



El pensamiento lógico, ¿es sencillo o complejo?

Rational Thinking. Is it Easy or Hard?

■ John F. Nash, Jr.

Resumen

El texto se basa en la conferencia que di en el Décimo Congreso de la Asociación Mundial de Psiquiatría (WPA), celebrado en Madrid en el verano de 1996. Se resume aquí lo expuesto en dicha ocasión, junto con algunas preguntas que se hicieron al finalizar; pero sin embargo, no se incluye la mayor parte de lo que se comentó sobre mi "caso clínico" personal. En la reunión de Madrid no me dirigía al público o a los lectores en general, sino más bien a un grupo específico en un momento especial. En mis notas escritas la conferencia dada entonces no tenía ningún título concreto, salvo el de "Resumen de la conferencia de Madrid".

Palabras clave

Pensamiento lógico. Inteligencia Artificial. Teoría de juegos.

Abstract

The present text is based on the talk given by the author at the 10th Congress of the World Psychiatric Association (WPA) in Madrid in the summer of 1996. However, the text here is condensed from all that was said on that occasion (where there were also some questions afterwards); and specifically most of what was said about my "case history" is not included. At the occasion of the meeting in Madrid, I was not simply communicating to the public or to readers in general but to a par-

El autor nació en 1928 en Bluefield, Virginia Occidental (EE.UU.). Estudió en la Universidad de Princeton y se doctoró en 1950 con la tesis titulada *Non-cooperative Games*. Matemático intuitivo y genial, en 1952 se incorporó como profesor al MIT. En 1994 fue galardonado, junto a Harsanyi y Selten, con el Premio Nobel de Economía por su trabajo sobre teoría de juegos. En la actualidad, interesado en temas de lógica, teoría de juegos, cosmología y gravitación, trabaja en el Fine Hall de la Universidad de Princeton. Su vida inspiró a Sylvia Nasar el libro *Una mente prodigiosa* y la película *Una mente maravillosa* dirigida por Ron Howard en 2001. Este artículo se incluyó en el epítome: *One World, One Language- Paving the Way to Better Perspectives for Mental Health*. Proceedings of the X World Congress of Psychiatry. Hogrefe & Publishers, 1996; y se publica con las debidas autorizaciones. La traducción es de Fernando Prieto Pérez. El autor quiere hacer constar que algunas de las notas insertadas a pie de página fueron ampliadas ligeramente en Madrid por medio de la proyección de transparencias que incluían más detalles en lo referente a algún tema específico, como los casos de Nietzsche o Cantor.

ticipants of a group at a particular time. The talk given then had no specific title in my written notes except "Madrid Talk Outline".

Key words

Rational Thinking. Artificial intelligence. Game theory.

■ Introducción

He pensado que sería mejor no hablar sobre una teoría que yo entendiera perfectamente pero que no fuera de mucho interés para la audiencia, del mismo modo que tampoco lo sería exponer temas sobre los que yo no tuviera un gran conocimiento. Y me ha parecido más adecuado no hablar sobre mí mismo simplemente como un "caso clínico" en el campo de la enfermedad mental. Por ello, quiero comenzar desarrollando un tema que resulta especialmente difícil para la mente humana, como es el de la introspección y el estudio del proceso del pensamiento; o, en general, el entender la propia mente humana como un mecanismo que lleva a cabo funciones útiles y oportunas.

Mi propio conocimiento como un matemático que ha vivido la mayor parte de su vida en, o cerca de Princeton, Nueva Jersey, me lleva a presentar como ilustrativo a un grupo de personas ya pasadas a mejor vida, que hicieron célebres contribuciones en el campo de la "lógica matemática", pero que también parecieron mostrar (y esto NO es una observación científica) un grado excepcional de signos o síntomas de alteración mental (en el sentido de "locura").

Hasta cierto punto esto parece una paradoja, ya que uno podría decir que el estudio de la lógica lleva a un pensamiento o a una conducta ilógica. (Cuando hablé en 1996 no lo sabía, pero más tarde descubrí que el profesor G.-C. Rota —matemático del MIT, EE.UU.— había publicado un artículo en el que señalaba esa misma "coincidencia" respecto a una sorprendente frecuencia de "problemas mentales" entre los lógicos matemáticos.) Un primer ejemplo interesante, como veremos, es el de un famoso filósofo.

