Alan Turing y el origen de la inteligencia artificial: la superación de la intuición[1]

Juan Antonio Valor Yébenes
Universidad Complutense de Madrid

Resumen: El presente artículo pretende mostrar que el objetivo del trabajo de Alan Turing fue, desde 1935 hasta su fallecimiento en 1954, encontrar un procedimiento sistemático para prescindir de la intuición en la matemática y, por extensión, en el conocimiento humano. Aunque los teoremas del Gödel y las conclusiones alcanzadas por él mismo en sus trabajos de los años 1936 y 1938 imponían limitaciones que no permitían eliminar la intuición, finalmente pudo mostrar en 1950 que era posible un procedimiento automático computacional que sustituyese al pensamiento.

Nos preguntamos cómo llegó Turing a esta nueva posición. La respuesta es que el trabajo desarrollado en Bletchley durante la guerra le condujo, en primer lugar, a una nueva lectura del teorema de Gödel, de tal manera que pasó de considerarlo una limitación para los sistemas axiomáticos formales a considerarlo una estrategia para la construcción de máquinas computacionales cada vez más potentes. En segundo lugar, la construcción de estas máquinas para descifrar los mensajes de las máquinas de codificación alemanas sin conocer su funcionamiento le llevó a considerar sistemas automáticos computacionales para simular el pensamiento humano.

Palabras clave: Turing, inteligencia artificial, máquina computacional, intuición, formalismo.

Alan Turing and the origin of artificial intelligence: overcoming intuition

[1] Este trabajo se enmarca en el proyecto de investigación PID2021-125822NB-I00 (IRENETIKA "Conflictos armados y crisis humanitarias: las Humanidades y las Ciencias sociales ante los desafíos de la seguridad multidimensional").

Summary: This article aims to show that the objective of Alan Turing´s work was, from 1935 until his death in 1954, to find a systematic procedure that would allow us to dispense with intuition in mathematics and, by extension, in human knowledge. Although Gödel's theorems and the conclusions reached by himself in his works of the years 1936 and 1938 imposed limitations that did not allow intuition to be eliminated, he ultimately showed in 1950 that an automatic computational procedure that replaced human thought was possible.

We wonder how Turing arrived at this new position. The answer is that the work carried out in Bletchley during the war led him, firstly, to a new interpretation of Gödel's theorem, in such a way that he transitioned from understanding it as a limitation of formal axiomatic systems to understanding it as a strategy for the construction of increasingly powerful computing machines. Secondly, the construction of these machines to decipher messages from german coding machines without knowing their operation led him to consider automatic computational systems to simulate human thought.
Keywords: Turing, Artificial Intelligence, Computing Machinery, Intuition, Formalism.

Recibido: 19 de marzo de 2024
Aceptado: 01 de mayo de 2024
DOI: 10.24310/nyl.18.2024.19496.

1. Un intento fallido de eliminación de la intuición

En 1938 presenta Turing su tesis doctoral, titulada *Systems of Logic based on Ordinals.* En el punto 11, el penúltimo, titulado «The purpose of ordinal logics», Turing intenta aclarar el sentido de su trabajo en torno a los conceptos de *intuición* e *ingenio*. Afirma que, a pesar de que su objetivo inicial coincidía con el del programa formalista anterior a Gödel, que buscaba la reducción del uso de la *intuición* para evitar la arbitrariedad de los matemáticos al elaborar teorías y fundamentar sus proposiciones, finalmente las conclusiones le han llevado en otra dirección, precisamente a reconocer la imposibilidad de una lógica formal que elimine por completo la necesidad de utilizar la intuición.

Debido a la influencia de Hilbert desde la publicación en 1899 de los *Grundlagen der Geometrie*, eran muchos los matemáticos que pensaban que el establecimiento de reglas formales de inferencia y su estricta aplicación en una demostración a través de pasos *mecánicos*, a lo que Turing denomina *ingenio*, permitiría incluso reemplazar los juicios intuitivos por un número finito de reglas, de tal manera que la intuición quedase eliminada de la matemática. Turing afirma que sus propias investigaciones tenían inicialmente esa orientación. Sin embargo, reconoce que el objetivo alcanzado en su trabajo no ha sido la eliminación de la intuición, sino del *ingenio*; es decir, concluye que no es posible reducir todos los pasos de una demostración a pasos *mecánicos*, sino que algunos de ellos necesitan del ejercicio de la intuición. Así lo explica en el referido punto 11, que cito por extenso:

> El razonamiento matemático puede considerarse, de manera bastante esquemática, como el ejercicio de combinación de dos facultades, que podemos llamar intuición e ingenio. La actividad de la intuición consiste en realizar juicios espontáneos que no son el resultado de un hilo de razonamientos conscientes. Esos juicios son a menudo, pero no siempre, correctos (dejando de lado la cuestión de qué se entiende por «correcto») ...
>
> El ejercicio del ingenio en matemáticas consiste en ayudar a la intuición mediante disposiciones adecuadas de proposiciones y quizás figuras o dibujos geométricos. Se pretende que cuando estos se encuentren realmente bien organizados no se pueda dudar seriamente de la validez de los pasos intuitivos que son requeridos.
>
> Los papeles desempeñados por estas dos facultades difieren, por supuesto, de una ocasión a otra y de un matemático a otro. Esta arbitrariedad puede eliminarse mediante la introducción de una lógica formal … En la época anterior a Gödel, algunos pensaban que probablemente sería posible llevar este programa

> hasta el punto de que todos los juicios intuitivos de las matemáticas pudieran ser reemplazados por un número finito de estas reglas. Entonces la necesidad de la intuición quedaría completamente eliminada.
>
> A pesar de que nuestro objetivo ha ido en gran medida en la misma dirección, sin embargo, en nuestras discusiones hemos ido al extremo opuesto y eliminado no la intuición sino el ingenio …
>
> Debido a la imposibilidad de encontrar una lógica formal que elimine por completo la necesidad de utilizar la intuición, naturalmente recurrimos a sistemas de lógica «no constructivos» en los que no todos los pasos de una demostración son mecánicos, siendo algunos intuitivos … (Turing, 1938: 57-58)[1].

Destaco la importancia de esta cita porque en ella Turing asume con naturalidad que no ha conseguido su objetivo inicial, enmarcado en el programa formalista. Es más, precisamente el valor de su trabajo radica en los argumentos ofrecidos en contra de la eliminación de la intuición. Larson se refiere a él diciendo que fue «un inteligente pero fallido intento de esquivar uno de los resultados obtenidos por Kurt Gödel» (Larson, 2022: 17). El resultado se refiere al *Entscheidungsproblem* y, efectivamente, se trata de un intento fallido reconocido como tal por el propio Turing. Sin embargo, y esto es lo que quiero mostrar en el presente artículo, Turing sigue pensando en la manera de hacer compatibles sus propias conclusiones y las de Gödel con el ideal de reducción -o incluso eliminación, si ello fuera posible- del uso de la intuición, tanto en el conocimiento matemático como en el conocimiento general.

2. La intuición en Kant

Comencemos por aclarar qué se entiende por intuición en el ambiente matemático en el que se sitúa Turing. En 1899 Hilbert había publicado *Die*

[2] La traducción es mía.

Grundlagen der Geometrie, que rápidamente fue reconocido por la comunidad de matemáticos como un nuevo abordaje del método axiomático formal en geometría y una de las contribuciones más relevantes a la matemática moderna. La filosofía de la matemática sobre la que se levanta la obra no aparece explícitamente sistematizada, pero podemos acercarnos a ella a través de los cursos sobre fundamentos que Hilbert impartió tanto en Königsberg como en Göttingen desde el año 1891 y de la correspondencia que mantuvo con Frege tras la publicación de la obra. La disputa entre Frege y Hilbert, así como la emergencia del nuevo programa formalista, acontecen sobre el trasfondo de la concepción kantiana de la matemática. Lo que nos interesa extraer es, en primer lugar, la concepción general del método axiomático que llega hasta Turing y su particular aplicación a la geometría y, en segundo lugar, el papel que se le asigna a la intuición en la nueva axiomática formal.

En la «Doctrina transcendental del método» expuesta en la *Crítica de la razón pura*, Kant caracteriza el conocimiento matemático como «conocimiento obtenido por construcción de los conceptos» (Kant, 1997: A731/B759), lo cual significa «exhibir a priori la intuición que le corresponde» (ibid.). Se trata de *construir* independientemente de la experiencia una representación del concepto en la intuición. Kant se refiere a una intuición no empírica que es conocimiento inmediato de un objeto singular, la cual, en cuanto construcción de un concepto, representa una validez universal en relación con todas las posibles intuiciones pertenecientes al mismo concepto. Es el caso, por ejemplo, de un triángulo representado en la intuición pura. El geómetra es capaz, guiado siempre por la intuición, de concluir que la suma de los ángulos interiores es igual a dos rectos.

La cuestión que pone en movimiento la *Crítica de la razón pura* es la de las condiciones que hacen posible que el conocimiento matemático, obtenido por conceptos en los que se considera lo general a partir de construcciones individuales, tenga validez en la experiencia. Ya en la *Dissertatio* de 1770 afirmaba Kant que «*el espacio no es algo objetivo* ni real, ni sustancia, ni accidente, ni relación; sino algo *subjetivo* e ideal, que brota de la mente, según ley establece, como un esquema coordinador de todo lo sentido externa-mente» (Kant, 1996: 14, D, 22). Es decir, el espacio no es un receptáculo absoluto e inmenso de todas las cosas posibles, como piensa Newton, ni la relación de las cosas existentes, como piensa Leibniz, sino un esquema puramente ideal que ordena y estructura todas nuestras sensaciones externas, de tal manera que las cosas se manifiestan a los sentidos «mediante esa virtud del espíritu, que coordina todas las sensaciones según la ley establece y propia de su naturaleza» (Kant, 1996: 14, D, 23). Por consiguiente, nada se da a los sentidos si no es como la geometría lo prescribe de acuerdo con sus axiomas primitivos, quedando así la naturaleza sometida a la geometría, de tal manera que «las leyes de la sensibilidad serán leyes de la naturaleza *en cuanto que ésta puede caer dentro de la esfera de los sentidos*» (ibid.). Esta doctrina reaparece en la *Crítica de la razón pura* depurada transcendentalmente. Aquí afirma Kant que el espacio es una forma *a priori* de la sensibilidad que actúa como condición de posibilidad de toda experiencia posible. Por ello «afirmamos, pues, *la realidad empírica*, del espacio (con respecto a toda experiencia externa posible), pero sostenemos, a la vez, la *idealidad transcendental* del mismo, es decir, afirmamos que no existe si prescindimos de la condición de posibilidad de toda experiencia y lo consideramos como algo subyacente a las cosas en sí mismas» (Kant, 1997: A28/B44).

De esta manera se explica la validez de la geometría en la experiencia. En cuanto a la construcción de la geometría misma, Kant afirma que descansa en la intuición formal pura del espacio entendido como idealidad transcendental. Es decir: ya sabemos que el conocimiento geométrico es obtenido por construcción de conceptos, lo cual equivale a su propia definición, porque en geometría (y, a juicio de Kant, en todo el conocimiento matemático) se define proporcionando las reglas de construcción. Pero resulta que al construir lo que se está haciendo es exhibir las formas puras de la intuición y, en el caso de la geometría, la forma pura del espacio. Kant lo expresa de la siguiente manera: «Una nueva luz se abrió al primero (llámese Tales o como se quiera) que demostró el triángulo equilátero. En efecto, advirtió que no debía indagar lo que veía en la figura o en el mero concepto de ella y, por así decirlo, leer, a partir de ahí, sus propiedades, sino extraer estas *a priori* por medio de lo que él mismo pensaba y exponía (por construcción) en conceptos. Advirtió también que, para saber *a priori* algo con certeza, no debía añadir a la cosa sino lo que necesariamente se seguía de lo que él mismo, con arreglo a su concepto, había puesto en ella» (Kant, 1997: B XI, XII).

