

Maquinaria computacional e Inteligencia

Alan Turing, 1950

Traductor: Cristóbal Fuentes Barassi, 2010,
Universidad de Chile.

1. El juego de la imitación.

Propongo considerar la siguiente pregunta: "¿Pueden pensar las máquinas?". Se debiera comenzar definiendo el significado de los términos 'máquina' y 'pensar'. Estas definiciones deberían ser elaboradas de manera tal que reflejen lo mejor posible el uso normal de estas palabras, pero una actitud así es peligrosa. Si el significado de las palabras 'máquina' y 'pensar' proviene del escrutinio de cómo son usadas comúnmente, se hace difícil escapar de la conclusión de que el significado y respuesta a la pregunta "¿pueden las máquinas pensar?" debiera ser buscado en una encuesta estadística, tal como la encuesta Gallup. Pero eso es absurdo. En vez de intentar una definición así, propondré reemplazar esa pregunta por otra, la cual se encuentra estrechamente relacionada y que se puede expresar en palabras relativamente poco ambiguas.

La nueva forma del problema puede ser descrita en términos de un juego, el cual llamaremos "el juego de la imitación". Se juega con 3 personas, un hombre (A), una mujer (B), y un interrogador (C) de cualquier sexo. El interrogador se encuentra en una habitación distinta a la de los otros dos participantes. El objetivo del juego para el interrogador es determinar cuál de los participantes es el hombre y cuál es la mujer. Él los identifica con las etiquetas X y Y, y al final del juego él dice si "X es A y Y es B", o "X es B e Y es A". Al interrogador se le permite hacer preguntas tanto a A como B del tipo:

C: X, ¿Sería tan amable de decirme el largo su cabello?

Ahora, suponga que X es de hecho A, por lo que A debe responder. El objetivo de A en el juego es tratar de que C haga una identificación falsa. Por lo que su respuesta podría ser:

"Mi pelo está cortado en capas, y los mechones más largos tienen unos 20 centímetros".

Para que los tonos de voz no ayuden al interrogador, las respuestas deben ser escritas, o mejor aún, tecladas. Las condiciones ideales deberían incluir un teletipo que comunique ambas habitaciones. De manera opcional, las preguntas y respuestas podrían ser repetidas por un intermediario. El objetivo de B en el juego es ayudar al interrogador. Probablemente, la mejor estrategia para ella sea dar respuestas verdaderas. Ella puede incluir en sus respuestas cosas tales como "Yo soy la mujer, ¡no lo escuches!", pero aquello no garantizaría nada ya que el hombre podría decir cosas similares.

Ahora hacemos la pregunta: "¿qué pasaría si una máquina asume el rol de A en este juego?" ¿Discriminaría equivocadamente el interrogador con la misma frecuencia con la que lo hace cuando el juego se juega con un hombre y una mujer? Estas preguntas reemplazan la pregunta original "¿pueden las máquinas pensar?".
(...)

3. Las máquinas involucradas en el juego.

(...)
La respuesta breve es que no estamos preguntando si todos los computadores digitales lo harían bien en el juego, ni tampoco si los computadores disponibles en la actualidad lo harían bien, sino que si hay computadores imaginables que lo harían bien. Pero ésta

es sólo una respuesta breve. Nos aproximaremos a esta pregunta desde otro ángulo después.
(...)

6. Perspectivas contrarias sobre la pregunta principal.

Ahora podemos considerar que el terreno ha sido delimitado, y nos encontramos listos para proceder con el debate sobre la pregunta "¿pueden pensar las máquinas?" y la variante de esta pregunta presentada al final de la última sección. No podemos abandonar del todo la versión original de problema, dado que las opiniones variarán con respecto a la idoneidad de la substitución y debemos al menos escuchar lo que se pueda decir en conexión con esto.