Hoy en día, muchos de los lógicos matemáticos encuentran su hueco académico en departamentos de filosofía, del mismo modo que expertos investigadores de la teoría de juegos encuentran su refugio en departamentos de economía.

Friedrich Nietzsche¹ fue un filósofo que, debido a su manifiesta brillantez, se incorporó como profesor en Basilea antes incluso de terminar su tesis doctoral. Más tarde, dejaría su

¹Friedrich Nietzsche: Profesor de Filología en la Universidad de Basilea. Dimitió y vivió (pobre) en Italia; autor de *Más allá del bien y del mal* y *Así hablaba Zaratustra*. Filósofo, escritor y poeta. Fue un gran adversario y crítico del cristianismo desde una perspectiva intelectual, más o menos en el sentido de "post-cristiano". Nietzsche era hijo de un pastor protestante. Su padre murió cuando Nietzsche tenía sólo cinco años. Señalé en Madrid que esas circunstancias eran importantes para explicar la posible tensión psicológica que podría haber dado lugar a los problemas mentales desarrollados tardíamente por Nietzsche.

cátedra de Basilea por motivos de salud y, aunque muy pobre, llegaría a tener más éxito como escritor. Sin embargo, en el fondo, y quizá influido por su idea del "superhombre", Nietzsche siguió desarrollando rasgos maníacos. Finalmente, fue internado en un asilo y allí murió. Nietzsche fue un estudioso de las ideas y de la orientación intelectual humana, no un lógico matemático, pero su caso es interesante. (No parece que haya una evidencia firme de la creencia popular de que Nietzsche fuera víctima de una sífilis terciaria.)

Cantor

La persona que ha dado a los matemáticos las ideas básicas de los infinitos, a través de las que, por ejemplo, podemos distinguir entre el infinito de los números y el infinito de los puntos matemáticos en el espacio, fue Georg Cantor².

Cantor nos proporcionó las notaciones de "aleph-cero" y "aleph-uno", en las que se utiliza una letra hebrea. Este uso del hebreo junto con su nombre podrían sugerir que Cantor era judío, aunque en realidad no lo fuera, sino que, como su mujer, fue un buen cristiano.

Las ideas de Cantor por un lado iban muy a la cabeza de su tiempo, y, por otro, dependían en gran medida del cuestionable rigor de sus argumentos, que levantaron una notable oposición por parte de algunos matemáticos. Cantor comenzó a sufrir depresiones (probablemente influido por tal oposición intelectual) y finalmente murió como un caso clínico psiquiátrico en el hospital de su universidad.

Descubrimientos negativos

En el campo de la Lógica Matemática, los años treinta fueron testigos de extraordinarios descubrimientos de carácter negativo; descubrimientos que mostraban las limitaciones de la lógica formalizada por sí misma. Whitehead y Russell habían publicado los *Principia Mathematica*³,

²Georg Cantor: Profesor alemán de matemáticas en el siglo XIX. Introdujo la teoría matemática de la comparación de los infinitos. Incorporó los símbolos "aleph-0", "aleph-1", etcétera, y conceptos como los de "número cardinal" y "número ordinal". En un principio fue oficialmente alumno de Kronecker; pero éste pasó a ser muy crítico con sus ideas y su trabajo. (Aquí reside otra posible clave psicológica. Los muchos enfrentamientos críticos con Cantor, pudieron influir en los estados depresivos de sus últimos años.)

³*Principia Mathematica* (PM): Se trata de un muy extenso libro escrito por A.N. Whitehead y Bertrand Russell. Los PM pretendían situar todos los razonamientos matemáticos con una sólida base dentro de la lógica simbólica. Los PM fueron ideados con el fin de evitar ciertas "paradojas" que surgían del razonamiento informal. Una de tales paradojas era la "paradoja de Russell", que aparecía con "la clase formada por todas las clases que no se contienen a sí mismas", (¿esa clase se contiene a sí misma?). Los PM fueron la diana específica de los trabajos de Goedel que le llevaron a su gran descubrimiento. Pero, la consecuencia de esos trabajos de Goedel llevaba inherente el resultado de que el gran objetivo (de los PM) había fracasado, y que todo el mundo del razonamiento matemático no se había reducido satisfactoriamente a unos razonables fundamentos lógicos.

texto que pretendía establecer unas bases en términos de elaboración de demostraciones a través de las que sería posible todo razonamiento matemático en una línea definitiva, precisa y lógica.

