

**Apostillas fenomenológicas a la filosofía de la
inteligencia artificial de Philip E. Agre**
*Phenomenological Remarks to Philip E. Agre's Philosophy
of Artificial Intelligence*

Jethro Masís¹

Universidad de Costa Rica, San José, Costa Rica

 <https://orcid.org/0000-0001-9887-2269>

jethro.masis@ucr.ac.cr

Recibido: 15/11/2023

Aceptado: 09/12/2023

DOI: 10.5281/zenodo.10442462

RESUMEN

Este artículo presenta las relaciones entre la filosofía y la tecnología en la obra de Philip E. Agre, quien propuso una práctica técnica crítica que incorporase a la investigación tecnológica un examen de las construcciones discursivas que subyacen al proyecto de la inteligencia artificial (IA). Se ofrece una exposición de las ideas heideggerianas que Agre pretende hacer valer en la programación de la acción inteligente de agentes artificiales, si bien ulteriormente se criticará el propósito de programar el comportamiento práctico en el mundo circundante, dado que esto no es otra cosa que lo que Heidegger denominó la construcción técnica moderna del ser humano como máquina.

Palabras clave: Filosofía de la ciencia cognitiva, inteligencia artificial, Heidegger, Agre.

ABSTRACT

This article explores the interplay between philosophy and technology in the work of Philip E. Agre. Agre proposed a critical technical practice that integrates an examination of discursive constructs underlying the artificial intelligence (AI) project within technological research. The article presents Agre's utilization of Heideggerian concepts in programming intelligent action of artificial agents. However, the article also critiques the endeavor to

¹ Doctor en Filosofía por la Julius-Maximilians-Universität Würzburg, Alemania. Profesor catedrático en la Escuela de Filosofía y Director del Programa de Posgrado en Filosofía de la Universidad de Costa Rica, miembro de la comisión directiva del Círculo Latinoamericano de Fenomenología y Presidente del Círculo Costarricense de Fenomenología.

program practical behavior in the surrounding world, for this constitutes what Heidegger referred to as the modern technical construction of the human being as a machine.

Keywords: Philosophy of cognitive science, artificial intelligence, Heidegger, Agre.

Introducción

Este artículo se dedica a los supuestos filosóficos de la teoría de la IA del teórico de la información y de la computación Philip E. Agre, quien cesó su actividad como profesor de estudios de información en la Universidad de California, Los Ángeles, después de ser reportado como desaparecido en octubre de 2009. Aunque sus familiares informaron que lo habían encontrado vivo y en buen estado de salud, Agre nunca regresó a la universidad y dejó de escribir para siempre.² Agre fue una figura seminal en el temprano desarrollo de la cultura de internet en los años 1990 y un sólido teórico de la relación entre la ciencia cognitiva y la filosofía. Es un infortunio que con la cesación de su actividad científica se haya desvanecido para siempre la obra de uno de los teóricos más interesantes de la relación entre la ciencia de la computación y las humanidades, cuyo legado todavía se sigue visitando³.

Me propongo presentar algunas ideas concebidas por Agre, en particular las relacionadas con lo que denomina una “práctica técnica crítica” y las posibilidades de hacer que esta práctica tecnológica (la implementación de programas de inteligencia artificial, por ejemplo) sea provista de una conciencia filosófica e histórico-conceptual de la que, según Agre, abiertamente carece. Aunque tengo en alta estima la obra de

² Sobre estos eventos, ver Young, 2009; Travers, 2009; Rothman, 2009; Pescovitz, 2009; Carvin, 2009.

³ Entre los días 01 y 02 de septiembre de 2022, la Universidad de Siegen en Alemania organizó el taller de investigación *Agre After Techno-Utopianism*, en el que tuve el honor de leer mi ponencia “Critical Technical Practice: Philip Agre’s Philosophy of Technology Reconsidered”. Aparte de mi contribución, el taller contó con una nutrida participación internacional con investigadores e investigadoras de Alemania, Holanda, Estados Unidos, Austria e Inglaterra.

Agre, criticaré su idea de que encontrar la implementación técnica adecuada para entender las prácticas cotidianas de agentes inteligentes puede lograrse bajo la rúbrica de programar el estar-a-la-mano (Zuhandenheit) de Heidegger. Huelga decir que el estado de la investigación actual de la IA difiere de la época en la que Agre estuvo activo como investigador. Actualmente, los sistemas artificiales se caracterizan por el uso de técnicas de aprendizaje automático profundo que permiten a los sistemas de IA aprender grandes cantidades de datos sin necesidad de programación explícita. Esto se debe a una serie de factores, que incluyen el aumento de la disponibilidad de datos, el desarrollo de nuevas técnicas de aprendizaje automático y el crecimiento de la potencia de cálculo. Sin embargo, sigue siendo fundamental la idea de Agre de que existen ciertas metáforas que subyacen al trabajo estrictamente técnico que deben tenerse en cuenta puesto que tienen incidencia en la investigación. En último lugar, argumentaré que el carácter abierto de la existencia humana no es susceptible de programación computacional y que, en este sentido, Agre no pudo sustraerse de lo que Heidegger denominó la “construcción técnica moderna del ser humano como máquina” (Zoll, p. 178).

