ἔλεγχος
Volumen II, N° 1
Septiembre 2018

HUBERT DREYFUS (1929-2017), RAYMOND SMULLYAN (1919-2017) Y EL SUEÑO DE LEIBNIZ (1646-1716)

Luis Camacho
Universidad de Costa Rica
lcn20032003@yahoo.com

Abstract. Two philosophers recently deceased, Hubert L. Dreyfus and Raymond Smullyan, deal with G.W. Leibniz´s ideas on formal systems and logical machines. Dreyfus acknowledges Leibniz´s importance in what he calls the long process of reduction of thinking to mechanical calculation by means of formal algorithms and Smullyan demonstrates the impossibility of devising a machine able to solve all mathematical problems, let alone all philosophical debates, despite Leibniz´s high expectations. Our contention is that Dreyfus´s vision of Leibniz´s influence can be complemented with the fact that Leibniz himself abandoned his wild expectations of youth on the possibilities of the *ars combinatoria* and *characteristica universalis*. Smullyan´s use of the theorems of incompleteness of formal systems cannot be refuted, but the undecidable proposition that fails to be proved or disproved in a formal system turns out to be rather trivial for the purposes of the leibnitian projects. However, we argue that it is important to note Leibniz´s influence on Gödel proof of the incompleteness theorem.

Key words: artificial intelligence, *characteristica universalis*, formal systems, Gödel´s Theorem, Leibniz.

Resumen. Dos filósofos recientemente fallecidos, Hubert L. Dreyfus y Raymond Smullyan, analizan las ideas de G.W. Leibniz sobre sistemas formales y máquinas lógicas. Dreyfus reconoce la importancia fundamental de Leibniz en lo que llama el largo proceso de reducción del pensamiento al cálculo mecánico mediante sistemas formales y Smullyan demuestra la imposibilidad de construir una máquina que resuelva todos los problemas matemáticos, y con mayor razón los debates filosóficos. La visión de Dreyfus sobre los proyectos de Leibniz se puede completar con el hecho de que Leibniz abandonó en su madurez las exageradas pretensiones de juventud. La incompletitud de sistemas formales utilizada por Smullyan y que se remonta a Gödel se considera irrefutable, pero la proposición no decidible que no puede ser probada ni refutada en un sistema formal resulta ser trivial para los propósitos de los proyectos leibnicianos. Sin embargo, es importante señalar la influencia de Leibniz en la prueba del teorema de incompletitud de Gödel.

Palabras clave: *characteristica universalis,* inteligencia artificial, sistemas formales, Leibniz, Teorema de Gödel.

1. Pérdidas recientes y conmemoración de muerte lejana

Las propuestas sobre procesamiento mecánico del conocimiento y el método aritmético para lograrlo aparecen por primera vez en G.W. Leibniz , cuya muerte hace poco más de 300 años se conmemoró en gran escala en el X Congreso Leibniz del 18 al 25 de julio de 2016 en Hannover, Alemania.[1] Muestra del gran interés que despiertan el pensamiento y las obras del famoso bibliotecario de esa ciudad en la Baja Sajonia, dicho congreso congregó a centenares de personas de 32 países. Solo los trabajos presentados, más de 300 sin incluir las conferencias, llenan cinco gruesos volúmenes publicados por la Olms Verlag.[2]

Un hecho fuera de lo común en congresos de filosofía fue la exhibición en este evento de aparatos construidos o diseñados por Leibniz: su famosa máquina de calcular, el molino de viento mejorado con un invento suyo y con el que intentó resolver el problema de la anegación de las minas del Harz , reconstruido por Stefan Kirschner (Li 2016, V, 413-426) y a la vista de todos durante el Congreso, así como una máquina encriptadora que fue construida recientemente por Erwin Stein con sugerencias de Nicholas Rescher, siguiendo el

1 La convocatoria para el X Congreso fue hecha por la Leibniz Gesellchaft y sociedades nacionales o regionales de América del Norte, España, Japón, Italia, Francia,Israel, Iberoamérica (Portugal, España, México, Costa Rica, Colombia, Brasil, Chile, Argentina), Rumanía y China. En el Congreso hubo, además, participantes de Polonia, Ucrania, Bulgaria, Hungría, República Checa, Suiza, Rusia, Canadá, Finlandia, Bélgica, Reino Unido, India, Austria, Lituania y Tunisia.

2 "Für Unser Glück oder das Glück Anderer", *Vörtrage des X.Internationalen Leibniz-Kongresses*, herausgegeben von Wenchao Li in Verbindung mit Ute Backmann, Sven Erdner, Esther-Maria Errulat, Jürgen Herbst, Helena Iwasinski und Simona Noreik. Hildesheim, Georg Olms Verlag, 2016.

diseño de Leibniz (Li 2016,V, 361-376). No puedo recordar ningún otro congreso de filosofía entre máquinas variadas, pero tampoco puedo recordar ningún filósofo que haya tratado de resolver problemas de ingeniería o que haya hecho propuestas para la solución mecánica de problemas teóricos. Las máquinas, la lógica y la epistemología conectan a Leibniz con una discusión que agita a filósofos , matemáticos e informáticos desde hace más de medio siglo y de la cual no se percibe el fin, aunque se pueden distinguir distintas etapas con ganadores y perdedores en el debate. En un nivel superficial la discusión tiene que ver con la posibilidad de producir máquinas capaces de pensar en el mismo sentido en que lo hacemos los seres humanos. En un nivel más profundo, se trata de la concepción misma del conocimiento y de la naturaleza del ser humano.

El tema de Dreyfus es la inteligencia artificial mientras el de Smullyan son los alcances de máquinas lógico-matemáticas, y en ambos casos se trata de los límites de sistemas formales y, por tanto, de lo que se supone es la propuesta leibniciana , aunque –como hemos señalado en otro lugar (Camacho 2016a) y mencionaremos más adelante– fueron varios y con resultados diferentes los proyectos defendidos por Leibniz. Las limitaciones que señala Dreyfus son externas al sistema, pues se basan en una concepción general de la naturaleza del pensamiento a su juicio inadecuada y que explica por qué no se cumplieron los pronósticos sobre la inteligencia artificial. En cambio, las limitaciones al proyecto señaladas por Smullyan son internas: una máquina lógica que genere respuestas a todos los problemas es imposible.

La palabra "sueño" en el título de este trabajo tiene que ver con la última sección de la parte final de la obra de Smullyan *La dama o el tigre* (1985)[3], la cual consiste en unas pocas páginas con el título "El sueño de Leibniz" y donde se expone por qué es imposible construir una máquina que resuelva todos los problemas de las matemáticas y –por extensión– las discusiones en filosofía, que Leibniz quería zanjar por medios mecánicos según la fórmula "pensar=calcular". En la última parte de la obra Smullyan pasa de los acertijos a las máquinas y hace la prudente advertencia de que "en rigor no se puede demos-

3 En inglés *The lady or the tiger? and other logical puzzles*.

trar o refutar la viabilidad de la esperanza de Leibniz porque no fue enunciada de una forma exacta" (268). Está claro, sin embargo, que en una etapa de su vida el gran filósofo y matemático esperó que su *characteristica universalis* resolviera los conflictos que azotaban el mundo de su tiempo. De la concepción leibniciana se sigue que resolver un problema matemático o de otra índole equivale a generar un número a partir de reglas. En términos también leibnicianos, quiere decir que una máquina programada con la característica universal podría generar cualquier número según reglas explícitas. Puesto que la máquina más avanzada en esta dirección que conoció Leibniz fue la calculadora que construyó, es difícil imaginar qué tenía en mente.