Se simplificarán ciertos asuntos para el lector si primero explico mis propias ideas con respecto al tema. Consideremos la versión más precisa de la pregunta. Creo que en un periodo de tiempo de 50 años será posible programar computadores, con una capacidad de almacenamiento de alrededor de 109, para que puedan jugar el juego de la imitación de tal manera que el interrogador promedio no pueda obtener más de un 70 por ciento de posibilidades de hacer la identificación acertada luego de cinco minutos de preguntas. Con respecto a la pregunta original, "¿pueden las máquinas pensar?", creo que no tiene mucho sentido como para merecer discusión. No obstante, creo que cuando llegemos a finales de siglo, el uso de las palabras y la opinión educada general habrán cambiado tanto, que uno podrá ser capaz de hablar de máquinas pensantes sin esperar ser contradicho. Creo además que ningún propósito útil se puede lograr al ocultar estas ideas. La visión popular que los científicos proceden inexorablemente desde los hechos bien establecidos hacia otros hechos bien establecidos sin nunca ser influenciados por alguna conjetura no probada es bastante equivocada. Dado que se hace claro cuáles son los hechos probados y cuáles son conjeturas, no puede haber ningún daño. Las conjeturas poseen una gran importancia debido a que éstas sugieren líneas útiles de investigación.

Ahora procedo a considerar opiniones opuestas a las mías.

6.1. La objeción teológica.

Pensar es una función del alma inmortal del hombre. Dios le ha otorgado un alma inmortal a cada hombre y mujer, pero no a otros animales o máquinas. Por lo tanto, ningún animal o máquina puede pensar.
(...)

Sin embargo, esto es mera especulación. No me impresionan mucho las discusiones teológicas, sea cual sea el tema. Tales discusiones han sido frecuentemente insatisfactorias en el pasado. En los tiempos de Galileo, se discutía que los textos "Y el sol se quedó quieto... y consideró no bajar por un día" (Josué Cap. 10, v. 13) y "Él puso los cimientos de la tierra, para que no se moviera en ningún momento" (Salmos Cap. 105, v. 5) eran una refutación adecuada para la teoría de Copérnico. Con nuestro actual conocimiento, tal argumento parece fútil. Cuando ese conocimiento no estaba disponible, causaba una impresión muy diferente.

6.2. La objeción de las "cabezas en la arena".

"La consecuencia de que las máquinas piensen sería demasiado espantosa. Esperemos y creamos que no lo pueden hacer".

Este argumento es rara vez expresado tan abiertamente como recién se plantea. Pero afecta a la mayoría de los que pensamos en ello. Nos gusta creer que el hombre es sutilmente superior al resto de la creación. Es mejor si se puede demostrar que es necesariamente superior, ya que entonces no hay peligro de que pierda su posición de mando.

La popularidad del argumento teológico está claramente conectada con esta sensación. Es probable que sea bastante fuerte entre los intelectuales, dado que ellos valoran el poder de pensar más que otros, y están más inclinados a basar su creencia en la superioridad del Hombre en base a este poder.

No creo que este argumento sea lo suficientemente sustancial como para requerir refutación. El consuelo sería más apropiado; quizás eso debiera buscarse en la trasmigración de las almas.

6.3. La objeción Matemática.

Hay varios resultados en la lógica matemática que pueden ser usado para mostrar que existen limitaciones para los poderes de una máquina de estados discretos. (...)
La respuesta breve a este argumento es que, aunque se establece que hay limitaciones para el poder de cualquier máquina dada, pero de nuevo, sin ningún tipo de prueba, que tales limitaciones no se aplican al intelecto humano. (...)
Con bastante frecuencia respondemos equivocadamente como para justificar algún tipo de satisfacción por tener evidencia de la falibilidad de las máquinas. Además, nuestra superioridad sólo se puede sentir en tales ocasiones en relación con la máquina sobre la que nos anotamos esa victoria pírrica. No habría la menor oportunidad de triunfar simultáneamente sobre todas las máquinas. En pocas palabras, puede que haya hombres más inteligentes que cualquier máquina dada, pero de nuevo, puede que haya otras máquinas aún más inteligentes, y así consecutivamente. (...)