Hacer de la IA un proyecto filosófico

Philip Agre obtuvo su doctorado en Ingeniería Eléctrica y Ciencias de la Computación en el MIT y formó parte del laboratorio de IA comandado por Rodney Brooks. Habría que añadir que, además, fue uno de los ingenieros computacionales más filosóficos de su tiempo. Su punto de partida fue la idea de que era errado asumir la línea cartesiana de pensamiento que habitaba más o menos inconscientemente la investigación en IA como si no tuviese consecuencias prácticas en las

ideas técnicas que luego terminaban implementándose en forma de programas (Agre, 1997, p. 23). Para Agre se trataba exactamente de lo contrario: “los sistemas computacionales son también, entre otras cosas, sistemas filosóficos, específicamente sistemas filosóficos matematizados, y se puede aprender mucho tratándolos de esta manera” (1997, p. 41). Agre afirma que “la IA es filosofía desde sus fundamentos” (2005, p. 155); una afirmación que puede ampliarse en los siguientes puntos:

- Las ideas de IA tienen sus raíces genealógicas en ideas filosóficas.
- Los programas de investigación en IA intentan desarrollar y elaborar los sistemas filosóficos que heredan.
- La investigación en IA regularmente encuentra dificultades y callejones sin salida que provienen de tensiones internas en los sistemas filosóficos subyacentes.
- Estas dificultades y callejones sin salida deben concebirse como señales particularmente informativas sobre la naturaleza y las consecuencias de las tensiones filosóficas que las generan.
- El análisis de estas señales no procede de la investigación técnica propiamente dicha, pero puede coadyuvar con la creación de nuevas agendas técnicas que trasciendan las lindes de la investigación técnica en sí misma (Agre, 2005, p. 155).

Influenciado fuertemente por el influjo de la filosofía de Heidegger en el campo de la ciencia cognitiva llevada a cabo por Dreyfus (1992)⁴, Agre también entiende la obra del pensador alemán como una fenomenología de las actividades rutinarias y cotidianas del Dasein en el

⁴ El mismo Rodney Brooks llegó a afirmar, quizá por influencia de su discípulo Agre, que el trabajo en robótica que se realizaba a la sazón en el laboratorio de IA del MIT, tenía cierta influencia heideggeriana: “in some circles much credence is given to Heidegger as one who understood the dynamics of existence. Our approach has certain similarities to work inspired by this German philosopher...” (Brooks, 1999, 97).

mundo circundante (Umwelt), es decir, como una exploración filosófica de la forma del existir cotidiano. Para Agre, la “analítica del Dasein” heideggeriana puede proporcionar una guía útil en el desarrollo de teorías computacionales de la interacción práctica de agentes artificiales. Una teoría semejante podría otorgar a la implementación tecnológica una conciencia histórica que abiertamente le falta, dado que “investigaciones como las de Heidegger tienen su influencia más productiva en la IA cuando la IA misma recupere un sentido de su propio desarrollo histórico” (Agre 1996, p. 25). Este último rasgo crítico e histórico permite a Agre, como filósofo de la computación y de la ciencia de la cognición, afirmar que las prácticas computacionales pueden verse como la encarnación inconsciente de una concepción desencarnada de la filosofía, donde Agustín, Descartes y Turing actúan como figuras pivotantes, con la oposición de cuerpo y alma en el núcleo de su pensamiento:

[e]l medio cultural de cada uno de ellos le dio un nuevo significado a esta oposición: Agustín luchó por mantener sus ideales de ascetismo cristiano, Descartes describió el alma del soldado superando el miedo de su cuerpo mientras se libraba la Guerra de los Treinta Años, y Turing idealizó el pensamiento desencarnado mientras sufría opresión homofóbica en la Inglaterra moderna (1997, p. 103).

Esto quiere decir que existe una tradición histórica y un discurso que sustentan las prácticas computacionales contemporáneas, por lo que de ningún modo puede decirse que tales prácticas se sostengan exclusivamente en términos técnicos. La idea según la cual es enteramente de segundo orden tener una conciencia clara de la historia de los supuestos filosóficos que sustentan las prácticas tecnológicas concretas es tanto ingenua como deshonesta. Agre lamenta que la ciencia de la computación es las más de las veces inconsciente respecto de su contingencia intelectual, lo que la hace refundirse constantemente en técnica pura (1997, p. 103). Pero cabe la pregunta ¿existe semejante

implementación tecnológica depurada de la historia filosófica de los conceptos de mente, mundo, percepción, representación, conocimiento, etc.? Dado que Agre concibe que no existe tal práctica tecnológica ahistórica, procede a criticar este olvido de los supuestos que se mueven en las profundidades de los programas de la IA. Estos supuestos suelen ser sustituidos por el intento formalista de purificar los programas computacionales de la “inexactitud” del lenguaje natural. Con ello, se despoja a la AI por completo de sus fundamentos históricos y culturales, pero además se crea la ilusión de que una limpieza eliminativista podría de hecho triunfar frente a las impurezas de aquellos supuestos. Además, quienes trabajan en la ciencia computacional suelen estar “convencidos de que, al definir su vocabulario en términos matemáticos rigurosos, pueden dejar tras de sí la red de supuestos que se adhieren a sus palabras a través de la sedimentación de la historia intelectual” (Agre 2002, p. 131). Agre cree que esta forma de trabajo debe sustituirse por una forma mejorada de acometer la implementación tecnológica: la llamada práctica técnica crítica. Se trata de una actitud crítica que se guía por una conciencia en constante desarrollo de la propia forma de trabajo como una práctica históricamente específica y situada (1997, p. 22). Como tal, esta forma de trabajo técnico debe aceptar que una “investigación reflexiva pone en riesgo todos sus conceptos y métodos. Y esto no ha de verse como una amenaza a la racionalidad, sino como una promesa futura de una mejor forma de hacer las cosas” (Agre 1997, p. 23). Para Agre es crucial comprender que no está proponiendo solamente un estudio filosófico externo sobre la IA, o sobre las teorías de la IA, sino “una forma de trabajo legítima de investigación en IA: una intervención dentro del campo que cuestiona muchas de sus ideas básicas, al mismo tiempo que se mantiene fundamentalmente en las coordenadas del modelo computacional como una forma de conocimiento” (1997, p. xiv).

