

42 preguntas para una Inteligencia Artificial.

Luis Montero

1. ¿Qué es la Inteligencia?

Hay muchas preguntas que una máquina nunca podrá contestar de forma satisfactoria o definitiva. Esta, seguramente, es la primera de ellas. Si en el equipo de programadores hay algún gracioso, la máquina tirará de ironía. Algo así como «*La inteligencia soy yo*». Respuesta, por otra parte, bastante cartesiana y que se ajustaría a los presupuestos metafísicos ilustrados, que a la pregunta «¿Qué somos?» responde con un quién. Pero para que eso se diera, la máquina tendría que superar el límite de lo pensable por una máquina y dar respuesta a una frontera que, de momento, es intraspasable para una máquina: el problema de la parada. Para una máquina no hay nada más terrible que un algoritmo que no finalice. Si el fin del algoritmo es el cumplimiento de la orden que lleva implícito, que no pueda cumplirse esa orden supone que el algoritmo no se detenga y que la máquina entre en bucle indefinidamente... Algo que, como demostraron Church o Turing, es magnífico para el conocimiento de las matemáticas pero una monstruosidad para la máquina que se pretende inteligente. En definitiva, la máquina sólo puede pensar dentro de lo que puede pensar. Está mucho más encerrada que nosotros dentro de una estructura lógica capaz sólo de garantizar el resultado del algoritmo, nunca de superarlo. Quizá la ironía no sea más que una de las formas del humano de dar el *salto combinatorio* del que hablaba Dennett, que no es más que un atajo, la capacidad del cerebro humano –y animal– de tomar una decisión sin la ejecución completa del algoritmo.

2. En serio, ¿qué es la inteligencia?

No tenemos ni idea porque, entre otras cosas, la inteligencia no es un algo determinado y por eso lo llamamos inteligencia sino que a los resultados de un conjunto dispar de procesos cognitivos los llamamos *inteligencia*. No hay una inteligencia absoluta salvo que se sea monoteísta. Y, dependiendo de qué resultados interesen más o menos, varía tanto la composición de ese conjunto como los gradientes de cada uno de sus componentes. La atribución de inteligencia siempre es un *ex post*.

De hecho, la inteligencia siempre ha sido un problema filosófico menor. No es comparable con las ideas de *mente* o de *conocimiento*, que han movido el pensamiento, al menos el occidental, durante siglos. La inteligencia es poco más que ese resultado al que, una vez alcanzado, llamamos inteligente.

Otra cosa es que ahora se haya puesto de moda; pero la moda no lo convierte en más relevante. Si acaso todo lo contrario. De hecho, y aunque se va un poco del asunto, la moda aparece cuando la evolución del objeto de moda se ha detenido. Pero esa es otra historia.

3. ¿Entonces nadie puede anticipar que una acción será inteligente?

Exacto. Nadie, salvo esa divinidad –cuya atribución de inteligencia no es más que una convención y que, irremediabilmente, constituye una petición de principio:

dios no puede ser sino inteligente y por eso es inteligente–, puede anticipar que el resultado de su acción, sea esta un acto o un enunciado, va a ser inteligente. Inteligente lo es después, una vez acometido ese acto. Y lo es, además, porque son los demás, lo otros, quienes califican esa acción como inteligente. La imposición de inteligencia es imposible o, por lo menos, muy difícil. De ahí la estrechísima relación de parentesco entre *inteligencia* e *inteligibilidad*. La inteligencia requiere comprensión. Que es penetrar en el asunto, sí; pero también hacerse penetrar con los demás. De ahí el prefijo *com-*. Com-prender. Quizá por eso todas las definiciones de inteligencia, que no pueden sino eludir el escollo que supone el otro, no pueden ser sino recursivas; como lo es la presunta respuesta de la máquina, «*la inteligencia soy yo*». Respuesta imposible o, más bien, inútil: porque es una respuesta sin los demás y sin contexto.

No, la inteligencia no es un absoluto. Una inteligencia absoluta sería incomprensible. Ininteligible. Inescrutable, que se decía antes.

4. **¿Inteligencia es un acuerdo?**

Más que un acuerdo es una convención. Pero no es una convención muy interesante. Por un lado, porque no es estática. Ninguna lo es, pero la inteligencia menos. ¿Y por qué no es estática? Porque la inteligencia, si se extiende demasiado, deja de serlo. La inteligencia general se niega a sí misma. Un acto inteligente, en el momento en que es un acto repetido por todos de forma automática, deja de serlo: se convierte en hábito o costumbre. Es inteligible pero no inteligente, no incluye la sorpresa, la imaginación o la fantasía, que son otros de los atributos imprescindibles de la inteligencia. Y, por otro lado, es mediante esa negación de sí misma por lo que avanza la inteligencia y, en su avanzar, consigue romper el estatismo de la convención. Lo que hace doscientos años llamábamos inteligencia hoy no lo consideraríamos inteligente, pero aún así seguimos considerando que existen comportamientos y proposiciones inteligentes.