Por consiguiente, los conceptos de la geometría (como los de punto, recta, plano, etc.) tienen como base una intuición *a priori*, no una empírica, que depende de la construcción del concepto -lo cual es a la vez su propia definición- con arreglo a la forma pura del espacio, que queda precisamente exhibida en la construcción. Pero no solo los conceptos, sino también los principios geométricos derivan de la intuición *a priori*. El conocimiento geométrico y, para Kant, todo el conocimiento matemático, ni es empírico ni resulta de un análisis de los conceptos, como el conocimiento filosófico, sino que requiere de la intuición, que hace posible ir más allá del concepto y establecer los juicios sintéticos *a priori* de la matemática.

En definitiva, el espacio caracterizado como forma pura de la intuición sensible es lo que le permite a Kant dar cuenta, por un lado, de la verdad de los juicios sintéticos *a priori* de la geometría, y por otro, de la estructura euclídea del espacio físico.

3. La intuición en Frege

En 1879 publica Frege su *Begriffsschrift* (*Conceptografía*). Siguiendo las tesis leibnizianas sobre un lenguaje universal y un cálculo de la razón, lo que pretende es una sistematización axiomática del cálculo lógico. De esta manera, obtiene los principios generales del sistema a partir de un determinado número de axiomas, nueve, y de cuatro reglas de deducción. Por primera vez se introduce una teoría de la cuantificación aplicando cuantificadores y variables ligadas, y se estable la distinción entre nombre y predicados, así como entre predicados de primer y segundo orden. Es central en la obra el análisis de las proposiciones en términos de funciones y argumentos, abandonando de esta manera el análisis kantiano en términos de sujeto y predicado. El objetivo último de Frege es la creación de un lenguaje lógico desvinculado de la gramática de los lenguajes naturales y de componentes psicológicos, que permita una estricta y definitiva fundamentación de las matemáticas alejada de las deficientes fundamentaciones realizadas hasta el momento.

La *Conceptografía* supone el primer paso de un camino que pasa por la definición lógica del concepto de número en los *Grundlagen der Arithmetik* de 1884 y por la deducción formal de las proposiciones de la aritmética en los *Grundgesetze der Arithmetik* de 1893. De esta manera intenta sacar adelante su proyecto logicista, que, en contra de Kant, pasa por considerar todas las proposiciones de la aritmética como proposiciones analíticas. Lo que significa la analiticidad no es lo mismo en Frege que en Kant. Para Frege una proposición es analítica si puede ser demostrada exclusivamente a partir de definiciones y leyes lógicas, y es sintética si la demostración exige recurrir a verdades de un campo particular irreductibles al ámbito de la lógica. Por otro lado, es *a posteriori* si su demostración requiere de verdades fácticas, es decir, de objetos particulares, y es *a priori* si puede demostrarse a partir de leyes universales indemostrables. La definición de los números en términos puramente lógicos y su reconocimiento como objetos lógicos llevada a cabo en los *Grundlagen*, así como la demostración de las proposiciones aritméticas a partir, exclusivamente, de leyes lógicas universales en los *Grundgesetze*, llevan a Frege a la conclusión de que los juicios de la aritmética no son, como Kant había indicado, sintéticos *a priori*, sino analíticos, en el sentido que él daba al término.

Kant afirmaba la intuición sensible de puntos o dedos para finalmente llegar a la conclusión de que «$2 + 3 = 5$». Como en el caso de la geometría, entendía que también el conocimiento aritmético se obtiene por construcción de conceptos tomando como base una intuición *a priori*, independiente de la sensible, y que en ello radicaba la condición que hacía posible que los juicios aritméticos fueran sintéticos *a priori*. En el punto 5 del capítulo

I de los *Grundlagen* dice Frege que con el recurso a los puntos y los dedos Kant corre el peligro de hacer que los enunciados de la aritmética aparezcan como empíricos, dado que la intuición de 37.863 dedos no puede ser una intuición pura. Parece que piensa exclusivamente en números pequeños, en cuyo caso solo las fórmulas para ellos resultarían inmediatamente evidentes para la intuición, pero no las fórmulas para números grandes, que tendrían que ser demostrables. Lo cual no puede ser admitido, sencillamente porque no es posible trazar la frontera entre números pequeños y grandes (Frege, 1884/1996: 46). Además, Kant apela a una intuición pura que, para el caso de la aritmética, es difícil caracterizar como espacial o temporal (Frege, 1884/1996: 55). En la conclusión de los *Grundlagen* vuelve Frege sobre este punto para insistir en que debe «contradecir la universalidad de la afirmación de Kant: la de que sin la sensibilidad no nos sería dado ningún objeto» (Frege, 1884/1996: 130). El cero, el uno, son objetos que no nos pueden venir dados por los sentidos. Por todo ello, Frege insiste al comienzo de los *Grundlagen* en que no es posible asumir la caracterización kantiana de los juicios de la aritmética como sintéticos *a priori*, ni el papel que se atribuye en este caso a la intuición.

Por otro lado, si atendemos a la estricta descripción de lo conocido, hay que asumir que «un punto geométrico, considerado en sí mismo, no se puede distinguir de otro cualquiera; lo mismo vale para rectas y superficies» (Frege, 1884/1996: 57). De tal manera que «si en la geometría se pueden obtener leyes generales a partir de la intuición, ello se explica por el hecho de que los puntos, rectas o superficies intuidos no son, en realidad, particulares, y pueden servir, por tanto, como representantes de toda su especie» (ibid.). Pero esto no ocurre en el caso de la aritmética, porque cuando tratamos con números hemos de reconocer que «cada uno tiene su peculiaridad. No puede decirse de buenas a primeras hasta qué punto un número determinado puede representar a todos los demás sin explicar todas sus particularidades» (ibid.).

Por consiguiente, Frege considera que Kant se precipita al sobrevalorar el parentesco entre la geometría y la aritmética. Los juicios de esta no pueden ser considerados sintéticos *a priori*. En cambio, coincide con él al afirmar el carácter sintético *a priori* de la geometría: «Al calificar las verdades geométricas de sintéticas y *a priori*, reveló su verdadera naturaleza» (Frege, 1884/1996: 130). Tal y como reconocimos en Kant, también Frege defiende que las verdades geométricas rigen en el dominio de lo espacialmente intuitivo, ya se trate de la realidad física o incluso de la imaginación. Por ejemplo, «las más extravagantes de las fantasías febriles, las invenciones más atrevidas de leyendas y poetas, en las que los animales hablan, las estrellas se quedan quietas, las piedras se hacen hombres y los hombres árboles, en las que se nos enseña cómo uno puede salirse de un cenagal tirando de su propia peluca, siguen estando sujetas, en la medida en que son intuibles, a los axiomas de la geometría» (Frege, 1884/1996: 58).

Matemáticos como Beltrami, Klein o Poincaré habían demostrado la consistencia de las nuevas geometrías no euclídeas, lo cual dejaba abierto un problema, a saber, la existencia de teorías matemáticas sobre un mismo ámbito internamente consistentes pero mutuamente excluyentes entre sí. El mismo Frege acepta que estas geometrías son útiles, pero insiste en que se apartan completamente del terreno de la intuición y quedan relegadas al «pensamiento conceptual». Este puede aceptar el opuesto de uno u otro axioma y extraer conclusiones sin caer en contradicciones, lo cual viene a mostrar que los axiomas geométricos son independientes entre sí y de las leyes lógicas primitivas, es decir, son sintéticos. Pero el pensamiento conceptual queda completamente al margen de la intuición (ibid.). Por consiguiente, la consistencia de las nuevas geometrías no prueba su verdad, mientras que la verdad de la geometría euclídea reconocida por intuición prueba su consistencia.

Aunque el propio Frege reconoce la proximidad de su concepción de la geometría a la de Kant, conviene no asimilarlas y explicitar las diferencias. Tanto para uno como para otro la intuición consiste en una representación particular e inmediata de un objeto simple o de una relación entre objetos a partir del ejercicio de la sensibilidad. Esta representación debe ser evidente, esto es, el objeto es aprehendido de manera inmediata, sin hacer uso de conceptos o de operaciones lógicas. Para Frege, lo característico de la geometría es que sus objetos son tomados como representantes de toda la especie debido a que, a diferencia de los números, no poseen propiedades peculiares que los definan. A partir de esta consideración da cuenta de la universalidad de los axiomas de la geometría sin recurrir, como hace Kant, al espacio como forma pura de la intuición.

4. La polémica entre Frege y Hilbert acerca del contenido intuitivo

En 1899 publica Hilbert los *Grundlagen der Geometrie*. Con la máxima austeridad lógica y el mínimo simbolismo, Hilbert convenció a los matemáticos del carácter abstracto y puramente formal de la geometría, haciendo que el programa formalista se extendiera a toda la matemática e incluso a las ciencias experimentales. El libro comienza con la siguiente cita de Kant: «Así, pues, todo conocimiento humano comienza con intuiciones, de éstas pasa a conceptos y termina con ideas» (Hilbert, 1996: 1). Sin embargo, la concepción de la geometría que tiene Hilbert está muy lejos de la de Kant y Frege. En la introducción aclara que el objetivo es el de «poner de manifiesto los axiomas de la geometría y el averiguar sus conexiones, problema que se encuentra discutido desde los tiempos de Euclides en numerosos y excelentes tratados de literatura matemática» (ibid.). Finalmente afirma que este problema queda reducido, en definitiva, «al análisis lógico de nuestras intuiciones espaciales» (ibid.).

La referencia a las intuiciones espaciales no debe llevarnos a sospechar una proximidad con el planteamiento de Frege. Por el contrario, el «análisis lógico» de las intuiciones busca la elaboración de un sistema completo de axiomas del que se puedan deducir los más importantes teoremas, de tal manera que en ese proceso «aparezcan con la máxima claridad la interpretación de los distintos grupos de axiomas y el alcance de las consecuencias que aisladamente se deriven de cada uno de ellos» (ibid.), y todo ello con independencia de la intuición. El «análisis lógico» no se asienta sobre el contenido intuitivo de los conceptos y las relaciones, sino que prescinde de ello, de tal manera que todas las definiciones se establecen a partir de las relaciones explícitas, en uno u otro nivel, entre los signos primitivos del sistema, que están vacíos de contenido (son *inhaltsleer*, como repite el propio Hilbert).

Después de leer los *Grundlagen*, Frege envió una carta a Hilbert el 27 de diciembre de 1899, en la que insistía en la justificación de los axiomas a partir de la intuición. La respuesta de Hilbert, del día 29, fue la siguiente (Frege, 1980: 39-40):

> Usted escribe: «Denomino axiomas a las proposiciones que son verdaderas pero que no son demostradas porque nuestro conocimiento de ellas surge de una fuente muy diferente de la lógica, una fuente que podría llamarse intuición espacial. De la verdad de los axiomas se sigue que no se contradicen entre sí». Encuentro muy interesante leer esta frase de su carta, ya que por lo que a mí se refiere siempre que he pensado, escrito, y dado clases o conferencias sobre estos temas, he dicho justo lo contrario: Si los axiomas dados arbitrariamente no se contradicen entre sí con todas sus consecuencias, entonces son verdaderos y las cosas definidas por los axiomas existen. Para mí, éste es el criterio de verdad y existencia.