Sí sabemos que las máquinas tienen límites. Los teoremas de la incompletitud señalan que para cualquier sistema formal correcto F existe una proposición P que no se puede probar en F, aunque intuitivamente sepamos que es verdadera. Aunque a veces se arguye que este y otros teoremas posteriores prueban que ningún sistema formal es consistente si es completo y viceversa –y está de moda extender esta interpretación a todo cuanto conocemos– el alcance original del Teorema de Gödel de 1931 tiene que ver con la completitud, no con la consistencia.

2. Axiomatización, formalización e inteligencia artificial

Puesto que hablaremos de los límites de la formalización, conviene precisar la diferencia entre *axiomatización* y *formalización*. En numerosos libros y artículos estos términos se confunden; la formalización puede ser un instrumento de la axiomatización, pero esta puede darse sin formalización.[4] En la axiomatización se parte de proposiciones sin definir que generalmente se aceptan como obviamente verdaderas sin necesidad de prueba, y a partir de ellas se derivan los teoremas. El trabajo se lleva a cabo en la lengua natural de quien emprende la axiomatización y obviamente se tiene en cuenta el significado de los términos. El ejemplo clásico de axiomatización es el de Euclides en sus *Elementos.* Históricamente la formalización viene después y presupone la construcción de lenguajes artificiales

4 Una cuidadosa explicación de la diferencia entre los dos conceptos se puede encontrar en van Heijenoort (1967, 349).

en los que cada símbolo se introduce con reglas precisas para su uso. Primero se establece un vocabulario con los símbolos que se van a usar, luego se determinan las leyes de combinación de estos de manera que se pueda distinguir entre frases bien formadas y mal formadas, y luego se precisan las reglas de manipulación de símbolos que permitan generar cadenas aceptables diferentes de las no aceptables. La posible interpretación del cálculo así creado es un asunto diferente.

Muy importante para la formalización es el uso de algoritmos, es decir, procedimientos fijos claramente establecidos que se pueden repetir cuantas veces se quiera o se necesite. La noción de *algoritmo* es muy amplia, pues obviamente incluye tanto fórmulas matemáticas (v.g. la que usamos para resolver ecuaciones cuadráticas) como recetas de cocina o instrucciones precisas para lograr que un aparato funcione. En el caso de la lógica y de las matemáticas, los algoritmos nos dicen cómo proceder cuando deseamos obtener algo a partir de una serie de símbolos . Lo que obtenemos es también un símbolo o un conjunto de símbolos, que pueden ser numéricos. La noción de algoritmo es muy anterior a las computadoras, pero la condición de ser computable se incorpora rutinariamente como parte de la definición desde hace por lo menos medio siglo. Ya en 1977 encontramos un ejemplo de este cambio. "An algorithm is a set of rules for getting a specific output from a specific input. Each step must be so precisely defined (that) it can be translated into computer language and executed by machine" dice Donald E. Knuth en un artículo de *Scientific American* (1977,63).

3. Dreyfus y los límites de la inteligencia artificial

Hubert Lederer Dreyfus nació en 1929, estudió filosofía en Harvard y pasó la mayor parte de su vida como profesor de esa materia en la Universidad de Berkeley, California. Su interés en la fenomenología de Husserl y Heidegger no hubiera resultado relevante para el enfoque de la computación e informática conocida en nuestros días como "inteligencia artificial" (IA) si no fuera por sus libros de 1965 *Alchemy and Artificial Intelligence* y 1972, *What Computers Can´t Do* (con la palabra "still" añadida a partir de la tercera edición).

Las ideas de Dreyfus, aunque tuvieran su origen en autores como Heidegger y Merleau-Ponty, son fáciles de entender y de aceptar para cualquier lector alarmado por algunas de las afirmaciones que se hacían hace varias décadas sobre las posibilidades de la computación: el programa de IA presupone que la inteligencia humana se caracteriza por la manipulación formal de símbolos (es decir, según reglas que se pueden explicitar), pero no es esto lo que ocurre en la vida cotidiana de los seres humanos, como enfatizan los autores en que se basa Dreyfus. Por ejemplo, los movimientos que hacemos al andar en bicicleta pueden ser descritos por un algoritmo que quizá se pueda usar para lograr que un robot consiga hacerlo, pero si un ser humano intenta aprender a andar en bicicleta aplicando conscientemente las reglas formales (en este caso ecuaciones) que describen el conjunto de movimientos para lograr el equilibrio, probablemente no lo logrará. Los procedimientos mentales que seguimos para aprender una habilidad práctica pertenecen a una esfera de la experiencia humana que no se puede reducir a la descripción teórica en términos de fuerzas, equilibrio y movimiento, como se hace en un texto de física.

Más en general, para distinguir en la vida cotidiana entre los pasos que conducen a un fin buscado y los que no lo logran no hace falta tener una teoría formal. Contra lo que afirman algunos filósofos, no es necesario estar en posesión de la teoría correcta para distinguir entre verdad y falsedad, justicia e injusticia, éxito y fracaso. Al no tener en cuenta la diferencia entre la experiencia cotidiana inmediata y la formalización de las acciones, y pretender conseguir efectos de un nivel mediante procesos de otro nivel, el programa de los partidarios de la IA según Dreyfus estaba condenado al fracaso si pretendía ir más allá de juegos triviales y prueba de teoremas. En una época en que la investigación en IA se apuntaba éxito tras éxito, semejante profecía parecía condenada al fracaso.

Hasta donde sabemos, Dreyfus no tuvo problemas con la publicación de su libro, pero este fue ampliamente rechazado apenas publicado. En el recuento de lo ocurrido a partir de la primera edición de su obra, Dreyfus señala cómo incluso se trató de impedir que publicara artículos que criticaban el modo de pensar imperante en esa época.[5] En nada ayudó al debate el tono a veces demasiado seguro de sí mismo con que Dreyfus expresó sus críticas ; el pronóstico que se le atribuyó de que una computadora nunca ganaría una partida de ajedrez con un ser humano fue refutado definitivamente el 11 de mayo de 1997 cuando la computadora IBM llamada "Deep Blue" ganó una partida al campeón mundial Gary Kasparov, después de una larga serie de partidas desde el año anterior.[6]

El mundo pareció olvidarse de *What Computers Can´t Do* por algún tiempo después de 1972, pero mientras tanto se acumulaban problemas en la marcha triunfal de la IA. La introducción a la segunda edición del libro hace un recuento de los puntos en los que los partidarios de la IA habían tenido que hacer correcciones en la dirección señalada por Dreyfus. [7] Cuando la editorial del MIT publicó una tercera edición del libro en 1992, un pequeño cambio en el título recoge la insistencia en haber tenido la razón todo el tiempo: *What Computers Still Can´t Do*. Esta nueva introducción[8] reclamaba el triunfo después del rechazo inicial: las pretensiones y expectativas de los partidarios de la IA se fueron reduciendo, y aunque muchos años después tengamos hasta automóviles que no necesitan conductor y pequeños teléfonos inteligentes con acceso a inmensas cantidades de datos y con gran cantidad de ingeniosas aplicaciones, numerosas afirmaciones que se oían en los últimos años del siglo pasado nos suenan ahora innecesariamente exageradas o ingenuas.

