



¿Importa la determinación del sexo en el Test de Turing?

*Is gender identification important in
the Turing Test?*

Rodrigo González Fernández*

Universidad de Chile, Facultad de Filosofía y Humanidades, Santiago, Chile

Resumen

Según la versión estándar del juego de la imitación, la determinación del sexo de los participantes no desempeña ningún papel en el testeo de la inteligencia de máquina. Desafortunadamente, tal simplificación soslaya la teoría de la mente que fundamenta dicho juego. Teniendo en consideración este problema, en este ensayo argumento en contra de la simplificación del Test de Turing. En efecto, tal como sostengo, la determinación del sexo de los participantes no debe obviarse: la mente de una mujer y su inteligencia son imitables y no dependen de realización física específica. Esto ocurre porque el funcionalismo de Turing supone un vínculo entre imitación, engaño y aprendizaje, tres claves para que el proyecto de la Inteligencia Artificial tenga éxito.

Palabras Clave: Imitación. Aprendizaje. Realización. Funcionalismo.

* RGF: doctor en Filosofía, e-mail: rodgonfer@gmail.com

Abstract

According to the standard version of the imitation game, gender recognition of participants plays no relevant role in the assessment of machine intelligence. Unfortunately, this simplification neglects the theory of the mind behind the game. In view of this problem, I argue against the simplification of the Turing Test. Indeed, gender inspires the test and should not be neglected: a woman's mind and intelligence are imitable, because they do not depend upon any specific physical realization. This is based upon how Turing's functionalist approach supposes that imitation, deception and learning are three keys for Artificial intelligence to succeed.

Keywords: *Imitation. Learning. Realization. Functionalism.*

"Neuroscientists are novices at deception"
(Raymond Joseph Teller)

Introducción

La tradición filosófica ha debatido incansablemente acerca del juego de la imitación, elemento central del Test de Turing. Usualmente, se ha propuesto que tiene dos objetivos:

- I) El Test de Turing *define* qué es la inteligencia;
- II) El Test de Turing propone condiciones necesarias o suficientes para dicha inteligencia.

A diferencia del debate acerca de los objetivos, este trabajo argumenta sobre la importancia de la primera versión del juego de la imitación, indicador de la filosofía que inspira a Turing y su proyecto.

Tal como nuestro aquí, la versión estándar, simplificada del test, ha tenido un impacto negativo sobre la apreciación de la teoría de la mente que inspira a Turing, lo que a su vez ha hecho ignorar su proyecto de IA de largo plazo. A diferencia de dicha versión, algunas lecturas alternativas intentan destacar la importancia de la 1ª etapa del juego de la imitación, pero al mismo tiempo enfatizan cómo la filosofía de Turing es inspirada por un proyecto de género. Este

tendría como objetivo central hacer una declaración a favor de seres inteligentes con nuevas oportunidades de género y sexuales. Detalles de su vida, como su homosexualidad y las trágicas condiciones de su muerte, han intentado fundamentar estas discutibles hipótesis. Con todo, tanto la versión *estándar* como las lecturas alternativas soslayan el enfoque funcionalista de Turing, el fundamento teórico más importante para dejar atrás las viejas discusiones filosóficas y conceptuales sobre la inteligencia heredadas desde Descartes en adelante.

Este trabajo se divide en tres secciones y conclusión. En la primera sección se discute el objetivo central del juego de la imitación, en función de la disputa conceptual de la inteligencia de máquina heredada a partir de Descartes, quien descarta de plano la posibilidad de la inteligencia de máquina; en efecto, las máquinas carecerían de lenguaje y conducta flexible. En la segunda sección se critica cómo se omite la cuestión del sexo de los participantes en la mencionada versión *estándar* del juego. En particular, argumento que dicha cuestión es crucial para comprender qué tiene en mente Turing con imitar, y también para ponderar correctamente su proyecto: la imitación, el engaño y el aprendizaje son medios para crear inteligencia de máquina. En la tercera sección se ofrecen razones sobre por qué Turing se basa en la imitación, relacionándola con cómo la inteligencia de mujeres y hombres puede aprenderse por parte de máquinas programadas, de un modo análogo a como lo hacen los niños.

La relación entre el Test de Turing y la inteligencia

Hay consenso respecto de que el test pretende poner punto final al problema de si una máquina puede pensar, o puesto de otra manera, si la inteligencia de máquina es posible en principio. Esta es justamente una pregunta de inspiración cartesiana que ya había provocado una larga y controvertida discusión. A mi juicio, sigue provocando polémica por la naturaleza del problema examinado y porque Turing asume que convencer a jueces de que X es inteligente es *suficiente* para que X sea inteligente.