En la respuesta del 6 de enero de 1900, Frege se opone frontalmente a que la consistencia de un sistema de axiomas implique la existencia de un modelo, precisamente porque no se puede demostrar la consistencia sin recurrir a un modelo previo. Efectivamente, en los propios *Grundlagen* se demuestra la consistencia del sistema de axiomas geométricos utilizando un modelo aritmético. Pero, si como el propio Hilbert mantiene, también la aritmética puede reducirse a una teoría axiomática formal, entonces la consistencia demostrada para la geometría es relativa a la aritmética, cuya consistencia será también relativa, de tal manera que solo saliendo de la axiomática formal y recurriendo a la intuición será posible encontrar una fundamentación absoluta que permita sustentar la cadena de fundamentaciones.

En la misma carta Frege insiste en que incluso la independencia de los axiomas que Hilbert demuestra queda reducida a una demostración de consistencia. Efectivamente, Hilbert prueba que cada axioma es independiente de los demás indicando un modelo que satisface a los demás axiomas, pero no aquel cuya independencia trata de probar. De esta manera, al tiempo que prueba la independencia del axioma de las paralelas prueba la consistencia de la geometría no euclídea. Sin embargo, todo ello es para Frege un ejercicio del pensamiento conceptual que pone de manifiesto mundos posibles, en tanto que no contradictorios, pero no efectivamente existentes ni verdaderos.

Como Kant, también Frege insiste en que el conocimiento es receptividad, esto es, presencia de lo presente en la intuición sensible o en la intuición matemática. Conocer no es crear, ni construir, en el espacio de lo no contradictorio, sino dejar aparecer, con arreglo a las condiciones que lo hacen posible. La objetividad y la verdad del conocimiento pasan precisamente por atenerse a dichas condiciones. Por ello, la única manera de demostrar la consistencia de los axiomas pasa por mostrar en la intuición matemática que existe un objeto que tiene las propiedades referidas en los axiomas. Frege lo dice así: «La única manera que conozco es la siguiente: mostrando un objeto que posea todas esas propiedades, señalando un caso en el que todos esos requisitos se cumplan. No me parece posible demostrar la falta de contradicción de ninguna otra manera» (Frege, 1980: 43). Y para demostrar la independencia de un axioma respecto de los demás también es necesario recurrir a la intuición, en la medida en que hay que mostrar un caso en el que el axioma en cuestión no se cumple cumpliéndose los demás. Frege insiste en que esto es imposible en el caso de la geometría euclídea, porque ateniéndonos a la intuición de lo dado se reconoce el cumplimiento de todos sus axiomas, esto es, su verdad.

Quiero destacar que la diferente concepción de Frege y Hilbert sobre la geometría radica en el papel que uno y otro atribuyen a la intuición. Según Frege, la matemática acontece en el ámbito de los pensamientos y no en el de los signos y las oraciones, que son simplemente el medio por el que los pensamientos son expresados. Y es a los pensamientos mismos a los que corresponde el establecimiento de axiomas y el seguimiento de unas determinadas reglas de inferencia que van de axiomas a teoremas. Por consiguiente, para Frege la consistencia y la independencia se predican de los pensamientos y no de las oraciones, entendidas por Hilbert como fórmulas vacías de contenido intuitivo.

A partir de 1892, fecha en que publica el artículo «Sobre sentido y referencia», Frege distingue entre signo (*Zeichen*), sentido (*Sinn*) y referencia (*Bedeutung*). El signo es la expresión lingüística; el objeto al que una expresión se refiere o designa es su referencia; la peculiar manera de referirse a él es su sentido. El sentido de un enunciado o de una sentencia es el pensamiento *objetivo* (*Gedanke*) por él expresado, que no hay que confundir con la representación subjetiva (*Vorstellung*) que se puede formar en la mente de quien use el enunciado. El sentido, o el pensamiento, no es el objeto, pero tampoco es algo subjetivo como la representación (Frege, 1892/1998: 89). Estas distinciones son aclaradas definitivamente en el año 1918 con la publicación de «El pensamiento: una investigación lógica». Se puede decir que «el pensamiento es el sentido de una oración, sin querer aseverar con esto que el sentido de toda oración sea un pensamiento. El pensamiento, imperceptible en sí, se viste con el ropaje perceptible de la oración, con lo que somos capaces de captarlo. Decimos que una oración expresa un pensamiento» (Frege, 1918/1998: 200).

Las representaciones constituyen nuestro mundo interior y son contenidos del yo. Hablamos de impresiones sensoriales, creaciones de la imaginación, estados de ánimo, inclinaciones, deseos, decisiones, etc. Se caracterizan por (Frege, 1918/1998: 209-211): 1) no pueden ser vistas, ni tocadas, ni olidas, ni gustadas, ni oídas; 2) se tienen, esto es, pertenecen al contenido de la conciencia de cada cual; 3) necesitan un portador, a diferencia de las cosas externas, que son independientes; 4) tienen solo un portador, de tal manera que no hay dos personas que tengan la misma representación. Ahora bien, no todo lo que puede ser objeto del conocer es representación, de tal manera que de ella hay que distinguir, por un lado, el objeto al que apunta y, por otro, el pensamiento. No somos portadores de los objetos, ni tampoco de los pensamientos. Unos y otros son independientes de la subjetividad y no dependen de los cambiantes estados de conciencia de los hombres. Sin embargo, hay una diferencia entre ellos: los objetos pertenecen al mundo exterior propio de las cosas sensorialmente perceptibles, y los pensamientos no son perceptibles sensorialmente, sino *captables*, o *aprehensibles*, por el pensar. La captación del pensamiento efectivamente presupone alguien que capta, alguien que piensa. «Ese es, pues, el portador del pensar, pero no del pensamiento. Aunque el pensamiento no pertenece al contenido de conciencia del que piensa, sin embargo, algo en la conciencia tiene que apuntar al pensamiento» (Frege, 1918/1998: 221). Por consiguiente, «el pensamiento no pertenece ni a mi mundo interior, como representación, ni tampoco al mundo exterior, al mundo de las cosas perceptibles por los sentidos» (ibid.).

El resultado es que, según Frege, debe admitirse un tercer reino de cosas, los pensamientos, que no pueden ser percibidos por los sentidos, sino captados por la intuición, y que son atemporalmente verdaderos independientemente de que alguien en algún momento los descubra y los tome por verdaderos. No necesitan portador, ni son verdaderos solamente desde que son descubiertos, sino que, como los planetas, ya estaban ahí. Además, permanecen intactos y el hombre no tiene ningún poder sobre ellos, de tal manera que, incluso el acto de captación, que provoca cambios en el mundo interior del que los capta, los deja inalterados. Por ello concluimos que son atemporales. El teorema de Pitágoras pertenece a este ámbito y, en general, toda la matemática.

La verdad es una propiedad de los pensamientos, o más precisamente, de los pensamientos reconocidos como verdaderos. Podríamos hablar, Frege lo hace en algún momento (Frege, 1918/1998: 200), de pensamientos falsos, aunque siendo estrictos la expresión es inadecuada, dado que en tal caso habría sentido, pero, en rigor, no habría pensamiento. Lo relevante al respecto es que la verdad no es tratada como una relación de correspondencia entre una figura, o una representación, o una oración, y el objeto figurado o representado. El argumento que ofrece insistentemente Frege es que representación y objeto no se pueden comparar porque son siempre diferentes, al pertenecer a distintas regiones ontológicas. La verdad es presencia, donación, manifestación de los pensamientos.

A partir de aquí es fácil comprender que Frege afirme que la lógica estudia «las leyes del ser verdad… Y, así, es posible hablar también de las leyes del pensamiento» (Frege, 1918/1998: 196). De la misma manera que hay unas leyes naturales que regulan y estructuran los acontecimientos naturales, hay unas leyes que estructuran el ámbito del pensamiento, y a la lógica corresponde la tarea de hacerlas explícitas. No trata, por consiguiente, de los procesos mentales del pensar y de sus leyes psicológicas, ni de explicar cómo se toma algo por verdadero, sino de la articulación del ámbito de la verdad, que es el ámbito del pensamiento. Lo primero, que pasa por estudiar las mentes y los contenidos de conciencia de los que el hombre individual es portador, corresponde a la psicología. Pero la frontera con la lógica es clara, dado que ésta se ocupa del estudio del tercer reino, el de los pensamientos. De esta manera tan rotunda expresa Frege su radical rechazo del psicologismo tanto en lógica como en matemáticas.

Así las cosas, se entiende la diferencia entre la concepción de la geometría de Frege y de Hilbert en relación con el papel que asignan a la intuición. Para Frege, ejercer la intuición es instalarse y aprehender el ámbito de los pensamientos y de la verdad. Desde la intuición es posible captar el sentido de conceptos y axiomas y reconocerlos como pensamientos, es decir, como verdades. A partir de aquí, una intuición guiada por los raíles de la estructura lógica del campo de los pensamientos es capaz de alcanzar los teoremas. Que el sistema axiomático finalmente reconocido, es decir, descubierto, a partir del ejercicio de la intuición, sea consistente e independiente, es algo que se infiere de su verdad, es decir, de la presencia misma de lo dado a la intuición. En la medida en que lo dado, tanto los pensamientos como la estructura lógica que los articula, se hace impertinentemente presente y es reconocido como independiente del acto de pensar y de los contenidos ingredientes de la conciencia, es posible hablar de una verdad ligada a la existencia. Por supuesto, siempre es posible abstraer de la existencia sus condiciones formales, es decir, prescindir del contenido intuitivo de los conceptos, del pensamiento que los define, y quedarnos con las relaciones que articulan los conceptos en axiomas y los axiomas con teoremas. Es, a juicio de Frege, lo que hace Hilbert en los *Grundlagen* utilizando lo que de forma despectiva llama «pensamiento conceptual». También es posible, una vez abstraídas las condiciones formales, jugar con ellas prescindiendo de unas o de otras, poniendo como regla del juego el respeto del principio de contradicción. De esta manera podemos prescindir de un axioma o de otro, entendidos ahora como meras fórmulas vacías de contenido. Lo único que se muestra con este proceder es, según Frege, que la geometría es sintética, y que lo formal define un ámbito de lo posible mucho más amplio que el ámbito de lo efectivamente existente. Por supuesto, no hay nada que objetar al juego mientras se tenga en cuenta, cosa que no hace Hilbert, que el paso de lo posible al reconocimiento de lo efectivamente verdadero y existente requiere por parte del matemático del ejercicio de la intuición, que es en el fondo la afirmación de la finitud del conocimiento humano. Por ello dice Frege, refiriéndose a las teorías formalistas en general y a la aritmética formalista de Hankel en particular, que se comportan «como un dios que con su pura palabra puede crear lo que necesita» (Frege, 1884/1996: 144).