6.4. El argumento desde la conciencia.

Este argumento se encuentra muy bien expresado en la disertación de la Medalla de Lister del profesor Jefferson en 1949, de donde cito: "Hasta que una máquina pueda escribir un soneto o componer un concierto debido a las emociones y pensamientos que tuvo, y que no sea debido al uso de símbolos al azar, podremos estar de acuerdo que máquina es igual a cerebro es decir, no sólo que lo escriba, sino saber que lo escribió. Ningún mecanismo podría sentir (y no sólo una mera señal artificial, cosa fácil de hacer) placer por sus éxitos, sentir pesar cuando se le funde una válvula, sentirse bien con un halago, sentirse miserable por sus errores, estar encantado por el sexo, estar enojado o deprimido cuando no consigue lo que quiere."
Este argumento parece ser una negación a la validez de nuestra prueba. De acuerdo a la forma más extrema de esta visión, la única manera con la cual uno podría estar seguro de que una máquina piensa es ser la máquina y sentir su propio pensamiento. Por tanto, se podría describir estas sensaciones al mundo, pero por supuesto nadie estaría justificado de poder percatarse. De manera similar, de acuerdo a esta visión, la única manera de saber que un hombre piensa es ser ese hombre en particular. Éste es el punto de vista de solipsista. Esa sería la visión más lógica de sostener, pero hace difícil la comunicación de las ideas. A está inclinado a creer "A piensa pero B no", mientras que B cree "B piensa pero A no". En vez de argumentar continuamente contra este punto, lo usual es tener la convención educada de que todos pensamos.

Estoy seguro de que el profesor Jefferson no quiere adoptar el punto de vista extremista del solipsismo. Probablemente él se encontraría dispuesto a aceptar el juego de la imitación como una prueba. El juego (con el jugador B omitido) es usado frecuentemente en la práctica con el nombre de viva voce para descubrir si alguien realmente entiende algo o ha "aprendido como perico". Escuchemos una parte de un viva voce:

Interrogador: En la primera línea de su soneto,

el cual dice "deberíais compararles con un día de verano", ¿no sería igual o mejor decir "un día de primavera"?

Testigo: no tendría la métrica correcta.

Interrogador: y si usamos "un día de invierno". Ese sí la tendría.

Testigo: sí, pero nadie quiere ser comparado con un día de invierno.

Interrogador: ¿se podría decir que el Sr. Pickwick le recuerda la navidad?

Testigo: de cierta manera.

Interrogador: pero la navidad es un día de invierno, y no creo que al Sr. Pickwick le importe la comparación.

Testigo: no creo que estés hablando en serio. Cuando uno dice día de invierno se refiere a un día típico de invierno más que a uno especial como la navidad.

Y se podría continuar. ¿Qué diría el profesor Jefferson si la máquina escritora de sonetos fuera capaz de responder de esta manera en el viva voce? No sé si él consideraría a la máquina como si estuviera mandando "meras señales artificiales" para responder, pero si las respuestas fueran satisfactorias y sostenidas como en el pasaje anterior, no creo que las describiría como un "cosa fácil de hacer". Creo que esta frase tiene por intención cubrir a dichos dispositivos en los cuales a una máquina se le incluye la grabación de alguien que lee un soneto, con un interruptor apropiado para prenderlo de cuando en cuando.

En pocas palabras, creo que aquellos que apoyan el argumento de la conciencia podrían ser persuadidos de abandonarlo en vez de forzar una posición solipsista. Probablemente, estarán dispuestos a aceptar nuestra prueba.

No quisiera dar la impresión de que creo no hay un misterio con respecto a la conciencia. Hay, por ejemplo, algo así como una paradoja en conexión con cualquier intento de localizarla. Pero no creo que estos misterios deban ser resueltos antes de que podamos responder la pregunta que concierne a este artículo.

6.5. Argumentos desde las discapacidades múltiples.

Estos argumentos toman la forma "Te aseguro que puedes hacer máquinas que hagan todas las cosas que dices, pero nunca podrás hacer una que sea capaz de X". Varias características de X se sugieren en relación a esto. Yo mismo ofrezco una selección:

Ser amable, ingenioso, hermoso, amigable, tener iniciativa, tener sentido del humor, diferenciar lo correcto de lo incorrecto, cometer errores, enamorarse, disfrutar las fresas con crema, hacer que alguien se enamore de él, aprender de la experiencia, usar las palabras apropiadamente, ser sujeto de sus propios pensamientos, tener tanta diversidad de conductas como un hombre, hacer algo realmente novedoso.

Usualmente no se ofrece ningún soporte a estos enunciados. Creo que la mayoría se basan en el principio de inducción científica. Un hombre ve miles de máquinas en su vida. Y de acuerdo a lo que ve, saca una cantidad de conclusiones generales. Son feas, se diseñan con un propósito muy limitado, y cuando se les requiere para una cosa ligeramente diferente, no sirven, la variedad de conductas es muy poca, y así sucesivamente. Naturalmente, un hombre concluye que estas son en general propiedades necesarias de las máquinas.

