



# Ensayos



## Ensayo

### ¿QUÉ PASÓ CON EL TEST DE TURING?

**Alberto Castillo Vicci**

Ciencia de la Computación. Universidad de Wisconsin (USA)  
Docente e Investigador de la Universidad  
Centroccidental Lisandro Alvarado.  
Venezuela  
E-mail: kikocastillovicci@gmail.com

#### RESUMEN

El llamado test de Turing—que su autor, Alan Turing, acuñó como Juego de la Imitación en 1950—, sirvió de estímulo a los científicos de la computación que investigaban como pioneros en la inteligencia artificial; algunos creen, desviando sus fines verdaderos. Pero, también, inició una larga controversia de medio siglo sobre su validez. Luego, abandonado por los científicos de la computación como pieza histórica. No así para los filósofos, que siguen buscando una prueba formal, científico-filosófica, de la relación mente-materia, pues se trata de un secular problema: ¿cuál es la naturaleza humana y si hay diferencias entre el hombre con todo lo demás en el universo? El presente trabajo es una crónica antológica de cómo ha sido tratado el problema por algunos científicos y filósofos, desde la propuesta de Turing hasta hoy.

**Palabras clave:** test de turing, inteligencia artificial, prueba de gödel, test mayziano, filosofía informacional.

**Recibido:** 4-10-2018

**Aceptado:** 8-11-2018

## WHAT HAPPENED TO THE TURING TEST?

**Alberto Castillo Vicci**

BA Computer Sciences. University of Wisconsin (USA)

Teacher and Researcher of the

University Centroccidental Lisandro Alvarado.  
Venezuela

E-mail: kikocastillovicci@gmail.com

### RESUMEN

The so-called Turing test-which its author, Alan Turing, coined as a game of imitation in 1950-served as a stimulus to computer scientists investigating as pioneers in artificial intelligence; some believe, diverting their true ends. But it also began a long half-century controversy over its validity. Then, abandoned by computer scientists as a historical piece. Not so for philosophers, who continue to seek a formal, scientific-philosophical proof of the mind-matter relationship, for it is a secular problem: what is human nature and if there are differences between man and everything else in the universe? The present work is an anthological chronicle of how the problem has been treated by some scientists and philosophers, from the proposal of Turing until today.

**Keywords:** turing test, artificial intelligence, gödel proof, mayzian test, information philosoph.

### 1. INTRODUCCIÓN

En el año de 1950, el matemático británico Alan Turing publicó, en la revista de filosofía *MIND*, una monografía con el título *Computing Machinery and Intelligence*. (Turing, 1964). Esta monografía resultó ser, en el mundo académico y científico, la mayor reimpresión, copiada, controvertida, discutida y con réplicas por célebres filósofos y científicos de todo linaje, durante la segunda mitad del siglo XX.

Aunque, la monografía en cuestión ha sido discutida hasta el agotamiento por Castillo (1993), vale la pena considerar, casi setenta años después de su primera publicación, sus predicciones para saber hasta qué punto se han dado y cuál es el estado del tema hoy en día, recapitulando algunos hitos en esos años de los argumentos a favor y en contra de la controversia.

En la presente monografía, volvemos a esas ideas y la extendemos a como la tratan, hoy en día, un grupo interdisciplinario de científicos y filósofos que toman a la información como instrumento más poderoso que la lógica y el lenguaje para la indagación filosófica, generando una nueva disciplina transdisciplinaria denominada Filosofía Informacional que fuera iniciada en las universidades de Oxford y Harvard en los días finiseculares del siglo XX y se ha consolidado en estas dos primeras décadas del XXI.

Intentamos, entonces, en este artículo, una crónica antológica de la pregunta ¿Puede una máquina pensar? Y una síntesis, por razones de espacio, del estado del arte de su formulación y prueba, según lo intentara Alan Turing en 1950, hasta la fecha.

### 2. DESARROLLO

#### 2.1. Alan turing y el juego de la imitación

En su monografía *Computing Machinery and Intelligence* Turing (1964), Turing argumenta, desde un punto de vista moderno, la posibilidad que algún día una máquina llegue a pensar. Un tema de interés no sólo para los profesionales de la inteligencia artificial y la computación, sino también de filósofos, estudiosos de las ciencias cognitivas y hasta para teólogos; pues se trata de la secular búsqueda en la historia del pensamiento occidental acerca de la naturaleza humana y si somos o no máquinas biológicas, aunque máquinas al fin, que piensan.

Tal especulación puede ser remontada a los versos de Homero, unos 850 años a.C. En el Canto XVIII de la *Iliada*, Homero nombra a los primeros autómatas que la imaginación humana haya creado, cuando Hefestos recibe agradecido en su taller a Tetis: “Vistió la túnica, tomó el fornido cetro y salió cojeando apoyado en dos estatuas de oro que eran semejantes a las vivientes jóvenes pues tenían inteligencia, voz y fuerza, ejercitadas en las propias obras de los inmortales dioses...” (Castillo, 1993).

La monografía de Turing publicada en 1950, podría tener su ancestro intelectual, entonces, en Homero y se le atribuyen las siguientes influencias, en la segunda mitad del siglo XX, a sus apenas casi 30 páginas de contenido:

- 1950-1965 Fuente inspiración de la disciplina académica conocida como inteligencia artificial.
- 1966-1972 Una distracción de lo que realmente debe investigar la inteligencia artificial.
- 1973-1990 De la que sólo se ocupan los filósofos y olvidada por los científicos.
- 1991- Consignada para la Historia.