A pesar de la crítica de Frege, los *Grundlagen* de Hilbert no supusieron la única interpretación formalista del método axiomático, pero sí la más depurada y la que más influencia tuvo en el desarrollo de la matemática y la lógica desde comienzos del siglo XX. Entender la geometría a partir de sistemas axiomáticos formales generaba problemas filosóficos de calado cuando se tomaba como trasfondo el kantismo, pero las nuevas corrientes positivistas en Europa y pragmatistas en USA se apartaban del kantismo con reflexiones ontológicas y epistemológicas más compatibles con el formalismo. Por otro lado, a los matemáticos les permitía sostener la coexistencia de varias geometrías no euclídeas, estudiadas a su vez con gran interés por los físicos, dado que veían en ellas la posibilidad de resolver problemas que habían quedado pendientes tras los últimos desarrollos de la mecánica clásica y el electromagnetismo. Además, el interés de Hilbert por hacer depender la normatividad matemática no de la intuición y la verdad, sino de la consistencia, la independencia y la deducibilidad, hizo que la geometría misma se tomara como objeto de estudio por parte de los matemáticos y que se desarrollase la *metamatemática* como estudio de las propiedades de los sistemas axiomáticos formales. Todo ello propició la generalización del formalismo no solo en la matemática, sino incluso en las ciencias empíricas.

La lectura formalista extrema de Hilbert ha sido la más extendida desde la aparición de los *Grundlagen*, si bien las revisiones de su obra realizadas especialmente a partir de los años 90 han destacado las distintas fases de su pensamiento y el diferente papel que le atribuye a la intuición en cada una de ellas[2]. En el origen de aquella lectura está la relación que, tras su publicación, se estableció entre los *Grundlagen* y el formalismo de Hankel, Heine y Thomae. El mismo Frege contribuyó a ello. En el artículo titulado «Sobre los fundamentos de la geometría», publicado en 1906 en *Jahresberichte* como respuesta al artículo publicado por Korselt en el número 12 de la revista, afirma Frege que Korselt tiene razón cuando dice que podemos llamar a la teoría formal pura de Hilbert «"juego de signos vacío, carente de significado" y cosas por el estilo; como rigurosa asociación legal de las proposiciones no precisa de ninguna otra "dignidad" especial» (Frege, 1906/1996: 317). A partir de la década de los 20 fue Hermann Weyl, discípulo de Hilbert convertido al intuicionismo, quien insistió en que el formalismo extremo de Hilbert conducía a concebir la matemática como un mero juego de signos primitivos arbitrariamente elegidos que solo se distinguen unos de otros por

[3] Un estudio riguroso de la evolución del pensamiento matemático y epistemológico de Hilbert es el artículo de J.M. Gambra: «La filosofía de David Hilbert», en *Thémata: Revista de Filosofía*, n.º 14, 1995, pp. 147-179. Un estudio más reciente es el de Giovannini, E.N. (2015): *David Hilbert y los fundamentos de la geometría (1891-1905)*. London: College Publications, Milton Keines.

su forma, y de fórmulas vacías de contenido intuitivo (Weyl, 1925, 1944). Este formalismo atribuido a Hilbert sin matices y con un escaso conocimiento de lo que había sido su reflexión epistemológica con posterioridad a 1922, cuando emprende lo que él mismo considera la construcción definitiva de su programa en oposición al logicismo por un lado y al intuicionismo de Kronecker, Poincaré, Brouwer y Weyl por otro[3], es el que explica Max Newman en Cambridge y el que recibe Turing cuando asiste a sus clases.

5. La conferencia de Newman en Cambridge

El trimestre de Cuaresma de 1935 finalizaba en Cambridge con la conferencia de Max Newman en el St. John´s College. El tema era el de los fundamentos de las matemáticas y, como era de esperar, convocó a pocos alumnos, entre los que estaba Turing. Se suponía que Newman hablaría de Gödel, sobre el que impartiría un curso en el trimestre que ahora comenzaba, pero no lo hizo; habló del formalismo de Hilbert y de la posibilidad de un procedimiento *sistemático* en matemáticas que no requiriese de la intuición.

La discusión en torno al problema de la intuición, que hemos expuesto al hilo de la polémica entre Frege y Hilbert, dejó de ocupar el interés de una buena parte de la comunidad de matemáticos en los años posteriores a la publicación de los *Grundlagen*, debido a la influencia de la lectura más extremadamente formalista de la obra de Hilbert. Ese ambiente formalista fue predominante en el II Congreso Internacional de Matemáticos celebrado en París, con presencia notoria de Hilbert por el anuncio de su famosa lista de 23 problemas abiertos en matemáticas. También Hilbert apareció como invitado estelar en los congresos de 1904, 1908 y, especialmente, en el celebrado en Cambridge en 1912. Fue Gödel, cuarenta años menor que Hilbert, el que puso límites al imperante programa formalista de esos años y obligó a retomar de nuevo la discusión en torno al papel de la intuición en la matemática.

[4] Véase al respecto Gambra (1995).

En su conferencia, Newman analizó el programa formalista de Hilbert al hilo del segundo problema de la famosa lista, el *Entscheidungsproblem*. Los matemáticos debían buscar un sistema completo, consistente y decidible, que consiguiese expresar «todo el contenido de pensamiento de las matemáticas de manera uniforme» (Hilbert, 1927/1967: 475). Se trataba, en fin, de prescindir de la intuición y atenerse estrictamente a un método sistemático que permitiese decidir para cada enunciado si es demostrable dentro del sistema. Ello permitiría eliminar la ignorancia y tener un control total y máxima objetividad en matemáticas, dado que, según Hilbert, se podría determinar la verdad o falsedad de cualquier enunciado al verificar si es demostrable en el sistema. En la conferencia impartida en Bolonia en el año 1928 Hilbert resume así la situación (Hilbert, 1930: 1)[4]. Son los mismos términos que recoge Newman en su conferencia (Copeland, 2013: 23).

> En una serie de presentaciones a lo largo de los últimos años, he... emprendido un nuevo camino para tratar cuestiones fundamentales. Con esta nueva fundación de las matemáticas, que se puede llamar convenientemente teoría de la prueba, creo que las preguntas fundamentales en matemáticas son eliminadas finalmente, al hacer que cada enunciado matemático sea una fórmula concretamente demostrable y estrictamente derivable...
>
> En matemáticas no hay *ignorabimus*, dado que siempre podemos responder preguntas significativas; y queda establecido, como quizás anticipó Aristóteles, que nuestra razón no involucra artes misteriosas de ningún tipo: más bien, procede según reglas formulables que están completamente definidas, y son también la garantía de la absoluta objetividad de su juicio.

Newman insistió en la importancia de alcanzar en matemáticas un procedimiento *sistemático* para prescindir de las «artes misteriosas» a las que se refería Hilbert, es decir, de cualquier forma de intuición y de inventiva. Newman era todavía más concreto en su exposición, puesto que se refería a

[5] La traducción es mía.

un método manual formado por reglas bien conocidas, que cualquiera puede llevar a cabo de forma mecánica, es decir, paso a paso, sin ninguna creatividad ni entendimiento. Es lo que en la época hacían las llamadas *computadoras*, miles de empleados que trabajaban en distintos negocios y agencias del gobierno, adiestrados para hacer cálculos siguiendo repetitivamente unas reglas sin necesidad de entender procedimientos ni objetivos. Newman resumió diciendo que el programa de Hilbert de búsqueda de procedimientos de decisión sistemáticos consistía en definitiva en la búsqueda de procedimientos mecánicos. Añade Cawthorne que «la palabra "mecánico", en su sentido original, se refería a la ocupación manual, al trabajo realizado por seres humanos. Sin embargo, en la década de 1930, mecánico significaba engranajes, rotores, tubos de vacío. Significaba una máquina. Turing tomó ambas definiciones en serio» (Cawthorne, 2014: 54)[5]. Efectivamente, esta originalidad, consistente en reducir los procedimientos lógico-matemáticos al funcionamiento automático de una máquina en el doble sentido señalado, fue lo que inspiró a partir de ese momento el trabajo de Turing. Años más tarde, en una entrevista con Christopher Evans en la que hablaba de la máquina universal inventada por Turing, Newman dijo: «Yo creo que todo comenzó porque asistió a una de mis conferencias sobre los fundamentos de las matemáticas y la lógica»[6].

6. Los teoremas de incompletitud de Gödel

El programa hilbertiano expuesto por Newman en su conferencia quedaba mal parado a la luz del resultado del teorema de incompletitud de Gödel de 1931. En su tesis doctoral, finalizada a mediados de 1929, Gödel respondió afirmativamente a las preguntas que Hilbert y Ackermann habían planteado en 1928, a saber, si el cálculo lógico de primer orden es comple-

[6] La traducción es mía.

[7] Copeland (2013: 24). El autor hace referencia a una entrevista entre Max Newman y Christopher Evans, titulada «The Pioneers of Computing: An Oral History of Computing», localizada en el Science Museum de Londres.

to y si sus axiomas son independientes. Para ello toma como referencia el cálculo lógico de primer orden contenido en los *Principia Mathematica* de Whitehead y Russell. Terminado su doctorado, en su *Habilitation* busca responder al segundo problema de Hilbert, a saber, el de encontrar una prueba finitista de la consistencia de la aritmética. En este momento, los partidarios del programa formalista extremo consideraban que se había alcanzado la axiomatización formal completa de la matemática clásica por medio de los *Principia Mathematica*, pero quedaba pendiente probar la consistencia del sistema formal. Pues bien, lo que Gödel acaba probando en su famoso teorema de 1931 es que *todos* los sistemas axiomáticos formales de la matemática clásica, incluido el de los *Principia Mathematica*, si son consistentes, entonces son incompletos. Aunque se amplíe el sistema añadiendo axiomas, el resultado obtenido sigue siendo incompleto. Además, Gödel demuestra que es imposible probar la consistencia de un sistema axiomático formal de la matemática clásica desde él mismo. Sí es posible una ampliación del sistema, de tal manera que desde el resultante se pueda probar la consistencia del anterior (Mosterín, 2000: 233).

Aunque Gödel no pretendía atacar el formalismo imperante, sino más bien avanzar en él, sus resultados suponían, como decía Weyl, definitivamente una catástrofe para el programa de Hilbert (Copeland, 2004: 48). El objetivo de su trabajo de *Habilitation* era el de buscar una prueba finitista de la consistencia de la aritmética, y se había encontrado con que ningún sistema axiomático formal de la matemática clásica consistente para la aritmética puede contener todas las verdades. Siempre quedan enunciados aritméticos indecidibles, cuya verdad conocemos por intuición pero que no son demostrables. Por consiguiente, el sistema axiomático formal no logra capturar todo el contenido de la aritmética, de tal manera que se hace imprescindible salir del sistema y recurrir a la intuición. Posteriormente, Gödel generalizó este resultado, gracias precisamente al trabajo de Turing, ampliándolo a todo sistema formal, que quedaba definido «como cualquier procedimiento mecánico para producir fórmulas, llamadas fórmulas demostrables» (Gödel, 1964/1965: 71).