5. **¿Y qué hace que convengamos que algo es inteligente?**

Que cumpla determinados atributos. Pero como decía antes, esos atributos varían con el tiempo y las necesidades de los tiempos. Por ejemplo, hace dos siglos, por seguir con el ejemplo, la inteligencia no tenía por qué ser eficiente. Hoy no consideraríamos que una respuesta fuera inteligente si, entre la combinación de atributos que la componen, la eficiencia no ocupara un lugar destacado. Pero no sólo eso. También la inteligencia, de alguna manera, anticipa la siguiente convención de la inteligencia, si no quiere verse condenada a convertirse en hábito o costumbre, que son su negación total. Para que un acto o un enunciado sean considerados inteligentes tienen que anticipar, en alguna medida, lo que las siguientes generaciones convengan que es inteligente. Al negarse a sí misma – aunque sea una negación parcial, blanda– se anticipa y consigue evolucionar. Esa es la dificultad, y la gracia, de la inteligencia.

6. **Entonces no tiene mucho sentido hablar de Inteligencia...**

Bueno, sí lo tiene; aunque quizá no en mayúscula... Inteligentes son todas esas acciones o enunciados que incluyen todos esos atributos mencionados y muchos otros. Tantos como sean exigidos, por los demás y por el contexto, en cada momento determinado. Es más, que varíen tanto los atributos como la graduación entre atributos, no quiere decir que no tenga sentido hablar de inteligencia, sino que

el contexto cambia y, con él, los demás y esa parte de los demás que es uno mismo. Y, entender ese contexto, es otra manifestación de inteligencia. Lo que no tiene sentido es hablar de una inteligencia en abstracto, absoluta, universal... La inteligencia es contingente y, de ahí, la posibilidad de que pueda existir una Inteligencia Artificial (IA, a partir de ahora).

7. **¿Por qué se llama Inteligencia Artificial?**

Porque es el producto de un arte humano, en contraste con lo producido por lo dado, ya sea el resultado de un acto creador por parte de una divinidad o de la evolución. En cualquier caso, la inteligencia humana ha dejado de ser natural hace mucho tiempo. O, quizá, una vez que cualquier inteligencia deja de soportarse en lo dado para empezar a habitar un espacio simbólico, entonces deja de ser natural. En ese sentido, la inteligencia de los delfines y chimpancés también sería artificial, aunque no fuera resultado del arte humano.

8. **¿Será realidad alguna vez la IA?**

Sí, pero con muchos *peros*... Es decir, será realidad esa inteligencia que hayamos decidido que *es* la inteligencia. Al menos mientras la IA dependa del humano para la definición de qué es inteligencia. Una vez que la máquina sea capaz de aprender y retroalimentarse, entonces quizá ampliará la definición de inteligencia de acuerdo a sus capacidades. Pero entonces ya no será tanto artificial, entendido como producto de un arte, sino autónoma. Se producirá a sí misma.

9. **¿Cuándo?**

Ya. Si la definición depende de los resultados conseguidos por esos procesos cognitivos entonces ya ha sucedido. Las máquinas hoy producen resultados con una rapidez y exactitud inaprensibles para el humano. Basta una calculadora de unos pocos euros para darse cuenta.

10. **¿Entonces ya se ha producido la *singularidad*?**

Sí y no. Sí, por las respuesta 9. No, porque aún no han tomado el poder. Como en el caso de la inteligencia, depende de lo que entendamos por *singularidad*. En cualquier caso, es muy posible que no exista una sola singularidad, un momento concreto en el que se de ese relevo entre inteligencias –de la humana a la de la máquina– sino que se produzcan muchas singularidades pequeñas. Quizá la singularidad sea la suma de avances menores de la IA. Y no nos demos cuenta de ello.

11. **¿Qué es la *singularidad*?**

El momento en que la IA sea mayor que la inteligencia humana. Sobre ese escenario se han trazado innumerables proyecciones, algunas más pesimistas que otras. Pero todas tienen algo en común: nunca una especie menos inteligente ha sometido durante mucho tiempo a una más inteligente. La posibilidad de una IA autónoma, independiente del humano, aterroriza al humano. Es normal.

12. **¿Ese es el miedo a la *singularidad*?**

«¿Qué es *miedo*?», debería contestar aquí la IA. Y haría bien. Si entre humanos es complicado definir el miedo, para una máquina es prácticamente impensable. O, mejor dicho, a día de hoy el humano, que es quien va a programar esa IA, no tiene

ni idea de cómo generar algo parecido a la sensación de miedo, si es que hiciera falta.