Nombres célebres contemporáneos en Princeton (Nueva Jersey, EE.UU.)

En esta relación pueden figurar Post, Goedel, Turing y Church.

Emil Post⁴ fue al City College en la ciudad de Nueva York, y estudió problemas de lógica a comienzos de los años veinte. Sus trabajos le permitieron cuestionar lo adecuado de la formulación de los *Principia Mathematica* de Whitehead y Russell, pero fracasó en su intento de obtener el avance crucial que alcanzaría Goedel alrededor de 1930.

Sin embargo, sus ideas fueron muy productivas en tanto que inspirarían más tarde el progreso en varias direcciones. Kurt Goedel demostró que los *Principia Mathematica* (y cualesquiera otros sistemas análogos) eran esencialmente *incompletos*. Ello significaba que existiría una verdad matemática expresable de hecho como una proposición verdadera sobre números, pero que no sería demostrable por medio de las reglas de los *Principia Mathematica* (u otros sistemas afines).

A mediados de los años treinta, A. M. Turing⁵ desarrolló un concepto preciso de computabilidad (o computabilidad efectiva) que se describía en términos de un modelo de máquina denominada "máquina de Turing", y mostraba la existencia de una "máquina universal de Turing" capaz de reemplazar a cualquier máquina particular de Turing. (Como un ordenador personal universal programado de tal forma que fuera capaz de funcionar como un determinado ordenador personal comercial, del tipo IBM, Apple, NeXT, etcétera.)

Esto permitió a Turing demostrar que no sería posible comprobar (o demostrar) computacionalmente una afirmación determinada, al contrario de lo que decían los *Principia Mathematica*. Pero esto le llevó a concluir que para alguna proposición no iba a poder ser demostrable tanto su veracidad como su falsedad y, por lo tanto, también concluía que el conjunto de axiomas dados para el sistema debía ser "incompleto", en el sentido de que no era suficiente para determinar correctamente la verdad o falsedad de todas las proposiciones candidatas a ser teoremas.

Turing pudo demostrar así que los *Principia Mathematica* eran por necesidad incompletos; lo que suponía una confirmación paralela al temprano descubrimiento de Goedel.

⁴Emil Post: Post desarrolló algunos conceptos lógicos sistemáticos que guardaban alguna relación con los *Principia Mathematica*. Comenzó a dudar de la validez absoluta de tales *Principia* como una reducción eficaz de la matemática a la lógica simbólica. Por tanto, pudo reclamar el haberse "anticipado" al avance de Goedel.

⁵A. M. Turing: Desarrolló en Inglaterra el concepto de "máquina de Turing". Poco después llegaría a Princeton y escribiría su tesis doctoral supervisada por el profesor Alonzo Church. Durante la Segunda Guerra Mundial Turing trabajó descifrando los códigos alemanes "Enigma".

Por su parte, el profesor Church, en la Universidad de Princeton, desarrolló el "cálculo lambda", gracias al que consiguió resultados también en esencia paralelos a los de Turing.

Características psiquiátricas de Post, Goedel, Turing y Church

Lugar: A principios de los años veinte Post tenía una beca de estudios en Princeton. Goedel llegó a esa ciudad a comienzos de los treinta; después volvió a Austria, para regresar de nuevo a Princeton antes de la guerra. A su vez, Turing llegó a Princeton a mediados de los años treinta, y terminó su tesis doctoral con Church como asesor.

Post mostró claros síntomas maniaco-depresivos y fue diagnosticado de tal enfermedad. Su familia y los médicos se mostraron a favor del tratamiento con electroshock, pero, tras un período de enfermedad y hospitalización, dijo que tenía que volver a aprender su matemática. Finalmente, Post falleció en 1954 de un ataque cardíaco después de un tratamiento con electroshock.