De modo que Dreyfus no estaba tan equivocado después de todo. Pero, como en otros muchos casos en la historia de la ciencia y de la tecnología, los dos bandos tenían parcialmente la razón: la IA nos ha dado resultados que incluso podrían sorprender a sus primeros promotores, aunque no de la manera que ellos defendieron. Las crí-

5 Dreyfus (1994, 307). Dreyfus recibió una invitación para reseñar libros de Minsky, Papert y Winston en la revista *Creative Computing,* pero entonces Papert amenazó con represalias por parte de la MIT Press si la revista publicaba la reseña.

6 El relato histórico aparece en Adaime (2011).

7 Dreyfus sigue la siguiente línea de razonamiento: sus estudios de Heidegger, Merleau-Ponty y Wittgenstein le permiten predecir el fracaso de la IA. Pero también se podría predecir lo mismo a partir de otras posiciones y el éxito de la predicción sobre el fracaso no implica que las ideas de sus autores queden confirmadas.

8 1994,ix-lii.

ticas de Dreyfus ayudaron a cambiar el rumbo de los años triunfales que siguieron a un simposio que tuvo lugar en Darmouth en el verano de 1956 en el que estuvieron presentes figuras de primera línea como Claude Shannon, Marvin Minsky y Allen Newell, donde por primera vez se habló de inteligencia artificial y de donde salieron los lineamientos para la creación de programas que permitirían a las máquinas realizar operaciones como el reconocimiento de símbolos y, en particular, el seguimiento de instrucciones mediante la manipulación de dichos símbolos.

Los rápidos avances que tuvieron lugar en los años siguientes motivaron a los fanáticos de la inteligencia artificial a hacer afirmaciones que parecían obviamente falsas a muchas otras personas. Así por ejemplo, a la objeción de que las computadoras no son capaces de emociones se respondía con frecuencia que estas no tienen nada que ver con la inteligencia y que, en todo caso, su ausencia en máquinas era una ventaja sobre los seres humanos. La insistencia en atribuir a las máquinas actos y estados intencionales o actitudes proposicionales en el mismo sentido en que se atribuyen a los seres humanos, tales como darse cuenta, tener conciencia, creer, saber, pensar, recordar, imaginar, inventar y conocer, generó reacciones en contrario que Dreyfus supo sistematizar en su obra mejor que otros. En particular, la idea de que no hay diferencia entre máquinas y seres humanos porque unos y otros pueden realizar algunas operaciones semejantes lleva a conclusiones alarmantes que muchos no aceptarían de buena gana: las computadoras también tendrían derechos y deberes, exigencias y responsabilidad, méritos y castigos.[9] Hace unos veinticinco años en una conferencia dada por un partidario de la IA escuché decir que debíamos acostumbrarnos a la presencia cotidiana en nuestras casas de otros seres inteligentes, y no se refería a otras personas ni a mascotas. Si bien Turing fue muy ingenioso cuando propuso el famoso test que lleva su nombre, conceder inteligencia a una máquina que lo apruebe pareció y sigue pareciendo exagerado a muchos, aunque no puedan explicar por qué.

9 A no ser –por supuesto– que tomemos el camino fácil de negar que los seres humanos tengan dichas atribuciones.

Sesenta años después del Simposio de Darmouth la retórica actual sobre la computación e informática es mucho más cautelosa, a pesar de que ha habido numerosos avances inesperados y asombrosos. El panorama en este año de la muerte de Dreyfus es entonces paradójico: hay programas de cómputo que pueden vencer al mejor jugador humano de ajedrez, pero muchas de las predicciones hechas en aquella primera y lejana etapa cuando Dreyfus inició su campaña contra las pretensiones exageradas siguen sin cumplirse o solo se cumplieron a medias, como la traducción simultánea de un idioma a otro o la corrección automática de un texto con contenido fáctico. Cualquiera que utilice medios sociales de comunicación electrónica en nuestros días puede atestiguar las irritantes interpretaciones equivocadas de los correctores y las frecuentes traducciones erróneas de los traductores automáticos. En muchas ocasiones más que ayuda son molestias. Dominar un lenguaje natural supone ser capaces de concretar el contexto de las expresiones y el uso cotidiano de las palabras. Aprender bien alguno de los aproximadamente seis mil idiomas naturales que se hablan actualmente supone conocer la cultura de quienes lo hablan; esto no ocurre cuando se aprenden lenguajes libres de contexto como los de las matemáticas o el esperanto. Estas diferencias y dificultades aparecen ampliamente discutidas en el texto original de Dreyfus, todavía interesante a pesar de los muchos años transcurridos.[10]

Según *What Computers Can´t Do* los que anticiparon el procesamiento mecánico del pensamiento no se percataron de las limitaciones inherentes al proyecto. Entre los autores más citados en esta categoría se encuentra Leibniz, de quien se dice en la primera cita que encontramos (Dreyfus 1994, 65) que para él todas las percepciones y prácticas requeridas por la inteligencia humana pueden ser representadas simbólicamente (Wiener (en adelante W) 18)[11], sin tener en cuenta que dichas percepciones y prácticas están situadas, es decir, ubicadas en las condiciones del ser humano concreto que percibe y

10 La discusión sobre la posibilidad de traducción simultánea de lenguajes naturales ocupa muchas páginas en la obra de Dreyfus.

11 Aunque hay una edición en español de los textos relativos a la característica, lengua universal y lógica (Velarde-Cabañas 2013), no aparecen en este volumen los textos de 1677 citados por Dreyfus, quien toma todas sus citas de Wiener (1951).

piensa. Sin el transfondo social e histórico de cada expresión resulta inadecuado aislarla y captar el trasfondo es una tarea que supera los alcances de una máquina basada en reglas formales, aunque esta pueda deslumbrarnos con otras tareas que no requieren percibir las condiciones.

En (69-70) la representación simbólica defendida por Leibniz se conecta con su sistema binario y con su "característica universal", en la que se asigna un número primo a cada uno de los conceptos básicos, de modo que mediante la multiplicación de números primos se puedan representar conceptos complejos y mediante la división de un número compuesto en sus primos se pueden recuperar los conceptos primitivos. Los textos leibnicianos citados o mencionados en estas páginas provienen del ensayo de 1677 "Hacia una característica universal" (W 17-25), uno de los más relevantes para este tema. La típica perspicacia de Leibniz, que lo hace tan interesante en nuestros días por sus curiosas anticipaciones, se muestra en un texto (W 23) recogido por Dreyfus (70): una vez que se establezcan los números característicos para la mayor parte de los conceptos , los seres humanos poseerán un nuevo instrumento que aumentará las capacidades de la mente en un grado mucho mayor que el aumento de la capacidad visual obtenido gracias a los instrumentos ópticos, y este instrumento nuevo superará al microscopio y al telescopio del mismo modo que la razón es superior a la vista. Aunque Leibniz está hablando aquí de un *lenguaje*, su utilización en máquinas está clara en otros textos. Ese "instrumento nuevo" está ahora presente por todas partes; hasta donde sabemos, fue Leibniz el único filósofo que anunció su llegada y el único que aportó varios de los elementos para su realización , particularmente la reducción de todas las nociones a números (W 17) y de todos los números a notación binaria , es decir, a 0 y 1 (W 598).