13. Me refería al miedo humano, no al miedo de la máquina...

Si quieres mi opinión, ese miedo a la singularidad no es más que un exceso de optimismo. Porque somos optimistas en relación a nuestras capacidades cuando pensamos que podemos crear una máquina como nosotros o, al menos, capaz de emularnos... Que podamos, en definitiva, ser demiurgos, por fin. Construir esa máquina es colocarnos en el atril en el que nos gustaría estar, pero eso no quiere decir que sea el atril que nos corresponda. Ni siquiera que tenga que haber atril alguno. ¿Y el miedo? El miedo. Todo nos da miedo. Tanto la posibilidad de poder construir esa máquina como la posibilidad de no poder hacerlo. Además, ese miedo también es reflejo de otra cosa: del miedo a nosotros mismos. Desde el renacimiento, una vez que el hombre se emancipó de dios, la filosofía ha jugado el papel de muro de contención del propio hombre. Algunos pensadores lo han hecho de forma suave, como Rousseau, que a pesar de confiar en la naturaleza del buen salvaje, pensaba que el grupo era la mejor forma de cohabitar para el humano, pero muchos otros lo han hecho de forma fuerte, como Hobbes, que popularizó la famosa locución de Plauto *homo homini lupus*, el hombre es un lobo para el hombre. Hoy la interpretación sigue siendo la misma: el humano necesita un muro de contención frente a otro humano. Y, más allá del otro humano, frente a sí mismo. O eso pensamos, sobre todo después de atrocidades como Auschwitz o Nagasaki. Pero esto no tiene por qué ser así. O parafraseando a Nietzsche, aunque él se refería al catolicismo: a base de pensar que el mundo es feo hemos hecho un mundo feo. Quizá, si las máquinas llegan a ser tan inteligentes que nos superen, no tengan la intención de someternos. El juego de poder entre especies de distinta inteligencia se debe, entre otras cosas, a que ninguna de ellas lo era mucho. La prueba son, otra vez, esos Auschwitz y Nagasaki. Quizá una especie mucho más inteligente no necesite dominar al resto de especies: simplemente las deje estar ahí, sin más.

Pero quizá esa sea la perspectiva más aterradora. No un mundo dominado por máquinas tipo Terminator, sino un mundo poblado por máquinas como ángeles, indiferentes al humano. Sería como estar enfrentados para toda la eternidad ante un espejo que sólo reflejara nuestra propia minusvalía. Un desenlace aterrador para el humano, acostumbrado a vivirse como el centro del universo, pero que no carece de lógica metafísica...

14. ¿Podría darse ese escenario?

Sí, en el caso de que la IA se desarrollara, es más que probable. Pero tendría que desarrollarse en todas sus consecuencias, que por cierto es la única forma de desarrollo posible. Si no, no sería inteligente. Y, sólo a partir de un desarrollo completo, podría darse ese escenario. O, mejor dicho, una vez que se diera ese desarrollo completo, que es muy improbable, se daría ese escenario casi obligatoriamente.

15. ¿Por qué?

Porque, como ya he dicho antes, la inteligencia evoluciona a base de negarse a sí misma. Una vez que la inteligencia se convierte en hábito, es la propia inteligencia la que genera respuestas alternativas produciendo así saltos en la comprensión. La

evolución de la inteligencia se podría dibujar como la negociación entre el hábito comprendido y la solución aún no comprendida del todo. Para que evolucione, la inteligencia tienen que eludir tanto la comprensión excesiva como la incomprensión total. Para que se de y evolucione tiene que moverse entre esos dos extremos. La inteligencia requiere de cierta ininteligibilidad para darse, pero no se da si existen demasiada comprensión y demasiada incomprensión. Así, si las máquinas llegaran a poseer IA real, esta empezaría irremediabilmente a evolucionar al margen de nuestra comprensión e incomprensión: esa inteligencia crearía su propio contexto, a partir del cual evolucionaría, independiente y autónoma. Es un escenario obligado: si la inteligencia es una atribución que otorgan los demás, es obligado que esos demás acaben siendo las demás máquinas. Y, si la inteligencia, crea su propio contexto, ese contexto será el de las demás máquinas. No el nuestro.

16. ¿Hay remedio?

Sí; pero antes: una reflexión sobre el contexto. La inteligencia, dado que es contingente, requiere contexto. Y no sólo eso: depende del contexto. Para que se pueda desarrollar algo parecido a una IA necesita que creamos el contexto para que pueda darse. Y, ese contexto, no es otra cosa que el mundo descrito por lenguajes formales. El mundo que ejecutan algoritmo y completan procesos. La IA necesita ese mundo formalizado para darse. Y llevamos casi tres siglos creando ese mundo, un mundo cuantificado y ordenado, capaz de ser descrito únicamente por lenguajes formales. Y, por ende, de ser entendido únicamente por lenguajes formales. Los ordenadores hacen eso, correr lenguajes formales en forma de código, y su actividad se resume a eso, a correr esos lenguajes formales en forma de código. Para que pudiera darse la posibilidad de la IA hemos reducido el mundo y esa reducción nos excluye: para nosotros los lenguajes formales son ininteligibles. No hay comprensión posible de ese mundo desde los lenguajes naturales y, más interesante, no es necesaria para que ese mundo siga ejecutándose. No podemos, por tanto, comprender el contexto en el que se mueve una IA: solo entendemos el resultado de su acción cuando incide en el mundo.