Newman se refiere a la situación de la siguiente manera (Newman, 1955: 256)[7]:

> El programa de decisión de Hilbert de las décadas de 1920 y 1930 tenía como objetivo el descubrimiento de un procedimiento general, aplicable a cualquier teorema matemático expresado en forma plenamente simbólica, para decidir la verdad o falsedad del teorema. Se asestó el primer golpe a las perspectivas de encontrar esta nueva piedra filosofal con el teorema de incompletitud de Gödel (1931), que dejó claro que la verdad o falsedad de A no podía equipararse a la demostrabilidad de A o no-A en ninguna lógica de base finita, elegida de una vez por todas; sin embargo, aún quedaba en principio la posibilidad de encontrar un procedimiento mecánico para decidir si A, o no-A, o ninguno de los dos, era formalmente demostrable en un sistema dado. Muchos estaban convencidos de que tal procedimiento no era posible, pero Turing se propuso demostrar rigurosamente esa imposibilidad. Evidentemente, el primer paso fue dar una definición de «procedimiento de decisión» lo suficientemente exacta como para formar la base de una prueba matemática de imposibilidad. A la pregunta «¿qué es un procedimiento "mecánico"?», Turing respondió con la respuesta característica: «Algo que puede ser hecho por una máquina», y se embarcó en la muy agradable tarea de analizar la noción general de una máquina computadora. Hoy en día es difícil darse cuenta de cuán audaz fue la innovación al introducir el discurso sobre cintas de papel y patrones grabados en ellas, en las discusiones sobre los fundamentos de las matemáticas. Vale la pena citar de su artículo [Turing, 1937] el párrafo en el que se presenta por primera vez la máquina computadora, tanto por su contenido como para dar una idea de los escritos de Turing.

La cita es significativa porque hace explícito el asunto central de la conferencia de Newman del año 1935, que definitivamente orientó la investigación de Turing. Efectivamente, Gödel había asestado un duro golpe al programa formalista y a la posibilidad de prescindir de la intuición y de «artes misteriosas» en la matemática. Esto lo reconoció el propio Hilbert cuando tuvo conocimiento de los teoremas de incompletitud, que entendió

[8] La traducción es mía.

y aceptó inmediatamente. Sin embargo, ni para Hilbert ni para los defensores del programa formalista estaba todo perdido, y es en esto en lo que insiste Newman en la conferencia. Y no lo estaba porque si de lo que se trataba era de desterrar la intuición del ámbito de la matemática, lo que Gödel había dejado claro es que tal objetivo no se puede conseguir con la axiomatización formal que recurre a los *Principia*, ni con la aritmética formal de Peano, ni con la teoría axiomática de conjuntos, ni, en general, recurriendo a lógicas de base finita; pero ello no cierra todas las posibilidades. Dado que de lo que se trata es, en definitiva, de encontrar un procedimiento general que permita decidir si A, o no A, o ninguna de las dos proposiciones, se puede deducir de un sistema axiomático formal dado, quizá pueda existir un procedimiento de decisión distinto. La clave que aportó Newman en la conferencia y que recogió Turing es que ese procedimiento de decisión distinto podría entenderse como un procedimiento *mecánico*, esto es, que opera siguiendo paso a paso unas reglas dadas, sin creatividad ni entendimiento, de la misma manera que operan esos empleados del gobierno a los que se les llamaba «computadoras». Tal y como señala Newman en la cita, Turing da un paso más al añadir otro sentido: precisamente porque se trata de un operar siguiendo reglas sin creatividad ni entendimiento, lo mecánico puede ser llevado a cabo por una máquina hecha de engranajes, rotores, tubos de vacío, amplificadores de estado sólido, o desarrollada con cualquier otra tecnología. Lo de menos es la tecnología que se use; lo importante es la explicitación de ese operar siguiendo reglas, sin ningún tipo de intuición ni de experiencia, que posteriormente se podrá implementar con una tecnología u otra.

Turing escucha la conferencia de Newman poco después de ganar una beca del King´s College, que le permitía elegir con total libertad su tema de investigación. Fueron las palabras de Newman sobre la posibilidad de un procedimiento mecánico las que le llevaron a investigar cómo debería ser ese procedimiento que, superando las limitaciones de los teoremas de incompletitud de Gödel, permitiese dar solución al *Entscheidungsproblem*, prescindiendo definitivamente de la intuición y llevando hasta sus últimas consecuencias el programa formalista. El trabajo ya finalizado se lo presentó a Newman con el título «On Computable Numbers with an Application to the *Entscheidunsproblem*», que inmediatamente se dio cuenta de su originalidad y ayudó a que se publicase a finales de 1936.

7. El artículo de 1936

En 1935 Church había comenzado a trabajar en el problema de la decisión a partir del cálculo lambda que acababa de construir. También trabajaban en ello dos discípulos suyos, Kleene y Rosser, así como von Neumann, todos en Princeton. Cuando Turing termina su artículo, Church ya había entregado uno sobre el mismo tema a la London Mathematical Society. Tuvo que intervenir Newman y explicar detenidamente la novedad del planteamiento de Turing frente a los de Gödel y Church para que finalmente se aceptase su publicación. Turing tenía fama de investigar en solitario y de espaldas al trabajo de otros matemáticos, de tal manera que no conocía los resultados que se estaban obteniendo sobre el problema de la decisión, excepto los dos artículos que había publicado Church en 1936, a los que hace referencia en su publicación.

«On Computable Numbers» establece el nexo entre los procedimientos formales algorítmicos y los dispositivos automáticos. Comienza distinguiendo dos tipos de máquinas, las *automatic machines* (*a-machines*), que son aquellas cuyo comportamiento está determinado completamente por su configuración, es decir, por las instrucciones que se le proporcionan, y las *choice machines* (*c-machines*, *oracle machines*), cuyo comportamiento está solo parcialmente determinado por la configuración, de tal manera que cuando alcanza una de las configuraciones ambiguas, un operador externo hace alguna elección arbitraria (Turing, 1937: 232). Soare afirma que las primeras se corresponden con las actuales máquinas *offline*, y las segundas con los actuales procesadores que se comunican con una base de datos externa, como puede ser la World Wide Web. Las *c-machines* proporcionan una teoría matemática para la información en línea interactiva, e incluso, como llegó a pensar el propio Turing, un modelo útil para la comprensión del cerebro humano (Soare, 2009).

En el artículo Turing define las a-máquinas a partir de: 1) una cinta con celdas que se extiende indefinidamente en ambas direcciones; 2) cada celda registra un único símbolo; 3) un cabezal, que es a la vez de lectura y de escritura, se comporta determinado por una configuración o un conjunto de reglas; y 4) el cabezal tiene tres comportamientos: ir hacia la derecha, ir hacia la izquierda, o permanecer en la misma celda una vez realizada la operación. La máquina opera sobre la cinta borrando el símbolo que se encuentra en la celda y sustituyéndolo por un blanco, o reemplazándolo por uno nuevo, o reescribiendo el mismo símbolo. En cada caso, toma la decisión de avanzar o retroceder una posición en la cinta o permanecer en la misma celda.

La formalización de la definición de una a-máquina puede ser la siguiente (Vargas, 2012: 53):

> Una a -máquina es una séxtupla $< Q, \Sigma, \Gamma, q_0, \delta, F >$, donde:
> • Σ un alfabeto de entrada, digamos, $\Sigma = \{a, b, c, d, \ldots\}$
> • Γ un alfabeto auxiliar, conteniendo tanto a Σ como a otros símbolos, en particular, B (el blanco), es decir, $\Gamma = \Sigma \cup \{B, \ldots\}$
> • Q un conjunto de estados, $Q = \{q_0, q_1, q_2, \ldots\}$
> • $q_0 \in Q$, denominado estado inicial
> • $F \subseteq Q$, el conjunto de estados finales
> • δ es una función de transición definida como $\delta: Q \times \Gamma \rightarrow Q \times \Gamma \times \{L, R\}$ (Left, Right)

Por consiguiente, «una descripción de estado, es decir, el comportamiento de la máquina en un momento determinado, puede ser descrito en términos de los siguientes cinco elementos: $(q_i, X, Y, q_h, \{L, R\})$, es decir, el estado en el que se encuentra, el símbolo que está leyendo, el símbolo que reemplazará, el nuevo estado al que cambiará y el movimiento que realizará a la derecha o a la izquierda» (ibid.).

Definida así la a-máquina, podemos decir que una función numérica es computable si y solo si hay una a-máquina que la computa. De la misma manera, una relación es decidible si y solo si hay una a-máquina capaz de computarla. Y un conjunto es recursivamente numerable si y solo si hay una a-máquina que lo genera, o dicho de otra manera, si resulta ser la imagen de una función numérica computable (por la a-máquina).

Después de remitir al lector al capítulo 3 de los *Grundzüge der Theoretischen Logik* de Hilbert y Ackermann con el fin de aclarar los términos del *Entscheidungsproblem*, Turing declara que su propósito es «mostrar que no puede haber un procedimiento general para determinar si dada una fórmula U del cálculo funcional K [de Hilbert], ésta puede probarse, es decir, que no puede haber ninguna máquina que, provista de una de estas fórmulas U, eventualmente diga si se puede probar» (Turing, 1937: 259)[8].

Indica el propio Turing que «Gödel ha mostrado que en el formalismo de los *Principia Mathematica* hay proposiciones U tales que ni U ni no-U son demostrables. Como consecuencia de esto, se muestra que dentro de ese formalismo no se puede dar ninguna prueba de consistencia de los *Principia Mathematica* (o de K)» (ibid.). Pero lo que Turing intenta demostrar es una tesis más amplia, a saber, «que no existe ningún método general que indique si una determinada fórmula U es demostrable en K o, lo que es lo mismo, si el sistema formado por K con no-U adjunto como un axioma extra es consistente» (ibid.).

La demostración fue inmediatamente aceptada por Newman, que la defendió insistiendo en que daba una respuesta negativa al problema de la decisión más extensa que la ofrecida por la demostración a través del cálculo lambda de Church, e incluso por el teorema de incompletitud de Gödel. Lo cual era debido, como reconoció el propio Gödel, a la precisa delimitación del concepto de lo «efectivamente computable» a través de la correcta definición de «computabilidad mecánica» (Soare, 2009: 14). Unos años más tarde, en 1939 en la universidad de Notre Dame, Gödel explicaba

[9] La traducción es mía.

que Turing había inventado un artefacto parecido a una máquina de escribir con una manivela en el que se podían teclear fórmulas matemáticas. Al teclear una fórmula en cálculo de predicados monádico y girar la manivela, la máquina hacía sonar su timbre si la fórmula era demostrable y permanecía en silencio si no lo era. Lo que Turing había demostrado es que saliendo de este cálculo era imposible construir una máquina computadora finita que pudiera decidir si las fórmulas eran o no demostrables.

Gödel concluía que, de esta manera, Turing confirmaba, a pesar del empeño de Hilbert y los formalistas, que la intuición y la mente humana nunca podrían ser reemplazadas por una máquina (Copeland, 2013: 31). Para Gödel la resolución del problema de la decisión tenía consecuencias no solo en el ámbito de la matemática, sino también en el de la teoría general del conocimiento, en la medida en que llevaba a considerar que la extensión del pensamiento es mayor que la de la axiomática formal y lo efectivamente computable. La intuición seguía siendo una capacidad exclusivamente humana que permitía el reconocimiento del amplio campo de la matemática más allá de los límites de la axiomática formal y la computación, así como el reconocimiento de una verdad que no se puede reducir ni a lo uno ni a lo otro. En un escrito muy posterior, titulado «A Philosophical Error in Turing´s Work» publicado en 1972, Gödel se refería a este momento insistiendo en que la mente humana es capaz de representar y evaluar proposiciones matemáticas utilizando recursos de orden muy superior a los utilizados por una máquina, y que por esta razón se debieran estudiar los procedimientos mecánicos y lo efectivamente computable sin ponerlos en relación con la mente, la teoría general del conocimiento o la ontología. Tal y como veremos, Turing está de acuerdo con esta perspectiva a partir de 1950. Sin embargo, en 1936 la perspectiva compartida era muy distinta: de una manera general, y Gödel en particular, se insiste en relacionar la mente y la máquina computadora cuando de lo que se trata es de poner límites a la capacidad de la máquina y así reconfortarnos en la especificidad y excelencia de la intuición y el pensamiento humano.