Goedel comenzó a sufrir períodos de depresión en los años treinta, ya después de sus tempranos y grandes descubrimientos. Tales períodos interfirieron en su trabajo y en sus planes de viajes. Ya en sus últimos años, cuando ambos estaban en el IAS (Institute for Advanced Study) de Princeton, Goedel pareció ser la única persona con la que intelectualmente Einstein podía hablar de igual a igual (presumiblemente lo harían en alemán). Goedel contribuyó también a la teoría de la relatividad, demostrando la existencia teórica de universos con "tiempo circular".

Goedel se perturbó definitivamente cuando su esposa y él empezaron a sufrir problemas de salud a la vez. De tal forma, manifestó en el hospital de Princeton su miedo a la comida, que creía envenenada, dejando de comer con normalidad. De hecho, por su pensamiento irracional llegó a ayunar hasta morir.

Turing se suicidó en 1954 después de haber tenido la mala experiencia de caer en manos de la policía y ser sentenciado posteriormente a unas condiciones de libertad bajo palabra muy desagradables. Pensando de forma lógica podía haber tenido varias opciones razonables. Pero, aunque la idea de que el suicidio es un acto irracional, a veces también puede ser un poderoso argumento en el sentido contrario.

Redes neuronales e inteligencia artificial

Algunos años atrás tuve cierta relación con estas áreas, gracias a un estrecho contacto personal en los primeros tiempos con Marvin Minsky y John McCarthy, que eran estudiantes de la licenciatura de matemáticas en la Universidad de Princeton, en una época en la que yo también era estudiante allí.

Uno puede considerar que este tema o campo de estudio se ocupa del "problema de la introspección", que se presenta cuando los hombres intentan mirar en su interior y analizar el auténtico funcionamiento del cerebro humano.

McCarthy lideró el desarrollo del lenguaje de computación denominado "LISP" y, asimismo, acuñó el término "inteligencia artificial" para expresar la idea de la "inteligencia" desarrollada por las computadoras. Minsky estudió el "perceptrón", un elemento utilizable en una red neuronal, y también escribió el libro *The Society of Mind*, en el que teoriza sobre la comprensión efectiva de la función mental llevada a cabo por la red neuronal que forma el cerebro humano.

Estas ideas estaban íntimamente relacionadas con los recientes enfrentamientos entre el hombre y la máquina en el juego de ajedrez, concretamente Kasparov (el hombre) contra Deep Blue (la computadora).

Cuando, en 1996, hablé en Madrid en el Congreso de la Asociación Mundial de Psiquiatría, me referí a ese hecho, en un momento en el que sólo se había disputado el primer enfrentamiento, ganado por Kasparov, no sin algunas partidas perdidas y otras agotadoras acabadas en tablas. Ahora sabemos que en el segundo enfrentamiento, en mayo de 1997, el ganador fue ¡Deep Blue!

Esto hace especialmente relevante el concepto clásico del test de Turing sobre el éxito de la inteligencia artificial a la hora de imitar a la inteligencia humana. La idea era que, bajo unas condiciones análogas a aquéllas mediante las que dos humanos podrían comunicarse por telégrafo (o por correo electrónico, que es lo mismo), una máquina inteligente sería indistinguible de un humano en lo referente a manifestaciones propias de una conversación, de preguntas y respuestas transmitidas por medio de la vía telegráfica. Y precisamente eso fue lo que sucedió cuando en el segundo enfrentamiento, incluso para Kasparov, ¡Deep Blue parecía jugar como un gran maestro humano extraordinariamente fuerte!

Tanto McCarthy como Minsky eran estudiantes que se graduaron en matemáticas en Princeton alrededor de 1950, y ambos fueron más tarde al MIT cuando yo estaba allí. En sus posteriores carreras universitarias se movieron en otros niveles, dentro del departamento, diferentes de aquéllos que explícitamente consideraban al catedrático como un matemático.