La dificultad, y esto es algo que los ingenieros de la IA tienen que resolver si quieren que la IA sea posible, es que ese contexto es estático. Y si es cierto que la inteligencia requiere un mundo dinámico, que plantee retos nuevos y exija soluciones nuevas, el mundo formalizado no lo es. Incluir el cambio en esas descripciones del mundo formales es el principal reto de la ingeniería si quiere que esa IA sea, de verdad, inteligente. Y, me temo, que eso no es un problema que los ingenieros puedan resolver: El movimiento tiene muy mal encaje en Lógica, como puso de manifiesto la crítica de Kierkegaard a Hegel y sus seguidores. Los ingenieros pueden soslayar esa dificultad, hacer como que está resuelto y esperar que todo funcione; pero no pueden evitar que, en el fondo, ese mundo formalizado sea el mismo una y otra vez. Ni siquiera una posible sucesión de mundos que son el mismo generarían cambio alguno. De no resolver ese problema, la IA corriendo sobre un mundo formalizado, no puede convertirse en un aparato artrítico, cuyo destino es oxidarse sobre sí misma, en una mera repetición perpetua de lo mismo conseguido tras lo mismo conseguido, que es, de todas formas, el destino de quien cree que el mundo ya ha sido completado y ha alcanzado su culminación. Y, ahí, la IA sería el factor de afianzamiento de una situación estática que, fijada por los propios resultados de la IA, continuaría inmóvil a perpetuidad. Dicho más claro, la

IA sería el ataúd de una realidad muerta. Que no sería sino la negación de la inteligencia, al menos en los términos en los que los hemos estado hablando.

Quizá sea su mayor amenaza. Y, pensándolo bien, quizá sea nuestra mayor amenaza, que la IA se convierta en soporte y validador de un mundo estático sin capacidad, no ya de progreso, sino de evolución alguna. Que el mundo que hoy vivimos no varíe en lo sustancial, cimentado en las lógicas formales que lo describen y los algoritmos que lo ejecutan y, por tanto, sin posibilidad de cambio alguno. Un mundo autorreferencial, creado y mantenido por sí mismo, mediante la eficiencia de esos lenguajes formales y algoritmos. Un mundo exactamente igual que el de ahora, con las mismas desigualdades, incertidumbres y desazones para nosotros, pero más enraizado en sí mismo gracias a la eficiencia de las máquinas. Sí, quizá ese mundo paralizado y paralizador sea el peor futuro que nos puede deparar la IA.

17. ¿Y cuál era el remedio?

El remedio es el mismo que para la escenario que mencionaba antes, de hecho ese mundo que nos excluye y ese mundo anquilosado de la IA y los lenguajes formales es el mismo: El ciborg. La combinación de la inteligencia humana y artificial en un cuerpo-máquina. Lo que nos llevaría a otra pregunta metafísica: ¿ese ciborg sigue siendo hombre o ya es otra cosa? Y, como todas las preguntas metafísicas, las posibles preguntas muchas veces producen pánico... Otra cosa, más allá del pánico, es que sí, en el fondo, la pregunta por la IA y por el ciborg no sean la misma pregunta: el miedo a dejar de ser lo que somos o, mejor dicho, el miedo a dejar de ser lo que creemos que somos: la inteligencia más avanzada. La posibilidad de la IA pone en cuestión años de antropocentrismo metafísico, y no tan metafísico, y eso nos da pavor.

Pero, al mismo tiempo, sólo el ciborg puede sumar la vida ahí donde no la hay. Sólo el ciborg puede incorporar el cambio, el movimiento al mundo formalizado. Aunque esa perspectiva abre otra pregunta sobre la tecnología y sus promesas...

18. Pero hay miedos más tangibles, menos metafísicos. Que parecen estar a la vuelta de la esquina. Como la posibilidad de que las máquinas tomen decisiones sobre la vida y la muerte de los humanos...

Aún queda mucho hasta que una máquina tome la decisión autónoma de matar a un humano. Y, sobre todo, que sepa que está matando a un humano. De momento, y eso ya se puede hacer, una máquina podría disparar contra un bulto que se está moviendo pero poco más. Es decir, podría ejecutar una tarea que, como resultado, conllevara la muerte de un humano. Pero no sé si a eso se le puede llamar una *decisión de matar*. De hecho, creo que es todo lo contrario de una decisión. Y mucho menos inteligente. Está más cerca del acto instintivo, que es lo que son los algoritmos, que del acto reflexivo capaz de preguntarse por el acto en sí y, por tanto, sus consecuencias. Nos matarán, sí, pero como nos mata un leopardo que tiene hambre. O ni siquiera...

Eso que llaman IA está aún muy lejos de ser mínimamente inteligente.

19. Sería eso que llaman IA débil...

Desde que Searle introdujo la distinción entre dos tipos de IA, la IA *fuerte* y la IA *débil*, la clasificación ha sido ampliamente aceptada por la academia y, sobre todo, por la industria y los medios de comunicación. Tanto que se ha convertido en uno de los tópicos de la categoría. Y, como todos los tópicos, tiende a vaciarse de contenido con cada nuevo uso y, por tanto, a explicar cada vez menos. La IA fuerte enuncia que todo lo que acontece en la mente de un humano es un programa y, como tal, puede ser realizado por una máquina. En las palabras de Searle, que no comparte: «una computadora apropiadamente programada con entradas y salidas correctas tendría con tal motivo una mente en exactamente el sentido que los humanos tienen mentes».