8. Maquinaria computacional e inteligencia

Siguiendo las indicaciones de Newman que Turing había escuchado en la conferencia de 1935, «On Computable Numbers» tiene por objetivo superar las limitaciones de los teoremas de incompletitud de Gödel a través de un procedimiento general mecánico de decisión. El artículo demuestra que no existe tal procedimiento, pero lo más importante es que ello exige a Turing definir lo que es un procedimiento general de decisión a partir de una máquina computadora. Dicho de otra manera: mientras tiene por objetivo encontrar un procedimiento general de decisión en la matemática que permita prescindir de la intuición, lo que logra es refutar que la matemática sea decidible a través de la invención de una máquina computadora determinista finita. Será John von Neumann en 1938 en Princeton el que le hable a Turing de un grupo de ingenieros eléctricos capaces de desarrollar una máquina universal con programa almacenado.

En septiembre de 1936 llega Turing al departamento de matemáticas de Princeton invitado por Church y de nuevo centra su trabajo, que presentará como tesis doctoral, en el problema de la intuición, pero ahora desde una nueva perspectiva. Su objetivo ya no es el de erradicar la intuición de la matemática, sino «controlarla» para evitar hacer un uso excesivo de ella, prohibiéndola en general excepto en circunstancias por determinar (Copeland, 2013: 42). Se trata, en definitiva, de entender hasta dónde se pueden superar las limitaciones de los teoremas de Gödel con la estrategia de añadir nuevos axiomas a la aritmética. Dado un sistema axiomático formal en el que se ha reconocido una fórmula indecidible, es posible su ampliación incorporando *ad hoc* algún axioma, de tal manera que la fórmula resulte decidible en el sistema ampliado. El problema es que en este nuevo sistema siempre encontramos una fórmula indecidible. Si ampliamos de nuevo el sistema añadiendo axiomas, encontraremos una nueva fórmula indecidible, de tal manera que no basta con añadir un número finito cualquiera de axiomas; habría que añadir infinitos axiomas que se corresponderían con los números ordinales transfinitos. Turing muestra de esta manera que no es posible superar las limitaciones impuestas por los teoremas de Gödel. Church revisó los resultados y Turing presentó su tesis en Princeton en 1938 con el título *Systems of Logic based on Ordinals*.

Me refiero de nuevo al punto 11 de la tesis doctoral. Turing insiste en que su objetivo iba en la dirección del programa formalista, la de encontrar una lógica formal que admita una variedad de pasos que hagan posible demostrar una proposición particular sin recurrir a la intuición, reemplazando así todos los juicios intuitivos de la matemática por un número finito de reglas. A esto es a lo que llama «ingenio». Sin embargo, acaba reconociendo que su trabajo le ha llevado al extremo opuesto, eliminando no la intuición, sino el ingenio. Por consiguiente, sigue siendo necesario recurrir a sistemas de lógica en los que no todos los pasos de una demostración son mecánicos, sino que algunos tienen que ser necesariamente intuitivos. Nos encontramos ante un nuevo intento fallido de prescindir de la intuición y las especiales capacidades del conocimiento humano -misteriosas capacidades, decía Hilbert- para hacer matemáticas.

En definitiva, a la altura de 1938 estaba abierta la posibilidad de construir una máquina computadora que llegase a las proposiciones alcanzables por procedimientos formales, es decir, a las proposiciones «mecánicas». Von Neumann estaba avanzado rápidamente en este proyecto. Pero quedaba claro que las «no mecánicas» ni eran demostrables, ni eran computables, y su verdad solo podía ser reconocida por la intuición.

En 1950 Turing publica el que se considera el artículo fundacional de la IA. Se titula «Computing Machinery and Intelligence». En la primera línea del artículo se explicita la pregunta que lo pone en movimiento: ¿pueden pensar las máquinas? Llegamos así al asunto que quiero destacar en este trabajo. Como hemos visto, hasta 1938 el debate sobre la especificidad del conocimiento humano se había reducido, sobre un trasfondo kantiano reinterpretado por Frege y criticado por Hilbert y los formalistas, al debate sobre la especificidad y superioridad de la intuición frente a la axiomática formal y lo efectivamente computable. La conclusión a la que se había llegado después del trabajo de Gödel primero y de Turing después es que el ámbito de lo intuitivo es más amplio que el ámbito de lo computable. Dado que se considera que la intuición es el rasgo más característico del pensamiento, el corolario es que el alcance del pensamiento es mayor que el de una máquina computadora. A la pregunta: ¿pueden pensar las máquinas?, en 1938 Turing hubiera respondido que no. Ahora, en 1950, la pregunta se plantea como una posibilidad. Dicho de otra manera: a partir de 1950 Turing trata de encontrar la manera de convertir las máquinas computadoras en «máquinas intuitivas» (Larson, 2021: 24), objetivo que se habría declarado imposible en 1938. ¿Acaso no lo es en 1950? ¿Qué le pasó a Turing entre 1936 y 1950 que le hizo replantearse el sentido de todo su trabajo?

9. El desarrollo de los ordenadores electrónicos

De hecho, no fue algo que le pasara solo a él, sino a toda Gran Bretaña: la Segunda Guerra Mundial. Los submarinos alemanes causaban estragos en el tráfico comercial del Canal de la Mancha. Como es sabido, a comienzos de septiembre de 1939 el gobierno británico reunió en Bletchley Park a un grupo de criptoanalistas y matemáticos, entre los que se encontraba Turing, para descifrar las comunicaciones de los submarinos. Los cifrados se generaban con la máquina Enigma, comercializada desde 1920 pero convenientemente modificada por los alemanes. El método de Turing para descifrar se basaba en eliminar un número muy grande de posibles soluciones para los códigos de Enigma buscando combinaciones en las que hubiera contradicciones. Ello se hacía en una máquina de origen polaco a la que llamaban la Bomba. El número de combinaciones posibles que había que comprobar para descifrar los códigos era abrumador para la intuición humana. Pero con los procedimientos adecuados aquella máquina podía cumplir con la tarea de simplificar posibilidades. Lo que no se podía resolver con la intuición se pudo resolver, finalmente, con el ingenio, es decir, a través de la mecánica de un algoritmo.

En 1941 comienza a operar una nueva máquina de cifrado alemana, Tunny, con una tecnología mucho más sofisticada, segura y potente que Enigma. En Bletchley se extiende la preocupación, porque justo cuando se está consiguiendo descifrar de manera acertada y rápida la codificación de Enigma, surge un nuevo código de naturaleza desconocida. En enero de 1942, en lo que se considera uno de los momentos más importantes del criptoanálisis durante la guerra (Copeland, 2013: 114-115), Bill Tutte consigue explicar el comportamiento de Tunny sin tener conocimiento del funcionamiento de la máquina. El trabajo de Tutte hizo posible que, a las pocas semanas, Turing ideara un método en papel para descifrar los mensajes de Tunny. El problema era que, conforme ampliaban los alemanes la red de Tunny, aumentaba la cantidad de cálculos necesaria para descifrar los mensajes y, consiguientemente, el tiempo que se necesitaba para descifrar un solo mensaje era cada vez mayor.

A mediados de 1942 Max Newman deja Cambridge para trabajar en Bletchley Park. En el laboratorio Cavendish de Cambridge había visto cómo se usaban circuitos electrónicos para medir emisiones radiactivas y propone emplearlos para aumentar la velocidad de cálculo de la nueva máquina que quiere construir para descifrar los mensajes de Tunny. Por otro lado, desde 1940 Tommy Flowers, especialista en ingeniería de válvulas electrónicas para procesamientos digitales a gran escala y velocidad, ayudaba a Turing para mejorar el comportamiento del descifrado que la máquina Bomba estaba realizando de los mensajes de Enigma. Aprovechando la tecnología de circuitos y de válvulas, Turing y Newman intentaron convencer al jefe de Bletchley, Edward Travis, para construir una máquina totalmente electrónica que tuviera la velocidad y la fiabilidad suficientes para vencer a Tunny. Pero, debido a que los filamentos incandescentes de las válvulas se rompían con facilidad, el proyecto fue finalmente rechazado. Flowers sabía cómo disponer de válvulas fiables, así que, en secreto, inició la construcción de la máquina en el laboratorio de ingeniería donde había trabajado. Este trabajo culminó con la construcción de Coloso, el primer ordenador electrónico digital. Cuando fue instalado en Bletchley Park en enero de 1944, los criptoanalistas se sorprendieron de su velocidad y fiabilidad, de tal manera que, en poco tiempo, el número de desciframientos se quintuplicó[9].

Junto con las máquinas Bomba de Turing, Coloso resultó de enorme utilidad en la preparación del desembarco de las tropas aliadas en Normandía. Las autoridades pidieron la construcción de más máquinas Coloso, y en junio de 1944 se puso en marcha Coloso II. Cuando terminó la guerra, Newman estaba al cargo de las diez máquinas de este tipo que había en Bletchley Park, de las cuales se destruyeron al menos ocho por orden, al parecer, del propio Churchill, con el fin de proteger secretos de estado. Continuó su trabajo con los ordenadores y su colaboración con Turing fundando el Royal Society Computing Machine Laboratory y desarrollando un ordenador electrónico con programa almacenado. Esta era una idea central en el planteamiento de Turing que, sin embargo, Flowers no supo entender ni implementar en Coloso, que se programaba a mano mediante clavijas e interruptores y laboriosos procesos de recableado.

Cuando Coloso era descargado en secreto en Bletchley Park, John Presper Eckert y John William Mauchly trabajaban en la Universidad de Pensilvania en el desarrollo de ENIAC (Electronic Numerical Integrator and Computer), una calculadora de gran velocidad para confeccionar las tablas de tiro de la artillería. Finalmente fue presentado en 1946 con todo el apoyo y publicidad por parte de von Neumann, que desconocía la existencia de Coloso. La referencia a ENIAC y a los avances tecnológicos de los americanos fue utilizada por John Womersley, que trabajaba en el National Physical Laboratory de Londres e ignoraba todo lo que había pasado en Bletchley Park, para convencer a su presidente de la necesidad de desarrollar un ordenador electrónico de programa almacenado siguiendo las directrices marcadas en «On Computable Numbers». Tras contactar con Turing a través de Newman, comenzaron a trabajar en el diseño de ACE (Automatic Compu-

[10] Un detallado estudio histórico de la construcción de Coloso se encuentra en Copeland (2013: VII).

ting Engine). Copeland insiste en el hecho de que «especificaba una memoria de alta velocidad con aproximadamente la misma capacidad que uno de los primeros Macintosh de Apple: algo descomunal para los estándares de la época. Turing entendía que la capacidad de memoria y la velocidad eran las claves de la informática» (Copeland, 2013: 152-153). Cuando finalmente se aprobó su construcción, los medios de comunicación publicaron: «ACE es superior al modelo estadounidense: tiene más capacidad de memoria»; «ACE acelerará los vuelos en reactor; resolverá los problemas de aerodinámica»; «puede que ACE sea el cerebro más veloz»; «el trabajo de un mes en un minuto» (Copeland, 2013: 154-155).