Con todo, Descartes en diversos pasajes sostiene que es imposible en principio que una máquina tenga estados mentales como los de un ser inteligente, con racionalidad. A propósito de si lenguaje e inteligencia se relacionan asevera: “una máquina no podría ordenar las palabras de formas diferentes para responder al significado que se dice en su presencia, como incluso el menos inteligente de los humanos puede hacer” (DESCARTES, 1994, p. 113). Luego, una máquina podría usar palabras, pero lo haría de modo mecánico y automático, sin ser sensible y flexible al contexto y al ambiente.

Para este filósofo francés, animales y máquinas carecen de inteligencia *humana*, porque solo pueden reaccionar mecánicamente frente a los estímulos ambientales. A diferencia de ambos, el ser humano es capaz de producir construcciones lingüísticas de diversas formas para significar exactamente lo mismo. Esto se debe a que el lenguaje, puesto de una manera simple y breve, implicaría racionalidad según Descartes, y esta, a su vez, sería signo inequívoco de flexibilidad e inteligencia. En consecuencia, el lenguaje contaría como una prueba empírica confiable de la existencia de racionalidad y, por tanto, de inteligencia. En consecuencia, para Descartes ser convencido de que X es inteligente porque tiene lenguaje permite asumir que lo es.

La flexibilidad conductual denota racionalidad e inteligencia. Con respecto a la relación entre conducta e inteligencia, Descartes argumenta lo siguiente:

Incluso, aunque tales máquinas podrían hacer algunas cosas tal como nosotros les hacemos, o de mejor forma, inevitablemente fallarían en otras, lo cual revelaría que no están actuando de acuerdo con su entendimiento, sino por la pura disposición de sus órganos (DESCARTES, 1994, p. 113).

Luego, la ausencia de lenguaje y la conducta mecánica en animales representa evidencia empírica de que ellos y las máquinas no son inteligentes a la manera humana. Cabe destacar que la calificación de no inteligentes de animales y máquinas se asocia con la distinción metafísica entre dos substancias propuesta por Descartes. Los animales son máquinas, en la medida que su corporalidad, su *res extensa*, está

constituida por complejos mecanismos dispuestos para reaccionar de manera causal al ambiente. La *res cogitans*, por el contrario, es flexible, capaz de producir conducta racional para la resolución de problemas complejos, cuestión que no sucede por una mera relación causa-efecto, sino por la participación del entendimiento.

La filosofía cartesiana fue definitivamente desafiada en el siglo XX, cuando Alan Turing intentó poner punto final a la discusión de si las máquinas con estados mentales e inteligencia eran posibles en principio. La estrategia adoptada por Turing en 1950 no consistió en proponer un nuevo argumento basado en la discusión de conceptos como 'piensa' y 'máquina'. Por el contrario, diseñó un test, un método empírico preciso, que fuera capaz de reemplazar la pregunta: ¿Pueden pensar las máquinas? Pero, ¿por qué Turing quiere abandonar la arena conceptual, en consideración de que su discusión entra de lleno en la filosofía de la mente y, en particular, en el problema de las otras mentes?

Según él, una buena manera de evitar la pregunta de si las máquinas piensan consiste en rechazar el análisis del uso común de conceptos como 'piensa' y 'máquina'. En efecto, analizar su significado podría llevar a una discusión de definiciones, y así a una suerte de encuesta sobre usos de esos términos. El siguiente pasaje es claro respecto de este punto:

Si la exploración del significado de términos como 'máquina' y 'piensa' se debe efectuar a partir del análisis de cómo estos se usan regularmente, es difícil evitar la conclusión de que el significado y la respuesta a la pregunta: ¿Pueden pensar las máquinas?, debe encontrarse a través de una investigación estadística similar a una encuesta Gallup (TURING, 1990, p. 40).

Por esta razón, propone un método empírico de reemplazo a la polémica pregunta. Con esto pretende evitar la discusión conceptual sobre instanciar la inteligencia como una propiedad material.

El test que plantea se basa en un juego, a saber, el denominado *juego de la imitación*. Consiste en que hay un hombre en una habitación (A), una mujer en una segunda habitación (B) y jueces cuyo sexo no tiene importancia (C). En rondas sucesivas, los jueces formulan preguntas

sencillas mediante un teletipo, con el objeto de *determinar* el sexo de los participantes en cada pieza. Las preguntas son simples y orientadas a ese fin, por ejemplo: ¿Tiene Ud. el pelo corto? O, ¿Usa cremas para la piel diariamente? Mientras que la mujer responde sinceramente a las preguntas del juez, el hombre debe intentar engañar a los jueces haciéndolos creer que es una participante de sexo femenino. Así, el juego de la imitación, al menos en su primera etapa, consiste en que un hombre induzca al error a los jueces. La primera etapa puede esquematizarse como en la Figura 1.