La IA débil no requiere esa profundidad, que la máquina sea el espejo de la mente del humano, bastaría con que la máquina fuera capaz de actuar como si fueran inteligentes. O, dicho de otra forma, los resultados de los procesos cerebrales pueden ser alcanzados por una computadora sin que sea necesario el isomorfismo mente-cerebro.

Obviamente, cuando la industria y los mercados hablan de IA hablan de esta última.

20. ¿Y tú por cuál optas?

La IA fuerte parte de una premisa falsa. O, dicho de modo más suave, es muy probable que el cerebro no sea un ordenador; o si lo es, se parece tan poco a los ordenadores tal y como hoy los conocemos que es muy imposible afirmar que es un ordenador. Hay infinidad de procesos que afectan a la inteligencia, o a lo que el humano entiende como inteligente, que no suceden estrictamente en el cerebro o, al menos, dentro del cerebro. El *cerebrocentrismo* –la tesis de que toda la vida mental sucede exclusivamente en el cerebro– es bonito, enmarca el problema y lo reduce a causalidades físicas, pero es muy probable que no sea cierto. Digamos que hay factores externos al cerebro que repercuten en un comportamiento inteligente.

21. ¿Cuáles serían?

Por un lado hay factores físicos, que se podrían resumir en una frase: «*no hay, no conocemos ningún sujeto inteligente que carezca de vida*». Piensa que cuando buscamos inteligencia extraterrestre hablamos de «vida inteligente», no sólo de inteligencia. Algo que llevado más allá, puede entenderse como: la inteligencia es un resultado de la vida, una respuesta. Un subproducto. Quizá se pueda definir lo inteligente como una respuesta adecuada a los problemas que plantea la vida. O, dicho de otra forma, sin homeostasis no hay inteligencia. Quizá porque sólo la vida es capaz de generar las condiciones para el desarrollo de respuestas inteligentes. Por un lado, porque es la vida la que plantea las preguntas necesarias para la generación de respuestas inteligentes. Si de un comportamiento depende de supervivencia, más vale que esa respuesta sea inteligente. Y, por otro, porque sólo la vida es capaz de producir la presión necesaria para que se den comportamientos inteligentes: fuera de los laboratorios la supervivencia aprieta bastante. ¿Es posible la inteligencia sin vida? Queremos pensar que sí, pero no lo sabemos. Y toda evidencia muestra que no. En el extremo, estamos ante una disyuntiva muy compleja: o bien creamos inteligencia de laboratorio o bien creamos vida de laboratorio. Y ambos logros, hoy, son muy poco probables. Porque tanto la vida como la inteligencia son dos conceptos de límites difusos, que no tenemos ni idea

dónde empiezan ni, por supuesto, dónde acaban. ¿En qué momento algo empieza a estar vivo? ¿Un virus, está vivo? ¿Y en qué momento una respuesta es inteligente? ¿Esta respuesta es inteligente? ¿O lo es la pregunta? O, dicho de otra forma, no tenemos ni idea de qué es lo uno ni lo otro. Y puede que no lo sepamos nunca: tanto la vida como la inteligencia se definen por convenciones. Es muy probable que para resolver el problema de la inteligencia haya que resolver también el problema de la vida. Quizá no sean si no la misma pregunta. O que la pregunta por la vida sea más interesante, siendo la pregunta por la inteligencia un componente de la primera.

22. ¿Y puede, por tanto, que no haya respuesta?

Puede. Pero eso sería irrelevante. O, dicho de otra forma, que no haya respuesta no va a impedir que sigamos buscándola.

23. Pero te he interrumpido...

Decía que quizá no haya una respuesta a la pregunta por la Inteligencia independiente de la respuesta a la pregunta por la Vida, que es algo semejante a lo que planteaba Bergson. Quizá no las haya nunca, al menos de modo definitivo. Pero eso no es más que otra oportunidad de ir afinando las preguntas.

Y, quizá una muy buena forma de afilarlas es cuando Chalmers habla de la *inteligencia extendida*. O la posibilidad de que la inteligencia no resida sólo en el cerebro y se encuentre distribuida a lo largo y ancho del cuerpo. Aunque su aproximación es una visión extrema de aquella vieja idea que casaba inteligencia con el sistema motor, es muy probable que la encarnación de la inteligencia no se ciña al cerebro únicamente. O, por reformular la pregunta en términos más fisiológicos: ¿dónde acaba exactamente el sistema nervioso? ¿En la punta de los dedos? ¿En las ramificaciones nerviosas de la epidermis?