Dreyfus acusa a Leibniz de anular la interpretación, el juicio y la intuición una vez que el pensamiento se vuelve mecánico gracias a la programación según reglas incorporadas en el mismo aparato (72). Afirma además (201) que las computadoras se basan en la idea "que Leibniz anunció por primera vez" de que se puede tener una teoría de la práctica. Para los propósitos de la argumentación, esta expresión significaría la posibilidad de expresar la práctica mediante una teoría que consiste en una colección de reglas. Dreyfus no menciona ningún texto de Leibniz en estos dos pasajes , pero si volvemos al ejemplo de cómo aprendemos a andar en bicicleta, la reducción de la práctica a una teoría significaría la expresión en fórmulas matemáticas de los movimientos requeridos para mantenerse en equilibrio. La formalización sería entonces la teoría de la práctica, que luego sirve por ejemplo para programar a un robot de modo que haga lo mismo.

En (211) Dreyfus vuelve al ensayo de 1677 "Hacia una característica universal" (W 20) , donde Leibniz habla de un "alfabeto del pensamiento humano" cuyos caracteres deben mostrar, en sus combinaciones, las mismas relaciones que se dan entre los objetos de los que se habla. Las letras del alfabeto se aplican a hechos separados, a ítemes de información que se pueden aislar. El isomorfismo aquí postulado es obviamente semejante al que Wittgenstein expone en su *Tractatus* y no es de extrañar que Dreyfus lo mencione en la misma página; además, este isomorfismo supone que la característica universal llega más allá de la representación de elementos con números primos, y de compuestos con números resultantes de la multiplicación de primos, recuperables mediante la descomposición de un producto en sus factores. Es por eso que Leibniz concibe la característica universal como el instrumento que permite no solo sistematizar el conocimiento, sino también ampliarlo mediante descubrimientos; esto es posible gracias al isomorfismo entre caracteres representativos y realidad representada. Dreyfus no cita uno de los textos más interesantes de Leibniz en relación con los éxitos logrados mediante el arte de combinaciones que expuso por primera vez en su obra de juventud *De Arte Combinatoria* (1666) [12]. Se trata de una carta al duque Johann Friedrich en octubre de 1671 (W xx-xxi). El texto es muy largo pero la idea es sencilla: según Leibniz gracias a su arte combinatorio ha sido capaz de inventar una máquina aritmética, otra máquina para solución de problemas geométricos, lentes y tubos en óptica, medidas de la perspectiva, descubrimiento de la

12 Traducción al español del original en latín por Manuel A. Correia . Ver Leibniz (1992).

longitud en alta mar, un traje de buzo , un compresor de aire y otra serie de aparatos admirables. Llama la atención en particular su pretensión de haber resuelto el cálculo de la longitud en alta mar, pues como sabemos este problema que atormentaba a los navegantes y para cuya solución se establecieron premios solo se resolvió con los relojes de John Harrison, muchos años después.[13]

Pero Dreyfus no señala que la insistencia en haber descubierto un método para la invención, ampliación del conocimiento, sistematización de la información y solución de conflictos pertenece a las primeras etapas de la vida de Leibniz. Es muy interesante que los ensayos que cita Dreyfus son de 1677. Unos años después, en un ensayo de 1685 titulado "El arte del descubrimiento", Leibniz refleja una actitud ligeramente más moderada en relación con los supuestos logros obtenidos con su *De Arte Combinatoria*, a la que considera aquí "una obra de juventud y de un aprendiz" (W 51). Sin embargo, en este ensayo (W 50-58) todavía se jacta de haber descubierto algo asombroso, a saber, que podemos representar toda suerte de verdades y consecuencias con números (W 50), lo que se logra haciendo uso –"como hacen los matemáticos"– de caracteres apropiados para fijar nuestras ideas, a los que se añaden pruebas numéricas (W 52).

En los años siguientes se desvanecen las referencias al maravilloso método capaz de resolver todos los problemas. En las obras que culminan su pensamiento filosófico, la *Monadología* y *Los principios de la naturaleza y de la gracia fundados en la razón* , ambas de 1714, no encontramos ninguna mención al arte combinatorio ni a la característica universal. Tampoco en los *Nuevos ensayos* de 1704 ni en la *Teodicea* de 1710. Quizá la prueba más interesante de este enfriamiento respecto de las exageradas expectativas suscitadas por el método maravilloso de la combinatoria y del lenguaje característico la tengamos en el último escrito que recoge la selección de Wiener (594-599), la carta de Leibniz al zar Pedro el Grande de Rusia redactada en 1716 poco antes de la muerte de su autor. En ella Leibniz insta al zar a convertirse en un promotor de las artes y de las ciencias ("verdadero tesoro de la humanidad") para el beneficio de todos los seres humanos. Leibniz menciona aquí otra vez "un nuevo y maravilloso

13 Sobel (1995).

descubrimiento" (W 598) pero no es el método de combinaciones, sino la interpretación de los 64 hexagramas del Li-King (I Ching) "del famoso Fohi" (también escrito Fu-Hsi) como representación en notación binaria de los números desde el 0 al 63 (Gardner 1974, 109). De este modo, los 64 hexagramas, numerados del 0 al 63, se transforman en los números binarios del 000000 al 111111. Un texto muy interesante al respecto se titula "Explicación de la aritmética binaria" de 1703, en el que Leibniz no solo explica la manera de representar todos los números usando únicamente la progresión de dos en dos, sino también da ejemplos de cómo llevar a cabo las cuatro operaciones aritméticas con el nuevo cálculo (de Mora 2014, 12-18). "Yo he empleado desde hace varios años la progresión más simple de todas, que va de dos en dos" dice el autor (de Mora 2014, 12). Tanto Leibniz como Joaquín Bouvet, el jesuita misionero en China con el que tuvo correspondencia y que le envió el material correspondiente, atribuyen a Fu-Hsi alguna participación en el descubrimiento de este sistema numérico al colocar los 64 hexagramas en el orden que permite la interpretación binaria (de Mora, 18) pero el matemático Martin Gardner (1974,108) considera que no hay fundamento para esta opinión, y que fue Leibniz el primero que se dio cuenta del isomorfismo entre los hexagramas y los números al interpretar en el I Ching la línea sólida como 0 y la línea dividida por la mitad como 1.

