

MENTE, COMPUTACIÓN E INTELIGENCIA

Gloria Mayela García Alcalá¹

¹Trabajo presentado durante la Licenciatura en Filosofía para la asignatura Filosofía de la mente.

Correo electrónico: bucletemporal@gmail.com

Resumen:

La oscuridad de la pregunta por la naturaleza de la *mente* exige un esfuerzo por convertirla en una más clara; es decir, requerimos de una traducción de nuestra pregunta a una forma más concreta que nos alcance medios para responder. El presente trabajo pretende exponer el razonamiento que atisba una forma de respuesta en el estudio de lo reproducible de la mente humana en términos computables; se trata de una presentación simplificada de algunos de los problemas vigentes de lo que hoy conocemos como Inteligencia Artificial (IA).

Palabras Clave: mente. inteligencia artificial, Turing, computación, procesos, cerebro.

Abstract:

The obscurity of the question about the mind's nature demands an effort to transform it into a much clearer one; i.e., we require a translation of our question into a more concrete form that brings closer the means to answer it. The present essay pretends to expose the reasoning that glimpses a form of answer within the study of that which can be reproduced of the human mind in terms of the computable; it seeks to be a simplified presentation of the current problems in what we know today as Artificial Intelligence (AI).

Key words: mind, artificial intelligence, Turing, computation, processes, brain.

La oscuridad de la pregunta por la naturaleza de la *mente* exige un esfuerzo por convertirla en una más clara, reclamando “¡precisión en el habla!”, al estilo del cuento infantil. Es decir, requerimos de una traducción de nuestra pregunta a una forma más concreta (análogamente al análisis de Turing sobre máquinas inteligentes, que veremos abajo), que nos alcance medios para responder. El presente trabajo pretende exponer el razonamiento que atisba una forma de respuesta en el estudio de lo reproducible de la mente humana en términos computables; se trata de una presentación simplificada de algunos de los problemas vigentes de lo que hoy conocemos como Inteligencia Artificial (IA).

La IA, considerada como la ciencia de la inteligencia en general, está intrínsecamente relacionada con disciplinas como la filosofía de la mente, filosofía del lenguaje, la epistemología, etc., y su consideración es fundamental entre las ciencias cognitivas, sobre todo la psicología de la computación. Resaltaremos de ésta última (como caso para señalar de una primera manera intuitiva la asociación mente-computadora), que considera a la mente desde un enfoque funcionalista, es decir, concibe los procesos mentales como rigurosamente especificables y los estados mentales como definibles por sus relaciones causales con la información sensorial recibida, el estado motor y otros estados mentales, todo con una ocurrencia material en la masa cerebral; desde esta perspectiva, el cerebro es como un sistema de computación que cuestiona qué tipo de relaciones funcionales incorpora.

Pero volvamos a lo nuestro: Documentalmente nos referiremos casi exclusivamente a las discusiones iniciales del tema (de la década de 1940), pero osamos proponer la generalizante idea de que aún con todos los más recientes descubrimientos y avances, tanto en las neurociencias como en computación, la estructura del razonamiento permanece aplicable, sobre todo en la medida en que las implicaciones de un supuesto más general, que quizá podamos catalogar como juicio (si es posible una máquina pensante), acarrear a deliberar una definición de inteligencia.

Así, comenzaremos con un breve despliegue de los descubrimientos de McCulloch y Pitts de estructuras lógicas en redes neuronales; luego, para mediar el salto a lo computable, las implicaciones según la proposición de Turing sobre que las funciones recursivas coinciden con las computables, basada en el concepto de Máquinas de Turing; en tercer lugar, y siguiendo un poco a Turing en otra de sus facetas, nos preguntaremos si es posible una máquina pensante; por último, propondremos que de la respuesta a la pregunta anteriormente planteada se sigue una postura con respecto a la definición de Inteligencia y de mente.