Muchas de estas limitaciones se encuentran asociadas a la poca capacidad de almacenamiento de la mayoría de ellas (asumo que la idea de capacidad de almacenamiento incluye de alguna manera a otras máquinas distintas a las máquinas de estados discretos. La definición exacta no importa dado que no se sostiene ninguna precisión matemática en esta discusión). Algunos años atrás, cuando no se sabía mucho sobre los computadores digitales, había mucha incredulidad con respecto a ellos, si es que alguien mencionaba sus propiedades sin describir su construcción. Presumiblemente, eso se debía a una aplicación similar del principio de inducción científica. Estas aplicaciones del principio son, por supuesto, inconscientes en su mayoría. Cuando un niño con quemaduras le teme al fuego y demuestra su miedo evitándolo, yo podría decir que él está aplicando inducción científica. (Por supuesto podría describir su conducta de muchas otras maneras). Los trabajos y costumbres de la humanidad no parecen ser material apropiado sobre el cual aplicar la inducción científica. Una gran parte del espaciotiempo debe ser investigada si se busca obtener resultados confiables. De otra manera, podríamos (como la mayoría de los niños ingleses lo hacen) decidir que todo el mundo habla inglés, y que por lo tanto es tonto aprender francés.

Sin embargo, se pueden hacer comentarios especiales acerca de las muchas discapacidades que han sido mencionadas.

La incapacidad de disfrutar las fresas con crema podría sorprender al lector dada su trivialidad. Posiblemente, se podría hacer una máquina que disfrute este delicioso plato, pero cualquier intento de crearla sería estúpido. Lo que es importante con respecto a esta discapacidad es que contribuye a otras discapacidades, por ejemplo, a la dificultad del mismo tipo de amabilidad entre hombre y máquinas como entre hombre blanco con hombre blanco, o entre hombre negro y hombre negro.

La declaración de que "las máquinas no pueden cometer errores" parece curiosa. Uno es tentado a responder "¿son peores por eso?". Pero adoptemos una actitud un poco más comprensiva, y veamos que quiere decir realmente. Creo que esta crítica puede ser explicada en términos del juego de la imitación. Se sostiene que interrogador podría distinguir simplemente la máquina del hombre al plantearle una cantidad de problemas aritméticos. La máquina sería desenmascarada dado su altísima eficacia. La respuesta a esto es simple. La máquina (programada para jugar el juego) no intentaría dar la respuesta correcta a los problemas aritméticos. Deliberadamente, cometería errores de manera calculada para confundir al interrogador. Una falla mecánica podría probablemente delatar a la máquina con una decisión poco apropiada con respecto al tipo de error que cometa en el cálculo. Incluso esta interpretación de la crítica no es lo suficientemente comprensiva. Pero no podemos dedicarle el espacio para ahondar en ello. Me parece que esta crítica se sostiene en una confusión entre dos tipos de errores. Podríamos etiquetarlos como "errores de funcionamiento" y "errores de conclusión".
Los errores de funcionamiento se deben a alguna falla mecánica o eléctrica que produce que la máquina se comporte de una manera distinta con respecto a la cual fue diseñada. En la discusión filosófica a uno le gusta ignorar la posibilidad de tales errores; uno se encuentra por lo tanto discutiendo sobre "máquinas abstractas". Estas máquinas abstractas son ficciones matemáticas más que objetos físicos. Por definición, incapaces de presentar errores de funcionamiento. En este sentido, podemos realmente decir que "las máquinas nunca cometen errores".

Errores de conclusión sólo se pueden producir cuando se adjunta algún significado a las señales de salida de la máquina. La máquina podría, por ejemplo, escribir ecuaciones matemáticas, u oraciones en inglés. Cuando escribe una proposición falsa, decimos que la máquina ha cometido un error de conclusión. Claramente, no hay razón para decir que la máquina no pueda cometer este tipo de error. Podría solamente escribir repetidas veces " $0 = 1$ ". Al considerar un ejemplo menos rebuscado, se podría tener algún método para dar conclusiones a través de la inducción científica. Debemos esperar que un método así nos lleve a resultados erróneos ocasionalmente.