Es verdad que en esas obras cercanas al final de la vida de Leibniz tampoco aparecen otros aspectos muy típicos de su pensamiento en etapas anteriores, en particular la teoría de la verdad como inclusión del sujeto en el predicado de la oración, que Bertrand Russell considera el fundamento de toda su filosofía , opinión que le parece confirmada con la publicación de los opúsculos lógicos hecha por Couturat (1977,9). Pero, a diferencia de lo que ocurre con el método cuyas bondades proclama con tanta insistencia entre 1677 y por lo menos hasta 1685, en el caso de la teoría de la verdad como inclusión encontramos que la reemplaza con la teoría de la correspondencia en *Nuevos ensayos*.[14] Tampoco aparecen en la *Monadología* los dos laberintos de los que habla en muchas ocasiones (el continuo y la libertad), pero hay una doctrina leibniciana que se mantiene hasta el últi-

14 Vease Camacho (2016b).

mo de sus escritos: la idea de que vivimos en el mejor de los mundos posibles.[15] El proverbial optimismo del bibliotecario de Hannover, ridiculizado por Voltaire en *Candide*, no se vio afectado por el fracaso de sus grandes proyectos de ecumenismo, conciliación y resolución automática de conflictos.

Si nuestra interpretación es correcta, Leibniz entonces resulta mucho más interesante de lo que vemos en *Lo que las computadoras no pueden hacer,* pues sería un antecesor tanto de la idea misma de la inteligencia artificial como del escepticismo posterior en cuanto a sus alcances. Las increíbles pretensiones y expectativas que tenía Leibniz en relación con su método combinatorio en el periodo que va desde *De Arte Combinatoria* (1666) hasta por lo menos 1685 tienen un obvio paralelismo con la situación trescientos años después y serían así un lejano anticipo de las exageraciones de los pioneros de la inteligencia artificial en la primera etapa , la que se inicia con el Simposio de Dartmouth en 1956. Así como Leibniz luego parece haberse olvidado del asunto , como muestra la ausencia de este tema en las obras de madurez y en la abundante correspondencia en sus últimos años[16], así también al periodo de efervescencia de la inteligencia artificial le sigue otro de enfriamiento que tuvo lugar cuando varias de las profecías sobre las capacidades de los programas inteligentes no se cumplieron. En nuestros días nos parecen asombrosas muchas de las cosas que se dijeron con gran seguridad en las últimas décadas del siglo pasado, de un modo parecido a como a Leibniz le parecieron cosas de juventud y de aprendiz las expectativas que tenía sobre su método.

Finalmente, y antes de terminar con nuestro recuento de las recriminaciones de Dreyfus contra Leibniz, conviene señalar que cuando el bibliotecario de Hannover habla de lenguajes característicos y del método de la combinación se pueden distinguir por lo menos ocho proyectos diferentes, como puede verse en el artículo reciente ya citado (Camacho 2016a, 51-54): un lenguaje pictórico o ideográfico tan transparente que podemos entenderlo sin estudiarlo, un lenguaje universal basado en la distinción entre conceptos simples y

15 *Monadología*, #85, *Principios,* #16.

16 Nicolás,J.-Cubells, M.R. (eds) (2007), Orio,B. (ed) 2011, Orio,B. (ed)2011b.

compuestos, un lenguaje artificial capaz de representar todo tipo de inferencia formal, un lenguaje artificial en el que todas las verdades de razón se reduzcan a un cálculo, un lenguaje capaz de representar cualquier contenido usando únicamente 0 y 1, un cálculo formal que use únicamente 0 y 1, una versión simplificada de un lenguaje natural (el latín) en el que se eliminen géneros, declinaciones y otras características gramaticales , y un lenguaje natural que se tome como modelo para otros lenguajes naturales.

4. Smullyan, los acertijos y las máquinas

Muy diferente al estilo de la argumentación de Dreyfus es la manera como Smullyan se aproxima al tema de los límites de sistemas formales. Su muerte el 6 de febrero de 2017 puso fin a una vida de dedicación a la lógica y a la filosofía que parece haber sido muy entretenida y divertida a juzgar por la mayoría de sus libros y por lo que cuentan de él sus biógrafos.[17] Smullyan tenía una combinación de habilidades que se repite en algunos de los autores más interesantes de tiempos recientes: lógica, matemáticas, ciencia, magia, música y filosofía. Fue un mago muy hábil y un buen pianista. La seriedad y grado de abstracción de su primera obra, *Theory of Formal Systems* (Princeton University Press, 1961) no anticipan la amenidad de *What Is the Name of This Book?* (1978) (*¿Cómo se llama este libro*? Cátedra, 1984), *Alice in Puzzleland* (1982) (*Alicia en el país de las adivinanzas,* Cátedra, 1991), *The Lady or the Tiger?* (1982) (¿La dama o el tigre?, Cátedra, 1991) y *5000 B.C. and Other Philosophical Phantasies* (*5000 años a. de C. y otras fantasías filosóficas* (Cátedra, 1989). La lista completa de sus obras llena páginas enteras; baste aquí con señalar que se distinguen dos conjuntos que sin embargo tienen mucho en común, el de las obras formales en lógica (entre ellas manuales universitarios) y las de entretenimiento. Lo admirable de las obras de Smullyan es su manera de presentar en forma sencilla y amena complicados problemas matemáticos y lógicos.

17 En el artículo aparecido el 11 de febrero de 2017 en *New York Times* con motivo de su muerte (Sandomir 2017) se incluyen varios de sus acertijos más ingeniosos.

Sus años de juventud muestran una combinación de estudios en música con formación autodidacta en lógica y matemáticas, al mismo tiempo que se ganaba la vida como mago. La lista de lugares donde llevó a cabo esta combinación es demasiado larga para citarla aquí y lo que nos interesa es señalar que al fin tuvo la suerte de encontrar en Chicago al famoso filósofo Rudolf Carnap, quien enseguida reconoció el talento del muchacho y lo recomendó para dar clases en Darmouth, aunque no tenía título universitario. Eran otros tiempos. Su primer título de educación superior fue posible gracias a que la Universidad de Chicago le reconoció créditos por cursos que había impartido aunque nunca se había matriculado en ellos como estudiante. En 1959 obtuvo el doctorado en Princeton bajo la dirección del gran lógico Alonzo Church. Este es uno de los autores en los que se basa su *Teoría de los sistemas formales* , junto con Goedel, Post y Rosser. Tema central en esta obra es el de las posibilidades, utilidad y limitaciones de los lenguajes artificiales desambiguados dotados de reglas para el cálculo. "Formal" y "mecánico" coinciden en esta obra: los sistemas formales se basan en reglas que se aplican recursivamente en operaciones que pueden llevar a cabo máquinas programadas para este fin. Las limitaciones de las máquinas, a su vez, reflejan lo descubierto en la teoría: algunas máquinas son físicamente imposibles (v.g. las de movimiento perpetuo) mientras otras son lógicamente imposibles, porque involucran contradicción. La máquina que resuelve todos los problemas es lógicamente imposible, pues entre sus resultados tendría que incluir la prueba de que es capaz de resolver *todos* los problemas, y para hacerlo cae en los problemas de la autorreferencia.