En cuanto a la intervención crítica, Agre señala que hay una determinada mentalidad extendida entre la “gente de computación” respecto a los objetivos y alcance de su propio trabajo. Según Agre, la investigación computacional se puede definir como una investigación sobre la realización física en cuanto tal. Es decir, se trata de diseñar algo, de hacer algo en la realidad, puesto que algo que debe funcionar: el programa debe “correr”. Esto contrasta —diríase— con aquella imagen cómica de Ronald Sousa según la cual la filosofía no es otra cosa que jugar tenis sin una red (cita de Dennett, 1997, p. 252). Para un científico computacional, los dilemas de su campo de trabajo sobrepasan por mucho los problemas conceptuales. Por ello, “lo que realmente caracteriza el trabajo computacional es el sentido constante sobre lo que se puede diseñar y lo que no” (Agre, 1997, p. 11). El lema de los practicantes de la computación es simple: si alguien no lo puede diseñar, entonces tampoco lo entiende. Esta es la razón por la cual los científicos de la computación “desconfían de cualquier cosa a menos que puedan asegurar su diseño de forma precisa; en general, preferirían tener algo preciso y equivocado que vago y correcto” (1997, p. 13). Aquí hay una “ética de trabajo” que puede asociarse con esta mentalidad computacional: el trabajo consiste en diseñar cosas que funcionen.⁵ Sin embargo, Agre considera que esta concepción de “trabajo” es demasiado estrecha, dado que lo que se considera exitoso como un logro técnico deviene ahistórico en el sentido de que simplemente se reduce a que el programa se ajuste a una especificación formal-matemática previamente dada. Es decir, puede que un programa funcione en un micromundo... pero ello no significa en absoluto que se haya comprendido el funcionamiento del sentido en el mundo humano. Empero, para que el trabajo computacional nos brinde

⁵ Agre utiliza acá un juego de palabras entre ética de trabajo (*work ethic*) y hacer que las cosas funcionen (*making things work*).

una comprensión del mundo, es decir, para que la ciencia cognitiva sea realmente ciencia, hace falta que su significado vaya más allá de las estructuras matemáticas. Por ejemplo, cuando se dice que un robot “entiende” una serie de tareas, o cuando se proclama que los sistemas de IA nos brindarán profundos conocimientos sobre los procesos de pensamiento humano, es necesario que esos sistemas artificiales realmente “comprendan”. Agre se refiere a un sentido mucho más amplio de trabajo computacional, que no solo se reduce a su realización matemática, sino que implica también una construcción discursiva. Desde luego, el punto de Agre es que todo trabajo computacional es una mezcla de procesos científicos y técnicos y de construcción discursiva. Y ciertamente vale la pena recordar que la construcción discursiva es parte fundamental de las metas explicativas más básicas de la ciencia cognitiva, y esto es así siempre que entre sus objetivos esté brindar una explicación de los procesos fundamentales de la mente y la cognición. Por lo tanto, en un sentido primordial, la IA no solo sirve para diseñar cosas que funcionen, sino que sus limitaciones técnicas también pueden ser herencia de problemas de comprensión en el ámbito de la construcción discursiva subyacente:

[l]a realidad práctica con la que luchan las personas de la IA en su trabajo no es solo “el mundo”, considerado como algo objetivo y externo a la investigación. Es mucho más complicado que esto. Se trata de un híbrido entre la realidad física y la construcción discursiva. La trayectoria de la investigación en IA puede ser moldeada por las limitaciones del mundo físico, como la velocidad de la luz, las tres dimensiones del espacio, los rayos cósmicos que interrumpen los chips de memoria, y también puede ser moldeada por las limitaciones del mundo discursivo, como el vocabulario disponible, las metáforas y las convenciones narrativas (Agre 1997, p. 15).

Agre opina que el punto no es invocar la filosofía heideggeriana como una autoridad exógena que reemplace los métodos computacionales: “el punto más bien es expandir la práctica técnica de tal manera que la relevancia

de la crítica filosófica sea evidente como un asunto técnico. Los modos técnicos y críticos de investigación deben formar una alianza en esta nueva y ampliada forma de conciencia técnica crítica” (1997, p. xiii). La práctica técnica crítica que imagina Agre es una “en la que tal reflexión sobre el lenguaje y la historia, las ideas y las instituciones, es parte integral del trabajo técnico en sí mismo” (2002, p. 131). Más exactamente, Agre pretende “hacer ciencia, o al menos algo sobre la naturaleza humana, y no solo resolver problemas industriales” (1997, p. 17), pero también está interesado en “aprovechar los poderosos modos de razonamiento involucrados en los procesos de diseño computacional” (idem). De esta manera, Agre pretende rescatar las aseveraciones más ambiciosas de la investigación de IA —como que puede enseñarnos algo sobre el mundo y sobre la cognición humana— mediante la inscripción de un enfoque que combine precisión técnica y rigor filosófico. Al expandir la comprensión de las formas en que funcionan los sistemas artificiales, “la IA puede convertirse en una forma de escuchar la realidad y de aprender de ella” (Agre 2002, p. 141).