Su autor, Alan Turing, fue un matemático que a la temprana edad de 24 años resolvió lo que se consideró un problema central de la lógica matemática, la existencia de un procedimiento mecánico — lo que los matemáticos llaman algoritmo— que decidiera por sólo su forma o sintaxis, si una proposición matemática era verdadera o falsa; valga decir, un teorema. Por eso recibió el nombre de problema de la decisión o en alemán como lo formulara el matemático David Hilbert “*Entscheidungs problem*” y así se conociera inicialmente. Turing demostró que tal algoritmo no existe y que no es posible mecanizar las matemáticas. Pero los matemáticos aunque reconocían un algoritmo al solo verlo (como se venía haciendo desde los tiempos de Euclides quien diseñó el más antiguo algoritmo que se conoce para calcular el máximo común divisor), no tenían una definición rigurosa de aquél. Turing inventó una máquina abstracta que llamó M-machine (M por matemática), pero hoy lleva su nombre: máquinas de Turing, y es la expresión rigurosa, matemática y precisa del concepto de procedimiento mecánico finito o algoritmo.

Se le recuerda a Turing también, por la tesis, compartida con su tutor Alonzo Church, que dice que todo algoritmo tiene una máquina de Turing que lo procesa. Hasta ahora aceptada por toda la comunidad científica.

También por su trabajo durante la Segunda Guerra Mundial en descifrar los mensajes en clave de los alemanes en sus máquinas llamadas Enigma, que redujo la guerra en Europa en dos años y posiblemente salvó 14 millones de vidas. Y, por haber construido uno de los primeros computadores electrónicos de programa almacenado, un verdadero autómatas, en la Universidad de Manchester. La carrera de Alan Turing se vio abruptamente interrumpida con su suicidio a los 42 años edad.

En la famosa monografía “*Computing machinery and intelligence*” Turing (1964), Turing retomó el llamado reto cartesiano. En el *Discurso del Método*, Descartes (1596-1650) afirma Descartes (1975) que si la industria del hombre llegara a construir, con apariencia idéntica a su espécimen natural, un animal artificial, no podríamos distinguir el artificial del natural, porque, como la máquina que lo replica, el animal natural es un autómatas también. No sería el caso, si se construyese un hombre artificial por los mismos medios. No siendo el hombre un autómatas, se diferenciaría de la máquina que lo imita, por lo menos con dos pruebas:

- Por la imposibilidad del artificial de expresar con palabras o signos compuestos, arreglando a éstas de manera tan diversa como para responder con sentido cuanto en su presencia se diga; esto es, con la creatividad con que el hombre expresa pensamientos nuevos y los entiende dentro de una lengua intuitiva, en el uso normal —cotidiano— del lenguaje; como puede hacer el más estúpido de los hombres.

- Porque, aunque la máquina hiciera tan bien o mejor que cualquier hombre, una gran cantidad de tareas sólo atribuibles a la raza humana, no podría hacerlas todas las que le son posibles a la razón, pues ninguna máquina tendría suficientes componentes para poder disponerlos en cada uno de los órdenes que requerirían la inmensa variedad de circunstancias que suelen presentarse en la vida humana.

La máquina a que se refería Descartes era una máquina física con un número limitado de componentes. Turing inventó una máquina abstracta con posibilidades infinitas de programación — de hecho hay tantas máquinas de Turing distintas como números naturales—; y cuya realización física es un computador electrónico. El número de programas posibles para un computador electrónico, como el de las máquinas de Turing, es infinito; basta cambiar el programa para que cada computador electrónico sea una máquina distinta en cuanto al problema que resuelve. Claro, hay problemas para los que no existe un programa como lo demostró Turing; valga decir, para los que no existe máquinas de Turing.

En su monografía, Turing enfrenta el reto cartesiano, reemplazando la pregunta ¿pueden las máquinas pensar?, en que los vocablos máquinas y pensar tienen las más variadas connotaciones y sentidos vagos, además de despertar sesgos y prejuicios (al punto de que para conocer su significado usual habría que recurrir a los métodos estadísticos de las encuestas), por otra que podríamos enunciar así (de hecho Turing la reemplaza por dos en lugar de una, como veremos en seguida): ¿puede una máquina imitar a un ser humano al grado de confundir a otro sobre su identidad un número significativo de veces? A tal efecto sugiere un juego que llamó de imitación y que ha venido a conocerse como el Test de Turing, pero que no es otra cosa que la versión moderna del reto cartesiano Participan en el juego tres jugadores: un hombre (A), una mujer (B) y un interrogador (C) que puede ser de cualquiera de los dos sexos.

Al interrogador se le aísla en cuarto aparte, de tal manera que no pueda ni ver ni oír a los otros jugadores; sólo tendrá comunicación con ellos a través de dos teletipos rotulados X e Y (hoy terminales). El juego consiste en que el interrogador adivine quién de los jugadores es el hombre y quién la mujer, diciendo que  $X=A$  e  $Y=B$  o, viceversa,  $X=B$  e  $Y=A$ . Para ello podrá hacerles las preguntas que quiera y evaluar las respuestas. El papel de A es engañarlo y el de B ayudarlo.