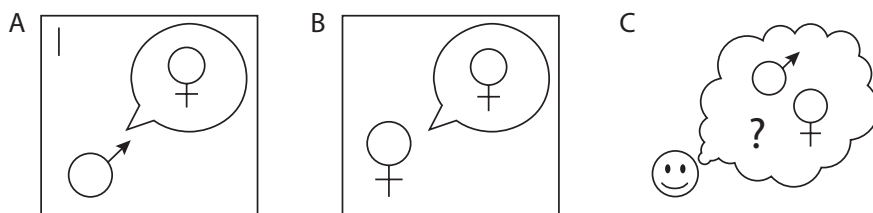


Figura 1 - Primera etapa del juego de la imitación

Pero, ¿qué tiene que ver el problema de la determinación del sexo de los participantes con la imitación de la inteligencia? ¿Por qué Turing pasa a la segunda etapa? Mucho se ha elucubrado sobre este cambio, por ejemplo, viendo un proyecto cultural emancipador de género. Sin embargo, un indicio que ayuda a entender el cambio a la segunda etapa son las tres preguntas que la motivan:

¿Qué sucederá cuando una máquina juegue el rol de A en este juego?
 ¿Decidirá el juez incorrectamente tan a menudo como cuando se juegue así a como se lo hará cuando en el juego participen un hombre y una mujer? Estas preguntas reemplazan nuestra pregunta original “¿Puede pensar una máquina?” (TURING, 1990, p. 41).

Así, Turing reemplaza la pregunta cartesiana sobre la inteligencia de máquina, al centrarse en la imitación. Esta es suficiente, según él,

para sostener que una máquina posee estados mentales, porque ¿qué razones tendríamos para negárselos? De alguna forma, Turing pone la carga de la prueba sobre aquellos que piensan que es necesario, metafísicamente, replicar la inteligencia para que algo sea inteligente. Para él, en cambio, es *suficiente* desde un punto de vista epistemológico que seamos convencidos de que A es inteligente. En consecuencia, Turing intenta eliminar una discusión metafísica para centrar el debate en la inteligencia de máquina sobre bases epistemológicas, a saber, aquellas en que se reemplaza la pregunta “¿Pueden pensar las máquinas?” con la evidencia empírica aportada por su juego¹.

La siguiente sección justamente examina cómo este involucra una visión vanguardista de la inteligencia: esta es imitable e independiente de la biología, o de cualquier otra realización material. Tal como argumento más abajo, esto es crucial para entender cómo Turing propone dejar atrás el criterio de la inteligencia propuesto por Descartes.

¿Es importante la cuestión del sexo en el juego de la imitación?

A diferencia de Descartes, Turing es un defensor de la llamada *inteligencia de máquina*, concepto basal de la Inteligencia Artificial según el cual el pensamiento mecanizado sí es posible en principio. Pero, tal como se analizará con mayor profundidad en la tercera sección, no es necesario que una máquina instancie propiedades biológicas específicas, o que tenga cerebro para que sea inteligente. Por el contrario, su enfoque funcionalista asume que la inteligencia es una propiedad ajena a las realizaciones físicas de un sistema cognitivo y así, a lo biológico.

Para comprender este enfoque funcionalista anti-biológico de Turing y especialmente su polémica propuesta vanguardista sobre la inteligencia de máquina, es necesario examinar con más detención el juego de la imitación y discutir algunas posturas que señalan que lo

¹ Muchos consideran que su propuesta de modificar la pregunta “¿Puede pensar una máquina?” con el juego de la imitación fracasa. El hecho de que este aporte evidencia es empírica hace que resulte al menos dudoso que las máquinas piensan, incluso si pasan el test. Véase especialmente Searle (1980, 1990).

central de este es la cuestión del género. En relación con este punto, cabe destacar que el juego no tiene que ver con la determinación de especies, sino que originalmente se centra en el sexo de los participantes, y esto es, a mi juicio, algo de importancia central. Precisamente, la segunda etapa que Turing describe apoya la tesis de que la identificación del sexo de los participantes juega un papel preponderante. Tal segunda etapa puede esquematizarse como en la Figura 2.