La aseveración de que una máquina no puede ser objeto de su propio pensamiento sólo puede ser respondida si se puede mostrar que la máquina posee algún pensamiento con algún objeto. No obstante, "el objeto de las operaciones de una máquina" pareciera significar algo, al menos para las personas que tratan con ella. Si, por ejemplo, la máquina estuviera tratando de encontrar una solución para la ecuación " x al cuadrado, menos $40x$, menos 11 ; es igual a 0 ", uno se sentiría tentado a describir esta ecuación como parte del objeto del pensamiento de la máquina en ese momento. En este sentido, la máquina puede sin lugar a dudas ser el objeto de su propio pensamiento. Podría ser usada para ayudar a crear sus propios programas, o para predecir el efecto de las alteraciones en su propia estructura. A través de la observación de resultados de su propia conducta, podría modificar sus programas de manera tal de alcanzar algún propósito de manera más efectiva. Éstas son más bien posibilidades para el futuro cercano que sueños utópicos.

La crítica que refiere a que una máquina no puede tener una gran variedad de conductas es sólo una manera de decir que no puede tener una gran capacidad de almacenamiento. Hasta hace poco, la capacidad de almacenamiento de incluso mil dígitos era muy rara.

Las críticas que estamos considerando acá son con frecuencia formas disfrazadas del argumento desde la conciencia. Generalmente, si uno sostiene que una máquina puede hacer una de estas cosas, y describe el tipo de método que la máquina podría usar, uno no produciría una gran impresión. Se cree que el método (cualquiera que sea, dado que debe ser mecánico) es realmente deshonesto. Compare con el paréntesis de la afirmación de Jefferson mencionada más arriba.

6.6. La objeción de Lady Lovelace.

La información más detallada de la Máquina Analítica de Babbage proviene de una de las memorias de Lady Lovelace (1842). En ésta, ella sostiene que "la Máquina Analítica no tiene pretensiones de originar nada. Puede hacer cualquier cosa que sepamos ordenarle que haga" (su cursiva). Esta declaración es citada por Hartree, que añade: "esto no implica que no sea posible construir un equipamiento electrónico que podrá 'pensar por sí mismo', o sobre el cual, en términos biológicos, uno pudiera construir un reflejo condicionado, que podría servir como la base del 'aprendizaje'. Si es que esto es posible en principio o no es una pregunta estimulante y apasionante, y es sugerida debido a algunos de los desarrollos recientes. Pero no parece que las máquinas construidas o proyectadas en ese momento pudieran tener esta propiedad".

Estoy completamente de acuerdo con Hartree. Uno podrá notar que él no asevera que las máquinas en cuestión no tuvieran la propiedad, sino que la evidencia disponible para Lady Lovelace no la alentaba a creer que la tuvieran. Es muy posible que las máquinas en cuestión tuvieran esta propiedad en cierto sentido. Suponga que alguna máquina de estados discretos tiene la propiedad. La Máquina Analítica era un computador

universal digital, así que, si la capacidad de almacenamiento y velocidad fueran adecuadas, se podría a través de una programación apropiada hacer que se imitara a la máquina en cuestión. Probablemente este argumento no se le ocurrió ni a la Condesa ni a Babbage. En cualquier caso, no tenían ninguna obligación de afirmar todo lo que se puede afirmar

Toda esta pregunta será considerada nuevamente bajo la perspectiva de las máquinas que aprenden.

Una variante de la objeción de Lady Lovelace afirma que una máquina "nunca hace nada realmente nuevo". Esto podría ser aludido desde la perspectiva de "no hay nada nuevo bajo el sol". Quién puede tener certeza que el "trabajo original" que alguien haya hecho no fue solamente el crecimiento de una semilla plantada en él a través de la enseñanza, o el efecto de seguir principios generales bien sabidos. Una variante mejor a la objeción dice que una máquina nunca puede "sorprendernos". Esta declaración es un desafío más directo y puede ser enfrentada más directamente. Las máquinas me sorprenden con gran frecuencia. Esto se debe en gran medida a que no hago el cálculo suficiente para decidir qué puedo esperar de ellas, o más bien porque, aunque hago una estimación, lo hago apurado, con descuido, tomando riesgos. Quizás me digo a mí mismo, "Y creería que el voltaje acá debiera ser el mismo que allá: bueno, supongamos que así es". Naturalmente, con frecuencia me equivoco, y el resultado es una sorpresa, pues para cuando el experimento se lleva a cabo, estos supuestos ya se han olvidado. Reconocer lo anterior me deja expuesto a críticas sobre mis modos descuidados de proceder, pero no arrojan ninguna duda sobre mi credibilidad cuando testifico las sorpresas que experimento.