La pregunta ¿pueden las máquinas pensar? sería reemplazada por: ¿qué pasará cuando una máquina tome la parte de A en este juego? Y ¿decidirá el interrogador erróneamente tantas veces cuando el juego sea conducido con la máquina como cuando es jugado por un hombre y una mujer? La máquina en cuestión es un computador electrónico digital.

La predicción de Turing es que para fines del siglo XX (cincuenta años después de publicada su monografía) se habrán construido máquinas con suficiente memoria (entre 107 a 109 bits) como para jugar con tal eficiencia que después de cinco minutos de juego el interrogador no tendrá más de 70% de probabilidades en acertar correctamente. Para entonces, quedará sin efecto la pregunta ¿pueden las máquinas pensar?, pues se hablará de máquinas pensantes sin esperar que nadie contradiga tal afirmación.

Qué se llegue a construir o no una máquina que pase el Test de Turing es materia de confirmación empírica; pero al dilema le han surgido objeciones a priori de las más variadas especies. Algunos afirman que ya la IA lo logró y se señalan casos en que algún programa le hace creer a una persona que no sabe que lo está usando que se trata de otro ser humano. Turing trató algunas y en la mencionada monografía las rebate; otras surgieron posteriormente en relación a la validez del Test de Turing.

Como las respuestas han sido muy numerosas e ingeniosas, nos limitaremos a las que se adelantó Turing en responder y otras pocas que han sido muy comentadas. De manera que vamos sólo a tratar las objeciones a priori a una máquina pensante según la monografía de Turing sobre el juego de la imitación y no a la prueba empírica.

#### A. La objeción teológica

“Puesto que sólo Dios puede crear el alma inmortal sin la cual no hay pensamiento, el hombre no puede construir un artefacto que piense” Turing (1964).

Turing advierte, al rebatir esta objeción, que a él particularmente no le impresionan los argumentos teológicos; ya antes han mostrado ser falsos, como los fueron cuando Galileo. Sin embargo, sugiere que puesto que los humanos constantemente estamos siendo instrumentos de Dios por vías naturales para que sean creadas nuevas almas, no es imposible para Dios darle alma algún artefacto humano que llene las condiciones materiales para que piense.

#### B. La objeción del avestruz.

“Las consecuencias de una máquina pensante son tan medrosas que es preferible no considerar su construcción” Turing (1964).

Con razón Turing opina que no es necesario refutar tal objeción desde un punto de vista científico; según él, le vale más a quienes respaldan tal objeción consolarse con la idea de la transmigración de las almas. Sin embargo, creemos que sí es conveniente considerar esta objeción desde un punto de vista ético al considerar las consecuencias de una máquina pensante.

#### C. La objeción matemática.

Turing, entre otros, demostró que hay limitaciones en los sistemas formales que únicamente el hombre conoce; puesto que un computador es un sistema formal, una máquina de Turing, está sujeto a tales limitaciones y hay verdades para el hombre que no son asequibles a la máquina; en consecuencia una máquina no puede imitar el pensamiento humano en su totalidad.

Turing admite que esta es una objeción fuerte, pero que si es cierto que hay preguntas que la máquina no puede contestar con un sí o un no, o dar una respuesta errónea que el interrogador sabe que lo es, tal máquina puede ser reemplazada por otra que si la conteste correctamente. Se da el hecho matemático, que a su vez ésta no podrá contestar correctamente otras preguntas, y el interrogador lo sabe, con lo que el humano se debería sentir superior a la máquina,

## ¿Qué Pasó con el Test de Turing?

pero, los hombres también dan respuestas falsas y el juego de la imitación sigue siendo válido. Este tema ha sido muy discutido en la literatura y en la opinión de sus críticos, Turing no es consistente en este punto, lo que para algunos filósofos descalifica el Test de Turing

### D. La objeción del consciente.

“Una máquina pensará cuando los símbolos que procesa tengan algún sentido para ella” Turing (1964).

Este es el argumento de la intencionalidad, del sentido o el problema semántico de la referencia: que nada material tiene intencionalidad o contenido; sólo el pensamiento trata acerca de algo, hay una diferencia entre la materia y el pensar de clase. Turing lo trató así: puesto que se arguye que una máquina no pensará hasta tanto no escriba un soneto y sepa que lo escribió; componga una melodía y sienta su música; en fin, hasta tanto no tenga conciencia de lo que hace. Entonces, habría que preguntarse ¿cómo sabemos que otro hombre tiene conciencia o que una máquina piensa? Únicamente, entonces, siendo ese hombre o esa máquina en particular y pensar que se piensa. Solamente quien sea un solipsista recalcitrante no aceptaría su Test, dice Turing; no hay otra forma de creer que un hombre piensa, siente dolor o alegría, etc., si no es por medio de cualquier tipo de diálogo o manifestación; particularmente si podemos conversar sobre lo que siente.

Al efecto, supóngase que una máquina imprime un verso como “Lo Fatal” de Rubén Darío (en su monografía, Turing recurre a William Shakespeare): “Dichoso el árbol que es apenas sensitivo y más la piedra dura porque esa ya no siente, pues no hay dolor más grande que el dolor de ser vivo, ni mayor pesadumbre que la vida consciente...”

