

Fenomenología Hermenéutica e Inteligencia Artificial: Otra Urbanización de la ‘Provincia Heideggeriana’

Masis, Jethro (Universidad de Costa Rica)

“Los matemáticos quieren tratar los asuntos de la percepción matemáticamente y, con ello, se ridiculizan a sí mismos. [...] La mente lo hace tácita, naturalmente, y sin reglas técnicas”.

— PASCAL, *Pensées*

La Discusión Abierta: La Irrupción de Heidegger en el Laboratorio de Inteligencia Artificial del MIT

En primer lugar, un hecho debe ser constatado: el laboratorio de Inteligencia Artificial (IA en lo sucesivo) del MIT se ha hecho heideggeriano. Por tanto, no tratamos acá de temas inventados, sino de un ámbito de discusión que se encuentra ya abierto. Incluso según Michael Wheeler en su más reciente obra, la influencia heideggeriana no se restringe al importante laboratorio de Massachusetts, sino que “una ciencia cognitiva heideggeriana está... surgiendo en estos momentos en los laboratorios y oficinas alrededor del mundo donde el pensamiento encarnado está bajo desarrollo e investigación activa” (2007: 285). Con todo, tiene razón Winograd cuando afirma en ‘Heidegger and the Design of Computer Systems’ que “para quienes han seguido la historia de la IA resulta irónico que el laboratorio del MIT se haya convertido en la cuna de IA heideggeriana” (citado por Dreyfus, 1992: xxxi). ¿No comprendería Heidegger, en efecto, el proyecto de la IA como el fastigio de la metafísica, como la tecnificación más nefanda?

La historia de cómo el pensamiento de Heidegger llegó a constituirse en un referente ineludible en las investigaciones del proyecto de investigación de la IA, se retrotrae a la publicación de *What Computers Can't Do* (1972), del profesor de la Universidad de Berkeley Hubert Dreyfus; una obra que, por cierto, estaba marcada en su subtítulo por el afán kantiano de imponer límites a una pretendida razón (en este caso, la de la simulación artificial de las competencias cognitivas humanas en una máquina digital), y que en su última edición de 1992 incluso se subtuló sin ambages ‘*A Critique of Artificial Reason*’.

No es este el lugar, por las comprensibles limitaciones que se nos imponen, de reseñar las etapas pioneras de la investigación en IA (1957-1962: simulación cognitiva y 1962-1967: procesamiento de información semántica), en que el proyecto se vio complicado en dificultades inextrincables. Ni siquiera es posible esbozar las etapas que constituyen los subsecuentes intentos correctivos de herencia heideggeriana (hasta cierto punto, los micromundos de Winograd, pero más decididamente los proyectos de Agre 1997, Brooks 1999 y Freeman 2000; este último más merleau-pontyano que otra cosa). Para tales efectos, la gran obra de Dreyfus y sus otros textos de menor rango prestan el servicio debido. Lo que más nos interesa es el contexto problemático y filosófico: en palabras de Minsky (cf. 2003), el que la IA haya sufrido una muerte cerebral desde los años setenta a partir de que descubrió el problema del conocimiento del sentido común.

Esto quiere decir que nuestro texto quiere solamente servirse tangencialmente de las contribuciones críticas de Dreyfus al proyecto de investigación de la IA y a la subsecuente generación de investigadores en esa área que reclaman explícitamente un linaje heideggeriano. Pero esto sólo con el fin de visitar una derrota que, sin embargo, resulta fructífera para aclarar qué significa la transformación hermenéutica de la fenomenología llevada a cabo por Heidegger en *Sein und Zeit* (1927, SZ en lo sucesivo), su *magnum opus*.

Que el laboratorio de investigación de IA del MIT (el más importante del mundo), otrora comandado por el racionalista y científico Minsky, se haya convertido en la cuna de un estilo de investigación de estos problemas marcado por el epíteto ‘heideggeriano’, no deja de ser notable. Sobre todo porque para quienes han sido educados filosóficamente en un contexto donde por años se ha forjado una actitud analítica y científica, aún resuenan las mofas de un Carnap (cf. 1981) contra ese lenguaje de Heidegger que había proferido aquellas infames preguntas respecto de la nada: “¿Sólo hay nada porque hay el no, es decir, la negación? ¿O es más bien al contrario? ¿Sólo hay la negación y el no porque hay la nada?” (H: 97). Y esto con la grave – para Carnap, ‘insensata’ – conclusión: “Pero nosotros afirmamos que la nada es más originaria que el no y la negación” (*idem*).