No espero que esta respuesta silencie a mi crítico. Él probablemente dirá que esas sorpresas se deben a algún acto mental creativo de mi parte, y que no otorga ningún crédito a la máquina. Esto nos lleva de vuelta a la discusión sobre la conciencia, y lejos de la idea de la sorpresa. Esta es una línea argumental que debemos considerar cerrada, pero quizás valga la pena destacar que la apreciación de algo como sorprendente requerirá un "acto mental creativo", sin importar si el evento sorprendente es originado por un hombre, un libro, una máquina o cualquier otra cosa.

La idea de que las máquinas no producen sorpresas se debe, creo yo, a una falacia a la cual se encuentran especialmente sujetos los filósofos y los matemáticos. Es el supuesto de que tan pronto como se presente un hecho a una mente, todas las consecuencias de ese hecho florecen en ella simultáneamente con el hecho. Es un supuesto muy útil en muchas circunstancias, pero uno olvida demasiado fácilmente que es falso. Una consecuencia natural de hacer esto es que uno asume que no hay mérito en la simple búsqueda de consecuencias a partir de datos y principios generales.

6.8. El argumento desde la informalidad de la conducta.

No es posible producir una lista de reglas que pretendan describir lo que un hombre debiera hacer en cada circunstancia concebible. Uno podría, por ejemplo, tener una regla en la que se debe parar cuando alguien ve un semáforo en rojo, y de seguir si alguien ve la luz verde; pero, ¿qué pasa si por alguna falla, ambas luces aparecen juntas? Uno quizás decidiría detenerse ya que es más seguro. Pero una nueva dificultad bien podría producirse más tarde por esta decisión. El intento de proveer reglas de conducta para cubrir cada eventualidad, incluso aquellas que se producen por los semáforos, parece imposible. Estoy de acuerdo con todo esto.

Dado lo anterior, se dice que no podemos ser máquinas. Trataré de reproducir el argumento, pero me temo que no podré hacerle justicia. Pareciera que es algo así: "**Si cada hombre tuviera un grupo determinado de reglas de conducta por las cuales él regulara su vida, no sería más que una máquina. Pero no existen tales reglas, así que los hombres no pueden ser máquinas**".

La falacia argumental es flagrante. No creo que el argumento se haya propuesto así alguna vez, no obstante creo que ese es el argumento que se da. Sin embargo, podría hacer una cierta confusión entre las "reglas de conducta" y las "leyes de comportamiento". Se entiende por "reglas de conducta" los preceptos tales como "detenerse si uno ve una luz roja", sobre las cuales alguien puede actuar, y de las cuales se está consciente. Por "leyes de comportamiento" me refiero a las leyes de la naturaleza aplicables al cuerpo de un hombre, tales como "si lo pinchas, va a chillar". Si substituímos "reglas de conducta que regulan su vida" por "leyes de comportamiento que regulan su vida" en el argumento citado con anterioridad, la falacia deja de serlo. Pues no solamente creemos que ser regulados por leyes de comportamiento implica ser cierto tipo de máquina (aunque no necesariamente una máquina de estados discretos), sino que recíprocamente, ser tal máquina implica ser regulado por tales leyes. Sin embargo, no podemos convencernos tan fácilmente de la ausencia de leyes completas del comportamiento, como de la ausencia de reglas completas de la conducta. La única manera que sabemos para descubrir tales leyes es la observación científica, y ciertamente conocemos de circunstancia alguna en la que podamos decir: "hemos buscado lo suficiente. No existen tales leyes".

Podemos demostrar más convincentemente que cualquier declaración de este tipo sería injustificada: Suponga que estuviéramos seguros de encontrar tales leyes cuando existieran. Entonces, dada una máquina de estados discretos, sería con certeza posible descubrir tales leyes a través de suficientes observaciones de ella, y así predecir su conducta futura, y todo esto dentro de un tiempo razonable, digamos, mil años. Pero este no parece ser el caso. He instalado en el computador de Manchester un programa pequeño usando solo 1000 unidades de almacenamiento, tal que cuando a la máquina se le provee con un número de 16 dígitos, responde con otro número de 16 dígitos en dos segundos. Desafiaría a cualquiera a aprender lo suficiente de las respuestas del programa para ser capaz de predecir cualquier respuesta a valores no probados con anterioridad.