En *La dama o el tigre* Smullyan empieza con una gran variedad de acertijos , procede luego a desarrollar "una novela matemática que presenta el gran descubrimiento de Gödel" (este es el subtítulo de la obra) y termina la parte cuarta, que lleva el título "¿Resoluble o irresoluble?", con el breve capítulo "El sueño de Leibniz". La mención de damas y tigres tiene que ver con una serie de acertijos en los que se trata de averiguar si al abrir una puerta cerrada aparecerá un tigre o una muchacha. Algunos de estos acertijos son variaciones de otros cuyos orígenes se remontan al pasado inmemorial. De todos es conocida, por ejemplo, la historieta de que al morir encontraremos dos ángeles idénticos frente a dos puertas idénticas, una que conduce al cielo y otra al infierno. El ángel en la puerta del cielo siempre dice la verdad y el otro siempre miente. Tenemos que descubrir la puerta correcta con una sola pregunta cuya respuesta sea sí o no según sea el ángel que responde, y es obvio que la única manera de preguntar con el resultado apetecido consiste en conseguir que uno de los interrogados necesariamente conteste "sí" y el otro "no", y que podamos distinguirlos de esta manera. "¿Miente Ud?" no nos servirá porque ambos dirán que no. Tampoco "¿Dice Ud la verdad?", porque ambos dirán que sí. Recordar que existen proposiciones necesariamente verdaderas cuya negación es imposible ayuda a resolver el acertijo.

Una variante de este acertijo es más interesante porque introduce la noción de límites. Si la mitad de los habitantes de una isla miente siempre y la otra mitad siempre dice la verdad, y tenemos que descubrir quién es quién a partir de lo que digan o mediante cualquier pregunta, entre las posibles frases de estos habitantes no encontraremos "soy un mentiroso". No la pueden decir los mentirosos porque sería verdadera y ellos solo pueden decir proposiciones falsas, ni los veraces porque sería falsa, contra lo que asumimos de que solo dicen verdad. Sin embargo, sabemos que dicha proposición es verdadera de la mitad de los habitantes. Si vemos el acertijo como un sistema formal, nos damos cuenta de que las premisas mismas del sistema hacen que el enunciado "soy un mentiroso" no se pueda emitir, aunque tenga cincuenta por ciento de probabilidad de ser verdadero. La diferencia entre verdad y falsedad se combina en los siguientes rompecabezas que aparecen en esta obra con otras oposiciones: cuerdo-loco, vampiro-humano, dormido-despierto, diurno-nocturno, lenguaje-metalenguaje. De aquí se pasa a la distinción entre acertijos y metaacertijos; en estos últimos hay que averiguar si con la información disponible se puede resolver el problema , cuál información sería necesaria para que el problema sea soluble, y –sobre todo– cuáles acertijos no tienen solución.

La "novela matemática" se encuentra en la parte tercera y tiene como título "El misterio de la cerradura de Montecarlo". Tiene que ver con una caja de seguridad cuya combinación se ha perdido. Urge abrirla y lo único que se llega a saber después de varias investigaciones preliminares es que funciona con combinaciones sujetas a ciertas reglas de relación entre secuencias de símbolos alfabéticos. Al combinar de cualquier manera las 26 letras mayúsculas del alfabeto obtenemos tres tipos de combinaciones: algunas abren la caja, otras la bloquean permanentemente y otras ni abren ni bloquean. Cómo se obtiene cada una de estas series se explica por las reglas formales que se aplican en las combinaciones; como resultado de dichas reglas algunas operaciones, por ejemplo, generan repeticiones de letras.

La caja de seguridad es una máquina y la novela introduce otras máquinas, lo que lleva al autor a hablar de demostrabilidad y verdad, de máquinas que hablan de si mismas, máquinas que nunca fueron construidas y del sueño leibniciano de construir una máquina capaz de resolver todos los problemas matemáticos e incluso de zanjar las disputas filosóficas. Así como las premisas de acertijos excluyen ciertas respuestas, de un modo parecido la estructura de los sistemas formales impide la construcción de una máquina que resuelva todos los problemas matemáticos. Smullyan basa su argumentación en los descubrimientos de Gödel, Rosser, Post y en los suyos propios. Gödel no habló de máquinas en su famoso teorema de 1931, pero otros después de él sí lo han hecho. Es ésta la conexión con Leibniz.

Al igual que en el caso de Dreyfus, la relación entre la argumentación de Smullyan y las ideas de Leibniz es más profunda de lo que parece a simple vista. El teorema de la incompletitud de sistemas formales, más conocido como el Teorema de Gödel de 1931, muestra el uso que hace Gödel del método de aritmetización utilizado en forma más simple por Leibniz, matemático por el que Gödel sintió una curiosa obsesión. No solo procuró leer todo cuanto encontró de nuestro autor en la biblioteca de Princeton, sino que llegó a la conclusión de que había existido una gran conspiración para impedir que se difundieran sus ideas por ser revolucionarias. [18]

A la idea de asignar un número propio a cada concepto se añade la diferencia en la utilización de números primos y compuestos. En

18 Mark van Atten (2015) trata en detalle la obsesión de Gödel por Leibniz y su convicción de que una gran conspiración se encargó de que no llegaran hasta nosotros los textos en los que Leibniz desarrolla su lenguaje característico.

Leibniz, los números primos se reservan para nociones simples y los compuestos para conceptos resultantes de la multiplicación de los primos, de modo que a partir de la descomposición de un número en sus primos podemos regresar a las nociones no definidas. Aunque ya Aristóteles (*Met*. VIII, 3, 1043b33-1044a15) había comparado las definiciones con los números compuestos, y los términos primitivos últimos en las que se resuelven, con los números primos (Sánchez-Mazas,1951), no hay pruebas de que Leibniz se inspirara en este antecedente; en cambio sí las hay de que Gödel apreció grandemente las ideas de Leibniz (van Atten, 2015).

Entre muchos otros escritos , en un ensayo de 1679 titulado "Elementos para fundar el cálculo" (Couturat 258, Velarde-Cabañas 43-47) se indica el método que sigue Leibniz: a cada término de una proposición se asigna un número, de tal modo que la proposición resulte de la multiplicación de los números característicos. Si se asigna el número 2 al concepto "animal" y el 3 a "racional", entonces el número característico del concepto "ser humano" es 6, que resulta de multiplicar 2 por 3. Al revés, de la descomposición del número 6 ("ser humano") en sus primos obtenemos 2 ("animal") y 3 ("racional"). El método empleado por Gödel es en principio el mismo de Leibniz pero con mayor complejidad, pues los números característicos de los símbolos primitivos se convierten en los exponentes de la serie de números primos, y a veces estos exponentes a su vez tienen otros. De esta manera el número característico de cada fórmula bien formada ("número goedeliano") resulta de la multiplicación de los primos sucesivos requeridos elevados a la potencia correspondiente al número característico de los primitivos involucrados, con tantos términos en la multiplicación como símbolos en la fórmula. Así por ejemplo (Nagel-Newman 1970,95) el número goedeliano correspondiente a 0=0 es $2^6 \times 3^5 \times 5^6$ debido a que el número goedeliano de 0 es 6 y el de "=" es 5. La serie de números primos (2,3,5,7,11…) es la misma para todos los números goedelianos, aunque por supuesto su extensión varía según sea la cantidad de símbolos que hay que representar (en este caso solo tres).