Pero, ¿de qué estamos hablando? ¿Qué tiene que ver la nada con el fracaso del proyecto de investigación de la IA? Tiene que ver mucho. Y es curioso que en el mismo texto que estamos citando, la lección inaugural *Was ist Metaphysik?* (1929), Heidegger advierta, como en una premonición que se adelanta a las objeciones de Carnap, que a la ciencia la nada no le puede parecer más que un espanto y una fantasmagoría, a saber, “[l]a ciencia no quiere saber nada de la nada” (cf. H: 95). Y nosotros afirmamos: lo que le hubiera tenido que interesar a la IA era, precisamente, la nada; y por no haber prestado atención al problema de la nada, en efecto, es que los investigadores de la IA no podrán ni por mil años entender por qué su proyecto no funciona ni puede funcionar nunca. Pero, ¿cómo podrían haberle prestado atención a la nada? No se trata sólo de que su educación analítica se los prohibiera, sino de que vale bien pensar qué se hace con la nada si es, precisamente, nada. Pero esta nada no es nada de nada, sino una nada que pertenece al ser.

La Nada y el Problema de la Programación del Sentido

El proyecto de investigación de la IA se encontró con un escabroso obstáculo. En palabras de Wheeler: “Dando por supuesto un mundo dinámico y cambiante, ¿cómo puede un sistema... dar cuenta de aquellos estados cambiantes del mundo... que importan, y de aquellos estados no cambiantes del mundo que importan, siempre que ignore aquellos que no importan? Y ¿cómo puede un sistema semejante reparar y (si es necesario) revisar, a partir de todas las creencias que posee, sólo aquellas creencias que son relevantes en un contexto particular de acción?” (2007: 179). Si la mente humana puede concebirse en parangón con una máquina recursiva, bastaría con añadirle a su memoria millones de enunciados, de los cuales, en su totalidad, estaría pretendidamente constituida. Pero los investigadores cayeron en la cuenta de que los enunciados necesitaban un contexto: por sí sola, una sentencia apofántica no tiene sentido, tiene que referir una totalidad bajo la cual esa sentencia adquiriera, por decirlo así, carne cognitiva, situacionalidad embebida e involucrada. A partir de este problema, lo que al parecer hacía falta era añadir a los enunciados un marco, es decir, un trasfondo contextual, el cual sería precisamente esa totalidad de sentido por lo pronto perdida. Respecto de este problema (el llamado *frame problem*) y de la definitiva degeneración del proyecto de la IA, ha afirmado Dreyfus:

Pero un sistema de marcos no está *en* una situación, así que en aras de seleccionar los hechos posiblemente relevantes de la situación actual, uno necesitaría marcos para reconocer situaciones como las fiestas de cumpleaños, y para diferenciarlas de otras situaciones tales como ordenar en un restaurante. Empero, aún me preguntaba, ¿cómo podría la computadora seleccionar, de entre esos pretendidos millones de marcos en su memoria, el marco relevante para seleccionar el marco de la

fiesta de cumpleaños en tanto marco relevante con el fin de que pudiera ver la relevancia actual de, digamos, el intercambio de regalos en vez de meramente dar dinero? Me parecía obvio que cualquier programa de IA que utilizara marcos para organizar millones de hechos sin significado para así categorizarlos dentro de los relevantes, caería en un regreso a marcos para reconocer marcos relevantes que reconocerían hechos relevantes, y que, por ende, el problema del marco no era solamente una dificultad sino un signo de que algo marchaba seriamente mal con todo el método (2007: 250).