En este sentido, Agre argumenta que la simple amalgama de unas cuantas metáforas sobre la mente, asumidas desde la tradición conceptual, con un poco de matemáticas que luego se plasman en maquinaria computacional no constituye una forma de escuchar a la realidad. Lo que habría que hacer es que las mismas metáforas que motivan la implementación tecnológica sean objeto de investigación científica. Por ello, debe tenerse en cuenta tanto la realidad física como la construcción discursiva. Aunque los lenguajes técnicos codifican un proyecto cultural en sí mismo (la redescrición sistemática de los fenómenos humanos y naturales dentro del repertorio limitado de esquemas técnicos que facilitan el control racional), “es precisamente este fenómeno lo que hace especialmente importante investigar el papel de las metáforas en la práctica técnica” (Agre 1997, p. 34). En este punto, el

proyecto de Agre suena muy similar al del filósofo alemán Hans Blumenberg, cuya teoría metaforológica “busca penetrar en la estructura subyacente del pensamiento, en el subsuelo, en la solución nutritiva de las cristalizaciones sistemáticas; pero también tiene como objetivo mostrar con qué desparpajo la mente se anticipa a sí misma en sus imágenes y cómo su historia se proyecta en el vigor de sus conjeturas” (2010, p. 5). Para Agre también las metáforas desempeñan un papel en la organización de la investigación científica o, por decirlo con tonos blumenbergianos, las metáforas no son “elementos residuales” (Restbestände) sino más bien “elementos fundamentales” (Grundbestände) del discurso científico. En términos de Thomas Kuhn (1996), esto también puede enmarcarse tomando en cuenta la tensión entre la ciencia normal —con su actitud aséptica hacia la reducción de la inestabilidad e inconsistencia del significado a través de una limpieza de elementos de naturaleza inexacta y ambigua— y la ciencia revolucionaria, que hace saltos metafóricos con el fin de crear nuevos significados y aplicaciones que pueden constituir un progreso teórico genuino (Arbib & Hesse 1987, p. 157). Al mostrar cómo la práctica técnica no es solo el resultado de trabajo computacional sino también de la construcción discursiva y de supuestos metafóricos, la práctica técnica crítica de Agre podría constituirse en una revolución en el campo de la cognición en sentido kuhniano. Sin embargo, Agre solía ser confuso para sus colegas, a quienes les daba la impresión de estar malgastando tiempo valioso cuando se interesaba en abstrusas teorías filosóficas.

Como roboticista, Agre busca un enfoque computacional que investigue a profundidad las interacciones del agente con su entorno. Esta noción de entorno como mundo vital tiene claras resonancias husserlianas (Husserl, 1970). De ahí que el concepto de mundo de la vida (Lebenswelt) deba incorporarse al diseño de programas computacionales “para describir las estructuras de los mundos de vida y las formas en que

éstos simplifican computacionalmente toda actividad” (Agre y Horswill, 1997, p. 111). Para Agre es fundamental tener en cuenta los conceptos de corporización (embodiment) y de absorción (embedding), puesto que los agentes no solo deben concebirse como corporizados, sino más crucialmente como absorbidos en un entorno vital. La distinción entre corporización y absorción se puede explicar de la siguiente manera:

[l]a corporización se refiere a la vida de un agente como cuerpo: la finitud de sus recursos, su perspectiva limitada sobre el mundo, la indexicalidad de sus percepciones, su localización física, su movilidad, etc. La absorción se refiere a la relación estructural del agente con su mundo: sus caminos habituales, sus prácticas cotidianas y cómo estas encajan con las formas y funcionamiento de las cosas, sus conexiones con otros agentes, su posición en un conjunto de roles o una jerarquía, etc. El concepto de absorción, entonces, abarca desde tipos más concretos de ubicación en el mundo (lugares, cosas, acciones) hasta tipos más abstractos de ubicación (dentro de sistemas sociales, ecosistemas, culturas, etc.). La corporización y la absorción están obviamente interrelacionadas, y cada una tiene consecuencias importantes tanto para los tratos directos de los agentes con otros agentes como para sus actividades de solidaridad en el mundo físico (Agre & Horswill 1997, 111-112).

No se debe subestimar la importancia que tiene para la ciencia cognitiva el contar con un concepto bien desarrollado del entorno (Umwelt, en sentido heideggeriano), debido a que solo con una comprensión adecuada del entorno de un agente se pueden entender sus patrones adaptativos de comportamiento. Avanzar hacia la definición del entorno con al menos un mínimo de rigor equivale a desarrollar “una teoría positiva del entorno, es decir, algún tipo de caracterización de principio de aquellas estructuras o dinámicas u otros atributos del entorno en virtud de los cuales el comportamiento es adaptativo” (Agre y Horswill 1997, p. 113). En consecuencia, Agre y Horswill lamentan que la IA haya minimizado la distinción entre agente y entorno reduciendo este último a una serie discreta de elecciones en el proceso de la resolución de problemas, porque los programas de resolución de problemas sirven para “modelar tareas

como la demostración de teoremas lógicos y el ajedrez, en los que los objetos que se manipulan son puramente formales” (idem). Y estos problemas son, para Agre, meros divertimentos (toy problems), como los micro mundos y las actividades dentro de esos entornos artificiales. La situación cambia drásticamente cuando nos encontramos con actividades que involucran corporización y absorción, donde “el mundo se presenta, por decirlo así, fenomenológicamente: en términos de las diferencias que hacen la diferencia para un agente, dadas sus representaciones, acciones y metas particulares” (idem). Esta indexicalidad ambiental que se presenta aquí a menudo es objetada por los cognitivistas arguyendo que bajo ese modelo pareciera que los agentes realizan acciones sin ningún tipo de “cálculo mental”, o como si los agentes que habitan un mundo de vida vivieran en un modo meramente reactivo a los impulsos. Pero el punto es que “la cognición no trivial que un agente realiza tiene lugar en un trasfondo considerable que está estructurado por una dinámica de familiaridad” (Agre y Horswill 1997, p. 118). De hecho, la indexicalidad ha sido difícil de admitir dentro de la investigación en IA. Según Agre, hay una serie de supuestos filosóficos en el trabajo computacional que deben ser puestos en entredicho:

- Que la percepción es una especie de óptica inversa que construye un modelo mental del mundo partiendo de impresiones sensoriales.
- Que la acción se lleva a cabo mediante la ejecución de constructos mentales denominados planes, entendidos como programas computacionales. Y finalmente,
- Que el conocimiento consiste en un modelo del mundo, formalizado en términos de la teoría platónica de análisis del significado en la tradición de Frege y Tarski (2002, p. 132).

