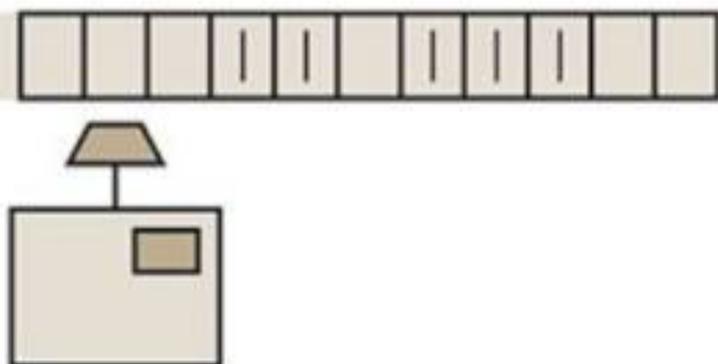


A.M. TURING

# ¿Puede pensar una máquina?

INTRODUCCIÓN DE MANUEL GARRIDO



En 1947 Alan M. Turing pronunció una conferencia ante un auditorio compuesto en su mayor parte por miembros del National Physical Laboratory de Londres en la que intentaba responder a la vieja y controvertida pregunta ¿Puede pensar una máquina?

Lo expuesto en ese acto apareció publicado tres años más tarde en *Mind* —una importante revista de filosofía británica— y es lo que ofrecemos aquí al lector en su traducción castellana. Este texto se convirtió enseguida en uno de los escritos fundacionales de la lógica informática y la inteligencia artificial, al presentar las líneas generales por las que debería discurrir una respuesta precisa y manejable (aunque no indiscutible) a la pregunta formulada.

Se trata del famoso Test de Turing, una prueba para decidir si una máquina es inteligente (o «piensa»). Para ello Turing diseñó un juego de imitación en el que participan una máquina y seres humanos; podemos decir que una máquina piensa si un ser humano que se comunica con la máquina y con otros seres humanos no logra distinguir cuando su interlocutor es una máquina y cuando un humano.

Una «máquina de Turing» como la que participa en el juego, es un dispositivo ideal de cálculo, capaz de resolver una función computable —una función cuya solución es susceptible de ser obtenida por un procedimiento mecánico—.

Pero lo más significativo es que Turing demostró que hay una máquina peculiar —la máquina universal de Turing— en la que se puede representar cualquier máquina que sea capaz de computar una función particular. De acuerdo con esto, una máquina universal de Turing sería una especie de sistema operativo en el que se implementan diferentes programas (máquinas de Turing especiales), un poco a la manera en que nos es familiar en los ordenadores personales. La denominada «metáfora del ordenador» como modelo

capaz de simular la mente humana y, por ende, el pensar, tiene aquí su fuente.

# ESPAÑOL

## 1. El juego de imitación

Propongo que consideremos la siguiente pregunta: «¿Pueden pensar las máquinas?». Para empezar, definamos el significado de los términos «máquina» y «pensar», pero es una actitud peligrosa. Si hemos de llegar al significado de las palabras «máquina» y «pensar» a través de su utilización corriente, difícilmente escaparíamos a la conclusión de que hay que buscar el significado y la respuesta de la pregunta «¿Pueden pensar las máquinas?» mediante una encuesta tipo Gallup. Pero es absurdo. En lugar de intentar tal definición, sustituiremos la pregunta por otra estrechamente relacionada con ella y que se expresa con palabras relativamente inequívocas.

El problema en su nuevo planteamiento puede exponerse en términos de un juego que denominaremos «juego de imitación». Intervienen en él tres personas: un hombre (A), una mujer (B) y un preguntador (C), indistintamente de uno u otro sexo. El preguntador se sitúa en una habitación aparte y, para él, el juego consiste en determinar quién de los otros dos es el hombre y quién la mujer. Los conoce por la referencia X e Y, y al final del juego determina si «X es A e Y es B» o si «X es B e Y es A». El preguntador puede plantear a A y a B preguntas como éstas: «Por favor X, ¿podría decirme cuán largo es su pelo?».

Supongamos que X es realmente A, entonces es A quien contesta. El objetivo de A en el juego es lograr que C efectúe una identificación errónea, por lo que su respuesta podría ser: «Mi pelo es corto, escalonado, y los mechones más largos son de unos veinte centímetros».