Pero no sólo hay factores físicos, también los hoy simbólicos, también externos al cerebro. La inteligencia estanca, abstracta es más un producto simbólico desarrollado por la creencia en una inteligencia absoluta que una realidad. La inteligencia, para desarrollarse como tal, requiere estar inmersa en una *sopa simbólica*, un medio de cultivo semiótico y semiológico, desarrollado a su vez por la propia inteligencia, pero ajeno al cerebro, un factor externo situado fuera del cerebro: ese mundo de signos y símbolos en el que nos movemos y que constatan que vivimos en una sociedad inteligente –que por cierto no es más que una autoatribución, puesto que hasta donde yo sé ninguna otra especie, ni terrícola ni extraterrestre, nos ha definido como tal– y al que Bartra se refiere como el *exocerebro*. Wittgenstein decía que conocer el proceso fisiológico de lo que sucede en el cerebro no iba a explicar las ideas que corren por ese mismo cerebro. Y, aunque esta ha sido una afirmación muy discutida, quizá haya algo de razón. Supongamos que algún día, y sin duda ese día llegará, sepamos con suficiente precisión qué sucede en el cerebro cuando un sujeto ve el color azul. Qué centros del cerebro se activan, qué conexiones neuronales se establecen y qué información produce. Sin embargo, hará unos cuatro siglos cuando alguien veía el color azul no lo veía como azul. El azul es, de todos los colores, el más reciente. Como prueba de ello, en la Odisea no hay ni una sola mención a ese azul. Hasta finales del medioevo no aparece como tal. Lo que no quiere decir toda la gente que vio el cielo hasta entonces no viera el cielo como lo vemos nosotros, pero no lo veía como azul. No, no basta con un conocimiento preciso de lo que sucede en el cerebro para

solucionar los problemas ni del cerebro ni de la inteligencia. La causalidad física exclusivamente nunca va a responderlo todo.

No, el *cerebrocentrismo* reduccionista no es la solución.

Y la IA es una solución extremadamente cerebrocéntrica. Quizá ahí radique su mayor debilidad.

¿Quiere esto decir que hay que volver a una perspectiva *mentalista*, más cartesiana? Rotundamente no. Pero quizá sí que hay que empezar a aceptar que el cerebro, y por tanto la inteligencia, no se acaba en las paredes craneales...

24. No has respondido a mi pregunta: ¿tú por cuál optas?

Bueno, a pesar de todas las interrupciones, yo creo que sí he respondido.

25. No.

¿Por qué crees que no?

26. ¿Por qué creo que no qué? ¿Qué no has respondido a mi pregunta o que por qué creo que no has respondido a mi pregunta?

¿Tú qué crees?

27. Que no has respondido a mi pregunta.

Yo creo que sí.

28. ¿Por qué crees que sí?

Yo he hecho la pregunta antes.

29. Pero soy yo el que hace las preguntas.

Y yo quien, como entidad supuestamente inteligente, las reformula, que es una de las capacidades que hoy se le exigen a la inteligencia. Y, a partir de ahí, decide cómo las contesta. Y si es mediante otra pregunta, así sea. Aunque, que yo me salte el protocolo, es probable que para ti no sea un indicio de inteligencia, sin embargo para mí, y espero que para otros lectores, sí es señal de un comportamiento inteligente.

Pero sigamos.

30. ¿Entonces, optas por la IA fuerte o la débil?

Obviamente, por ninguna. O no de momento. ¿O cómo vamos a emular algo que no sabemos qué es? Y aún en caso de generar una emulación satisfactoria, ¿cómo vamos a poder definirla como inteligente si no sabemos qué es la inteligencia? Claro, que este razonamiento es aplicable a nuestras propias acciones, así de esquivo es el concepto. Así, hay especialistas que afirman que basta con conseguir la representación de un comportamiento inteligente para que éste sea inteligente. Así un termostato es capaz de reaccionar a la temperatura ambiente y tomar decisiones... luego es inteligente.

Quizá. Aunque yo diría que más que inteligencia hemos emulado al instinto. La IA muchas veces se parece más al instinto que a la inteligencia. Es más un *instinto*

*artificial que una IA. Y, si volvemos a Bergson, hay un mundo de diferencia entre uno y otro: lo que el francés llamaba *lo geométrico*. Y, como partícipes de mundos distintos, comparten pocos puntos en común.*

31. ¿Por qué «no de momento»?

Porque no contamos con el conocimiento necesario. Eso lo sabe cualquiera. Y, derivado de este, porque no contamos con las metáforas adecuadas para que se de la IA, o una IA lo suficientemente potente como para que, de verdad, sea considerada como inteligente de una forma general. La metáfora del cerebro de la IA fuerte falla, eso es obvio. La emulación de inteligencia de la IA es eso, una emulación. Y como tal, puede ser útil para la ejecución de determinadas tareas, que es lo que están haciendo hoy las máquinas: desde la toma de decisiones financieras en el *High Speed Trading* hasta ganar el concurso *Jeopardy*, como ha hecho Watson, o al campeón del mundo de ajedrez, como ha hecho *Deep Blue*. Sí, todo son progresos muy llamativos, casi milagrosos, pero conceptualmente estamos donde estábamos con la máquina de calcular de Leibniz, no mucho más allá.

32. Sin embargo, la respuesta generalizada ha sido la contraria.

Sí, y ese es un punto muy revelador. Si para la filosofía esas máquinas no suponen un avance magnífico, tanto la industria como los mercados no paran de contarnos que esa IA es el futuro, que es un avance fundamental para la especie. Estamos en un escenario curioso: la inteligencia ya no es definida como tal por la inteligencia, sino por quienes tienen puestos sus intereses económicos en que eso sea así. Quieren, necesitan que eso sea así para seguir ganando dinero que más que por eso sea así. Si escuchas a muchos emprendedores de Silicon Valley, su discurso es una renovación del discurso milagrero del producto milagroso.