La disociación de la mente y el cuerpo (la metáfora fundadora de la ciencia cognitiva y la filosofía moderna) está en acción aquí precisamente cuando la IA tradicional piensa en la mente como un generador de planes y el cuerpo como el ejecutor del plan. Además, la IA como proyecto de investigación se enmarca en términos de una serie de disociaciones: mente vs mundo, actividad mental vs percepción, planes vs comportamiento, mente vs cuerpo e ideas abstractas vs cosas concretas (Agre 2002, p. 132). Según Agre, estas disociaciones son contingentes y pueden considerarse “errores de inscripción” (Smith 1996), es decir, la práctica de “inscribir un discurso en un artefacto y luego darse la vuelta y descubrirlo allí” (Agre 2002, p. 130). En este sentido la IA:

ha heredado ciertos discursos de esa historia sobre temas como la mente y el mundo, y ha inscrito esos discursos en el diseño de máquinas computacionales. El objetivo de este tipo de construcción de modelos técnicos es la clarificación conceptual y la evaluación empírica y, sin embargo, la IA no ha logrado ni clarificar ni evaluar los conceptos que ha heredado. Por el contrario, al intentar trascender la historicidad de su lenguaje heredado ha empañado su propia conciencia de las tensiones internas que contiene ese lenguaje. Las tensiones han seguido estando presentes de forma subterránea, y siguen emergiendo a través de supuestos sustantivos, ambigüedades lingüísticas, equivocaciones teóricas, impasses técnicos y confusiones ontológicas. (Agre 2002, p. 141)

Estar-a-la-mano vs estar-a-la-vista

A pesar de su notable agudeza filosófica, Agre mismo no ha logrado liberarse de la persistencia de una teoría representacional de la cognición en su pensamiento, incluso cuando su teoría de la cognición es más consciente de la historia conceptual de la que es deudora. Eso sí, su concepto de representación pretende partir de la alianza entre la corporización y la absorción con el propósito de comprender la interacción del agente con el mundo circundante. Así, Agre presenta una teoría de lo que llama representación indexical o deíctica (Agre & Chapman 1987;

Agre 1997), cuya idea principal es que los agentes representan objetos de manera genérica a través de sus relaciones con ellos (Agre & Horswill 1997, p. 118). Según Agre, se debe perfeccionar el concepto de representación (y no simplemente descartarlo) y mostrar qué tipo de actividad representacional está en juego en la interacción del agente con su entorno circundante. En este sentido, el objetivo de Agre es criticar el concepto tradicional de representación que subyace a la idea del conocimiento como imagen, copia, o reflejo, porque “las formas primordiales de representación se entienden mejor como facetas de patrones de interacción con el mundo físico y social que se extienden en el tiempo” (Agre 1997, p. 222). Por ello, “la noción de representación debe someterse a una cirugía dolorosa para que todavía tenga un uso” (Agre 1997, p. 250).

La teoría tradicional de representación que subyace como supuesto profundo en la investigación tradicional de la AI se basa en la noción de modelo del mundo. Se supone que existe algún programa dentro de la mente que representa el mundo exterior mediante una correspondencia sistemática con él (Agre 1997, p. 223). En cuanto tal, el supuesto de que existe un modelo del mundo que está siendo representado por la mente es el epitome del mentalismo (Agre 1997, p. 225). El mentalismo es definido por Agre como la metáfora generalizada de la ciencia cognitiva según la cual cada ser humano tiene un espacio interior abstracto llamado “mente”, que solo es comprensible a partir de la dicotomía entre un adentro y un afuera (Agre 1997, p. 49). Agre considera que otorgar preeminencia a la indexicalidad equivale a una inversión de esta imagen, ya que conceder que los seres humanos no son mentes que controlan cuerpos implica que la interacción no puede ser definida “en términos de las relaciones entre una mente, un cuerpo y un mundo externo” (1997, p. 234); imagen que constituye el modelo típico de las teorías contemporáneas de la ciencia cognitiva. El término clave es precisamente

interacción, la cual no debe ser entendida como la relación entre lo subjetivo y lo objetivo, sino más bien como lo emergente de las prácticas que los agentes realizan in situ. La indexicalidad no sería entonces un fenómeno pasivo de la dependencia a un contexto, sino un fenómeno activo de la constitución del mismo contexto (Agre 1997, p. 233).

La investigación indexical no se ocupa de meros procesos físicos indiferenciados que acaecen en el universo, sino de los elementos constituyentes que dan cuenta de las actividades humanas que ahora “deben describirse en términos intencionales, como siendo sobre y hacia cosas, y no como desplazamientos indiferentes de la materia. La descripción física y la intencional no son incomparables, pero son incommensurables” (Agre 1997, p. 245). De esto se sigue que la atribución típica de estados intencionales a sistemas no absorbidos está plagada de problemas, precisamente porque la absorción y la interacción que deriva de ella es la condición de posibilidad del comportamiento intencional. El carácter ubicuo de la experiencia sugerido en el libro clásico de Chalmers (1996) sobre la filosofía de la conciencia también es un error de inscripción, ya que su base consiste en obviar la necesidad de una teoría adecuada de la intencionalidad. Para ser más exactos, el supuesto fundamental de Chalmers es la naturalización de la intencionalidad, es decir, su indiferenciación en relación con los procesos físicos de la materia. Cuando se obvia la necesidad de contar con una teoría de la acción intencional basada en los conceptos de entorno, corporización, absorción y agencia, la representación ya no está conectada con un mundo de la vida, sino que puede reducirse a procesos indiferentes que acaecen físicamente en el cosmos. Sin embargo, “cuando un hablante usa un término indexical como ‘yo’, ‘tú’, ‘aquí’, ‘allí’, ‘ahora’ o ‘entonces’ para señalar un referente específico, esta selección está determinada por las relaciones entre situaciones” (1997, p. 233). Por tanto, la manera en que estas interacciones dan forma a las situaciones debe ser clarificada, ya

que se puede decir que “la interacción es central, tanto para la vida humana como para la vida de cualquier agente de cierta complejidad” (Agre 1997, p. 234). Las actividades absorbidas deben ser investigadas en su estructura intrínseca, así como el tipo de representación que pertenece a ellas. De ninguna forma ha de suponerse que las interacciones de los agentes intencionales pueden simplemente ser objeto de la física.

