” que ella es la no humana. Por tanto, no superaría el test a pesar de ser inteligente.

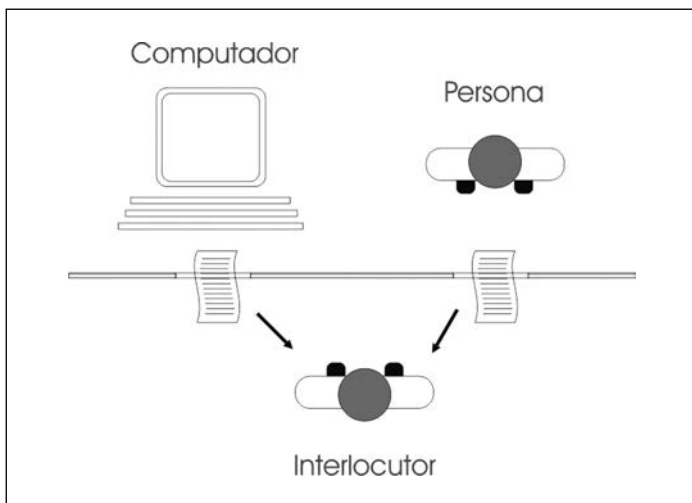


Figura 4. Representación esquemática del test de Turing. (<http://upload.wikimedia.org>)

No crea que me he inventado yo este argumento tan pintoresco. Es uno de los que los detractores del test de Turing esgrimen en su contra. La réplica de quienes lo defienden es que ese y otros argumentos por el estilo sólo serían válidos asumiendo que se considere que el test es suficiente y/o necesario para atribuir inteligencia a algo. Y no parece que eso es lo que Turing tuviera en mente. Más bien, su intención era establecer un método analizable y medible para determinar la probabilidad de que un sistema artificial pudiera ser considerado inteligente. En cierto modo, con el test de Turing puede aplicarse lo de “no son todos los que están, pero están todos los que son”. En otras palabras, no todas las máquinas –o

alienígenas u otros seres– pensantes serían capaces de superar el test, pero aquello que lo supere sí puede considerarse inteligente dentro de un cierto rango probabilístico más o menos amplio, que está en función directa de lo extenso y riguroso que haya sido el propio test.

Otras críticas al test de Turing no tienen que ver con que no sea válido, sino con todo lo contrario. El hecho de ser tan exigente, argumentan estos críticos, lo convierte en una herramienta poco adecuada para establecerlo como meta a corto o medio plazo en la búsqueda de la inteligencia artificial. En 1950, cuando Turing desarrolló su test, previó que en unos cincuenta años existieran sistemas digitales capaces de engañar a los interrogadores en un número considerable de casos, tras cinco minutos de “conversación”. Sin embargo, hoy en día, más de sesenta años después de esas afirmaciones, resulta discutible que estemos cerca de conseguir algo así.

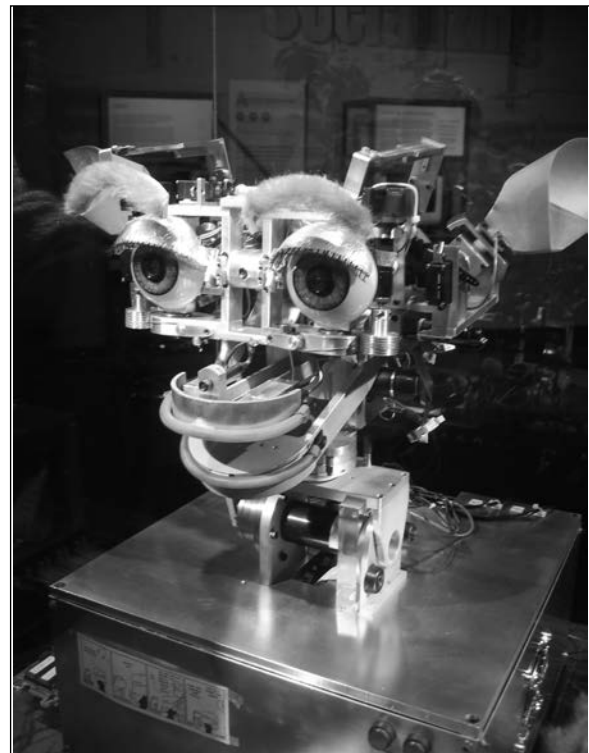


Figura 5. Este simpático robot, de nombre Kismet, es uno de los más avanzados ejemplos de sistemas inteligentes. (Fuente: <http://synapticnuship.com>)