Turing mantiene que si sobre un verso como éste, una máquina puede sostener un diálogo (omitiendo el jugador B en el juego de la imitación), tal como éste, por ejemplo:

C: Muy hermoso su verso, pero ¿no se contradice al afirmar simultáneamente que árbol y piedra son dichosos y a la vez que no sienten?

A: En tal contradicción está el sentido.

C: ¿Podría explicarse mejor?

A: Ud. no puede esperar que la fría lógica de la razón rija la poesía. La intención de “Lo Fatal” es expresar la angustia existencial de tener conciencia. Así se contrasta el sentido humano con el no sentir de la materia: el poeta ve detrás de las formas externas otras cosas; ese algo misterioso no es aprisionado por la razón sino por la actividad creadora. Para el poeta, árboles y piedras pueden no sentir y por ello ser dichosos” (Castillo, 1993).

Nadie debería dudar de que sabe lo que ha respondido— afirma Turing.

### E. La objeción de la variedad de inhabilidades.

La máquina puede hacer lo que Turing le atribuye pero nunca podrá hacer X (donde se sustituye a X por cualquier comportamiento: ser creativa; tener iniciativa; sentido del humor; amar y que alguien la ame; etc.). Según Turing, pueden ser hechas éstas y más cosas por una máquina que apruebe su Test, la objeción surge por inducción de las máquinas que se conocían para su época; en cincuenta años, se producirán máquinas con suficiente capacidad de memoria (109 bits) como para imitar cualquier comportamiento humano, ya que la objeción de las limitadas potencialidades de variedad en comportamiento mecánico es simplemente una objeción a la capacidad de memoria donde programarlos.

### F. La objeción de Lady Lovelace.

(Primera programadora de la historia, con la máquina analítica de Charles Babbage en 1870.)

“Una máquina nada más puede hacer lo que está programado.” (Castillo, 1993).

Suele presentarse también este argumento bajo la idea de que una máquina no puede sorprendernos. Turing advierte que él ha sido sorprendido con frecuencia por los resultados de sus programas; decir que una máquina no puede sorprendernos es pensar que podemos figurarnos todas las consecuencias que pueden tener los axiomas de un sistema formal, por caso; y esto no es así.

### G. El argumento de la continuidad del sistema nervioso

“El sistema nervioso no es una máquina discreta” Turing (1964).

Aunque Turing lo rebate con la posibilidad de simular un computador analógico en uno digital, el paradigma algorítmico de los sistemas de procesamiento de información en que el hardware no es necesario y comienza con Turing en la monografía que tratamos, hoy en día ha vuelto con mayor fuerza bajo el paradigma conexionista o neural. En todo caso, que una máquina imite el pensamiento humano de manera algorítmica o de forma continua y holística como se da en las redes neurales, no hace menos máquina ni menos física a la mente. Ya sea una máquina de Turing o una red neural la que logre imitar la mente humana.

### H. La objeción de la percepción extra-sensorial

“El pensamiento es un fenómeno en que lo extrasensorial – telepatía, clarividencia, etc. es relevante” Turing (1964).

Turing tomó muy en serio esta objeción. Pensaba que la percepción extra-sensorial sustentada por muchos casos estadísticos, se resiste a un tratamiento científico, a pesar de que en la práctica las teorías científicas lo pasaban bien sin tomarla en cuenta. Sin embargo, el Test de Turing sería burlado por un interrogador con poderes extra-sensoriales; sólo se le ocurría a Turing modificar el juego de la imitación aislando al jugador “C” en un cuarto a prueba de tales facultades. Pocos han comentado sobre tal versión del juego.

### I. La objeción del comportamiento informal

*Esta objeción que en la práctica ha creado la barrera más difícil de superar que ha enfrentado la inteligencia artificial: la representación y manipulación del conocimiento cotidiano, de la vida diaria, del sentido común: la información acumulada por el hombre dentro de una cultura.*

*Formular una teoría sobre el conocimiento mayormente compartido por los seres humanos, sobre las casi infinitas circunstancias que enfrentamos con tanta habilidad y sin esfuerzo en la vida diaria, de manera que podamos programarlo en un computador, es lo que ha venido a conocerse en inteligencia artificial como el problema del sentido común; el cual se resiste hasta la fecha a una total solución.*

*Para Turing, como para los primeros investigadores de la inteligencia artificial, el problema se reduce a descubrir todas las reglas de conductas humanas: aquellas que no podemos dejar de seguir, como las que adoptamos a voluntad. De manera que podemos programarlas en una máquina para que resuelva, como lo hacemos los humanos, qué hay que hacer en cada situación de la vida cotidiana: tomar un ascensor, planificar una ruta que nos lleve de la casa al aeropuerto, esperar en una antesala para acudir a una entrevista de trabajo y así, casi al infinito.*

*Quienes sostienen la objeción creen que no hay tales reglas; para Turing y sus inmediatos seguidores, es posible que no se haya indagado lo suficiente. Posiblemente un centenar de miles de hechos, creencias y procedimientos serían insuficientes para disponer de patrones con que enfrentar casi todas las posibles circunstancias. La investigación adecuada y la construcción de máquinas con suficiente memoria deberían de resolver el problema. Para Turing no deberían de pasar cincuenta años antes de alcanzar su solución.*

*Al estudiar el avance tecnológico de la inteligencia artificial, aunque las máquinas han alcanzado casi la capacidad de memoria prevista por Turing, el problema del sentido común no da señales de solución; sin resolverlo, difícilmente pueda construirse una máquina que apruebe el test de Turing". Aunque en escenarios limitados la robótica muestra alcances asombrosos.*