Brevemente expondremos los supuestos que permitieron a McCulloch y Pitts pensar por primera vez (en 1943) en el cerebro como un mecanismo computacional, dando inicio y base para el estudio de las redes neuronales. Se trata de los siguientes:

1. La transmisión entre neuronas es un proceso regido por lo que ellos llaman el “Proceso de Todo o Nada”; esto es, su función es de carácter binario (o transmiten o no). Esto hace posible representar la actividad neuronal como una proposición.
2. Para poder excitar a una neurona en cualquier momento (o, mejor dicho, para hacerla rebasar el nivel de excitación en que se encuentra, suficientemente como para generar un impulso), se requiere la excitación de un número fijo de sinapsis, independiente de la actividad y posición previa de la neurona.
3. El único intervalo significativo dentro del sistema nervioso es el sináptico.
4. La sinapsis inhibitoria evita absolutamente que la neurona se excite en ese momento.
5. La estructura de la red no se altera con el tiempo. (McCulloch et al, 1943)

Lo anterior permite probar que el comportamiento de cualquier red es calculable y que es posible encontrar una red que se comporte de cierta manera, en tanto que exista. Es decir, permite describir la operación de una red a partir de la definición del estado de una neurona mediante un algoritmo, dado el estado de otra con la cual esté unida y haga sinapsis. El sistema nervioso está conformado por un sinnúmero de circuitos, cuya actividad se regenera de tal modo que la excitación de cualquier neurona participante refiere a un momento pasado. En términos lógicos, se trata de encontrar un método que permita obtener un conjunto de números para cuyos predecesores sea válida cierta propiedad: una función recursiva primitiva. (McCulloch et al, 1943)

Haciendo un pequeño paréntesis de lógica matemática: Una función f es *recursiva primitiva* si existe una sucesión de funciones f_1, \dots, f_n tales que f_n es f , y para todo número natural i entre 1 y n , la función f_i es recursiva elemental o bien f_i está definida por composición o recursión a partir de funciones anteriores de la sucesión. (Ivorra, s/f:121)

Puesto que las funciones elementales se pueden calcular mediante algoritmos y las funciones definidas son también calculables mediante algoritmos, ya que se obtienen por composición o recursión a partir de funciones calculables

mediante algoritmos, resulta entonces claro que toda función recursiva es calculable mediante un algoritmo. Más concretamente, si f es una función recursiva, una sucesión de funciones f_1, \dots, f_n , según la definición, determina un algoritmo para calcular f (en el sentido de que conociendo la sucesión es fácil diseñar el algoritmo correspondiente) (Ivorra, s/f:122). El argumento de Turing para probar que las funciones recursivas coinciden con las funciones calculables mediante un algoritmo se basa en el concepto de máquina de Turing. (Turing, 1936)

En términos generales, una máquina de Turing es un modelo teórico de ordenador con infinita memoria disponible. Consta (teóricamente) de un cabezal tanto lector como escritor, y una cinta infinita en la que el cabezal lee el contenido, borra el contenido anterior y escribe un nuevo valor. Las operaciones que pueda máquina analítica no pretende *crear* nada. Puede hacer *lo que sea que sepamos ordenarle* ejecutar son: avanzar el cabezal por la izquierda o la derecha de manera repetida, después de haber leído o escrito algo en la casilla de la cinta, pero de tal modo que el signo que se imprime, el movimiento que se hace y el estado al que se pasa, son función exclusivamente de la configuración de la máquina en ese instante. (Turing, 1936)

Pues bien, si cada red se equipara con una cinta, ésta sólo puede computar tantos números como una máquina de Turing, y cada uno de estos números puede ser computado mediante una red similar, donde por computación es entendido el manejo formal de símbolos no interpretados mediante la aplicación de reglas formales (o sea, una operación recursiva).

Surge así la noción de “proceso eficiente”, como proceso de computación estrictamente definible.

Entonces, si es posible explicarla en términos generales desde el punto de vista de procedimientos eficientes ejecutados en el cerebro (los cuales ocurren, según descubrieron McCulloch y Pitts), la inteligencia podría simularse con la máquina universal de Turing o algún mecanismo semejante.

La relevancia ahora de considerar las máquinas de Turing es que, de acuerdo con tal conjetura, de todo problema de cómputo que sea resoluble en una máquina de Turing se considera que también lo será en una computadora, y viceversa: lo mismo que ocurre con las redes neuronales, si son supuestas como explicables en términos de funciones recursivas.