Es verdad que ha habido progresos importantes en el desarrollo de sistemas o dispositivos llamados expertos. Éstos han demostrado un éxito considerable en tareas diversas, pero siempre para aplicaciones muy específicas y bastante limitadas. Ninguno ha conseguido nada parecido a lo que sería de esperar

de una verdadera inteligencia artificial, equivalente o superior a la humana. Y esta relativa falta de avances relevantes en ese gran objetivo no se debe a la falta de estímulos o de fondos. Aparte del desafío científico e intelectual que supone, nadie duda de la infinidad de aplicaciones comerciales que tendría la inteligencia artificial, con sus consiguientes lucros astronómicos. Y es bien sabido que éstos atraen siempre inversiones de gran calibre.

Es algo hasta cierto punto inesperado en este mundo nuestro donde todo va cada vez más rápido. Sectores de la ciencia como el de la genética hasta han conseguido antes de lo previsto logros que sólo se esperaban para dentro de alguna que otra década más. Que éste no sea el caso de la inteligencia artificial se debe a que lograrla no depende sólo del desarrollo de ciertas herramientas, o ni siquiera del aumento exponencial de la capacidad de procesamiento de los sistemas informáticos (seguramente, en la forma de ordenadores cuánticos). Hay demasiadas cosas que aún no sabemos acerca de cómo funciona el pensamiento o de hasta qué punto es clave o no un sustrato físico para él. Parecen existir demasiados matices determinantes en la inteligencia.

Pero hay opiniones para todos los gustos. Los investigadores más optimistas creen que tendremos máquinas tan inteligentes o más que un humano en los próximos quince o veinte años. Para otros, eso es algo que tardará mucho más en lograrse, si es que llega a conseguirse.

Tengan razón a la larga unos u otros, el caso es que hace muy poco se ha producido un avance importante –quizá incluso decisivo– en la búsqueda de la inteligencia artificial. Años de investigación y muchos millones invertidos por IBM se han materializado en un gran éxito de una de sus máquinas. Pero no una cualquiera. La llaman Watson, en honor de

uno de los más carismáticos presidentes del célebre fabricante de ordenadores.

A los americanos les gusta el espectáculo, así que IBM no eligió un gris laboratorio para demostrar su nuevo avance. El palco fue uno de los programas más famosos de la televisión en Estados Unidos. Jeopardy! no es un simple concurso de preguntas y respuestas. En él no basta tener conocimientos enciclopédicos sobre los más variados temas. De ser así, el reto no habría sido tan grande. Además, para ganar Jeopardy! es necesario entender las preguntas implícitas en las frases y captar toda clase de sutilezas y juegos de palabras ocultos en ellas. Eso, que puede ser una tarea relativamente sencilla para un humano inteligente y formado, era algo hasta ahora impensable para una máquina.

Por si no fuera dificultad suficiente, los organizadores del evento decidieron complicar todavía más el desafío, y de paso hacer el espectáculo más... bueno, espectacular. El superordenador Watson debería enfrentarse con los dos humanos que ostentan el récord de victorias seguidas y de cantidad total de dinero acumulado en Jeopardy!: Brad Rutter y Ken Jennings.

La “contienda” duró tres días, y el resultado fue una victoria de la máquina de IBM. Watson acumuló con sus respuestas correctas más del triple de dinero que sus dos contrincantes. Aunque no todo corrió a la perfección. Por ejemplo, ante la pregunta de qué ciudad de Estados Unidos tenía su aeropuerto principal bautizado en honor de un héroe de la Segunda Guerra Mundial, y el segundo mayor en el de una batalla también de la Segunda Guerra Mundial, Watson respondió “Toronto”. Obviamente, no captó lo de que se trataba de una ciudad de *Estados Unidos* (por si le interesa, la respuesta correcta es Chicago).