Para que el preguntador no se guíe por el timbre de voz, las respuestas deben ir por escrito o, mejor aún, mecanografiadas. Lo ideal es disponer de un impresor telegráfico que comunique las dos habitaciones. Otro procedimiento consiste en que un intermediario repita pregunta y respuesta. El objeto del juego para el tercer jugador (B) es ayudar al preguntador. La mejor estrategia para la jugadora es probablemente responder la verdad, añadiendo quizás a sus respuestas cosas como ésta: «¡Soy la mujer, no le haga caso!», pero de nada sirve, ya que el hombre puede hacer observaciones similares.

Ahora planteemos la pregunta: «¿Qué sucede cuando una máquina sustituye a A en el juego?». ¿Se pronunciará el preguntador en este caso tan erróneamente como lo hace cuando en el juego participan un hombre y una mujer? Estas preguntas sustituyen a la original: «¿Pueden pensar las máquinas?».

## 2. Crítica del nuevo problema

Del mismo modo que preguntamos: «¿Cuál es la respuesta a este nuevo tipo de pregunta?», podemos preguntar: «¿Merece la pena resolver esta nueva pregunta?». Resolvamos esta última pregunta sin plantear más objeciones para cortar una regresión infinita.

El nuevo problema presenta la ventaja de que traza una línea definida entre las aptitudes físicas e intelectuales de una persona. Ningún ingeniero o químico puede atribuirse la capacidad de producir un material que no pueda distinguirse de la piel humana. Quizá sea posible algún día, pero, aun suponiendo la viabilidad de semejante invención, nos parece que de poco serviría tratar de hacer una «máquina pensante» más humana, forrándola con esa epidermis artificial. El modo en que hemos planteado el problema refleja el obstáculo que impide al preguntador ver o tocar a los otros concursantes, oír su voz. Otras ventajas del criterio propuesto pueden resumirse en un modelo de preguntas y respuestas. Por ejemplo:

P: Por favor, escriba un soneto sobre el tema del Cuarto Puente.

R: Hágame otra pregunta; la poesía no es mi fuerte.

P: Sume 34957 con 70764.

R: (Pausa de unos 30 segundos) 105621.

P: ¿Juega al ajedrez?

R: Sí.

P: Tengo el rey en la casilla 1R y ninguna otra pieza. Usted tiene sólo el Rey en la casilla 6R y la Dama en 1D. Le

toca mover. ¿Qué juega?

R: (Pausa de unos 15 segundos) La Dama a D8, mate.

El método de preguntas y respuestas parece adecuado para introducir casi todos los campos de actividad humana que queramos. No vamos a sancionar a la máquina por su incapacidad para destacar en concursos de belleza, del mismo modo que no castigamos a una persona por perder una carrera en una competición aérea. Las condiciones del juego hacen irrelevantes esas torpezas. Los «testigos» pueden alardear, si lo creen conveniente, tanto como deseen con respecto a sus encantos, su fuerza o su heroísmo, pero el preguntador no puede exigir demostraciones fehacientes.

El juego quizá provoque críticas porque la máquina tiene demasiados factores en contra. Si una persona lo intentara haciéndose pasar por la máquina, sin duda haría un papel deplorable. Quedaría rápidamente eliminada por lentitud e inexactitud aritmética. ¿No harán las máquinas algo que permita la definición de pensamiento, pero que es muy distinto a lo que hace una persona? Se trata de una objeción de peso, pero cuando menos podemos decir que, dado que es posible construir una máquina que realice satisfactoriamente el juego de imitación, la objeción no viene al caso.

Podría alegarse que la mejor estrategia en el «juego de imitación», para la máquina, es posiblemente algo distinto a la imitación de la conducta humana. Puede, pero yo no creo que esto influya demasiado. En cualquier caso, no nos proponemos aquí analizar la teoría del juego y supondremos que la mejor estrategia es tratar de dar las respuestas que una persona daría con toda naturalidad.