Hay aquí un supuesto de analiticidad según el cual se puede acceder al sentido a partir de lo apofántico. Las proposiciones enunciativas vendrían a ser como los átomos del sentido, de forma que la mente humana podría comprenderse como una máquina recursiva: máquina que, cuando comprende, *re-flexiona*, por decirlo así, sobre los propios recursos apofánticos que ya tiene ‘almacenados’. Para hacer que una máquina logre adquirir afrontamiento (*coping*), es decir, la experiencia humana de arreglárselas, habérselas conductual, corporal, comprensiva y cognitivamente (esto es, de forma *encarnada* y prácticamente *situada*) en un contexto determinado en el que se está envuelto de manera ejecutiva y praxica, nada más habría que programarla – según esta falsa suposición – con millones y millones de datos enunciativos, de los que pretendidamente estaría conformada la mente humana. Pero no es sino Heidegger quien ha advertido que “el sentido no puede ser definido como algo que se encuentra ‘en’ el juicio, junto con el acto de juzgar” (SZ, § 33: 154). Y ¿dónde se encuentra el sentido? Propiamente, en ningún lugar.

La suposición abstracta del análisis quiere, en efecto, servirse de la parte para dar con el todo y, en el mismo acto, ya interpreta el todo del sentido como algo total que podría descomponerse en sus partes enunciativas. Consecuentemente, hay también una concepción de la filosofía en relación con la gramática. Es la concepción que expresa el filósofo analítico Peter Strawson: “[A]sí como el gramático modélico trabaja para elaborar una explicación sistemática del sistema de reglas que observamos sin ningún esfuerzo cuando hablamos gramaticalmente, el filósofo lo hace para conseguir una explicación sistemática de la estructura *conceptual general* de la que nuestra práctica diaria muestra que tenemos un dominio tácito e inconsciente” (1997: 50). Según Strawson, el trabajo filosófico consistiría en hacer explícita la teoría que ya dominamos cuando, verbigracia, pensamos, es decir, se trata de descubrir la teoría implícita a todo nuestro pensamiento. Es la suposición de que hay algo así como un sistema implícito de reglas (que, por supuesto, no ha dado entretanto pruebas incontestables de su existencia), lo que siempre dan por sentido los investigadores de la IA cuando se afanan por programar enunciados. Pero, ¿es cierto que hay una teoría anterior o implícita que el filósofo debe traer a explicitar? ¿En esto realmente consiste la labor de filósofo: en sonsacar la estructura apriórica del pensamiento que se halla ya interpretada como reglamentación implícita?

Si se comprende bien la empresa que Heidegger se trae entre manos en la década fenomenológica (1919-1929), hay que conceder, con Figal, que se trataba principalmente de la pregunta de “cómo es posible la filosofía sin ser teoría” (1996: 35). El llamado *Problem des Theoretischen* o ‘problema de lo teórico’ surge de forma explícita en la primera lección universitaria dictada por Heidegger en la Universidad de Friburgo en Brisgovia en 1919 (cf. GA 56/57). Además, está presente como insumo instrumental y metodológico en SZ, e incluso puede entenderse que la insatisfacción de Heidegger respecto del lenguaje utilizado en aquel gran tratado – al que ha denominado, en *Brief über den Humanismus* (1947, cf. H: 313-364), ‘fracaso’ en la exposición (*Versagen*) – se debe

asimismo al afinamiento del acervo protoconceptual que provee un sometimiento a las dificultades de un posible decir en filosofía que, a su vez, no sea teórico. El problema hermenéutico de SZ y, en correspondencia con ello, su nivel de experiencia y de meditación, sin los cuales toda recepción es inefectiva (y no toca en lo más mínimo esa supuesta ‘filosofía de Heidegger’, que no hay), es el que atañe a la diferencia ontológica (*ontologische Differenz*). Es, en efecto, porque el ser ha sido tradicionalmente interpretado como ente, que SZ no puede regodearse en las magníficas cualificaciones de un ser ontificado. Toda cualificación, es cualificación de ‘algo’. Pero si es del ser de lo que tratamos, que no del ente, cabe bien cuestionarse qué puede decirse del ser sin que de inmediato caigamos en la susodicha ontificación nefanda. Aquí la valoración de lo que signifique el trabajo que está siendo llevado a cabo en SZ se bifurca: o bien Heidegger se ocupa del ser y es – como lo afirma una, a nuestro juicio, inexacta recepción contemporánea – el pensador que restauró los derechos de la vetusta ontología, o más bien Heidegger de ninguna manera se ocupa del ser (siempre que comprendamos ese ‘de’ del ser en el sentido objetivo del genitivo, porque precisamente fue la tradición metafísica la que se ocupó del ser en ese sentido). En esto hay que ser claros: la pregunta de SZ no es qué es el ser, sino que esa fue la pregunta de la metafísica, cuyo decurso, muy radicalmente, Heidegger traza desde Parménides hasta Nietzsche.