La "teoría de la red neuronal" es considerada en la actualidad como un campo especial dentro de las matemáticas, a falta de otra ubicación más adecuada; una ubicación que, por supuesto, puede cambiar en el futuro. La teoría de juegos fue vista originalmente como un área especial del estudio matemático, pero hoy en día la mayor parte de aquéllos que son eruditos "teóricos del juego", encuentran su refugio académico en otros departamentos no matemáticos.

Yo mismo escribí un artículo sobre la idea de la computación en paralelo, que no se llegó a publicar completo. Se titulaba *Parallel Control*, y en él presentaba el término "máquina pensante" para describir un tipo de ordenador que utilizara un potente sistema en paralelo, y con un tipo de "red neuronal" en su arquitectura que pudiera desarrollar funciones comparables al pensamiento humano. (Hay que decir aquí que Deep Blue utilizaba un poderoso sistema en paralelo para luchar con la auténtica red neuronal humana del cerebro de Kasparov; y esos

procesadores en paralelo usados por Deep Blue en 1997 eran exactamente el doble de potentes que los empleados en 1996.)

Es una coincidencia que mucho más tarde (aquel trabajo lo escribí en 1954), apareciera una compañía de ordenadores llamada Thinking Machines, que fabricaba ordenadores capaces de trabajar en paralelo de forma extraordinariamente coordinada. Dicha compañía fue fundada por D. Hillis, que era un alumno de doctorado de Marvin Minsky. Desgraciadamente, Thinking Machines se estancó después de que los grandes productores de ordenadores empezaran a competir en la estructura de los ordenadores en paralelo. (Y, por supuesto, un ejemplo de esto es el desarrollo por IBM del mismo tipo de componentes utilizados en la construcción del *hardware* de Deep Blue.)

Mis tempranas llamadas a la racionalidad como un ideal

Antes de hacer algunas observaciones sobre mi historia intelectual y psicológica, expuse en Madrid la idea de si los humanos podían, o no, llegar a ser más lógicos si hicieran esfuerzos específicos en esa dirección.

En el programa de televisión *Star Trek*, en EE.UU., había un personaje llamado Spock que representaba a un vulcaniano; pertenecía, por tanto, a una de las especies de humanoides extraterrestres que poseían una fanática obsesión por la racionalidad y el pensamiento lógico. Por supuesto, los diferentes personajes no humanos de *Star Trek* estaban representados por actores humanos, al igual que Spock.

En relación con la enfermedad mental uno se podría preguntar, en el caso de que existiera una sociedad de humanos que diera una prioridad especial a lo racional y al pensamiento lógico, si dicha cultura sería capaz, o no, de reducir la incidencia de la locura y de los estados alucinatorios, que son característicos de la enfermedad mental y tradicionalmente descritos como locura, y que poseen implicaciones legales con respecto a las acciones humanas trascendentes.

Como alguien que sufrió una fase de "locura", o de pensamiento delirante e irreal, yo mismo también soy una persona que en algunos momentos tempranos de mi carrera científica ha prestado atención al concepto del comportamiento "idealmente racional". Este ideal fue comentado en mis trabajos sobre la *teoría de juegos*, tanto para la *teoría de los juegos no cooperativos con cualquier número de jugadores*, como para la *teoría de los juegos cooperativos de dos jugadores*, incluyendo "problemas de negociación".

La teoría no dependía por completo de conceptos lógicos, pero estos se introducían en ella de una forma muy natural. Y en el amplio contexto de la teoría económica en general se ha desarrollado la idea del *Homo oeconomicus*. Éste es un concepto de hombre visto esencialmente como si fuera un hombre de negocios movido exclusivamente por lo material, el beneficio o la pérdida, consideraciones de tipo práctico para un hombre de nego-

cios preocupado sólo por su negocio. Y podemos observar que en 1995 el Premio Nobel de Economía le fue otorgado al profesor Robert Lucas por su trabajo sobre "expectativas lógicas".