33. ¿Una especie de argumento de venta?

Sí, y ni siquiera muy novedoso. Por mucho que la industria y los mercados insistan, necesitamos mejores metáforas. Y estas no pueden darse sino a partir del conocimiento, que hoy es promovido por la ciencia. Y, prueba de que se está agotando la fuerza expresiva de la metáfora del cerebro como un ordenador, basta recordar un hecho al que se la ha dado muy poco pábulo: la última década del siglo pasado fue declarada por el gobierno de Estados Unidos como la Década del Cerebro. Después del éxito de la Década del Genoma Humano –que, por cierto, conviene no olvidar que esa década fue declarada como tal contra los intereses de la industria y, en cierta medida, de los medios que viven de esa industria: el genoma humano no podía ser patentado por una empresa privada– se pretendió alcanzar los mismos resultados pero, esta vez, centrados en el cerebro. ¿Qué se consiguió? Bueno, básicamente nada. Tras diez años de investigación, y después de invertir presupuestos astronómicos, la única conclusión es que nadie sabe nada y, lo que es peor, que nadie sabe quién tiene que saberlo: los científicos acabaron concluyendo que el cerebro es un problema filosófico y los filósofos que es un problema científico. Aún no sabemos nada; ni siquiera sabemos cómo deberíamos saberlo. Necesitamos metáforas pero ni siquiera sabemos dónde buscarlas. Esa necesidad de metáforas es, quizá, uno de los mayores problemas a los que se enfrenta la inteligencia humana, mucho mayor que el de la IA que, por otra parte, hoy se ha visto reducido a un problema básicamente operativo.

El problema es aún mucho peor que lo mejor que pueda hacer una máquina hoy. E, incluso, en el futuro.

Y, por eso, es un problema mucho mejor.

34. ¿Por?

Porque abre la puerta a un nuevo modelo, más allá del modelo de progreso ilustrado, basado en una epistemología que, a través de la cuantificación del mundo, lo ordena. El agotamiento de las metáforas no es el ocaso de las mismas sino el amanecer de la búsqueda de nuevas. Pero quizá aún ni conozcamos los paradigmas desde los que haya que buscarlas.

Quizá, que sólo hablen de IA la industria y los medios, mientras la ciencia y la filosofía intenten echarse pasarse la patata caliente una a la otra no sea sino la primera muestra de que esa metáfora del cerebro como un ordenador está gastada. O que está empezando a gastarse. Y, dado que a la industria y a los medios lo que le interesa, mucho más que la propia IA, es ganar dinero, la metáfora sólo resistirá mientras produzca dinero. Pero como no hay promesa irredenta eterna, puede que el fracaso de la IA puede convertirse, en realidad, en una victoria. Una que nos libere tanto de la metáfora del cerebro como ordenador y de su estructura de soporte: el dinero. Porque quizá, esa derrota, sea la demostración definitiva del final del sueño cuantificador racionalista. Lo que no estaría mal.

Y ya, que ya estoy cansado. Y los lectores, si es que alguno ha llegado hasta aquí, supongo que más

35. ¿Ya?

Sí. Ya.

36. ¿Pero no iban a ser 42* preguntas? Bueno, tómatelo como una prueba de que hasta la IA más desarrollada puede equivocarse. Empieza a aceptarlo.

37. ¿En serio me vas dejar así?

¿Así cómo?

38. A cinco preguntas del final...

¡Qué pesados sois los humanos, siempre insistiendo en conseguir redimir todas las promesas! Esa necesidad vuestra de ver cualquier promesa hecha realidad os va a matar... Pero, venga, una pregunta más. Sólo una. Así que aprovecha.

39. ¿Cuál es el futuro de la IA?

Singularidades aparte, digamos que, en el medio plazo, hay dos futuros probables, y ambos implican que la IA ha triunfado. Y no son excluyentes.

En el primero, al que ya hemos sido invitados, es muy probable que asistamos a una depreciación de la idea de inteligencia. O, cuando menos, a una instrumentación. Pero no es un fenómeno nuevo. Es algo que lleva pasando décadas y que tiene su epítome en los malogrados Test de Inteligencia y su corolario en el estrambótico

* Si no sabes qué es 42 es que no has leído The Hitchhiker's Guide to the Galaxy. Ya estás tardando.