¿Qué pregunta, pues, la pregunta planteada en la gran obra heideggeriana? Si queremos partir de ese *Besinnungsebene* o del mismo nivel de meditación de la obra, hay que conceder que nuestra pregunta es inexacta porque, propiamente, la pregunta no pregunta por ‘algo’, no es un ‘qué’ lo que aparece cuestionado. Y la clave para entender este asunto paradójico, aparece en la primera página del libro, el ‘prólogo en el cielo’ (Safranski *dixit*), con que se abre la gran obra. Se trata, a no dudarlo, de un extraño proemio. Pero toda la extrañeza que suscita lo tratado en SZ se debe a que, de alguna forma, es intratable: nos hallamos ante una introducción a lo atemático. Y eso atemático es lo que, justamente, merece el título de ‘lo hermenéutico’: no el ser a secas, ni el ente, sino el ser del ente, es decir, el sentido del ser, que, justamente, no puede aparecer como ‘algo’ ni puede ser tematizado en sentido estricto. De alguna forma – y aquí hacemos recurso a prestaciones gráficas – lo que hace Heidegger no es filosofía ni no-filosofía. Heidegger hace ~~filosofía~~ y no se ocupa el ser, sino del ~~ser~~.

Ha habido un giro fenomenológico-hermenéutico en los estudios cognitivos evidenciado en el grupo de investigadores en el campo de la IA que reclaman para sí un linaje heideggeriano. Pero ello coadyuva más a una crítica de los supuestos apofánticos y teóricos que en pocos años de investigación comenzaron a testimoniar cómo el proyecto de la IA se hallaba en franca degeneración. Motivado por las investigaciones en neurodinámica de Walter Freeman, Dreyfus concluye que tal vez “se puede avizorar una suerte de IA animal inspirada por Heidegger y Merleau-Ponty, pero esto no da pie para creer, y hay muchas razones por las cuales dudarlo, que tal dispositivo fungirá como un primer paso hacia un *continuum* que se dirija al forjamiento de una máquina que pueda estimular el afrontamiento humano con lo que es significativo.” (2007: 268).

Por otra parte, lo que ha acaecido con la irrupción heideggeriana en los estudios cognitivos merece al nombre de nueva ‘urbanización de la provincia heideggeriana’ (*pace* Habermas). Puesto que el ser sea propiamente nada (no nada de nada, sino, como la expresión inglesa *no-thing*, nada en el sentido de que no se trata de *algo*), es ahora empíricamente comprensible: el estar-a-la-mano (*Zuhandenheit*) de lo útil (*Zeughaftigkeit*) que mantiene al Dasein en ocupación circunmundana, no es ningún ente, y como no reside en lo apofántico,

pues es pretemático y preontológico, no puede ser de ninguna forma programable o simulable. Lo que hay en todo afrontamiento es una “ejecución autónoma y expresa de la comprensión de ser que desde siempre pertenece al Dasein y que está ‘viva’ en todo trato con los entes” (SZ, § 15: 67). Esta *Seinsverständnis* es lo hermenéutico y es el “‘en cuanto’ [Als] originario de la interpretación circunspectivamente aprehensora” (SZ, § 33: 158). Empero, hablamos de ‘lo hermenéutico’ (que no de la hermenéutica como una suerte de ‘teoría de la interpretación’). Abordar ‘lo hermenéutico’ por medio de un rodeo óntico en torno al fracaso empírico de la IA, tiene la finalidad de señalar, además de la miseria de lo apofántico en lo que más importa desde el punto de vista filosófico, los peligros de toda pavimentación de un terreno tanto agreste como lleno de trampas: el de ‘lo hermenéutico’. Haríamos bien en precavernos de tematizar lo hermenéutico como si se tratase de un ámbito óntico, es decir, pasible de ser tematizado. Esto quiere decir, finalmente, que una teoría de la interpretación es lo que, de hecho, no hay en Heidegger, pero solamente porque lo que no hay en absoluto, y lo que desde siempre está siendo minado, son los prestigios falsos de lo teórico en cuanto tal.