Racionalidad → Irracionalidad

Los lógicos matemáticos han estudiado problemas sobre cómo los humanos pueden razonar de una forma precisa, lógica y sistemática con respecto a las tareas de demostraciones matemáticas y sistemas adecuados de axiomas. Algunos de esos lógicos enloquecerían más tarde. En mi caso, yo había estudiado algo de lógica durante mi período de licenciatura en Princeton, y había tratado en profundidad el tema de la racionalidad (en paralelo con R. Lucas), desarrollando conceptos teóricos del juego que serían aplicables al campo del *Homo oeconomicus*. A principios de 1959 caí en la irracionalidad, en el pensamiento delirante, y presenté pronto mi dimisión como profesor del MIT. Fui entonces hospitalizado involuntariamente y acusado de actuar como un demente... (En Madrid, en el Décimo Congreso de la Asociación Mundial de Psiquiatría, hablé sobre mis ideas incluidas en este escrito, y también me referí con detalle a mi caso personal durante aquella etapa de mi vida del ingreso en el hospital, la psicoterapia, etcétera; y el patrón general de los delirios interpretados sin ninguna base o evidencia objetiva. Pero me presenté personalmente en la reunión de psiquiatras de todo el mundo y me pareció del todo adecuado profundizar en este tipo de detalles. Sin embargo, ahora, al redactar esta versión de lo que expuse en Madrid para que sea publicado y quede registrado, me gustaría ser más discreto y facilitar menos información sobre los detalles de mi "caso clínico" personal.)

Preguntas, debate

Con este título encabezé la última parte del resumen de mi conferencia en Madrid. Hubo allí algunas preguntas y una o dos tuvieron el valor de servir de "preguntas de fondo", que me permitieron extenderme en un tema: cómo evaluar la calidad de la recuperación de la enfermedad mental. Era una pregunta lógica. Si una persona tuviera tuberculosis y fuera enviada al hospital de la *Montaña mágica* para un tiempo de tratamiento, y después fuera dada de alta, ¿en qué medida estaría sana esa persona?

Respondí que con respecto a la enfermedad mental parecía haber una gran tendencia a aceptar niveles de recuperación bastante bajos. Si una persona sencillamente se vuelve manejable (tal vez mediante prescripciones farmacológicas) fuera de los hospitales mentales propiamente dichos, en tal caso podría considerarse como un caso de recuperación. Puse como ejemplo que si Mozart, que había enloquecido a los 36 años, y en vez de morir joven hubiera

recuperado la cordura a los 50 años, pero sin ninguna capacidad posterior para componer música, ¿podríamos considerar entonces el caso como una auténtica recuperación de la enfermedad mental?

Sin embargo, esas líneas encerraban más enjundia. Yo podía haber mencionado la idea de que lo que a menudo convencionalmente se consideran criterios de recuperación de la enfermedad mental, en el fondo podrían tener una base económica. Si el "caso" llega a convertirse en barato para la sociedad y/o la familia como para poder manejarlo, será entonces clasificado como recuperación, aun incluso si alguna capacidad o talento especial del individuo pudiera haberse perdido, y el estilo de vida de la persona se hubiera visto transformado en el sentido de una mayor limitación como, por ejemplo, la forma de vida de un monje o de una monja.

Quizá hubo también allí un cierto debate sobre las perspectivas de la evolución de la psicoterapia con los nuevos medicamentos que están apareciendo constantemente, y con el nuevo conocimiento científico que sigue desarrollándose acerca del cerebro, la psicología y las funciones cerebrales.

Por lo tanto, y aunque el proceso real del tratamiento de los casos de enfermedad mental esté en gran medida influido por la "tradicición", como lo fue ayer, las costumbres han evolucionado y racionalmente podemos prever grandes e impresionantes cambios en el futuro.

Una idea interesante en la IA (Inteligencia Artificial) es la de un "cyber-cuidador", que podría tratarse de un programa de ordenador mediante el que los enfermos mentales trabajarían en su propio beneficio psicoterapéutico. Si el programa fuera lo suficientemente bueno como para superar un adecuado test de Turing, al que lo usase le podría parecer entonces como si estuviera trabajando por medio de un videoteléfono con un auténtico terapeuta humano. Y, por supuesto, aquí son interesantes y relevantes las consideraciones económicas, al igual que lo son las consideraciones de en qué medida los usuarios (los pacientes) podrían estar, o no, satisfechos de sus sesiones con un "cyber-cuidador".