Cociente Intelectual que, como todo el mundo sabe, no miden tanto la inteligencia y sus capacidades como esa pequeña parte de la inteligencia y sus capacidades que hemos creído saber cuantificar. Así, una IA menor, dado que no vamos a poder emular, en el medio plazo, el comportamiento inteligente no ya de un humano sino de cualquier otra especie considerada inteligente, va a horadar aún más ese concepto tan desgastado que es la inteligencia. O, dicho de otra forma: va a suceder que, de pronto, tengamos una definición delimitada de lo que es la inteligencia y que, curiosamente, coincida justo con aquello que saben hacer las máquinas: que no es otra cosa que gestionar un mundo formalizado que responde únicamente a algoritmos y procesos. Es lo que parecen traslucir todos esos reportajes que anuncian que el siguiente paso en la evolución de la IA es ayudarnos a sobreponer las deficiencias de nuestro aparato cognitivo... Como no entendemos un mundo formalizado necesitamos una máquina que nos ayude a movernos por él. O eso es lo que anuncia el titular del enésimo artículo que anuncia el enésimo avance de la IA: *Cognitive Computing is Transforming the Way We Work. It also Offers a Window to the Limitations of the Mind to Help Us Overcome Them*. Con un par. ¿Y qué es lo que nos van a ayudar a sobreponer? Hablan de sesgos cognitivos, como el de ambigüedad o el de confirmación, y cómo la máquina nos ayudará a tomar mejores decisiones. Pero no reparan en lo importante: esos sesgos, responsables en buena medida de una cognición, la humana, que nos ha traído hasta aquí, tienen dos pies, y cada uno está apoyado en entidades distintas. Por un lado, sí, están apoyados en nuestro aparato cognitivo, es en nuestra cognición donde se producen esos sesos, pero, por otro lado, esos sesgos no se dan en abstracto: son generados por la realidad que vivimos. De hecho, es gracias ese arsenal de capacidades, que es lo que son ahora y no vicios como pretenden que creamos, como hemos podido navegar –tomar decisiones– en el mundo hasta hoy. Y nos han permitido evolucionar porque mediante esos sesgos comprendíamos el mundo, lo hacíamos inteligible y nos hacíamos inteligibles en él. ¿Por qué los atacamos ahora? Porque hemos creado un mundo, formalizado, en el que esos sesgos heurísticos no garantizan la toma de decisiones eficientes que predica el racionalismo. Ese ataque a la cognición es, en el fondo, una imposición ideológica –de esa ideología de la peor especie, que es la que se niega a sí misma– cuya meta es la imposición de un mundo racionalista ininteligible, y el que no podemos hacernos inteligibles. Así, la IA es una inteligencia secuestrada que pretende rescatarnos de un mundo secuestrado. Pero no es el único ataque que sufre nuestra cognición: es muy relevante que ese argumento, la inteligencia de las máquinas como una especie de prótesis cognitiva, sea el mismo que se encuentra detrás de eso que llaman *behavioral economics*, esa confluencia entre psicología y economía que pretende explicar nuestro comportamiento en términos racionalistas, que también pretende desnudar los sesgos de nuestra cognición para ayudarnos a navegar por el mundo... A ver si va a resultar que ambas disciplinas –o los ideólogos de ambas disciplinas– no son sino el síntoma de algo mucho peor: Después de haber creado un mundo que nos excluye, que no podemos comprender porque no contamos con las capacidades ni las habilidades, lo que necesitamos es un perro lazarillo mecánico, que cuente con conocimientos en lenguajes formales y estadística. En resumen: como no podemos cambiar el mundo, lo *inteligente* es cambiarnos a nosotros mismos. Que lo consigan o no ya se verá según evolucione este primer escenario futuro, uno que ya estamos viviendo.

El segundo, al que no hemos sido invitados y es muy probable que no lo seamos salvo como meros comparsas, es una extensión del primero y entronca con esa visión de máquinas-ángeles de las que hablaba antes. Quizá muy pronto la IA del primer escenario repare que, a pesar de todos sus esfuerzos por ayudarnos, somos incapaces de vivir en un mundo absolutamente racionalista. O, en sus términos, que nuestros sesgos lo sean tanto que seamos incapaces de *overcome* nada, que esos sesgos, y aún otros muchas más atribuciones de nuestra cognición que, en un mundo racionalista, pueden ser considerados vicios, sean rasgos estructurales de nuestro aparato cognitivo y de mi difícil, si no imposible, solución. Y, entonces, ya sea porque aumenten sus capacidades o porque ganen autonomía y se desconecte de nosotros –ambos procesos no son excluyentes y hasta puede que sean complementarios–, puede que las máquinas se centren no tanto en instarnos a que llevemos a cabo comportamientos inteligentes y dediquen sus capacidades a profundizar aún más la formulación del mundo. Si no podemos cambiar el mundo ni cambiarnos a nosotros mismos, hagamos que el mundo sea más de lo mismo. Crearían un mundo tan distinto, tan cerrado sobre sí mismo, que bien parecería inmóvil, inaccesible al cambio. Sería un mundo muerto a nuestros ojos, cadáver, un mundo inerte, ajeno a la vida, porque sólo lo inerte no cambia, en el que quizá pasaran cosas o hubiera cambios, pero seríamos ciegos a ellos. No podríamos no ya comprenderlos, ni tan siquiera conocerlos. El secuestro del mundo terminaría en destierro, el nuestro. Un mundo periclitado en el que seríamos ininteligibles y para el que seríamos ininteligibles. Y ahí, en ese escenario, la IA culminaría su proyecto: la IA se convertiría en el dinero. Porque, como decía Hegel, «el dinero es la vida de lo que está muerto, moviéndose en sí misma.»

Pero eso no sería inteligencia, sería otra cosa.