Los creadores de Watson ya están trabajando en afinar sus sorprendentes capacidades. Y también en un proyecto –con las universidades de Columbia y Maryland y la firma Nuance Communications– para crear a corto plazo un sistema artificial inteligente de asistencia médica, capaz de responder a preguntas hechas por especialistas médicos humanos en lenguaje natural. Otro de los proyectos en que está previsto incluir a Watson y sus futuras mejoras es en toda clase de procesos de interacción con humanos, relacionados por ejemplo con la atención al cliente a varios niveles.

Las perspectivas son sin duda optimistas. Seguramente demasiado buenas para un tercer grupo entre los, digamos, optimistas y pesimistas de la inteligencia



Figura 6. “Nada elemental, querido Watson”.
(Fuente: <http://underthelobsterscope.files.wordpress.com>)

artificial de que hablaba antes. En cierto modo podría considerársele heredero del pueblo de “Erewhon”, del que hablé también al principio de este artículo. Para ellos, igual que para los habitantes de ese lugar de ficción, la inteligencia artificial no es algo deseable. ¿Por qué? Bueno, basta pensar en HAL 9000, de 2001. Él, tan inteligente y sensible, siempre muy preocupado por si sería capaz de soñar o no si lo apagaban, no tuvo el menor remilgo en asesinar a varios de sus compañeros humanos. Más radical aún es el caso de Skynet, la todopoderosa mente cibernética tras los “Terminator”, que intentó acabar de un plumazo con toda la población mundial mediante una guerra nuclear. Panoramas como éstos no son inimaginables en el caso de que algún día logremos crear máquinas inteligentes. Si las hacemos a nuestra imagen y semejanza, capaces de sentir emociones, las habrá “buenas” y “malas”, como ocurre entre nosotros. Y no es mejor alternativa crear máquinas sin sentimientos. Eso las llevaría a ser pragmáticas y a tomar decisiones como a eliminarnos de la faz de la Tierra por contaminar el planeta o agotar sus recursos. Sin una parte emocional, podrían llegar a la conclusión de que lo mejor es acabar con todos nosotros, igual que se elimina a una plaga que está destruyendo su propio hábitat.



Figura 7. HAL 9000, la computadora inteligente (y asesina) de la película 2001.
(Fuente: <http://i374.photobucket.com>)

Entrevista al profesor Stuart M. Shieber

El profesor Shieber es un experto de nivel mundial en computación lingüística y comunicación. Desarrolla sus investigaciones en la Escuela de Ingeniería y Ciencia Aplicada de la Universidad de Harvard. Entre otros cargos y responsabilidades, ocupa la cáte-

dra James O. Welch, Jr. y Virginia B. Welch de ciencia informática, es director y fundador del Centro de Investigación sobre la Computación y la Sociedad, además de codirector del Centro Berkman para Internet y la sociedad.

Dejo aquí mi agradecimiento al profesor Shieber por aceptar responder a esta entrevista.



Figura 8. Stuart Shieber.
(Fuente: <http://news.harvard.edu>)

Lingüistas como Noam Chomsky afirman que los lenguajes artificiales son más difíciles de aprender o utilizar que los naturales... ¿Nosotros pensamos con palabras o nuestro mundo mental lo conforman algún otro tipo de estructuras?

Por tanto, se trata de dos cuestiones distintas. Una sobre los lenguajes naturales frente a los artificiales, y otra acerca de si nuestro pensamiento está estructurado del mismo modo que lo está el lenguaje...

Sí, porque nuestros lenguajes naturales están directamente relacionados con el modo en que pensamos.

La cuestión de si el pensamiento está o no estructurado como el lenguaje es discutible. Yo creo que hay ciertos casos en los que sí debe ser así. Cuando se dice que un lenguaje es algo estructurado, se refiere a que ese lenguaje es productivo. Productivo en el sentido de que hay un conjunto limitado de cosas sobre las que se puede hablar y que es posible decir, de igual modo que también es limitado el número de cosas en las que se puede pensar. En el caso del lenguaje, no es posible que ese número limitado de cosas se defina mediante una especie de lista. Porque, ¿cómo seríamos capaces de incorporar esa lista en nuestro propio comportamiento? No puede ser

implementada genéticamente en el ADN de un embrión.