*Un argumento muy interesante, porque se refiere a los resultados del juego no considerados en su total implicación por Turing, es el que un ser humano pueda simular a un computador y el interrogador (C) confundirlo con una máquina; entonces, ¿afirmaría el interrogador que no piensa? Después de todo, los humanos venimos por siglos ejecutando algoritmos que nos hacen más eficientes con relación a generaciones anteriores; el gran poder del computador es precisamente ejecutarlos a velocidades muy superiores a las nuestras*

*La ejecución de esos algoritmos es mecánica, tanto en el hombre como en la máquina y no se está necesariamente pensando durante su ejecución. Supóngase, entonces, que un número significativo de personas, por caso, 1.000.000, reunidas en algún lugar limitado, de manera que cada una represente un carácter y que estas personas puedan ejecutar el programa con que la máquina (A) juega a imitar a un ser humano. El millón de personas no está pensando, sino que*

*en un momento dado el conjunto de ellas tiene el estado X del algoritmo. Su comportamiento es idéntico al del jugador (A), pero ninguna de estas personas está pensando sino actuando mecánicamente (pudieran estar pensando cada una en distintas cosas, pero ello resulta irrelevante, para el algoritmo); si sus resultados son los de (B), entonces habría que atribuirle siguiendo a Turing un solo pensamiento a todo el conjunto de la multitud, cuando esto, evidentemente, no es así.*

*Otro posible resultado del Test de Turing es que, a pesar de las predicciones, el interrogador siempre llegue a identificar correctamente a la máquina y al ser humano; entonces, ¿estaría Turing dispuesto aceptar su derrota? Según su propio test no se le debería atribuir pensamiento a la máquina en este caso, puesto que el interrogador sabe quién piensa y quién no; esto es: quién es la gente y cuál la máquina. ¿Cuántas veces habría que repetir el juego con distintas personas para justificar resultados?*

*En síntesis, en opinión de críticos al juego de la imitación, como dirimente de la controversia de quién o qué piensa, Turing le ve un solo resultado: que el interrogador engañado por la máquina admita que ésta piensa. Como hemos visto, pueden darse otros resultados no considerados por Turing: que un ser humano pueda ser confundido por una máquina en el mismo juego, por lo tanto hay que afirmar que no piensa; que la mayoría de las veces se dé más bien el caso que el interrogador hábilmente logra diferenciar a la máquina de la persona, por lo tanto Turing y sus seguidores admitirán que la máquina no piensa; y, en general, que a pesar del comportamiento aparentemente inteligente de la máquina, el interrogador buscado afirme: "Bien, he sido engañado, creía que era una persona, como es una máquina creo que no piensa; puesto que una cosa es comportamiento y otra conciencia" (Castillo, 1993).*

*Un artificio muy interesante que justificaría la última posición es el ideado por John Searle y que se conoce por el cuarto chino según Castillo (op. cit): a una persona que no habla chino ni siquiera conoce los caracteres chinos, se le entrega un juego de cartas con símbolos chinos y reglas en su idioma de como compilarlas de tal manera que produce oraciones con sentido en chino, pero con una total ignorancia acerca de su significado por parte del compilador. Así mismo, el computador que pase el Test de Turing producirá resultados significativos sin ninguna conciencia de ello y por lo tanto no piensa. El juego de la imitación, por consiguiente, no es prueba que determine que una máquina piense.*

*Las objeciones al juego de la imitación, rebatidas o no, serían posiblemente, como sugiere Turing, no resueltas sino disueltas cuando se construya una máquina que apruebe el test; ya que para entonces las palabras máquinas y pensantes tendrán distintas connotaciones y no serán una contradicción en término: se podrá hablar de máquinas pensantes sin lugar a contradicción o caer en un sin sentido. Cumplido el plazo propuesto por Turing, ya nadie toma como relevante el test para decidir si una máquina piensa o no.*

*Turing creía, como hemos antes dicho, que el problema es el de complejidad y capacidad de la máquina. El cerebro humano contiene de 10<sup>10</sup> a 10<sup>14</sup> neuronas, y, según Turing, bastaría nada más que 10<sup>9</sup> o quizás sólo 10<sup>7</sup> para llevar a cabo la mayoría de las*

## ¿Qué Pasó con el Test de Turing?

funciones inteligentes de diálogo, aprendizaje, comprensión, planificación, etc., ya que gran parte de la corteza cerebral está dedicada a atender las percepciones visuales especialmente.

Entonces, un computador con tan sólo 107 bits sería suficiente para jugar a la imitación (las estimaciones hoy en día arrojan una cifra mayor  $4 \times 10^{15}$  en base al conocimiento actual de la neurofisiología). En cuanto a la velocidad requerida, Turing creía que un milisegundo por ciclo era suficiente, ya que las neuronas se disparan en tal intervalo; pero el aumento en velocidad facilitaría la ejecución de los programas y compensaría los procesos seriales del computador con los paralelos del cerebro. Por otra parte, no era, según Turing, necesario esperar medio siglo para comenzar a programar la máquina. Podría comenzarse con las máquinas disponibles a principio de los años cincuenta, partiendo de un computador cargado con el conocimiento que tiene un niño (calculado hoy en día en medio millón de hechos), para luego por un proceso de aprendizaje, a través de los años y a medida que se aumentara la capacidad de memoria con máquinas más avanzadas que estuvieran disponibles, el programa iría añadiendo conocimiento evolucionado hacia el nivel de un adulto y entonces sería posible practicar el juego de la imitación.