En el cuadro siguiente vemos los parecidos y diferencias entre los tres autores hasta ahora mencionados.

Aristóteles (*Met.* VIII, 3)	Las definiciones son como los números compuestos	Al resolverlas llegamos a conceptos primitivos indefinibles, que son como los números primos
Leibniz	Las proposiciones tienen un número característico compuesto	Que se descompone en los números primos característicos del sujeto y predicado de la proposición (v.g.6=2x3)
Gödel ("Sobre la incompletitud de sistemas formales", 1931)	Cualquier fórmula matemática o metamatemática tiene un número propio y único	Que resulta de la multiplicación de los primos con los exponentes que corresponden a los números asignados a los símbolos (v.g. $2^1 x 3^2=18$)

En el Teorema de Gödel la incompletitud se genera al incluir la metamatemática o teoría de la prueba dentro del lenguaje formal. Dado que los sistemas formales a los que se aplica el teorema –como el de Russell y Whitehead en *Principia Mathematica* y el de Zermelo-Frankel– buscan fundamentar la aritmética derivando todos los números a partir de axiomas lógicos , lo que busca la prueba del teorema es mostrar que por lo menos un número no podría generarse de esa manera. Sabemos que la prueba empieza con una variante de lo propuesto por Leibniz mediante la diferencia entre números primos y compuestos. A los símbolos primitivos del sistema se le asignan números que se convierten en los exponentes de la serie de números primos cuando se generan los números de fórmulas variadas (incluyendo la fórmula que expresa la relación de demostración entre dos números) , como hemos visto antes y podemos ver con más detalle en el siguiente ejemplo (Nagel-Newman 1970,97):

- Fórmula lógica : $(p \vee p) \supset p$
- Número goedeliano correspondiente: $2^8 x 3^{121} x 5^2 x 7^{121} x 11^9 x 13^3 x 17^{121}$
- Explicación: en la fórmula aparecen 7 símbolos, por lo cual se ne-

cesitan 7 números primos sucesivos (2, 3,5 ,7 ,11 ,13 ,17). Cada uno de los símbolos que aparecen en la fórmula tiene su propio número característico , a saber: (=8, p=121, v=2,)=9, ⊃=3, que aparece como exponente en la serie de números primos, de acuerdo con el orden en que aparecen en la fórmula. El parecido con el método propuesto por Leibniz en los ensayos de 1677 citados por Dreyfus y otros de esa época es obvio. De este modo llegamos a nuestra observación principal: *los límites del cálculo leibniciano se obtienen mediante aplicación de métodos inspirados en el mismo autor.*

Para ver de qué manera se introduce la autorreferencia hay que tener en cuenta que la demostración de una proposición es también una relación entre números. La proposición indecidible que no puede ser probada ni refutada es justamente una proposición que se refiere a si misma, a saber "esta proposición no puede ser probada"(van Heijenoort 1967, 352; Nagel-Newman 1970,114) , a la que se le asigna un número siguiendo el mismo proceso aplicado en el caso de todos los símbolos introducidos y las fórmulas resultantes de sus combinaciones. Puesto en términos de máquinas lógicas, habría un número que no podría ser generado a partir de las premisas del sistema.

Un mérito de Smullyan es mostrar la incompletitud mediante máquinas, de las que se pueden imaginar muchas variaciones. Por ejemplo, una máquina diseñada para generar números impares que funcione aplicando la regla x + 1, donde x es un número par, no tendría límites para generar cualquier número impar. Y si pensamos en dos máquinas coordinadas, la primera de las cuales genera números pares mediante alguna fórmula y la segunda números impares aplicando x + 1, donde x es un número generado por la primera máquina, tampoco encontraríamos ninguna limitación en cuanto a la generación de cualquier número par o impar ni confusión en cuanto a cuál máquina genera cuál número. Ambas máquinas, y otras muchas que podemos imaginar (como las sumadoras) son completas dentro de sus límites, sin que necesiten recurrir a la autorreferencia.

En las explicaciones de Smullyan al final de *La dama o el tigre* podemos ver que el problema empieza a surgir cuando se introduce una máquina (o más bien metamáquina, pues el ámbito de su función son los resultados de las otras máquinas y de sí misma) que asigna un número a los resultados de cada una de las anteriores y también asigna un número a sus propios resultados. De esta manera se introduce la autorreferencia.

Una salida ad hoc al problema es negar la condición de proposición, es decir, de enunciado que puede ser verdadero o falso, a la colección de palabras “esta proposición no puede ser probada” y afirmar que no existe manera de probar o refutar dicha serie de términos, a diferencia de las proposiciones auténticas. Por esta razón la calificaríamos de pseudoproposición. Ciertamente hay una diferencia entre decir “ 2+3=0 no puede ser probada” y “esta proposición no puede ser probada” cuando la referencia de “esta” es la colección de palabras “esta proposición no puede ser probada”. En este enfoque se niega la similitud entre enunciados como “esta proposición está escrita en alemán” y “esta proposición es falsa.” Podemos determinar que “esta proposición esta escrita en alemán” es falsa con solo mirar las palabras que integran el conjunto, pero no podemos hacer lo mismo con el segundo conjunto.

Esta solución se parece a la que se propone para otra supuesta paradoja, la de la proposición “Desobedece esta orden”. Puesto que no sabemos a qué orden se refiere, tampoco tenemos la menor idea de cómo sería desobedecerla. En nuestra vida cotidiana no existe ninguna experiencia que corresponda a un comportamiento generado por la frase aislada de todo contexto que ordene desobedecerse a si misma. También se parece a la oración “¿Cuál es la respuesta a esta pregunta?” que para muchos solo podría ser una pseudopregunta, pues no hay manera de encontrarle respuesta. Sería necesaria otra investigación para saber si la anterior manera de ver las cosas habría sido del agrado de Leibniz, aunque provisionalmente podemos señalar que no hemos encontrado textos sobre la paradoja del mentiroso ni sobre la autorreferencia en los volúmenes de obras de Leibniz a nuestro alcance.

Por este motivo el hecho de que no se pueda derivar un número que corresponde a una autorreferencia de los sistemas formales a los que se refiere el teorema de la incompletitud no parece ser objeción a los planes y proyectos de Leibniz. Hasta donde sabemos nuestro autor no incluye la autorreferencia en su arte combinatorio ni en su característica universal; ninguno de los dos habla de sí mismo. Tiene razón Smullyan (1982,268) cuando dice que en relación con las máquinas lógicas de Leibniz tenemos que asumir muchas cosas que no preocuparon a su autor, pues solo recientemente se han precisado las nociones y los alcances de esta noción. Por otra parte, las afirmaciones de Leibniz sobre el arte combinatorio y la característica universal son tan escuetas como ambiciosas; con los pocos datos que incluye en sus escritos resulta imposible justificar las afirmaciones de grandes éxitos que encontramos en varios de los textos de juventud.