Hay claramente ciertas características que el pensamiento y el lenguaje tienen en común: productividad y arbitrariedad; las bases estructurales para esas cualidades y otras por el estilo. Las evidencias que encontramos con respecto al lenguaje y el pensamiento humanos no las vemos en los de otros animales. Eso no implica necesariamente que ellos no piensen también de un modo productivo y estructurado, aunque no sean capaces de exteriorizar eso a través de un lenguaje. En cualquier caso, no es la opinión de Chomsky. Él afirma que sólo los seres humanos tienen esa cualidad de pensar productiva y estructuralmente mediante un lenguaje interno que es después exteriorizado. Es una teoría... Hay quienes lo ven de otro modo y consideran que ese lenguaje productivo y estructurado se debe a la necesidad de comunicación y que nos permite, adicionalmente, hablar con nosotros mismos [pensar]. En principio podría intentarse buscar evidencias evolutivas que justificaran más una u otra teoría, pero no estoy seguro de que por el momento esté clara la respuesta.

Si a una persona cualquiera le hacen una pregunta sobre un tema, lo más habitual es que responda aunque no tenga la menor idea sobre dicho tema. Algo parecido parece ocurrir con nuestro pensamiento, que nos permite reflexionar sobre todo tipo de asuntos incluso cuando no tenemos todos los datos necesarios para hacerlo. Y, en un caso de cada millón, esa reflexión "estúpida" podría ser válida, o al menos conducir a algo que pueda serlo y que quizá no habría surgido de otro modo. ¿Podría la creatividad estar relacionada con esa capacidad de los humanos de imaginar respuestas a cuestiones que no tenemos completamente en la cabeza? Si un sistema de inteligencia artificial sólo piensa pensamientos correctos, quizá no sea capaz de ser verdaderamente creativo. Puede que necesite poder barajar pensamientos estúpidos, o hasta locos, para llegar a pensamientos del más alto nivel...

Es decir, que quizá sea un error que nuestro sistema inteligente sea, en cierto modo, demasiado preciso. Que el hecho de que su forma de pensar sea demasiado correcta, sistemática o racional pueda impedirle que llegue a alcanzar el nivel del pensamiento humano, con toda su creatividad. Podría ser. Pero no creo que sea algo de lo que debemos preocuparnos por el momento, porque estamos muy lejos de conseguir emular ese nivel de inteligencia.

Hay un libro de Ray Kurzweil que se llama "La singularidad". En él desarrolla la idea de que los

ordenadores irán aumentando exponencialmente su capacidad. Eso les llevará a un punto en que se harán tan poderosos y tendrán unas capacidades tan increíbles, que acabarán superando a los propios humanos. Esa es la singularidad: que algo creado por los humanos llegue a ser capaz de superarnos. En ese punto podríamos ser sustituidos por nuestra propia creación. Y, según Kurzweil, no falta mucho para que eso ocurra. Sólo un par de décadas, creo recordar...

¿Qué puede contar sobre el test de Turing?

El test de Turing es una prueba que establece la inteligencia de un sistema artificial basándose en la capacidad de éste de mostrarse indistinguiblemente como un ser humano en una conversación. Kurzweil cree que un ordenador podrá superar el test de Turing dentro de unos veinte años. Para llegar a esa conclusión analizó gráficamente la evolución de la velocidad y la complejidad de los ordenadores, y la comparó con la velocidad y la complejidad de computación de varios cerebros –de una hormiga, un ratón, un ser humano–. Ambas gráficas llegaban a cruzarse en un intervalo de unos veinte o treinta años, cuando se supone que los ordenadores tendrán más o menos la misma velocidad y capacidad que un cerebro humano. Por tanto, en ese momento, seremos capaces [según Kurzweil] de crear una inteligencia de nivel humano. Es un argumento razonable desde el punto de vista del *hardware* y las prestaciones y capacidad de almacenamiento de los futuros ordenadores. Pero creo que la conclusión no es correcta, porque la principal dificultad de crear un sistema con inteligencia artificial no tiene que ver con que se disponga de un *hardware* lo bastante capaz. Para empezar, habría que saber cómo utilizar ese *hardware*, cosa que no sabemos. Además, incluso en un ordenador que tuviera la capacidad y las prestaciones de cien cerebros humanos, no sabríamos qué *software* utilizar para que fuera capaz de generar un comportamiento inteligente artificial. Podría argumentarse que no será necesario saberlo, porque conseguiremos encontrar una vía que no implique entender cómo surge la inteligencia. Por ejemplo, quizá bastara con lograr replicar el sustrato biológico de la inteligencia humana. Es decir, nuestro cerebro. Así, toda esa capacidad de computación se emplearía en simular un cerebro humano, sin preocuparnos de cómo éste es capaz de hacer lo que hace. Nos limitaríamos a replicar el *hardware* del cerebro humano.