La tecnología de la computación ha sido más que generosa en el incremento de memoria y velocidad, relativamente la tecnología de mayor crecimiento y productividad en la historia (si el transporte hubiere evolucionado igual, se viajaría de una ciudad a otra en minutos a un costo menor de un dólar), así que las predicciones de Turing han sido confirmadas en cuanto al poder computacional por él esperado antes de finalizar el siglo. Pero, en cuanto a la posibilidad de programarla con el conocimiento de un niño y la capacidad de aprender para que evolucione, como lo hace un ser humano intelectualmente en adulto, la historia es diferente.

El considerado mayor lógico desde Aristóteles, Kurt Gödel creía que mente y cerebro eran diferentes en cuanto a que la mente tiene la capacidad infinita de posibles estados y el cerebro no, según Wong (1998). El escritor Jostein Gaarder en su hermosa historia de la filosofía para jóvenes, afirma que "si nuestro cerebro fuera tan simple que pudiéramos comprenderlo, seríamos tan estúpidos que no podíamos comprenderlo" (Gaarder, 1954). Si unimos esta paradoja a la afirmación de Gödel, concluimos que si entendemos el cerebro es porque lo hacemos con la mente, no con el cerebro; cerebro y mente son distintos. El cerebro es un instrumento de la mente: el fantasma en la máquina.

### 2.2. El Problema sigue en pie y las dos tesis de la Inteligencia Artificial

El caso es que el test de Turing intenta resolver el problema desde el punto de vista de un observador; es decir externo, esto es según el comportamiento de lo observado, no intrínsecamente, interno. En tal sentido hay dos tesis para la inteligencia artificial:

- La tesis débil: que admite que la inteligencia y el pensamiento pueden simularse o imitarse artificialmente en una máquina, sin considerar que la máquina tenga en sí o no autoconciencia ni piense.

- La tesis fuerte: la cual postula que si una máquina se comporta en tal grado de inteligencia que no podemos diferenciar tal comportamiento del de un ser humano cuando realiza las mismas tareas... entonces, debemos admitir que la máquina piensa y no podemos negarle autoconciencia como lo hacemos con los demás seres humanos, además de nosotros.

El afamado físico, Roger Penrose (1994), en dos libros bestsellers *The emperor new mind* y *The shadows of the mind* Penrose (op. cit) amplía las tesis débil y fuerte así:

1) Todo pensamiento es computacional. En articular, el conocimiento de nuestro pensamiento consciente es evocado sólo por el proceso computacional apropiado.

2) El pensamiento consciente es una expresión de la acción física del cerebro. Y aunque cualquier acción física puede simularse computacionalmente, la simulación computacional por sí sola no puede evocar el pensamiento consciente.

3) La acción física adecuada del cerebro puede evocar el pensamiento consciente. Pero, esta acción física no puede ser simulada computacionalmente de manera apropiada.

4) pensamiento consciente no puede ser explicado por procesos físicos computacionales o por ningún otro termino científico.

Obviamente, 1) es la tesis fuerte y 2) la tesis débil; mientras que 3) es la negación de la tesis débil y 4) de la fuerte.

De acuerdo a Penrose (op. cit), el teorema de Gödel o "Gödel Proof" con que ha probado formalmente la existencia de funciones no computables decide a favor del punto 4, puesto la comprensión matemática tiene un acceso a la verdad matemática, es decir que la conciencia humana aprehende la verdad matemática de manera platónica, por algún acceso directo al mundo platónico de las ideas, lo que el filósofo Henry Bergson llamaba la "intuición", pero que nuestra ciencia actual no sabe explicar; necesitamos de una nueva ciencia de la mente basada en una nueva física.

Los científicos de la computación programan la IA sobre la base de la tesis débil. No consideran el problema de la tesis fuerte para sus avances. Y hay que hacer notar que los programas científicos, tesis débil o fuerte, de la IA y los modelos de la relación mente-materia o mente-cerebro de las ciencias cognitivas que toman al computador como modelo, y al software-hardware como paradigma de relación mente-cerebro, se fundamentan en una tesis filosófica: el Computacionalismo. Que dice que el significado o el pensamiento se generan por un proceso formal sintáctico esto es automático; que en el procesamiento de símbolos está la base del pensamiento consciente y hasta del auto-conocimiento.

Tesis sostenida originalmente por Thomas Hobbes, materialista en la tradición empirista británica del siglo XVII, que afirmó que el pensamiento son marcas en el cerebro, y podrá ser explicado por las leyes físicas como el comportamiento de la materia es explicado (en su tiempo) por las leyes descubiertas por Newton.



Los filósofos modernos no lo ven así. Numerosas monografías y ensayos se han publicado en revistas de filosofía tratando la materia y presentando soluciones de acuerdo a la Escuela y sus tendencias. A continuación vamos a tratar la propuesta por el filósofo venezolano doctor Ernesto Mayz Vallenilla de la Escuela Fenomenológica de Hursell-Heidegger, y el tratamiento que se le dado dentro del método de la Filosofía Informacional; porque se considera una propuesta original y novedosa, y presenta una solución intrínseca superior a las otras propuestas hechas para sustituir el Test de Turing por un procedimiento con mayor validez.