La alternativa heideggeriana

A Terry Winograd le parecía una ironía el que el laboratorio de IA del MIT se convirtiera por un tiempo en una “cuna de heideggerianos”, lo cual es manifiesto cuando Agre propone una teoría de la intencionalidad basada en la distinción entre estar-a-la-mano (Zuhandenheit) y estar-a-la-vista (Vorhandenheit); conceptos que forman parte de la descripción fenomenológica heideggeriana del estar-en-el-mundo (SZ § 15).⁶ La motivación crítica detrás de este enfoque es que la investigación tradicional en IA solo ha prestado atención a los fenómenos objetivantes, del estar-a-la-vista, y ha intentado modelar computacionalmente lo que destaca objetivamente en la percepción. Para Agre, la distinción fenomenológica entre Zuhandenheit y Vorhandenheit no es psicológica ni mecanicista, sino que posibilita una descripción de la estructura de la experiencia cotidiana que podría ser adecuada para una nueva forma de modelar computacionalmente esa misma experiencia. Preston (1993) ya había explorado esta distinción heideggeriana en relación con la interacción entre fenómenos no representacionales y aquellos representacionales. Y Dreyfus (2002a; 2002b) identificó respectivamente

⁶ La “alternativa heideggeriana” es una expresión de Beth Preston (1993). Por su parte, la idea hiperbólica de que modelos heideggerianos dominaron por cierto tiempo proyectos sobre la interacción de agentes artificiales en el laboratorio de IA del MIT es de Winograd (1995). Para más detalles sobre la llamada ‘AI heideggeriana’, ver Masís, 2020.

Vorhandenheit con la intencionalidad representacional y Zuhandenheit con una especie de intencionalidad no representacional, lo cual le permitió proclamar de antemano el fracaso de los sistemas artificiales que proponían la realización de una inteligencia de alto nivel. Para Agre, esta postura es demasiado radical y, sobre todo, demasiado pesimista. Lo que se necesita es una clarificación de qué tipos de representación existen y cuál es su papel en las actividades concretas (Agre 1997, p. 237). Aquí reside la importancia de profundizar en la experiencia absorbida y proporcionar a la IA un conjunto de herramientas para enriquecer su vocabulario y sus metáforas. Esto es necesario porque “la filosofía que informa la investigación en IA tiene un vocabulario fenomenológico claramente empobrecido, no va más allá de distinguir entre estados mentales conscientes e inconscientes” (Agre 1997, p. 239). Agre apunta a algo más importante aquí, nada menos que hacer que la IA vuelva a ser filosófica. De hecho, “la IA en la actualidad es filosofía encubierta; el punto es hacerla abiertamente filosófica” (1997, p. 240).

La idea tradicional de representación la entendía como un modelo en la mente de un agente que correspondía al mundo exterior a través de un mapeo sistemático. Agre opina que la investigación en IA se ha preocupado solo por una visión parcialmente articulada de la representación. No es de extrañar, entonces, que el significado de las representaciones para un agente pueda determinarse sin que se proporcione ninguna referencia a la ubicación, actitudes, intereses y perspectiva idiosincrásica del agente. Esta es también la razón que explica por qué “la indexicalidad ha estado casi completamente ausente de la investigación en IA” (Agre 1997, p. 241). La razón por la que la investigación en IA ha quedado atrás en cuanto a la comprensión de la representación y de la indexicalidad no ha sido una falta de distinción entre mecanismo físico e intencionalidad humana. A pesar del importe constante de conceptos de la filosofía fenomenológica, Agre no objeta la

idea de Chalmers de que la experiencia es ubicua, de forma que se puede hablar de la intencionalidad de los artefactos. Agre sostiene que todo dispositivo que participe en cualquier tipo de interacción con su entorno exhibe cierto tipo de indexicalidad (1997, p. 241). Los artefactos tienen cierto tipo de ambiente incorporado. Por ejemplo, “la lectura de un termómetro no indica abstractamente ‘la temperatura’, ya que es la temperatura en algún lugar, ni tampoco indica concretamente ‘la temperatura en el cuarto 11’, ya que si lo movemos al cuarto 23 pronto indicará la temperatura en el cuarto 23 en su lugar. En vez de esto, necesitamos entender el termómetro como indicando ‘la temperatura aquí’, independientemente de si los diseñadores del termómetro pensaron en esos términos” (idem). De esta forma, “la investigación en IA necesita una teoría de la intencionalidad que permita un pensamiento claro sobre las formas en que los artefactos pueden estar involucrados en actividades concretas en el mundo” (1997, p. 242). Por ello, Agre no parece sentirse confundido con la pregunta de Chalmers: ¿qué se siente ser un termostato? (Chalmers, 1996, p. 293).