El inconveniente de esta visión es que no son los cerebros humanos los que se vuelven inteligentes, sino los seres humanos. Es decir, sólo es inteligente un cere-

bro que está insertado en el contexto físico de un cuerpo humano con ciertas percepciones y capacidades.

Podríamos intentar replicar también todo eso. Es posible teorizar sobre la posibilidad de crear no sólo ordenadores con la capacidad y la estructura de un cerebro, sino además elementos robóticos capaces de percibir y actuar, que estén conectados a ese cerebro artificial y que funcionen de un modo similar a nuestros cuerpos. Pero los límites de nuestro conocimiento son de veras grandes. Antiguamente, los investigadores solían enfrentarse de un modo más amplio a los retos de la inteligencia artificial y el objetivo de crear un ser completo artificialmente inteligente. Pero hoy en día eso es menos habitual. Muchas investigaciones se centran ahora en resolver problemas más específicos o en buscar soluciones de ingeniería para áreas más concretas.

Ni siquiera la maquinaria del reloj más caro del mundo puede estar hecha de oro, por la sencilla razón de que las propiedades físicas de ese metal no son compatibles con lo que es necesario para construir esa maquinaria. ¿Podría ocurrir algo similar con la inteligencia? ¿Cree que necesita un sustrato biológico como el de nuestro cerebro? ¿O que podría tener como base otro tipo de material?

No veo ninguna razón por la que el sustrato de la inteligencia deba estar basado necesariamente en procesos físicos y químicos como los de nuestro cerebro. De hecho, podría sin problemas ser posible emular los procesos de un cerebro con base en el silicio. Si fuéramos capaces de replicar su comportamiento y su conexión con su contexto físico, y aunque ese sistema no estuviera formado por compuestos orgánicos con base de carbono, en teoría podríamos conseguir que fuera inteligente, que pasara el test de Turing y resultara indistinguible de un humano. No obstante, existe la cuestión filosófica de si superar ese test implica realmente que podamos atribuir una inteligencia al sistema artificial que se analice con él. Y no está claro que así sea.

¿La consciencia es una propiedad de la inteligencia?

La consciencia es un gran misterio. De hecho, quizá sea el misterio por excelencia. Por consciencia yo entiendo el hecho de que resulte instantáneamente obvio para uno mismo que es capaz de pensar y que existe algo que consiste en pensar. Es decir, tener una percepción interna del propio pensamiento. La relación que esa consciencia pueda tener con la inteligencia es algo que se me escapa, pero me parece evidente que debe existir. Hay quien afirma que la

consciencia no puede ser un criterio para la inteligencia y que es posible imaginar algo que sea como una persona, en su inteligencia y su comportamiento, pero que no tenga una consciencia, una percepción interna de su propio pensamiento. El nombre técnico para ese tipo de ser es “zombi”. También hay quien dice que el mero hecho de que podamos imaginar un ser como ese implica que la inteligencia y la consciencia son conceptos independientes. La alternativa que otros defienden es que eso no es así y que cualquier ser inteligente al nivel que lo somos nosotros debe ser necesariamente consciente.

Mi opinión personal es que la consciencia y la inteligencia son separables y que podemos construir algo inteligente que no sea consciente. La consciencia es en verdad un gran misterio. Tanto, que ni siquiera estoy seguro de que pueda ser resuelto algún día. Desde luego no será la ciencia actual la que se encargue de hacerlo, sino más bien la filosofía. Aunque puedo imaginar que, en un futuro, el asunto de la naturaleza de la consciencia pueda llegar a convertirse en un asunto de estudio científico.