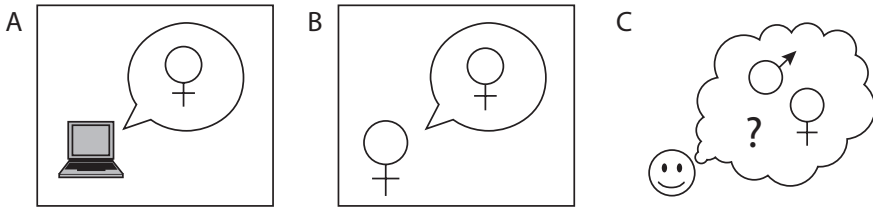


Figura 2 - Segunda etapa del juego de la imitación

Esta descripción ha motivado la denominada “versión de género” del test, y así contrariamente a los que consideran que el género es una pista falsa [*red herring*] que puede ignorarse del todo. Hay cuestiones crípticas en el rol que juega el género en el test, porque Turing describe de modo algo ambiguo que en la etapa 2 la máquina programada debe jugar “tal como [A]” [*like this*], de la Figura 1 (SHIEBER, 2004, p. 101). Esto ha motivado, precisamente, el debate entre la versión de género y la “versión estándar del test”. De hecho, en 1952 el propio Turing describe una versión simplificada de su test, dejando de lado la cuestión del sexo de los participantes. Tal versión propone el reemplazo de la pregunta sobre si las máquinas piensan así: “Quisiera proponer un tipo particular de *test* que uno podría aplicar a una máquina [...] La idea del test es que la máquina trate y pretenda ser un hombre” (apud COPELAND, 2001, p. 6). Con base en esta simplificación, filósofos y científicos cognitivos han tomado la versión estándar como la esencia del test, ignorando del todo la cuestión del sexo de los participantes. La influencia de la simplificación puede apreciarse en este pasaje:

“Turing propuso un test (conocido como el Test de Turing) que involucra a un juez humano, el que trata de descubrir mediante respuestas verbales si quien conversa es un humano o una máquina” (CLARK, 2001, p. 21).

La versión estándar supuestamente cuenta con la ventaja de centrarse solo en la imitación, tal como se ilustra en la Figura 3.

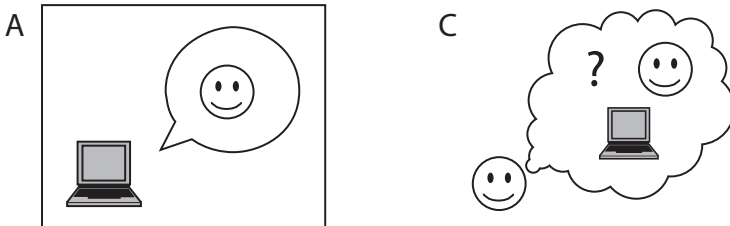


Figura 3 - Versión estándar del juego de la imitación

Pese a esta versión, parece central en el juego de la imitación cómo un hombre puede imitar la conducta lingüística inteligente de una mujer, y no solo como una máquina puede hacerse pasar por un humano al imitar su conducta lingüística (o como una máquina puede hacerse pasar por cualquier especie). El sesgo epistemológico del enfoque de Turing puede apreciarse en su predicción: para el año 2000 el computador digital, con más velocidad y capacidad de almacenamiento, debería haber engañado al menos al 30% de los jueces (TURING, 1990, p. 49).

Sin embargo, la versión estándar ha dado lugar a toda clase de interpretaciones incorrectas sobre su filosofía y proyecto de largo plazo. Aunque parece más “simple y neutral” que las etapas 1 y 2, dicha versión soslaya la determinación del sexo de los participantes, lo cual también ha propiciado interpretaciones discutibles.

Por ejemplo, Genova (1994, p. 313-316) sostiene que podría existir una abierta declaración de Turing a favor de seres inteligentes con nuevas oportunidades de género y sexuales, y esto podría estar inspirado por el pensamiento de Charles Darwin, quien también esperaba la evolución de

nuevas formas dentro de un contexto naturalista. La hipótesis de Genova es sintetizada por Saygin et al. (2003, p. 59) de la siguiente manera:

Genova cree que las nociones sexistas acerca de la mujer siendo menos inteligente, por sí mismas, no explican el diseño particular del juego. Ella sostiene que al complicar el juego de esta manera, Turing cuestiona la existencia de categorías discretas. En otras palabras, al referirse al tópico mujer/hombre, él intenta demostrar que el género mismo es un concepto socialmente impuesto que no es “natural” de la forma en que usualmente pensamos que es.

Genova considera el juego de la imitación como parte de la filosofía general de Turing de “transgredir los límites” (Genova, 1994, pp. 313-16). Bajo el supuesto de que Turing admirara tales transformaciones que no encajan con las categorías discretas, Genova sugiere que Turing podría estar discriminando a la mujer como un ser pensante de categoría inferior por ser incapaz de engañar a un hombre.