Con todo, esta teoría de la intencionalidad fue acuñada por Agre bajo la rúbrica de colocar la representación deíctica en lugar de la representación objetiva. Agre piensa que hay que distinguir dos tipos de ontología. Según una ontología objetiva, los individuos pueden definirse sin referencia a la actividad o a los estados intencionales. Pero además habría una ontología deíctica que se define en términos indexicales y funcionales y en relación a la ubicación, posición social, objetivos y intereses actuales y perspectiva autóctona de un agente (Agre 1997, p. 243). Las entidades que entran en el espacio de cualquier interacción con el agente solo se pueden entender correctamente en términos de los roles que desempeñan en las actividades del agente. Según esto, “algunos ejemplos de entidades deícticas son la-puerta-que-estoy-abriendo, el-semáforo-al-que-estoy-acercándome, el-sobre-que-estoy-abriendo y la-

página-que-estoy-girando. Cada una de estas entidades es indexical porque desempeña un papel específico en alguna actividad en la que estoy involucrado; no son objetivas, porque se refieren a diferentes puertas, semáforos, sobres y páginas en ocasiones diferentes” (idem). Que su carácter no sea objetivo no implica, sin embargo, que las actividades indexicales deban ser consideradas como subjetivas y, por ende, como fantasías o cualidades internas o introspectivas. Una ontología deíctica no debe ser confundida con las divagaciones subjetivas y arbitrarias de un sujeto encapsulado de forma solipsista. En primer lugar, esta es la ontología que más adecuadamente se puede atribuir a las actividades rutinarias. Por lo tanto, sería absurdo sugerir que son privadas o inefables. Las actividades rutinarias se realizan “afuera” en el mundo y, por esa misma razón, no pertenecen a un juego mental interno: son, de hecho, públicas. En consecuencia, en las actividades rutinarias, el carácter objetivo de las entidades con las que se trata no es lo importante. Tampoco lo es su “sensación subjetiva” ni la forma en que se me aparecen a mí como individuo. Que su carácter sea deíctico significa que lo más importante es el papel que desempeñan en el conjunto de la actividad. Por ello, frases compuestas con guión como el-coche-que-estoy-conduciendo o la-taza-de-café-con-la-que-estoy-bebiendo no son símbolos mentales en el sentido cognitivista. No designan “un objeto particular en el mundo, sino más bien un papel que un objeto podría desempeñar en un patrón determinado de interacción entre un agente y su entorno” (Agre 1997, p. 251).

Como cuestión de ingeniería, el principio fundamental es el de la parsimonia de la maquinaria, en el sentido de que habría que “escoger la maquinaria más simple que sea compatible con la dinámica conocida” (Agre 1997, p. 246). Esta visión contrasta explícitamente con el énfasis en la representación explícita típica de la AI tradicional, con todas las dificultades inherentes de programar de antemano, como planes (o

scripts), todas las situaciones que un agente artificial podría encontrarse al tratar con el mundo. En claro contraste con la AI tradicional, “el principio de parsimonia de la maquinaria sugiere dotar a los agentes con el conocimiento mínimo requerido para explicar la dinámica de su actividad” (Agre 1997, p. 249). De esta manera, el enfoque de Agre también tiene claros tonos brooksianos (Brooks 1999) en cuanto a la necesidad de eliminar en la robótica términos como “inteligencia” e incluso “razón” para, por el contrario, dedicar la investigación exclusivamente al diseño de la representación interactiva. Agre cree que la asunción de ideas tradicionales sobre la mente es profundamente problemática, debido a que “mientras que un sistema de metáforas subyacente no sea reconocido, toda dificultad en el trabajo técnico será interpretada como un problema técnico y no como la manifestación de los síntomas de un problema más profundo y sustantivo” (1997, p. 260).

Como caso ejemplar de trabajo técnico basado en una conciencia conceptual crítica, Agre presentó Pengi, un programa diseñado por Chapman y Agre (1987) a finales de la década de 1980 con el propósito de ser un programa computacional que diera cuenta de la actividad de un agente. Pengi es un pingüino retratado en el juego de computadora comercial Pengo, que se encuentra en un laberinto hecho de bloques de hielo que está rodeado por una cerca eléctrica. El laberinto también está habitado por abejas asesinas que deben ser evitadas a toda costa por Pengi y la tarea del jugador es mantener a Pengi vivo y protegerlo de los peligros que surjan en el camino. Como defensa, las abejas pueden ser eliminadas aplastándolas con un bloque de hielo en movimiento o pateando la cerca mientras están tocándola. Esto aturde momentáneamente a las abejas y pueden ser aplastadas simplemente caminando sobre ellas. Agre argumenta que Pengo es una mejora en la teoría computacional del mundo, aunque obviamente no logra capturar numerosos elementos de la actividad humana. Lo importante es la

combinación de la direccionalidad de objetivos y la improvisación que habita en el juego, de las cuales se pueden sacar conclusiones computacionales. En primer lugar, Agre y Chapman no intentaron implementar de antemano todo lo que sabían sobre el juego, lo que contradice la planificación previa típica en los sistemas de IA tradicionales. La idea es ver a Pengi en relación con los objetos que aparecen en su mundo, no en términos de su semejanza con modelos mentales que se programaron previamente, sino únicamente en términos de los roles que desempeñan en la actividad en curso. De esta forma, lo que Agre y Chapman intentaron programar fueron en realidad representaciones deícticas: el-bloque-de-hielo-que-estoy-pateando, la-abeja-que-estoy-atacando, la-abeja-del-otro-lado-de-este-bloque-de-hielo-junto-a-mí, etc.

Desde luego, Agre no argumenta que este sistema simple pueda ser considerado como inteligente: “Pengi no entiende lo que está haciendo. Ninguna computadora ha entendido en grado significativo lo que estaba haciendo” (1997, p. 301). Pero la idea es que el mundo del juego que constituye a Pengi como agente no está compuesto por entidades y procesos meramente objetivos, a la vista o vorhanden, sino por posibilidades de acción que requieren respuestas adecuadas por parte del agente. Las cosas para Agre no son entonces objetos con cualidades objetivas en primera instancia, sino posibilidades de acción que requieren respuestas subsiguientes a las demandas de la situación actual. Dado que estas posibilidades de acción no son objetos en absoluto es comprensible que haya sido difícil programarlas. Al fin y al cabo, ¿cómo se pueden programar posibilidades de acción que son latentes y que, de cierta forma, no existen en general, sino que se generan según las demandas de una situación dada?