### 2.3. Una prueba formal y el test mayziano

El Dr. Ernesto Mayz Vallenilla— en una prueba bautizada como test mayziano por Castillo y Marante (2002)— propone un nuevo enfoque para diferenciar las opciones 1) y 2); esto es que una máquina tenga pensamiento o solamente lo simule. Según Mayz, una máquina tendrá pensamiento si disfruta de libre albedrío; valga decir, si es capaz de fijarse — de manera autárquica y autónoma— fines distintos a los que le impuso su programador; esto es, si por sí misma permite y propicia la negación de sus propios sustentáculos. Dice Mayz:

“...la autoconciencia entendida como una conciencia consciente de sí misma por sí misma... cuyo único responsable y libre ejecutor es cada hombre por mandato expreso de su propia libertad. Entendida así la autoconciencia...ella no puede ser poseída por ningún artefacto, sea cual fuere su índole o nivel, a menos que semejante artefacto despliegue por sí mismo la requerida autarquía y autonomía que distinguen a la libertad humana. Un libertad programada, en cualquier caso no es sinónima de genuina libertad.... A menos que desde sí misma permita y propicie la negación de su propio sustentáculo, erigiendo en lugar suyo un reino de fines, opuestos a los de aquél, elegidos y queridos espontáneamente. (Mays, 1990).

En términos computacionales, puesto que una máquina es un sistema formal, este test se aprobaría formalmente si cuando una máquina, a partir de sus axiomas (estados iniciales) no alcance a demostrar teoremas (estados objeto) que le son vedados por sus deducciones, se modifique a sí misma y por sí misma— de manera autárquica y autónoma—, para que con tal cambio se den las demostraciones buscadas y se alcancen los estados objetos perseguidos.

En otras palabras, una máquina se fijará sus propios fines si puede escoger sus futuros estados y tendrá libertad si tales estados no son deducibles de los axiomas que son sus sustentáculos; en cuyo caso, tendrá que sustituir los estados actuales por otros escogidos. La máquina negará, entonces, sus sustentáculos, sus axiomas, se fijará nuevos fines para los cuales no fue diseñada, y gozará de libre albedrío; en consecuencia tendría autoconciencia de lo que hace...y pensaría.

Demostrándose así la tesis fuerte de la IA. Pero, si por los teoremas de las limitaciones de los formalismos demostrados por Turin, Post, Church, Kleene, Gödel... y otros lógicos matemáticos en los años 30 del siglo XX; no es posible, matemáticamente razonando, tal proeza; se elimina la tesis fuerte y queda solo la débil, a la que se le puede atribuir el gran éxito práctico, no teórico de la IA.

Puesto que, hasta el momento, no hay teorías consistente que puedan predecir el comportamiento futuro de un sistema dinámico mecánico arbitrario; valga decir, una máquina de Turing, un computador, en cualquiera de sus versiones clásicas o cuánticas.

En Castillo (1997), Castillo y otros (2000), Castillo (2001), el ya citado Castillo y Marante (2002), y Castillo (2012) y Castillo (2018) se busca el tratamiento formal de la Meta-técnica como Nanotecnología, dentro de un nuevo paradigma para enfocar problemas tradicionales de la filosofía con el concurso de especialistas en el dominós multi y transdisciplinario de las ciencias de la información y la filosofía que se denomina Filosofía Informacional, a la que dedicamos las últimas palabras de esta monografía

### 2.4. La filosofía informacional

Desde los años finiseculares del siglo XX, un grupo de científicos y filósofos que se han autodenominado “information philosophers”,— denominación que traducimos por “filósofos informacionales”—, abordan los problemas tradicionales de la filosofía, con un nuevo método filosófico basado en la física moderna, la biología, la neurociencias y las ciencias de la información; en cuanto a la base común de estas ciencias para explicar la creación y desarrollo de cuanto existe (ontología) y conocemos (epistemología) por la información que se crea, se mantiene y aumenta en lo existente.

Afirman que el origen y organización del cosmos y su evolución (cosmogénesis); de la vida (biogénesis); y del hombre (antropogénesis); parten del caos en forma creativa y emergente hacia ordenes cada vez más complejos denominados ergódicos, porque son sistemas que evolucionan, una vez creados, con entropía negativa, o neguentropía, en oposición a la Segunda Ley de la Termodinámica que predice el desorden total, entropía positiva, y el fin del universo.

Una filosofía o concepción de un universo abierto, indeterminista, continuamente en emergente creación y mayor complejidad en algunos de sus componentes, como son los seres biológicos autoconscientes, que en la especie del homo sapiens son sistemas dinámicos impredecibles, indeterministas, poseen creatividad y libre albedrío, y en consecuencia intención y propósito o telos; en lugar de otra filosofía o concepción del mundo, concebida antes de la física cuántica y las teorías de la información del siglo XX— según la cual el universo es un sistema dinámico que de un estado inicial de alta información, reduce permanentemente la organización y la información de manera determinista, sin propósito y valor alguno— ; ofrece soluciones experimentales, verificables o refutables, a algunos problemas metafísicos de qué existe (ontología), cómo lo representa en la mente y conoce la conciencia humana (epistemología), y cómo lo usa. En la filosofía informacional se presentan tres teorías filosóficas seculares para las que ofrece un nuevo enfoque: la teoría del conocimiento, la teoría de la mente y la teoría del libre albedrío.