Frente a la efectiva creación de computadoras digitales que simulaban algunos aspectos de la inteligencia (hacia 1950) surge la pregunta de si es posible que una máquina piense. Sin embargo, evadiendo la ambigüedad de la pregunta bajo esa forma (calificándola como de opinión y rechazando que

se trate de decidir lo que “pensar” signifique), Turing prefiere plantearla en términos de si una posible máquina sería capaz de realizar lo que nosotros hacemos requiriendo inteligencia, la llamada *Prueba de Turing*, y a la que éste se refirió como el “juego de la imitación”. (Turing, 1950)

Se trata de un juego hipotético donde en principio participan un hombre, una mujer y un tercero cuyo sexo no interesa y que será el examinador de los otros dos. Los primeros dos participantes son anónimos al examinador, quien les plantea preguntas (como “¿qué tan largo es su cabello?”, etc.) a las que uno de ellos responde ayudando a quien interroga y otro intentando hacer que los identifique erróneamente. La cuestión es si cuando una computadora ocupa el lugar de este último (el jugador que pretende inducir al examinador en el error) haría evidente alguna diferencia que le identificara; es decir, si la máquina puede hacerse pasar por, y conseguir los mismos resultados que si en su lugar estuviera la mujer, por ejemplo. (Turing, 1950)

El problema no radica en las facultades de la ingeniería, sino en decidir si en las estructuras de lo computable se alcanza a reproducir lo que con inteligencia se es capaz de hacer. Particularmente, se piensa en una computadora digital, o de estado discreto. Se considera que éstas constan de tres partes:

1. Almacén. Que guarda información; se corresponde a la memoria y al espacio necesario para realizar sus cálculos.
2. Unidad operativa. Que se encarga de realizar las diversas operaciones individuales involucradas en el cálculo.
3. Control. No permite que se opere incorrectamente ni en desorden; el control está construido de tal forma que esto es necesario. (Turing, 1950)

La idea de una computadora digital es más antigua: Charles Babbage (en 1830, aproximadamente) proyectó una, a la que bautizó como la Máquina Analítica. Lo interesante de aquél proyecto es que se trataba de una máquina completamente mecánica; así, nos damos cuenta de que, a diferencia de como se estaría tentado a creer, el uso de la electricidad no tiene ninguna relevancia teórica, lo que las máquinas hacen no está en la electricidad, no obstante que ésta favorece bastante en lo que a velocidad de transmisión toca. Y a pesar de que a nivel cerebral el medio de transmisión es el mismo, semejanzas reales sólo se encuentran con notabilidad a nivel de analogías matemáticas en el funcionamiento. (Turing, 1950)

Dijimos que las computadoras digitales son de estado discreto, lo cual significa que funcionan mediante saltos repentinos de un estado bien definido a otro (evidentemente distinto), de la misma manera que la transmisión entre neuronas.

Precisamente por esta propiedad, en una máquina digital, el conocimiento preciso del estado en un determinado momento proporciona con la misma precisión el de algunos estados futuros inmediatos. Si este cálculo se efectúa con suficiente rapidez, la computadora (digital) podría imitar el comportamiento de cualquier máquina de estado discreto (propiedad denominada *universalidad*); si el cerebro es considerado una de ellas, una computadora digital bien programada y con suficiente espacio de almacenamiento, puede perfectamente imitarlo. Una máquina, efectivamente puede superar la prueba del *juego de las imitaciones*. (Turing, 1950)

Es digno de mencionar el hecho de que, gracias a la equivalencia entre máquinas de Turing y computadoras, se haya determinado que existen cálculos que no pueden ser resueltos en un tiempo razonable en ninguna computadora imaginable, o incluso, que no pueden resolverse en lo absoluto, por ejemplo el problema de correspondencia de Post o el problema de predecir si una máquina de Turing cualquiera va a llegar a un estado final (conocido como el problema del *halting* en inglés, o problema de *la parada*). Una sencilla respuesta a esto es que no hay ningún tipo de prueba que evidencie que la mente carezca de estas limitaciones; por el contrario, una máquina correctamente programada puede presumir de mucho menor nivel de falibilidad.