### 3. Las máquinas que intervienen en el juego

La cuestión que planteábamos en el apartado 1 carece de precisión si no especificamos qué entendemos por el término «máquina». Es lógico que deseemos que nuestras máquinas estén dotadas de cualquier tipo de ingeniería mecánica. Del mismo modo que aceptamos la posibilidad de que un ingeniero o un equipo de ingenieros construya una máquina que funcione, pero cuya modalidad operacional no pueden describir satisfactoriamente sus constructores porque se han servido de un método fundamentalmente experimental. Finalmente, excluirémos de la categoría de máquinas a las personas nacidas del modo habitual. Es difícil adaptar las definiciones de modo que cumplan estos tres requisitos. Se puede insistir, por ejemplo, en que el equipo de ingenieros sea de un solo sexo, lo cual no sería satisfactorio, ya que probablemente se puede crear un individuo completo a partir de una simple célula epidérmica de un hombre (pongamos por caso). Esto sería una proeza de biogenética merecedora de máxima admiración, pero no por ello la calificaríamos de «construcción de máquina pensante». Esto nos obliga a descartar el requisito de permitir cualquier tipo de técnica, y con mayor razón dado que el interés actual por las «máquinas pensantes» se ha suscitado gracias a un tipo particular de máquina, generalmente denominada «computadora electrónica» o «computadora digital». Con arreglo a esto, sólo permitiremos que tomen parte en el juego las computadoras digitales.

A primera vista esta limitación parece muy drástica, pero intentaré demostrar que no es así. Para ello es necesario un breve resumen sobre la naturaleza y las propiedades de estas computadoras. Podría también aducirse que esta identificación de las máquinas con las computadoras digitales, al igual que nuestro criterio sobre el término «pensar», son insatisfactorias si (en contra de lo que creo) resulta que las computadoras digitales son incapaces de hacer un buen papel en el juego.

Existen ya varias computadoras operacionales, y es lógico que se diga: «¿Por qué no realizar el experimento ahora mismo? No resultaría difícil cumplir los requisitos del juego. Se pueden utilizar varios preguntadores, compilando unas estadísticas para comprobar cuántas veces se produce la identificación correcta». La respuesta inmediata es que no se trata de plantearse si todas las computadoras digitales actuarán bien en el juego, ni de si las actuales computadoras actuarán bien, sino de si existen computadoras imaginables que actúen bien. Pero esto es sólo la respuesta inmediata, más adelante consideraremos la cuestión bajo otra perspectiva.

## 4. Computadoras digitales

Podemos explicar el concepto de computadoras digitales diciendo que son unas máquinas ideadas para realizar cualquier tipo de operación propia de un computador humano. El computador humano sigue unas reglas determinadas sin opción a desviarse de ellas bajo ningún concepto. Supongamos que esas reglas figuran en un libro que cambia cada vez que el computador acomete un nuevo trabajo. Dispone también de una cantidad ilimitada de papel para efectuar cálculos y hace las multiplicaciones y sumas pertinentes con una «máquina de bolsillo», pero esto no tiene importancia.

Si utilizamos como definición la anterior explicación, corremos el riesgo de caer en una argumentación circular. Para evitarlo, esbozaremos los medios con los que se logra el efecto deseado. Suele considerarse que una computadora digital consta de tres partes:

1. Almacenamiento
2. Unidad procesadora
3. Control

El almacenamiento es el acopio de información y corresponde al papel sobre el que se efectúa la computación humana, ya sea el papel en que la persona realiza los cálculos o aquél en el cual está impreso el libro de reglas. Del mismo modo que el computador humano efectúa sus cálculos con su cabeza, parte del almacenamiento corresponde a la memoria de la máquina.

La unidad procesadora es el sector que realiza las distintas operaciones de cálculo. La naturaleza de estas operaciones varía de una máquina a otra. Generalmente pueden efectuar operaciones bastante largas, tales como «Multiplicar 3540675445 por 7076345687», pero en algunas máquinas sólo pueden llevarse a cabo operaciones muy simples, tales como «Escribe 0».

Hemos mencionado que el «libro de reglas», de que se vale el computador, se sustituye en la máquina por una parte del almacenamiento. Esta se denomina «tabla de instrucciones». Corresponde al control comprobar que las instrucciones se sigan correctamente y en su debido orden. El control está construido de tal manera que es infalible.

La información almacenada suele estar dividida en paquetes de tamaño relativamente modesto. En una máquina concreta, por ejemplo, el paquete puede constar de diez dígitos decimales. Se asignan números a las partes del almacenamiento en que se guardan los diversos paquetes de información, con arreglo a una modalidad sistemática. Un ejemplo de instrucción corriente podría ser: «Suma la cifra almacenada en la posición 6809 a la situada en la 4302 y devuelve el resultado de la última posición de almacenamiento». Ni que decir tiene que la operación no se desarrolla en la máquina expresada de este modo, sino que se lleva a cabo siguiendo una codificación como 6809430217. La cifra 17 indica cuál de las posibles operaciones hay que efectuar con las dos cifras. En cuyo caso la operación es la anteriormente descrita: «Suma la cifra...». Se advertirá que la instrucción consta de diez dígitos y, por lo tanto, constituye exactamente un paquete informativo. El control suele captar las instrucciones a seguir en el orden de posición en que están almacenadas, aunque a veces pueda surgir una instrucción como ésta: «Sigue ahora la instrucción almacenada en la posición 5606 y continúa», o bien: «Si la posición 4505 contiene 0, sigue la instrucción almacenada en 6707; en caso contrario continúa».