6.9. El argumento desde la percepción extra-sensorial.

Asumo que el lector se encuentra familiarizado con la idea de percepción extra-sensorial, y del significado de sus cuatro términos, que son telepatía, clarividencia, precognición y psicoquinesia. Estos fenómenos perturbadores parecieran ir en contra de todas nuestras ideas científicas usuales. ¡Cómo nos gustaría desacreditarlos! Desafortunadamente, la evidencia estadística, al menos para la telepatía, es abrumadora. (...)

Si se admite la telepatía, será necesario afinar nuestra prueba. La situación podría ser considerada como análoga a lo que ocurriría si el interrogador estuviera hablando consigo mismo en voz alta y uno de los competidores estuviera escuchando con su oído en la pared. Poner a los competidores en una "habitación a prueba de telepatía" bastaría para cumplir todos los requerimientos.

7. Máquinas que aprenden.

(...)

La analogía de las "capas de cebolla" ...es útil. Al considerar las funciones de la mente o el

cerebro, encontramos ciertas operaciones que se pueden explicar en términos puramente mecánicos. Decimos que esto no corresponde a la mente real: es un cierto tipo de capa que debemos sacar si queremos encontrar la mente real. Pero luego en lo que queda, encontramos otra capa que se puede remover, y así sucesivamente. Si se procede de esta manera, ¿podemos alcanzar en algún momento la mente real, o eventualmente llegamos a una capa que no contiene nada? En este último caso, toda la mente sería mecánica. (...)

El único fundamento realmente satisfactorio que se puede dar para la visión expresada al principio de la sección 6 será que, debemos esperar el final de siglo y recién entonces, hacer el experimento descrito. Pero, ¿qué podemos decir mientras tanto? ¿Qué pasos deberían darse para que el experimento sea exitoso? Como he explicado, el problema es esencialmente un problema de programación. También tendrán que haber avances en la ingeniería (...)

En vez de tratar de producir un programa similar a la mente adulta, ¿por qué no tratar en vez de producir una que simule la mente de un niño? Si ésta fuera luego sujeta al curso apropiado de educación, uno obtendría el cerebro adulto. Presumiblemente, el cerebro de un niño es algo así como un cuaderno que uno compra en una tienda. Un mecanismo más bien simple, con muchas hojas en blanco. (Mecanismo y escritura son casi sinónimos para nuestro punto de vista.) Nuestra esperanza es que haya tan poco mecanismo en el cerebro del niño, que algo así pueda ser programado fácilmente. Podemos asumir en una primera aproximación que la cantidad de trabajo en la educación sería muy similar a la de un niño humano.

Hemos por tanto dividido nuestro problema en dos partes: el programa - niño y el proceso educativo. Estos permanecen estrechamente relacionados. No podemos esperar encontrar un buen niño-máquina al primer intento. Uno debe experimentar enseñando a una máquina así y ver qué tan bien aprende. Luego, se puede intentar con otra y ver si es mejor o peor. (...)

La idea de una máquina que aprende podría parecer paradójica para ciertos lectores. ¿Cómo pueden las reglas de funcionamiento de una máquina cambiar? Estas debieran describir completamente cómo la máquina reaccionará sin importar cuál sea su historia, o los cambios por los que pudo haber pasado. Las reglas son entonces bastante invariantes en el tiempo. Esto es muy cierto. La explicación de la paradoja es que las reglas que son cambiadas en el proceso de aprendizaje son de un tipo menos pretencioso, afirmando solamente tener una validez efímera.
(...)

Podríamos esperar que las máquinas eventualmente compitan con los hombres en todos los campos puramente intelectuales. Pero, ¿cuáles son los mejores para comenzar? Incluso eso es una decisión difícil. Mucha gente cree que una actividad muy abstracta, como jugar ajedrez, sería lo mejor. También se puede sostener que lo mejor es proveer a la máquina con los mejores órganos sensoriales que el dinero pueda comprar y enseñarle a comprender y hablar inglés. Este proceso podría seguir la enseñanza normal de un niño. Las cosas podrían ser señaladas y nombradas, y así sucesivamente. Nuevamente, no sé cuál sea la respuesta correcta, pero creo que ambas aproximaciones debieran intentarse.