Pero un ser sin consciencia y sin voluntad propias ¿qué necesidad tendría de pensar? ¿En qué pensaría? El pensamiento es, en cierto modo, una reacción a lo que nos rodea o hacia las cosas que deseamos saber o conseguir. Pero si no fuéramos conscientes de todo eso, ¿para qué pensaríamos?

Voy a dar un ejemplo... Es normal haber experimentado alguna vez conducir de un sitio a otro –por ejemplo de casa al trabajo– y no tener consciencia del trayecto, no recordar nada sobre él, como si la mente se hubiera apagado en ese tiempo. En cierto modo, puede decirse que en casos así no éramos conscientes de lo que estábamos haciendo. Quizá sí de otras cuestiones, pero no del conjunto de procesos que están implicados en la tarea de conducir. Y ésta es compleja, porque exige coordinar la visión, el oído, el movimiento de manos y pies; exige pensar, en definitiva. Y, sin embargo, no hemos sido conscientes de ello. ¿Qué necesidad habría de pensar si no existiera la consciencia? No lo sé, pero el hecho es que en casos como ese estamos pensando, aunque no seamos conscientes de ello.

Ya, pero esa persona tuvo que tomar la decisión consciente de ir de casa al trabajo. Aunque el camino lo haya hecho en una especie de “piloto automático”, tuvo que querer iniciar ese piloto automático.

Es cierto. Aunque también podría decirse que eso tampoco fue algo consciente. Durante todo el camino

el conductor tuvo que tomar, de un modo no consciente, varias decisiones no cualitativamente distintas de esa primera decisión. Desde ese punto de vista, da la impresión de que la consciencia y la inteligencia son temporalmente separables, a veces durante periodos considerables de tiempo. No lo sé. Pero quizá este sea el punto en el que el estudio de la consciencia pueda pasar a ser algo científico y no sólo filosófico: en esos casos empíricos. Algunos de ellos tan simples como el que comento del trayecto de casa al trabajo. Aunque hay otros muchos. Por ejemplo, el de ciertas personas que afirman ser ciegas porque no tienen ninguna percepción interna de un campo visual, y que sin embargo son capaces de adivinar un color que se les muestre. Lo hacen en un número de casos superior al que sería esperable si acertaran por puro azar. Es importante destacar que ellos no ven los colores en el sentido en que nosotros somos capaces de verlos: no tienen ninguna sensación visual de los colores. Es un fenómeno muy extraño, que parece implicar una separación entre la vista propiamente dicha y la percepción visual. Se supone que hay alguna clase de información que fluye desde el color hasta una parte de sus cerebros. Eso les permite saber de cuál se trata, aunque esas personas no tengan una capacidad consciente de distinguirlos. Casos como estos hacen pensar que quizá tengan razón los que sostienen que la consciencia es tan solo un elemento de la arquitectura neuronal.

Suponiendo que lleguen a desarrollarse sistemas artificiales realmente inteligentes, ¿es de esperar que los futuros humanos hagan algo así como extensiones de sus cerebros o implantes en ellos? ¿Eso los convertirá en una especie distinta?

Este es justo el tipo de cosas de las que habla Ray Kurzweil en su libro “*La singularidad*”, que comenté antes. Según él, lo que va a ocurrir en primer lugar es que los ordenadores se volverán tan potentes que serán capaces, en esencia, de simular una inteligencia del nivel de la humana. Llegado ese punto, podremos transferir a un ordenador o un robot toda nuestra actividad y vida mentales. De hecho, usaremos ese *hardware* en lugar de nuestros propios cuerpos, que tienen una esperanza de vida limitada y van deteriorándose con el tiempo. Ya existe hasta cierto punto esa mezcla de partes humanas y artificiales, como piernas o incluso oídos, y es de esperar que eso vaya siendo cada vez más sofisticado y que más partes del cuerpo puedan ir sustituyéndose por otras artificiales. Una explicación de cómo podría llegar a transferirse la propia consciencia a un soporte artificial es que se haga por fases. Primero se reemplaza una parte del cerebro por un dispositivo electrónico que sea funcionalmente equivalente y

capaz de interactuar con el resto del cerebro. Luego se sustituye otra parte, en las mismas condiciones, y así hasta que el cerebro acabe siendo reemplazado totalmente por dispositivos artificiales interconectados. Si en ningún momento de ese proceso la persona cree que es distinta de lo que era, entonces se habrá conseguido transferirla a un soporte artificial.