BIBLIOGRAFÍA

- Agre, Philip (1997) *Computation and Human Experience*. Cambridge: Cambridge University Press.
- A. J. Ayer (comp.) (1981) *El Positivismo Lógico*. Trad. L. Aldama, U. Frisch, C.N. Molina, F.N. Tourner y R. Ruiz Harrel. Madrid: Fondo de Cultura Económica.
- Brooks, Rodney (1999) *Cambrian Intelligence. The Early History of the New AI*. Cambridge, MA: The MIT Press.
- Carnap, Rudolf (1981) “La Superación de la Metafísica Mediante el Análisis Lógico del Lenguaje”. En: A. J. Ayer (comp.) *Op. Cit.*, 66-87.
- Dreyfus, Hubert (1972/1979) *What Computers Can't Do. A Critique of Artificial Reason*. -- Revised edition. New York: Harper and Row.
- _____. (1992) *What Computers Still Can't Do*. Cambridge, MA: The MIT Press.
- _____. (2007) “Why Heideggerian AI Failed and How Fixing it would Require Making it More Heideggerian”. *Philosophical Psychology*. 20 (2), 247-268.
- Figal, Günter (1996) *Der Sinn des Verstehens. Beiträge zur hermeneutischen Philosophie*. Stuttgart: Philipp Reclam.
- Freeman, Walter (2000) *How Brains Make Up Their Minds*. New York: Columbia University Press.
- Habermas, Jürgen (1979) “Urbanisierung der heideggerschen Provinz. Laudatio auf Hans-Georg Gadamer”, 392-401. En: *Philosophisch-politische Profile*. Frankfurt am Main: Suhrkamp. 1981.
- Heidegger, Martin (56/57) *Zur Bestimmung der Philosophie*. [KNS 1919]. Gesamtausgabe Bde. 56/57. Ed. B. Heimbüchel. Frankfurt am Main: Vittorio Klostermann. 1987. [Versión castellana de una de las tres lecciones: *La Idea de la Filosofía y el Problema de la Concepción del Mundo*. Trad. J. A. Escudero. Barcelona: Herder. 2005].
- _____. (SZ) *Sein und Zeit*. [1927]. Tübingen: Max Niemeyer Verlag. 1979. [Versión castellana: *Ser y Tiempo*. Trad. J. E. Rivera. Santiago: Editorial Universitaria. 1998.]
- _____. (H) *Hitos*. Trad. H. Cortés & A. Leyte. Madrid: Alianza.
- Herrmann, Friedrich-Wilhelm von (1987) *Hermeneutische Phänomenologie des Daseins. Eine Erläuterung von »Sein und Zeit«*. Bd. I: *Einleitung: Die Exposition der Frage nach*

dem Sinn von Sein. Frankfurt am Main: Vittorio Klostermann.

_____. (2005) *Hermeneutische Phänomenologie des Daseins. Ein Kommentar zu »Sein und Zeit«*. Bd. II: *Erster Abschnitt: Die vorbereitende Fundamentalanalyse des Daseins*, § 9-§ 27. Frankfurt am Main: Vittorio Klostermann.

Minsky, Marvin (2003) "Why IA is Brain-Dead". *Wired* 11 (8). URL: <http://www.wired.com/wired/archive/11.08/view.html?pg=3> [Consultado 9/4/09].

Safranski, Rüdiger (1997) *Un Maestro de Alemania. Martin Heidegger y su Tiempo*. Trad. R. Gabás. Barcelona: Tusquets.

Strawson, Peter (1997) *Análisis y Metafísica*. Trad. N. Guasch. Barcelona – Buenos Aires – México: Paidós.

Wheeler, Michael (2007) *Reconstructing the Cognitive World: The Next Step*. Cambridge: The MIT Press.