Conclusión

Tiene razón Dreyfus cuando asigna una posición prominente a Leibniz en el largo camino hacia la mecanización del pensamiento. Sin embargo, hubiera sido más justa su apreciación de la riqueza de su pensamiento si también hubiera tenido en cuenta la ausencia de los proyectos de formalización universal en la etapa de madurez del famoso bibliotecario de Hannover. De esta manera habría podido fortalecer su argumentación : así como Leibniz abandonó las exageradas pretensiones juveniles sobre su método de las combinaciones y características, de un modo semejante lo que Dreyfus llama "Good Old-Fashioned Artificial Intelligence (GOFAI)" dejó de estar de moda al no cumplirse las predicciones iniciales. Más en concreto, la retórica asociada a ese proyecto cambió con el tiempo. Hoy cargamos teléfonos "inteligentes" pero no oimos afirmaciones de que son inteligentes en el mismo sentido en que lo somos quienes los usamos.

En cuanto a Smullyan, su referencia a Leibniz es más matizada por cuanto reconoce que el planteamiento actual se basa en varias nociones que no estaban disponibles en el siglo XVII. Cuando utiliza ampliamente el teorema de incompletitud de sistemas formales de Gödel para explicar por qué una máquina no podría generar todos los números correspondientes a las soluciones de todos los problemas matemáticos (y no digamos nada de los filosóficos), hubiera sido interesante ver de qué modo habría explicado el uso de métodos leibnicianos en la construcción de la prueba de dicho teorema.

Paradójicamente, la prueba contra la pretensión leibniciana se basa en un procedimiento propugnado por el mismo contra quien se emplea. En todo caso, la insistencia de Smullyan en concretar difíciles problemas abstractos mediante el uso de acertijos y máquinas debería ser imitada por todos los que estamos dedicados a estudios de lógica y filosofía, y aprovechada en la enseñanza de las matemáticas.

En los dos casos de filósofos reseñados, Leibniz se encuentra en ambos lados de la discusión aunque de distinta manera. Quizá esto ayude a entender el interés actual por el pensamiento de alguien que murió hace poco más de 300 años y que, sin embargo, parece más contemporáneo que muchos autores más recientes.

Bibliografía

Adaime, I. (2011) *Kasparov vs. Deep Blue : la conflictiva relación hombre máquina* . Universidad de Buenos Aires. Disponible en pdf en búsqueda hecha el 13 de julio de 2017 : http://newpagecomunicacion.sociales.uba.ar/files/2013/02/Adaime.pdf.

Camacho,L. (2016a) Lógica de Aristóteles en Leibniz: el largo camino desde la gramática hacia la teoría de la inferencia, en Nicolás,J.-Öffenberger,N.(eds) (2016) *Beiträge zu Leibniz´ Rezeption der Aristotelischen Logik und Metaphysik*, 41-62. Hildesheim, Georg Olms Verlag.

(2016b) Leibniz´s Perspectives on Truth, en Li (2016), III,98-110.

Couturat,L. (ed)(1903) *Opuscules et fragments inédites de Leibniz.* París.

de Mora Ch., M.S. (ed) (2014) *G.W. Leibniz, Obras filosóficas y científicas,* vol. 7A Escritos matemáticos. Granada, Comares.

Dreyfus,H. (1965) *Alchemy and Artificial Intelligence*. Rand Corporation.

(1992) *What Computers Still Can´t Do*. Cambridge, Mass. The MIT Press. En las ediciones anteriores (1972 y 1979) no aparece la palabra "still" en el título.

Gardner, M. (1974) Mathematical Games, en *Scientific American*, 1/1974,108-113.

Kirschner, S.(2016) "Der wind selbst fast allezeit drehn, wenn die hauptflugel nicht im winde stehn"—Der "Leibniz-Regler", ein Mechanismus zur automatischen Ausrichtung von Windmühlenflügeln, en Li (2016),V, 427-448.

Knuth, D.E. (1977) Algorithms, en *Scientific American*, 4/1977, 63-80.

Leibniz, G.W. (1992) *Disertación acerca del Arte Combinatorio.* Traducción de Manuel A. Correia. Ediciones Universidad Católica de Chile.

Li, W. ,ed. (2016) "Für Unser Glück oder das Glück Anderer", *Vörtrage des X.Internationalen Leibniz-Kongresses*, herausgegeben von Wenchao Li in Verbindung mit Ute Backmann, Sven Erdner, Esther-Maria Errulat, Jürgen Herbst, Helena Iwasinski und Simona Noreik. Hildesheim, Georg Olms Verlag.

Nagel,E.-Newman,J.R. (1958) *El teorema de Gödel*. Madrid, Tecnos.

Nicolás, J.-Cubells, M.R.(eds) (2007) *Gottfried Wilhelm Leibniz , Obras filosóficas y científicas,* vol.14, Correspondencia I. Granada, Comares.

Orio, B.(ed) (2011) *G.W. Leibniz, Obras filosóficas y científicas,* vol. 16A, Correspondencia III. Granada, Comares.

G.W.Leibniz, Obras filosóficas y científicas, vol.16B. Correspondencia III. Granada,Comares.

Russell, B. (1977) *Exposición crítica de la filosofía de Leibniz, con un apéndice integrado por los pasajes más importantes.* Buenos Aires, Ediciones Siglo Veinte.

Sánchez-Mazas, M. (1951) Sobre un pasaje de Aristóteles y el cálculo lógico de Leibniz, en *Revista de Filosofía* (Madrid), 38, 529-534.

Sobel,D. (1995) *Longitude, The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time.* New York, Penguin Books.

Stein, E.(2016) Design and Construction of Leibniz´s Proposed *Machina Deciphratoria* en Li (2016), V, 361-376.

Smullyan, R. (1961) *Theory of Formal Systems* . Princeton University Press.

(1984) ¿Cómo se llama este libro? Madrid, Cátedra.

(1989) *5000 años a. de C. y otras fantasías filosóficas* .Madrid, Cátedra.

(1991a) *Alicia en el país de las adivinanzas,* Madrid, Cátedra.

(1991b) ¿La dama o el tigre?, Madrid, Cátedra.

Sandomir, R. (2017) Raymond Smullyan, Puzzle-Creating Logician, Dies at 97, *New York Times* , 11 de febrero 2017 (búsqueda hecha el 5 de junio 2017) en https://www.nytimes.com/2017/02/11/us/ray mond-smullyan-dead-puzzle-creator.html?_r=0

Wiener, P.P. (1951)(ed.) *Leibniz Selections.* New York, Charles Scribner´s Sons.

van Atten, M. (2015) *Essays on Gödel's Reception of Leibniz, Husserl, and Brouwer*. Springer Verlag.

van Heijenoort, J. (ed) (1977) *From Frege to Gödel, A Source Book in Mathematical Logic, 1879-1931.* Cambridge, Mass. Harvard University Press.

(1967) Gödel´s Theorems, en Edwards,P.(ed) (1967) *The Encyclopedia of Philosophy* , 3-4, 348-357. New York, Macmillan.

Velarde,J.-Cabañas, L. (eds) (2013) *G.W.Leibniz, Obras filosóficas y científicas,* vol. 5, Lengua universal, característica y lógica. Granada, Comares, 2013.