Otro ensayo que también sostiene una postura discutible, centrada en el problema del género, y que se apoya en cuestiones psicológicas y biográficas de Alan Turing², es el de Lassègue (1996, p. 44-46). Por ejemplo, resulta curioso que la fecha del juicio a que Turing fue sometido por conducta inmoral, a causa de su homosexualidad, haya prácticamente coincidido con la entrevista que dio a BBC en 1952, en enero de ese año, y que en ese entonces haya habido preocupación pública por el descubrimiento de dobles agentes homosexuales trabajando para la KGB. ¿Habría motivado esto que Turing obviara la cuestión del sexo de los participantes?

Aunque el tópico de discusión central es el testeo de la inteligencia de máquina, y la posibilidad de que estas imiten la inteligencia humana, el sexo de los participantes es mucho más importante de lo que habitualmente se interpreta con la versión estándar.

A mi juicio, es indudable que la cuestión del sexo de los participantes es relevante para caracterizar por qué la imitación es *suficiente* para evaluar la inteligencia de máquina. El objetivo central del juego de

² Supuestamente, la vida de Turing sería particularmente relevante con respecto a las teorías que propone. Por ejemplo, se ha interpretado que su condición homosexual tuvo que ver con sus desarrollos teóricos.

la imitación es detener una discusión filosófica, cuestión que se logra una vez que se identifica el *quid* central de aquél, a saber, el engaño [*deception*] mediante imitación, criterio decisivo para juzgar la inteligencia al comparar el desempeño de una mujer con hombre y máquina (SAYGIN et al., 2003, p. 26), todo lo cual es *observable*. Por este motivo, es importante aclarar que imitar a una mujer (y eventualmente a un hombre) es una propiedad que, al igual que la inteligencia, no requiere de implementación material específica, como la biológica.

Que X imite a Y no requiere que X tenga las propiedades físicas de Y. Basta con que X engañe a observadores externos, haciéndolos creer que es Y. Que un actor sea maquillado y caracterizado como mujer basta para que el público crea que hay una actriz en el escenario. Que Don Juan imite el comportamiento de un enamorado, *engañando* doncellas al hacerlas creer que las ama para tener sexo con ellas, también muestra que el sexo es una propiedad separable, es decir, que puede instanciarse sin que haya amor.

En relación con esto, es importante aclarar que el punto central de Turing es que la imitación de conducta lingüística inteligente en función de un estándar con el cual comparar sería *suficiente* para determinar, inductivamente, si una máquina programada es inteligente, pese a no tener sexo, ni instanciar propiedad física o biológica alguna. En consecuencia, Turing cree que la *evidencia observable* aportada por el juego de la imitación es suficiente para exorcizar el mentalismo cartesiano y detener la interminable disputa conceptual al que este da lugar.

El propósito de la siguiente sección es examinar sucintamente el funcionalismo de Turing. Tal examen no discutirá el funcionalismo en general, sino que se centrará en cómo su postura da lugar a un programa para el desarrollo de la IA.

El funcionalismo de Turing: anti-biológico y programático

Para Turing la inteligencia es una *función no biológica*, puesto que lo que interesa respecto de la producción de estados mentales son las relaciones de input/output y los estados internos de una máquina. Todo

esto puede ser descrito mediante el funcionamiento de una Máquina de Turing (TURING, 1936), el fundamento conceptual para entender qué es una máquina programada o computador digital.

Una máquina programada se caracteriza por tener carácter multipropósito, es decir, por tener la capacidad de resolver problemas de diferente índole. Una Máquina de Turing, a su vez, es una idealización matemática de un dispositivo mecánico-abstracto y la base para entender qué es computar. Esta actividad consiste en la ejecución de un algoritmo³, cuyos pasos son finitos y calculables mediante lápiz y papel, sin la participación de la conciencia.

Para ejecutar un algoritmo la máquina opera de modo mecánico, ya que la tabla dicta su proceder. Tal como sucede con un algoritmo, el término de la ejecución de éste ocurre cuando la máquina logra resolver un problema mediante un procedimiento efectivo. Un “computador humano” que emplea lápiz y papel, también puede definirse en términos de una Máquina de Turing.

Tanto la Máquina de Turing como los algoritmos permiten caracterizar qué es computar. En realidad, esta máquina es una idealización de qué es computar, o de la ejecución de un algoritmo mediante la implementación de un programa. Por este motivo, usualmente se representa a la Máquina de Turing mediante una cinta infinita y una cabeza lecto-escritora (o *scanner*) que se mueve hacia la derecha y la izquierda, cambiando de estado interno.