Hay otras múltiples objeciones a esta perspectiva (de las que, de hecho, se desprenden diferentes ramas de la IA), de entre las que, particularmente, resulta ahora interesante presentar la de Lady Lovelace, con fines de pasar concienzudamente a un punto importante en la exposición de la máquina pensante de Turing. Se trata de la afirmación (que señalaba en alusión a la máquina analítica de Babbage, en 1842) de que “la máquina analítica no pretende *crear* nada. Puede hacer *lo que sea que sepamos ordenarle*”. En otras palabras, si bien en términos conductistas una computadora digital puede imitar las funciones cerebrales, lo que ésta hace no es en absoluto inteligencia, sino una suerte de instrucciones ejecutadas con precisión. (Turing, 1950)

En respuesta, propone Turing lo siguiente: algunas de las funciones mentales son estrictamente operaciones explicables en términos mecánicos; esto no corresponde a *La Mente* propiamente, sino una especie de cubierta que hay que quitar para desentrañar la mente real. Pero nos encontramos con más y más funciones de este tipo. Ocurre lo que con una cáscara de cebolla que no es *La Cebolla*, a la que si procedemos quitando cada cáscara que no es lo que buscamos, nunca llegaremos. Desde este punto de vista, la mente es concebida, funcionalmente, como toda mecánica. (Turing, 1950)

Parece que la objeción de Lady Lovelace reclama de esta analogía la capacidad de aprendizaje (y no sólo la de ejecutar operaciones determinadas previamente) como fundamental en la inteligencia. Podríamos esbozar la mente humana (inteligente) como formada por:

1. Un estado inicial (como se nace).
2. Educación a la que se ha sometido.
3. Otras experiencias, distintas de la educación.

Pues bien, de una computadora podríamos señalar que a estos tres componentes se les corresponden:

1. Un lenguaje y una serie de axiomas.
2. Reglas de un cálculo.
3. Proposiciones que a modo de premisas se van introduciendo como datos. (Turing, 1950)

Así, podríamos decir que una computadora también es capaz de aprender (y además lo hace). Una pregunta que habría que plantearnos es si esto es suficiente para derrumbar el argumento de la condesa, o si, en cambio, a la manera de Russell es factible pensar que dadas las reglas de un cálculo y los axiomas, todo lo que pueda incluir el sistema de que formen parte está también, por añadidura, dado.

No hemos llegado a una determinación sobre si lo que una máquina puede hacer es o no inteligente. Pero tampoco es nuestro objetivo. Es más, los argumentos, al menos tal como se han desarrollado aquí, no apuntan a nada concluyente, considerando que si de ellos hiciéramos sostén para afirmar algo con seguridad, estaríamos incurriendo en una terrible falacia de razonamiento por analogía: “La mente cuenta con tal propiedad, una computadora digital también; lo que la mente hace es pensar, lo que la máquina también”, etc. En cambio, lo interesante de el repaso que hemos hecho sobre la posibilidad y la discusión al respecto, es evidenciar cómo del reconocimiento o no de que lo que ocurre en estos casos (algunos reales, otros hipotéticos) es efectivamente inteligencia, estamos implicando (o suponiendo) una definición de lo que sea la mente. O, como una manera inversa de analizar las consecuencias de la definición que aceptemos. Si decidimos que la mente es la abstracción de los *procesos* mentales, entonces estamos dispuestos a afirmar que la inteligencia, propia de la mente, es predicable también de otro tipo de mecanismos que operen con *procesos* equivalentes. En cambio, si no aceptamos esto último, asumimos una postura que nos obliga a pensar qué sea entonces la característica propia de lo mental si no la estructura de su funcionamiento. **Ψ**

BIBLIOGRAFÍA:

Ivorra Castillo, Carlos (s/f) *Lógica y Teoría de Conjuntos* (Archivo electrónico).

McCulloch, Warren y Pitts, Walter 1943 (1990) *Un cálculo lógico de las ideas inmanentes en la actividad nerviosa* en Boden, Margaret A. (comp.) *Filosofía de la Inteligencia Artificial* (México : Fondo de Cultura Económica).

Turing, Alan M. 1950 (1990) “La maquinaria de computación y la inteligencia” en Margaret A. Boden (comp.) *Filosofía de la Inteligencia Artificial* (México : Fondo de Cultura Económica).

Turing, Alan M. 1936 *On computable numbers, with an application to the entscheidungsproblem* (Archivo electrónico).

Recibido: Marzo 24, 2011; Aceptado: Mayo 20, 2011.