Las instrucciones de este tipo son muy importantes porque permiten la repetición de una secuencia de operaciones una y otra vez hasta que se cumple un determinado requisito, pero, al hacerlo, la máquina sigue en cada repetición, no nuevas instrucciones, sino las mismas indefinidamente. Recurramos a una analogía casera: supongamos que mamá desea que Tommy pase por el zapatero cada mañana camino del colegio para ver si han arreglado sus zapatos; puede decírselo cada mañana, o puede dejar una nota permanente en el vestíbulo para que el niño la vea al salir y recuerde que tiene que pasar por el zapatero, y luego, al volver, si trae los zapatos, rompa la nota. El lector debe aceptar como un hecho la construcción de computadoras digitales que, efectivamente, se han construido con arreglo a los principios expuestos y que realmente mimetizan con gran fidelidad los actos de un computador humano.

El libro de reglas que, según hemos señalado, utiliza el computador humano es, naturalmente, una ficción convencional. Los computadores humanos recuerdan en realidad lo que tienen que hacer. Si queremos hacer una máquina que mimetice el comportamiento de un computador humano en operaciones complicadas, hay que preguntarle a éste cómo lo hace y luego transferir la respuesta en forma de tabla de instrucciones. La elaboración de tablas de instrucciones suele denominarse «programación». La «programación de una máquina para que efectúe la operación A» significa insertar en la máquina la tabla de instrucción adecuada para que lleve a cabo A.

Una variante interesante de la idea de computadora digital es la «computadora digital con un elemento aleatorio». Estas máquinas disponen de instrucciones en las que interviene un dado o un proceso electrónico equivalente; una instrucción de este tipo puede ser, por ejemplo: «Arroja el dado y almacena la cifra resultante en 1000». A veces se las denomina máquinas de libre voluntad (aunque perso-

nalmente yo no utilice esta expresión). Normalmente no se puede determinar por simple observación de la máquina si ésta posee un elemento aleatorio, ya que se logra un efecto similar con dispositivos cuya elección depende de los dígitos de los decimales de  $\pi$ .

La mayoría de las computadoras digitales poseen un almacenamiento finito, aunque no existe dificultad teórica en la concepción de una computadora de almacenamiento ilimitado. Naturalmente, sólo podría utilizarse una parte finita de cada fase. De igual modo se habría podido construir una cantidad finita, pero cabe imaginar que sucesivamente fueran añadiéndose otras. Estas computadoras presentan especial interés teórico y las denominaremos computadoras de capacidad infinita.

El concepto de computadora digital es antiguo. Charles Babbage, profesor de matemáticas en la Universidad de Cambridge entre 1828 y 1839 concibió una a la que denominó Máquina Analítica, pero no la terminó. Aunque Babbage expuso los principios fundamentales, la máquina no representaba en aquella época gran interés. Su rapidez habría sido mucho mayor que la de un computador humano, pero unas 100 veces inferior a la de la máquina de Manchester, que a su vez es una de las máquinas modernas más lentas. El almacenamiento era puramente mecánico y se efectuaba por medio de ruedas y tarjetas.

El hecho de que la Máquina Analítica de Babbage estuviera concebida de forma totalmente mecánica nos ayudará a despejar cualquier superstición. Muchas veces se atribuye importancia al hecho de que las computadoras digitales modernas son eléctricas, igual que el sistema nervioso. Como la máquina de Babbage no era eléctrica, y como todas las computadoras digitales son en cierto modo equivalentes a ella, el empleo de la electricidad no es teóricamente relevante.

Siempre que se trata de señalización rápida interviene, claro, la electricidad. Por lo tanto, no es de extrañar que és-

ta se halle relacionada con ambos conceptos. En el sistema nervioso los fenómenos químicos son, cuando menos, tan importantes como los eléctricos. En ciertas computadoras el sistema de almacenamiento es fundamentalmente acústico. Por lo tanto, el empleo de la electricidad como propiedad no deja de ser una similitud muy superficial. Para establecer similitudes reales debemos más bien buscar analogías en el funcionamiento matemático.