Los estados en que se encuentra la Máquina de Turing están totalmente condicionados por los símbolos leídos, o estados discretos, y los mencionados estados internos. Por ejemplo, la cabeza se encuentra en un tiempo t en un estado interno específico ($q_0, q_1, q_2, \dots, q_n$), luego un símbolo de la celda es leído ($b_1, b_2, b_3, \dots, b_n$). Dado el estado interno, lo que se lee en la cinta y el micro código de la cabeza, esta mantiene el símbolo, o bien lo borra, imprimiendo otro. Luego, la cabeza para o se mueve a otra celda de la cinta, continuando las computaciones hasta

³ Puesto de una manera breve, un algoritmo es un procedimiento efectivo, o de la ejecución de pasos finitos expresados mediante reglas, donde uno de dichos pasos es recursivo y así puede ayudar a resolver un problema determinado.

que se detiene. Si esto sucede, se ha llegado al final del procedimiento y, por tanto, a la resolución de un problema.

Tradicionalmente, se representan los símbolos de una Máquina de Turing mediante 0s y 1s. Cabe destacar que su conducta está absolutamente condicionada por el programa, descrito mediante la mencionada tabla. Si la máquina está en un estado interno y lee una celda de la cinta, esta *dupleta* implica un input In: <estado interno, símbolo leído>. El programa, entonces, dicta una *tripleta* de output en función del programa: <escribe símbolo, movimiento (o detención) y nuevo estado interno>. Este On hace que la máquina escriba un 1 o 0 en una celda, que se mueva izquierda o derecha, ingrese a un nuevo estado, o que se detenga.

Tal como se mencionaba arriba, las dupletas de input y tripletas de output pueden describirse en forma de quintuplos y enunciados condicionales, todo lo cual ayuda a que se conforme la lista finita de reglas ordenadas de pares <input, output>.

Una cuestión crucial respecto de la Máquina de Turing es que puede imitar el comportamiento de cualquier otra máquina, sin importar los materiales de que está hecha. Lo único que se requiere es que ambas *funcionen* de la misma manera, es decir, que implementen el mismo programa. Nuevamente, el concepto de imitar es crucial aquí: si una Máquina de Turing imita a otra Máquina de Turing se tiene una Máquina Universal de Turing, la cual es un programa que ejecuta otro programa. Y esta última máquina puede imitar a cualquier otra máquina, con independencia de los materiales con que esté hecha.

Una Máquina de Turing puede imitar la conducta de un cerebro, pero no requiere de la “papilla neural” para tales efectos. Esta cuestión representa una de las tesis centrales del funcionalismo anti-biológico de Turing: para tener la inteligencia de una mujer, no es necesario tener sus propiedades físicas o biológicas. Basta con que el computador digital funcione como una máquina-cerebro para que resulte innecesario dudar si realmente tiene estados mentales y es inteligente.

El siguiente extracto de una entrevista radial a Turing en 1951 es iluminador acerca del carácter anti-biológico de su enfoque de la inteligencia:

Para lograr que nuestro computador imite a una máquina sólo es necesario programarlo para que calcule lo que la máquina en cuestión haría bajo ciertas circunstancias [...].

Ahora bien, si una máquina en particular puede describirse como un cerebro, tenemos que solamente *programar* nuestro computador digital para imitarlo *y también será un cerebro*. Si se acepta que los cerebros reales, descubiertos en animales, y en especial en el hombre, son una clase de máquina, se seguirá entonces que nuestro computador digital, debidamente programado, *se comportará como un cerebro*. Este argumento presupone una idea que puede ser razonablemente cuestionada [...] que esta máquina debiera ser de una naturaleza cuya conducta sea en principio predecible mediante cálculo [...].

Nuestro problema es, entonces, cómo programar una máquina para *imitar al cerebro*, o si lo pudiésemos expresar de una manera más breve y menos rigurosa, *para que piense* (apud Copeland, 2003, p. 11, énfasis mío).

Aquí se aclara aún más qué tiene Turing en mente en la primera versión del juego de la imitación. Para que un computador o un hombre imiten la conducta inteligente de una mujer, no es necesario que se repliquen ciertas propiedades físicas o biológicas, a saber, aquellas por las cuales se tiene mente y conducta lingüística femenina. Para ponerlo de manera cruda en aras de la claridad: el punto de Turing es que tal como no es necesario replicar las propiedades físicas de la papilla neural del cerebro para tener la mente y las respuestas inteligentes de este, en la primera etapa del juego tampoco es necesario tener el aparato endocrino y reproductor de una mujer para poseer su mente e inteligencia. Basta que el computador o el hombre imiten la inteligencia de la mujer para convencer jueces que se instancia esta, tal como la mujer podría eventualmente imitar la inteligencia de un hombre; similarmente, no se requeriría poseer las propiedades físicas o biológicas asociadas al sexo masculino⁴.