Mientras todo esto sucedía en Londres, en EEUU continuaba el desarrollo de los ordenadores electrónicos. Von Neumann había entrado tarde en el proyecto ENIAC, cuando la construcción de la máquina ya había comenzado, de tal manera que su participación fue la de consultor. Rápidamente señaló que el defecto principal de la máquina era que no disponía de programa almacenado, e insistió en trabajar en ello. A partir de ese momento comenzaron una serie de reuniones entre von Neumann, Eckert y Mauchly, que se extendieron durante 1945, para desarrollar EDVAC (Electronic Discrete Variable Arithmetic Computer).

10. ¿Piensan las máquinas? 1950

Con estas referencias históricas sobre el desarrollo de la computación pretendo dar cuenta del cambio de posición de Turing que se aprecia en su trabajo de 1950. Él ha sido parte activa y una continua referencia en el desarrollo de máquinas computacionales cada vez más veloces y fiables, cuyo ámbito de resolución de problemas se ha ido ensanchando. Los códigos irresolubles para una máquina Bomba eran resueltos por una máquina Coloso, y la velocidad de resolución aumenta con Coloso II. Por otro lado, ENIAC hace posible que los cálculos requeridos por el cuerpo de artilleros, que necesitaban de cientos de hombres y mujeres computadoras y de largos periodos de tiempo para su resolución, se realicen en pocos días. Después de la guerra, se desarrolla el proyecto ACE con el objetivo de construir una máquina universal de programa almacenado que se pueda utilizar no solo en el ámbito militar, sino también en la industria civil, en los laboratorios de investigación, en las universidades o en las distintas administraciones del estado. Los problemas que en un determinado momento parecen irresolubles por las máquinas y solo al alcance del conocimiento humano amplificado a través del trabajo colaborativo de cientos de personas, resulta que pueden ser reducidos a problemas computables que se resuelven a través de algoritmos o programas ejecutados por máquinas con sistemas computacionales cada vez más potentes. Por supuesto que hay un límite a la deducibilidad y a la computabilidad, tal y como habían mostrado Gödel y Turing respectivamente con anterioridad a 1938 y, por consiguiente, para cada nueva máquina computacional encontraremos problemas irresolubles que requerirán del concurso del pensamiento humano. Pero después del desarrollo de los sistemas computacionales que se ha realizado durante la guerra la cuestión es: ¿dónde está el límite cuando parece que la frontera de lo irresoluble se va desplazando?

En el punto 6(3) de «Computing Machinery» Turing insiste, en primer lugar, en que Gödel «demuestra que en cualquier sistema lógico lo suficientemente poderoso se pueden formular aseveraciones que no se pueden ni probar ni desaprobar dentro del sistema, a menos que el sistema en sí sea inconsistente» (Turing, 1950/2004: 450); insiste, en segundo lugar, en que él ha descrito los sistemas lógicos en términos de máquinas computacionales digitales y éstas en términos de sistemas lógicos; y concluye, y esto es lo que quiero destacar en el presente trabajo, haciendo una relectura del teorema de Gödel, según la cual «las preguntas que no podrían ser respondidas por una máquina podrían ser respondidas satisfactoriamente por otra» (Turing, 1950/2004: 451). Turing ya no entiende que el teorema de Gödel marque un límite para los sistemas axiomáticos formales, sino que, por el contrario, define una estrategia para superar los límites de cada sistema axiomático formal en particular. En consecuencia, y dada la relación por él establecida entre sistemas lógicos y máquinas computacionales, también ofrece una estrategia para construir sistemas computacionales más potentes capaces de resolver los problemas que quedaron pendientes en los sistemas computacionales de partida.

Si esta interpretación no se había hecho en los años 30 es porque, tal y como hemos visto, el debate estaba centrado en el papel de la intuición y formulado en los términos que aparecen en la polémica entre Frege y Hilbert primero y entre intuicionistas y formalistas después. Lo que se había entendido es que el teorema de Gödel ponía un límite a la deducción y al reconocimiento de la verdad *frente a la supuestamente ilimitada* capacidad de la intuición. El trasfondo del asunto es, como ya hemos examinado, kantiano: los humanos somos seres privilegiados con la capacidad de obtener conocimientos verdaderos por intuición, y el intento de Hilbert y el formalismo extremo de reducir lo verdadero a lo deducible queda frustrado por el teorema de Gödel. Recordemos que, en este contexto, la conclusión que extrae el propio Gödel de «On Computable Numbers» es que Turing ha confirmado que la intuición y la mente humana nunca podrán ser reemplazadas por una máquina. Y el propio Turing explica en su tesis doctoral que, aunque el objetivo de su trabajo era el de llevar el programa formalista hasta el final y reemplazar la intuición por el ingenio, finalmente ha acabado en el extremo opuesto, eliminando no la intuición, sino el ingenio, y reconociendo la necesidad del uso de la intuición para resolver los problemas que no pueden resolver las máquinas.

Si a la altura de 1950 ni los teoremas de Gödel ni los trabajos sobre computabilidad de Turing se leen como la confirmación de los límites del formalismo y la computación frente a la capacidad *ilimitada* de la intuición, es porque después de la guerra han quedado desconectados de la problemática epistemológica que, sobre el trasfondo kantiano, había surgido con la publicación de los *Grundlagen* de Hilbert y la profunda crítica de Frege. Desconectada la metafísica kantiana que, como hemos visto, Frege inyecta a través de su filosofía de la matemática, es posible entender que lo que han hecho Gödel y Turing es, en primer lugar, reconocer los límites de la deducibilidad lógica y de la computabilidad respectivamente, y en segundo lugar -y lo que es más importante- aportar una estrategia para superar dichos límites, ya sea a través de axiomas *ad hoc* en el caso de los sistemas axiomáticos formales, o a través de sistemas computables más potentes en el caso de las máquinas.

En definitiva, solo suponiendo que el conocimiento humano tiene la *ilimitada* capacidad de alcanzar la verdad a través del ejercicio de la intuición primero y la deducción después, es posible decir que Gödel y Turing han mostrado que ningún sistema formal y ninguna máquina computacional podrán superar el rendimiento del pensamiento. Porque pusimos así las cosas acabamos argumentando -dice Turing ya en 1950- «que hay una discapacidad en las máquinas que el intelecto humano no posee» (ibid.). Ahora bien, si eliminamos la fundamentación filosófica de origen kantiano y fregeano sobre la que se soporta tal supuesto y, consecuentemente, estudiamos los procedimientos mecánicos y lo efectivamente computable sin ponerlos en relación con la mente, la teoría general del conocimiento o la ontología -tal y como propone Gödel en 1972-, entonces podemos hacer la relectura del teorema de Gödel que hace Turing, según la cual los problemas que no pueden ser resueltos por una máquina computable podrán ser resueltos por otra.

¿Por qué -se pregunta Turing- supusimos que el intelecto humano no está discapacitado, e incluso que la capacidad de su intuición es ilimitada? Efectivamente, «se establece que hay limitaciones para el poder de cualquier máquina, [y] solo se ha dicho, sin ningún tipo de prueba, que tales limitaciones no se aplican al intelecto humano» (ibid.). Este es el fondo del asunto, la permanente e implícita vigencia de una metafísica kantiana que no nos ha permitido preguntarnos por la discapacidad del conocimiento humano, sencillamente porque nos ha instalado en una «sensación ilusoria de superioridad», que creemos confirmada cuando se pone en evidencia la falibilidad de las máquinas[10]. Sin embargo -continúa Turing- «nuestra superioridad solo 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» (Turing, 1950/2004: 451). Efectivamente, desconectados de la metafísica kantiana, los trabajos de Gödel y Turing nos llevan a la siguiente conclusión: puesto que siempre es posible un sistema axiomático formal más amplio y un sistema computacional más potente, siempre habrá una máquina más inteligente que el hombre más inteligente.

11. El juego de la imitación

La nueva interpretación que ofrece Turing en «Computing Machinery» del trabajo de Gödel y del suyo propio no pasa solo por el abandono de las tesis kantianas a propósito del conocimiento y del ejercicio de la intuición, sino también por el abandono de una definición del pensamiento desde la perspectiva de la primera persona, que atribuye a modo de ejemplo a un reputado neurólogo contemporáneo suyo, Sir Geoffrey Jefferson. Según esta concepción, pensar no es solo presencia del objeto, sino, además, pre-

[11] La «sensación ilusoria de superioridad» surge porque, al considerar que el conocimiento es la condición que hace posible la emergencia del mundo, Kant ha fusionado los límites del ser y los del pensar. Si, saliendo del marco kantiano, se plantea la posibilidad de que los límites del ser sean más amplios que los del pensar, entonces adquiere pleno sentido la pregunta por un acceso al ser que no sea desde el conocer. Desde luego, se pueden plantear accesos desde la mística o la religión. Dennett propone acceder desde la *acción*, desplegada por algoritmos recursivos biológicos que desarrollan *competencias sin comprensión* (véase Dennett, 2017). Es importante señalar en este punto que el propio Turing mantuvo hasta 1936 especulaciones espiritualistas que en algún momento fundamentó en la mecánica cuántica y el principio de indeterminación de Heisenberg. Es a partir de esa fecha cuando se interesa por el materialismo y, especialmente, por el biologicismo y el evolucionismo. De tal forma que, cuando en 1947 regresa a Cambridge, estudia los últimos desarrollos en biología y neurociencia y comienza a pensar en el crecimiento de formas biológicas a partir de ecuaciones no lineales (puede verse al respecto su trabajo «La base química de la morfogénesis», de 1952), así como en el desarrollo de redes de neuronas artificiales a partir de algoritmos recursivos (véase Copeland, 2013: cáp. XI).

sencia del acto psíquico dirigido al objeto. Así las cosas, una máquina puede resolver problemas o dejarlos sin resolver, pero no puede sentir placer por sus éxitos, ni sentirse miserable por sus errores, ni estar enojada o deprimida cuando no consigue algo. Este punto de vista es calificado por Turing de solipsismo extremo, en virtud del cual solo se puede estar seguro de que una máquina piensa siendo la propia máquina, y solo se puede estar seguro de que un hombre piensa siendo el propio hombre (Turing, 1950/2004: 452).

Lo que hace Turing en el punto 6(4) es negar que la perspectiva de la primera persona sea la privilegiada para definir lo que es el pensamiento. De hecho, no es la que se adopta cuando nos dirigimos a otra persona y nos comunicamos con ella, atribuyéndole en todo momento pensamiento. Lo que solicita Turing es, sencillamente, no dejar de adoptar esta perspectiva en tercera persona cuando nos dirigimos a las máquinas o tratamos con ellas. Lo que hace que juzguemos que una persona piensa es el resultado de su acción, y este mismo criterio es el que debiéramos utilizar para juzgar si una máquina piensa. Ahora se entiende que «Computing Machinery» comience reemplazando la pregunta sobre si las máquinas piensan por el juego de la imitación: diremos que una máquina piensa si un humano, comunicándose por escrito con ella y con otro humano, es incapaz de distinguir entre ambos.