El Test Mayziano descrito antes informalmente, es un ejemplo de cómo puede aplicarse el enfoque interdisciplinario y transdisciplinario de la Filosofía Informacional para el avance en la

solución de problemas filosóficos, cuando filósofos trabajan en concierto con científicos de las ciencias de la información; en este caso los proyectos de la Meta-técnica y la Computación clásica y cuántica llevados a cabo por investigadores del Departamento de Filosofía de IDEA y la Unidad de Investigación en IA de la UCLA; que se realizaron entre 1997 y el 2005 y cubrieron tópicos tales como la nanotecnología como meta-técnica; pensamiento complejo y transaridad; fundamentos matemáticos de la meta-técnica; paradigmas clásicos y cuánticos del libre albedrío y otros más.

Es de esperarse, como se hace en centros avanzados de investigación sobre sistemas computacionales, que los fundamentos filosóficos de la computación sean incluidos en los proyectos nacionales de investigación con el apoyo de la Sociedad Venezolana de Computación (SVC).

### 3. EPÍLOGO

En el siglo XXI, como se dijo al comienzo de este artículo, sólo los filósofos a quienes se les acusa de no avanzar en sus investigaciones, se ocupan de la búsqueda de un test, prueba o argumento, como el que sugirió Turing hace siete décadas, para responder a la pregunta ¿puede una máquina pensar?; que fue la forma en que el siglo XX se planteó el secular problema mente y materia, ya con 25 siglos de antigüedad — estos es, si materia y mente son dos entes distintos y diferenciados o una misma cosa—; entonces, si son diferentes, una máquina no puede pensar; pero, si son iguales, sí puede; entonces el Test de Turing sigue siendo una propuesta a considerar como una posible respuesta empírica.

Quiere decir que está vigente para los filósofos, y como antes expusimos hay nuevas formas de considerarlo partiendo de los avances en las ciencias, particularmente las de información y sus efectos en la nueva escuela filosófica conocida como filosofía informacional.

Pero, el problema de máquinas que asumen tareas en la sociedad, en el trabajo y el empleo, antes sólo posibles al ser humano, como son las que requieren funciones cognitivas de percepción y conocimiento, sí está en la palestra de los científicos socioeconómicos y políticos.

Centenares de artículos en revistas especializadas, tesis de postgrados y ensayos como libros sobre la materia que son publicados anualmente, tratan el asunto con gran preocupación y es tema de la agenda en los congresos profesionales de especialistas sobre los efectos de la inteligencia artificial en la organización social, política y económica del mundo.

Obviamente deberá ser enfrentado en los próximos años de manera ineludible. Pero, ya entonces para nada servirá preguntarse si una máquina puede o no pensar, sino como evitar que nos quiten el empleo y pasar a ser una masa de desempleados e indigentes sustituidas por las máquinas.

### 4. REFERENCIAS

Castillo Vicci, A. (1997). *Crítica a la Teoría Computacional de la Mente*. UCLA-IDEA, Mayo 1997.

Castillo Vicci, A. (1993). *Machina Ratiocintarix. En busca del Razonamiento Automático*. Fundaciones Banco Consolidado y Universidad Yacambú. Caracas, Venezuela. Pp: 575-660, Diciembre, 1993.

Castillo Vicci, A. y Marante, G. (2002). *Inteligencia Artificial y Libre Albedrío en Paradigmas Clásicos y Cuánticos*. Revista de la Sociedad Argentina de Filosofía. TomoVIII, Córdoba, Argentina. Pp\_347-368., Diciembre 2002.

Castillo Vicci, A., Jiménez, D. y Vallota, A. (2000). *Técnica y meta-técnica de la computación*. UCLA-IDEA, Septiembre 2000.

Castillo Vicci, A. (2001). *Inteligencia Artificial y Meta-técnica*. Revista de la Sociedad Argentina de Filosofía. Año X-XII, 2001.Pp: 151-174. Córdoba, Argentina, diciembre 2001.

Castillo Vicci, A. (2012). *Fundamentos Matemáticos de la Meta-técnica*. Editorial Publica, Alemania-España, Mayo 2012.

Castillo Vicci, A. (2018). *La Lógica de la Filosofía Informacional*. EAE. Mayo 2018.

Descartes, R. (1975). *El Discurso del Método*. Editorial Vosgos, S.A., Pp: 80-83, Abril 1975

Gaarder, J. (1954). *El Mundo de Sofía*. Siruela/Norma. Bogotá. Colombia. Diciembre 1994

Mayz Vallenilla, E. (1990) *Fundamentos de la Meta-técnica*. Monte Avila Editores. Caracas pp. 84.

Penrose, R. (1994). *Shadow of the Mind: a Search for the Missing Science of Consciousness*. Oxford University Press. N.Y. Oxford., pp. 12, Abril 1994.

Turing, A. (1964) *Computing Machinery and Intelligence*, reproducido en *Mind and Machines*. Anderson Editor, Prentice Hall, N.J. USA, .pp. 4-40, Febrero 1964

Wang, H. (1988). *Reflections on Kurt Gödel*. MIT Press. Cambridge, Mass. Mayo 1988.