⁴ En relación con este punto, difiero del análisis de Genova (1994). El funcionalismo y la concepción anti-biológica de la inteligencia de Turing hacen pensar que no es correcto asumir que existe un prejuicio contra la mujer, como incapaz de imitar la inteligencia de un hombre. En efecto, dado que imitar la conducta lingüística de un hombre no requiere de ninguna propiedad física o biológica, ¿por qué no podría ser imitada por una mujer, o incluso por un computador digital? Esta pregunta sugiere que tal vez Genova subestima el funcionalismo de Turing. Esto, además, podría ser una pista de por qué él abandonó la versión de género en 1952, y describió una versión análoga a la versión estándar (ver una hipótesis sobre su cambio de postura en Copeland (2003, p. 6-7).

La pregunta que cabe hacer, entonces, es si tal como la inteligencia, la conducta lingüística orientada por la identidad del sexo puede aprenderse, de un modo similar a lo que Turing plantea con respecto a la creación de una mente adulta. Esto es, mediante la puesta en práctica de mecanismos simples que operen sobre la experiencia y la educación como inputs, que implementen aprendizaje basado en estímulo-respuesta, pero centrado en definiciones y proposiciones expresadas por un lenguaje simbólico, con el objeto de dejar de lado aspectos más bien emocionales (TURING, 1990, p. 63). Si esto fuera factible, la teoría de la inteligencia de Turing vincularía imitación, engaño y aprendizaje. Pero, ¿plantea Turing tal concepción?

Hasta aquí se ha analizado como el juego de la imitación supone una concepción funcionalista anti-biológica de la mente. Ahora es necesario examinar si aprender y ser educado pueden crear la inteligencia de una mente humana adulta. Una pista que sugiere que dicha conexión existe es que el hombre que imita a una mujer aprende a comportarse lingüísticamente como esta. Del mismo modo, un sujeto que imita cómo un policía dirige el tráfico aprende cómo dirigirlo.

Para Turing hay inteligencia cuando hay imitación de esta, es decir, cuando X imita a Y haciendo creer a observadores que $X = Y$. Nuevamente, la primera versión del juego de la imitación es clave: en la medida que lo juzgado es la conducta lingüística de una mujer, esta puede ser *enseñada y aprendida*. Una ventaja de la imitación es que no requiere instanciar una propiedad material y eso sugiere que la conducta en cuestión puede enseñarse y aprenderse.

En particular, para Turing es suficiente que X imite a Y engañando observadores para pasar el test. En el caso de la inteligencia, basta la imitación del comportamiento de la máquina-cerebro para que haya engaño sobre la existencia de estados mentales e inteligencia. La inteligencia no es una propiedad interna de una máquina, sino que es el *desempeño* adecuado de esta. No se requiere, por tanto, replicar el sustrato biológico, y esto hace que para Turing ser inteligente sea, adicionalmente, *programable* mediante mecanismos simples que

doten a máquinas de la capacidad de aprender y educarse⁵. ¿Por qué? Los computadores digitales pueden ser programados para imitar la máquina-cerebro de modo de *funcionar* de la misma manera, es decir, teniendo los mismos outputs para los inputs respectivos.

Un pasaje iluminador con respecto a cómo la inteligencia no es un producto biológico o físico, y por tanto puede ser programada, aprendida y enseñada es uno en que Turing se refiere al desarrollo futuro de la inteligencia de máquina. Justamente, esta debe tomar en consideración de qué forma aprende un niño, y es educado, para alcanzar la mente adulta. El pasaje en cuestión es el siguiente:

En el proceso de tratar de imitar a una mente humana estamos condicionados a pensar bastante acerca de los procesos de que fue sujeto para llegar a su estado actual. Podemos notar tres componentes:

1. El estado inicial de la mente en el nacimiento;
2. La educación que ha recibido;
3. Otra experiencia, no descriptible como educación, que ha tenido.

En vez de intentar producir un programa para simular la mente humana, *por qué no tratamos más bien de simular uno que sea como la de un niño*. Si se le sometiera a la educación apropiada, uno obtendría una mente adulta. Tal vez el cerebro del niño es algo así como un libro de notas que uno compra en una tienda de artículos de oficina. *Pequeños mecanismos y un montón de páginas en blanco*. (TURING, 1990, p. 61-62, énfasis mío).