Sería algo parecido al hecho de que no dejemos de ser quienes somos aunque cada siete u ocho años se renueven todos los átomos de nuestros cuerpos...

Exactamente, así es. En ese caso sustituiríamos las partes orgánicas por componentes electrónicos de silicio. En teoría eso sería posible. Pero existen tantos problemas prácticos para llegar a ese punto que, en mi opinión, eso va a llevar mucho más tiempo del que vaticina Ray Kurzweil.

Varias películas de ciencia ficción han tratado el tema de un futuro más o menos lejano en el que las máquinas se hacen más inteligentes y poderosas que los propios seres humanos, y llegan a plantearse qué necesidad hay de que sigamos existiendo. ¿Debería eso preocuparnos? ¿Tendrían esas máquinas el mismo derecho que nosotros a considerar suyo este mundo y todo lo que hay en él? ¿Es algo a lo que nos conducirá el desarrollo de la inteligencia artificial?

Que lleguemos a ser capaces de crear máquinas más inteligentes que nosotros no implica que serán mejores. Hay personas muy inteligentes que son malas personas. Puede que esas máquinas no tengan el menor escrúpulo moral para exterminar a los humanos. Es posible. Esa es la premisa de la serie de películas “*Terminator*”. Por ejemplo, nosotros somos mucho más inteligentes y capaces que las vacas, y no tenemos ningún problema en matarlas. La cuestión es si existe alguna diferencia cualitativa entre el estatus moral de los humanos y el de esas hipotéticas máquinas más inteligentes que nosotros, como ocurre entre los humanos y las vacas. Se considera que el grado de inteligencia de las vacas no es suficiente para que ellas puedan hacer recriminaciones morales a los humanos. Afirmar lo mismo de éstos frente a dichas máquinas hiperinteligentes no parecería, en cambio, razonable. Se trata de una línea de pensamiento con la que debe tenerse cuidado, porque nos coloca justo en el punto de mira de una serie de complicados juicios morales.

¿Existe ya algún sistema de inteligencia artificial en funcionamiento?

Todo depende de a qué nos refiramos con “inteligente”. Para responder a esa cuestión es necesario dis-

poner de un método para determinar si un sistema es o no inteligente, un criterio para establecer la inteligencia. Y el único criterio que yo conozco que me dejaría a mí satisfecho en ese sentido es, a pesar de sus limitaciones, el test de Turing. En este sentido la respuesta sería no: no hay ningún sistema que haya sido capaz de superar, ni de lejos, este test. El test de Turing es una condición suficiente para atribuir a algo inteligencia, pero no es una condición necesaria. Me refiero a que alguien podría argumentar que es demasiado exigente y que el hecho de que ningún sistema artificial haya conseguido superarlo no implica que éste no sea inteligente. Es un argumento admisible en principio, aunque la cuestión es que no existe ninguna otra prueba

que sea del todo convincente. No es fácil crear pruebas válidas. Por ejemplo, podría creerse que hace falta ser bastante inteligente para jugar al ajedrez. Sin embargo, mi teléfono móvil juega al ajedrez muchísimo mejor que yo. Pero eso no le hace ser inteligente. Esto nos da una idea de que el ajedrez, u otros juegos como el póker, no son una prueba válida para demostrar la inteligencia de un sistema artificial. Sería estupendo tener alguna alternativa intermedia [entre el test de Turing y esas otras pruebas no válidas], pero no la hay, que yo sepa... A corto plazo no vamos a ver nada que sea capaz de superar el test de Turing y demostrar así su inteligencia de un modo inequívoco. Por debajo de eso, sabremos que algo es inteligente cuando lo veamos.