Efectivamente, no podemos asegurar que la máquina y el humano piensan de la misma manera, como no podemos asegurar que dos interlocutores humanos distintos piensan de la misma manera. Sin embargo, el hecho de alcanzar el mismo resultado y que ello no se produzca por casualidad, nos obliga a reconocer algún tipo de isomorfismo entre los pensamientos de dos personas distintas, o entre el funcionamiento de una máquina y el pensamiento de un humano. Recordemos que esto mismo es lo que pasó con Tunny. En 1942 Tutte consigue explicar desde Bletchley el comportamiento de Tunny sin tener conocimiento del funcionamiento de la máquina -con una tecnología muy superior a la de Enigma-, y a partir de este trabajo

Turing idea un método en papel para descifrar sus mensajes. Después de la guerra Turing y Flowers fueron enviados a Alemania para estudiar los sistemas criptológicos y de comunicaciones de su ejército. Un ingeniero alemán les habló de una máquina de cifrado de doce rotores que habían utilizado en la guerra y les enseñó una máquina Tunny. Al verla comenzaron a reírse; nunca habían tenido una ante ellos y, sin embargo, de ella lo sabían todo (Copeland, 2013: 144-145). De la misma manera, afirma Turing que no hace falta descubrir los secretos de la conciencia para entender su comportamiento e idear máquinas que piensen. Precisamente, serán las máquinas más resolutivas las que nos permitan una mejor comprensión del pensamiento humano estableciendo isomorfismos entre sus funcionamientos respectivos. Esto fue lo que llevó a Turing durante los últimos años de su vida a estudiar algoritmos recursivos biológicos y su implementación en máquinas.

12. Conclusión

Lo que he querido defender es que, desde que Turing se adentró en la nueva matemática formalista, primero a través de la conferencia de Newman al finalizar el trimestre de Cuaresma de 1935 y después asistiendo al curso sobre Gödel que impartió en el trimestre posterior, el objetivo de su trabajo siempre fue el de encontrar un procedimiento sistemático en matemáticas y, por extensión, en el conocimiento general, que no requiriese de la intuición ni de las «artes misteriosas» que se atribuyen al pensamiento humano. La definición que asume de lo que es un procedimiento sistemático es la que proviene del formalismo de Hilbert, que se desarrolla frente a las críticas de Frege y sobre el trasfondo de la filosofía kantiana. Toda esta carga metafísica es la que genera una particular lectura de los teoremas de incompletitud de Gödel del año 1931, según la cual ningún sistema axiomático formal consistente de la matemática clásica -que recurra a los *Principia*, a la aritmética formal de Peano, a la teoría axiomática de conjuntos o, en general, a lógicas de base finita- para la aritmética puede contener todas las verdades, de tal manera que siempre encontraremos enunciados aritméticos verdaderos por intuición sobre los que no podremos decidir si son o no demostrables. Si el objetivo era del de alcanzar las verdades por deducción prescindiendo así de la intuición, resulta que no podemos prescindir de ella.

En la citada conferencia, Newman entiende que los procedimientos sistemáticos son mecánicos, a saber, que siguen unas reglas dadas sin creatividad ni entendimiento. Es en ese momento cuando Turing piensa en dispositivos automáticos capaces de llevar a cabo todo tipo de procedimientos formales algorítmicos. Escribe «On Computable Numbers» precisamente para definir una máquina computadora que consiga superar el teorema de Gödel, solucionar el problema de la decisión y prescindir definitivamente de la intuición. Sin embargo, el trabajo acaba demostrando todo lo contrario, a saber, que ningún procedimiento algorítmico computable permite la solución del problema de la decisión. Con ello se extienden las limitaciones del teorema de Gödel desde los sistemas axiomáticos formales de la matemática clásica a la totalidad de lo efectivamente computable. Todavía sobre un trasfondo kantiano, la lectura que se hace -y Gödel el primero- es que Turing ha confirmado, a pesar del empeño de Hilbert y los formalistas, que la intuición y la mente humana nunca podrán ser sustituidas por una máquina.

Cuando dos años más tarde Turing escribe su tesis doctoral, sigue respondiendo al objetivo formalista de sustituir la intuición y, en general, el pensamiento, por sistemas axiomáticos formales, reformulando el objetivo de una manera más amplia al considerar los sistemas algorítmicos computacionales. El mismo Turing reconoce en el trabajo que tampoco en esta ocasión ha alcanzado su objetivo, porque en lugar de eliminar la intuición y hacer matemática desde la mecánica del ingenio, ha acabado eliminando el ingenio.

Lo que he pretendido mostrar es que la guerra no cambia el objetivo del trabajo de Turing, que vuelve a ser, formulado ahora en términos más generales, la sustitución del pensamiento humano por máquinas computacionales. Pero lo que sí cambia en ese periodo es la lectura que hace del teorema de Gödel y el sentido que da a sus investigaciones, de tal manera que la sustitución que antes no parecía posible acaba siéndolo. Dos son las causas que motivan el cambio. La primera tiene que ver con la construcción de máquinas computacionales cada vez más potentes capaces de resolver problemas que las anteriores, menos potentes, dejaban sin resolver. Esto hace que el teorema de Gödel ya no se lea como una confirmación de los límites de las máquinas frente a la capacidad ilimitada de la intuición y el pensamiento humano, sino como una estrategia para la superación de los límites de cada máquina particular. Tal estrategia es la que se pone en ejercicio con la construcción de las máquinas Bomba, Coloso y Coloso II. Ello tiene una importante deriva, y es que la metafísica de Kant inyectada por Frege deja de ser el marco de interpretación del teorema de Gödel y, de modo general, el terreno en el que se debate sobre filosofía de la matemática.

La segunda causa a la que me refiero es la construcción de Coloso para descifrar los mensajes de Tunny. De Tunny solo se conocían codificaciones, es decir, las consecuencias de su comportamiento. Y, sin embargo, a partir del estudio de las consecuencias, Tutte y Turing crean una máquina, Coloso, capaz de descifrar los mensajes de Tunny. Cuando en 1945 Turing viaja a Alemania y ve una máquina Tunny, poco le interesa ya su funcionamiento, debido a que Coloso tiene un comportamiento equivalente, aunque un funcionamiento diferente. Sin conocer cómo funciona Tunny, ni cuántos rotores tiene, ni de qué material está hecha, se ha construido una máquina, Coloso, capaz de sustituirla por completo. De la misma manera, acaba reconociendo Turing a la altura de 1950, para sustituir el pensamiento por una máquina no hace falta conocer cómo funciona el pensamiento, ni su base neuronal, ni saber si el cerebro tiene -en expresión del propio Turing- «la consistencia de las gachas frías» (Copeland, 2004: 495). Lo importante es tener un sistema computacional capaz de resolver los problemas que resuelve el pensamiento humano, o dicho de otra manera, una máquina capaz de simular su comportamiento. Cuando se argumenta que esta simulación no tiene nada que ver con el pensamiento humano que se puede describir en primera persona, Turing responde criticando el privilegio que se le concede a la perspectiva en primera persona e insistiendo en que la atribución de pensamiento siempre se hace en tercera persona, a partir del comportamiento observable, ya sea de humanos o de máquinas.

Por consiguiente, es la relectura del teorema de Gödel, que ahora se entiende como una estrategia para la construcción de sistemas computacionales cada vez más potentes, junto con una concepción del pensamiento humano desde la tercera persona, lo que lleva a Turing a partir de 1945 a hablar de la posibilidad de construir máquinas que simulen el pensamiento y nos eviten, ¡por fin!, recurrir a la intuición y a las «artes misteriosas» del conocimiento humano.

Bibliografía

Cawthorne, N. (2014): *Alan Turing: The Enigma Man*. Arcturus Publishing Ltd.

Copeland, B.J. (2004): *The Essential Turing. The ideas that gave birth to the computer age*. Oxford: Clarendon Press.

Copeland, B.J. (2013): *Alan Turing. El pionero de la era de la información*. Madrid: Turner.

Dennett, D. (2017): *From Bacteria to Bach and Back. The Evolution of Minds*. New York: W.W. Norton and Company.

Frege, G. (1873): «On a Geometrical Representation of Imaginary Forms in the Plane», en B. Mc Guinness (ed.) (1984), *Gottlob Frege: Collected Papers on Mathematics, Logic, and Philosophy*. New York: Basil Blackwell, pp. 1-55.

Frege, G. (1884): «Los fundamentos de la aritmética», en J. Mosterín (ed.) (1996), *Gottlob Frege: Escritos filosóficos*. Barcelona: Crítica, pp. 31-146.

Frege, G. (1892): «Sobre sentido y referencia», en G. Frege (1998), *Ensayos de semántica y filosofía de la lógica*. Madrid: Tecnos, pp. 84-111.

Frege, G. (1906): «Sobre los fundamentos de la geometría», en G. Frege (1996), *Escritos filosóficos*. Madrid: Crítica, pp. 279-334.

Frege, G. (1918): «El pensamiento: una investigación lógica», en G. Frege (1998), *Ensayos de semántica y filosofía de la lógica*. Madrid: Tecnos, pp. 196-225.

Frege, G. (1980): *Philosophical and Mathematical Correspondence*. Chicago: University of Chicago Press.

Gambra, J.M. (1995): «La filosofía de David Hilbert», *Thémata: Revista de Filosofía*, (14), pp. 147-179.

Giovannini, E.N. (2015): *David Hilbert y los fundamentos de la geometría (1891-1905)*, London: College Publications, Milton Keines.

Gödel, K. (1934): «On Undecidable Propositions of Formal Mathematical Systems», en M. Davis (ed.) (1965), *The Undecidable: Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions*. New York: Raven Press, 41-71.

Hilbert, D. (1927): «The Foundations of Mathematics», en J. van Heijenoort (ed.) (1967), *From Frege to Gödel: A source Book in Mathematical Logic, 1979-1931*. Boston: Harvard University Press.

Hilbert, D. (1930): «Probleme der Grundlegung der Mathematik», *Mathematische Annalen*, (102), pp. 1-9.

Hilbert, D. (1996): *Fundamentos de Geometría*. Madrid: CSIC.

Kant, I. (1996): *Principios formales del mundo sensible y del inteligible (Disertación de 1770)*. Madrid: CSIC.

Kant, I. (1997): *Crítica de la razón pura*. Madrid: Alfaguara.

Larson, E.J. (2022): *El mito de la inteligencia artificial. Por qué las máquinas no pueden pensar como nosotros lo hacemos*. Schackleton Books, S.L.

Mosterín, J. (2000): *Los lógicos*. Madrid: Espasa Calpe.

Newman, M. H. A. (1955): «Alan Mathison Turing, 1912-1954», *Biographical Memoirs of Fellows of the Royal Society*, vol. 1, issue 1, pp. 253-263.

Soare, R. (2009): «Turing oracle machines, online computing, and three displacements in computability theory», *Annals of Pure and Applied Logic*, vol. 169, 3, pp. 368-399.

Turing, A. (1937): «On Computable Numbers with an application to *Entscheidungsproblem*», *Proceedings of the London Mathematical Society*, s. 2, vol. 42, pp. 230-265.

Turing, A. (1938): *Systems of Logic based on Ordinals*. Princeton: Seeley G. Princeton, Mudd Manuscript Library.

Turing, A. (1950): «Computing Machinery and Intelligence» (*Mind*, 49, pp. 433-460), en B.J. Copeland (2004), *The Essential Turing*. Oxford: Clarendon Press, pp. 442-464.

Vargas, C. (2012): «Alan Turing: máquinas e inteligencia. En conmemoración de los 100 años de su nacimiento», *Aporía. Revista Internacional de Investigaciones Filosóficas*, (4), pp. 43-63.

Weyl, H. (1925): «Die heutige Erkenntnislage in der Mathematik», *Symposion*, (1), pp. 1-32.

Weyl, H. (1944): «David Hilbert and his Mathematical Work», *Bulletin of the American Mathematical Society*, (50), pp. 612-654.

Juan Antonio Valor Yébenes
jantonio.valor@filos.ucm.es