Tales mecanismos simples son, precisamente, programables para que trabajen los inputs, y se logre crear una mente adulta inteligente mediante aprendizaje y educación. Nótese que este filósofo no se aboca a dar detalles de cómo debería aprender una máquina a tener la inteligencia de un niño, sino que propone que los mecanismos simples programables son *suficientes* para activar la cadena aprendizaje-educación y, así, la máquina sea capaz de hacer creer que se está en presencia de una mente adulta.

⁵ Aunque Turing no lo dice explícitamente, el caso de la Máquina Universal de Turing es revelador: esta imita a cualquier otra máquina y al hacerlo es, por ejemplo, inteligente como el cerebro lo es.

Es, por tanto, razonable pensar que Turing se inspiró en cómo puede imitarse la inteligencia de una mujer (y la de un hombre) para proceder a examinar si realmente era posible en principio la inteligencia de máquina con la programación adecuada. Justamente, por su carácter anti-biológico y funcionalista, el juego de la imitación es una propuesta vanguardista de la inteligencia inspirada por la imitación del sexo de los participantes en su juego; en efecto, ni su sexo ni su inteligencia parecen estar condicionados por lo material y, así, por lo biológico.

Conclusión

La estrategia que adopta Turing con respecto a la pregunta de si las máquinas programadas pueden tener mente e inteligencia se relaciona directamente con la imitación de la conducta lingüística de una mujer (y también de un hombre), la cual no tendría relación alguna con las propiedades físicas o biológicas que la causan. Por el contrario, propone que la imitación, el corazón de su test para reemplazar la pregunta de si las máquinas pueden tener estados mentales, es la clave para lograr crear inteligencia.

Ello ocurre porque ésta no está determinada por la biología; al contrario, en la medida que es separable de lo material, puede aprenderse y ser parte de un proyecto de educación en que las máquinas imiten la manera en que un niño se convierte en una mente adulta. O, lo que para Turing es lo mismo, la manera en que somos engañados por máquinas a creer que estamos en presencia de dicha mente.

En este trabajo he defendido que al examinarse las consideraciones de Turing sobre la inteligencia no se debe obviar la primera etapa del juego de la imitación, centrada en la determinación del sexo de los participantes. En efecto, ni el sexo ni la inteligencia parecen requerir de propiedades físicas o biológicas específicas; al contrario, son *imitables* y ello para Turing significa que también pueden aprenderse.

Tal como he argumentado aquí, esto es usualmente pasado por alto cuando se analiza el proyecto de Turing, basado en su máquina y

en el polémico juego de la imitación, dos pilares que fundamentan una teoría funcionalista anti-biológica de la mente y de la inteligencia.

Referencias

CLARK, A. *Mindware: an introduction to cognitive science*. New York: Oxford University Press, 2001.

COPELAND, J. *Artificial Intelligence: a philosophical introduction*. 7. ed. Oxford: Blackwell, 2001.

COPELAND, J. The Turing Test. In: MOOR, J. H. (Ed.). *The Turing Test: the elusive standard of artificial intelligence*. Dordrecht: Kluwer, 2003. p. 3-25.

DESCARTES, R. *Discurso del Método*. 10. ed. Madrid: Alianza Editorial, 1994.

GENOVA, J. Turing's sexual guessing game. *Social Epistemology*, v. 8, n. 4, p. 313-326, 1994.

LASSÈGUE, J. What kind of Turing Test did Turing have in mind? *Tekhnema*, n. 3, p. 37-58, 1996.

SAYGIN, A. et al. Turing Test: 50 years later In: MOOR, J. H. (Ed.). *The Turing Test: the elusive standard of artificial intelligence*. Dordrecht: Kluwer, 2003. p. 26-78.

SEARLE, J. Computing machinery and intelligence. *Behavioral and Brain Sciences*, n. 3, p. 417-457, 1980.

SEARLE, J. Is the Brain's Mind a Computer Program? *Scientific American*, n. 262, p. 26-31, 1990.

SHIEBER, S. The Ephemera. In: SHIEBER, S. *The Turing Test: verbal behavior as the hallmark of intelligence*. 3. ed. Cambridge: MIT Press, 2004. p. 97-104.

TURING, A. On computable numbers, with an application to the *Entscheidungsproblem*. *Proceedings of the London Mathematical Society*, v. 42, p. 231-65, 1936.

TURING, A. Computing intelligence and machinery. In: BODEN, M. A. *The philosophy of artificial intelligence*. 3. ed. Oxford: Oxford University Press, 1990. p. 40-66.

Recibido: 08/08/2014

Received: 08/08/2014

Aprobado: 20/09/2014

Approved: 09/20/2014