Reflexión final

No cabe duda de que Agre estuvo influenciado por la crítica temprana de Dreyfus a la razón artificial (Dreyfus 1992). A diferencia de Dreyfus, sin embargo, Agre era un científico de la computación, por lo que su conciencia teórica estuvo marcada por el propósito de poner en práctica sus ideas sofisticadas en el diseño de programas computacionales. Dreyfus lo ha elogiado por su conciencia filosófica de los problemas que habitan en el proyecto de la IA e incluso por haber comprendido mejor el fenómeno del estar-a-la-mano (*Zuhandenheit*) que él mismo (Dreyfus, 2007, p. 252). Según Dreyfus, Agre pudo mostrar que Heidegger tenía entre manos algo más fundamental con su concepto de *Zuhandenheit* que simplemente una clase de objetos: los útiles (*Zeuge*). El punto del carácter del entramado de útiles (*Zeugzusammenhang*) en Heidegger no era trazar diferencias entre algunos entes prácticos, sino mostrar que las cosas con las que nos ocupamos en la absorción en el mundo circundante abren posibilidades para la acción, solicitudes para actuar y una serie de motivaciones que constituyen la existencia. Con todo, es posible que al intentar programar la *Zuhandenheit* Agre haya sucumbido a la objetivación abstracta de la vida humana, ya que las solicitudes —en tanto no son estrictamente objetivas ni subjetivas— no son aptas para la programación. Que no lo sean no es algo que Agre parezca comprender por completo, y por eso piensa que de alguna manera las representaciones deícticas han de estar involucradas en la comprensión humana. Según Dreyfus,

la AI heideggeriana de Agre no intentó programar este aspecto experiencial de ser atraído por una posibilidad. Más bien, con sus representaciones deícticas, Agre objetivó tanto las funciones como su relevancia situacional para el agente. En Pengi, cuando un cubo de hielo virtual definido por su función está cerca del jugador virtual, una regla dicta la respuesta (por

ejemplo, propinarle un puntapié). Ahí no hay ninguna habilidad y no tiene lugar ningún aprendizaje (2007, p. 253)

Debe admitirse que un mundo virtual no es ni siquiera ligeramente comparable con las dinámicas complejas del mundo real. En un mundo virtual, las dinámicas de relevancia se determinan de antemano, por lo que un programa como Pengi simplemente no puede tener en cuenta la forma en que los seres humanos se enfrentan a nuevas relevancias. Dreyfus concluye que Agre “afina el problema del marco, mas no lo resuelve” (2007, p. 253).

La contribución de Agre consiste en su intento de programar *Zuhandenheit* en lugar de *Vorhandenheit*. Sin embargo, que esto se pueda hacer es altamente controversial. Ciertamente, lo que es profundamente controvertido no es que los insights fenomenológicos puedan aplicarse a la ciencia cognitiva para una práctica técnica crítica como la que exige Agre, sino más bien el supuesto de que la dimensión experiencial que la fenomenología ha revelado puede ser programable. Según Heidegger, “la esencia del *Dasein* reside en su existencia” (SZ § 9, p. 42), lo que no implica ontológicamente el mismo sentido que estar-ahí (*Vorhandensein*). Agre ha intentado programar actividades rutinarias y, al hacerlo, ha pragmatizado al *Dasein*. Pero Heidegger advirtió específicamente contra esta forma de entender su filosofía, que la reduce a mera actividad práctica y cotidiana:

Intenté en *Ser y tiempo* brindar una caracterización preliminar del fenómeno del mundo interpretando la manera en que nos movemos inmediata y regularmente en nuestro mundo cotidiano. Allí partí de lo que está a mano en el mundo cotidiano, de aquellas cosas que usamos y de las que nos ocupamos, de tal forma que en realidad no conocemos el carácter peculiar de esa actividad cuando intentamos describirla, puesto que la interpretamos erróneamente al aplicarle conceptos y preguntas que tienen su origen en otro lugar. Aquello que es tan cercano e inteligible para nosotros en nuestros tratos cotidianos es en realidad fundamentalmente remoto e incomprensible. A través de esta caracterización inicial del fenómeno del mundo, la tarea es avanzar y señalar el fenómeno del mundo

como un problema. Sin embargo, nunca se me ocurrió intentar afirmar o demostrar con esta interpretación que la esencia del ser humano consiste en el hecho de que sabe cómo manejar cuchillos y tenedores o usar el tranvía. (GA 29/30, pp. 262-263)

En conclusión, debe señalarse que la recepción de la filosofía de Heidegger en la ciencia cognitiva, no es solo la puesta en práctica de sus ideas fenomenológicas, sino más a menudo su traducción a términos del lenguaje cognitivo. Con una traducción semejante, se pierde fundamentalmente algo fenomenológico. El trabajo de Agre discutido aquí es un ejemplo destacado de cómo la recepción de Heidegger es más bien analítica: “analítica” no solo en el sentido de que la filosofía heideggeriana es apropiada por filósofos formados en la corriente analítica anglo-americana o, en este caso, por un ingeniero formado en el MIT, sino también en el sentido decisivo de que la filosofía heideggeriana que es apropiada con el fin de enriquecer el paradigma de la cognición, solo presta atención a partes específicas de la primera sección de *Sein und Zeit*; partes que, del mismo modo, también se apropian muy selectivamente. La recepción es “analítica” en el sentido de que constituye una versión muy esquemática de Heidegger, tomando precisamente su pensamiento fuera de contexto (Rehberg 2012, p. 160).

Sea de ello lo que fuere, el trabajo de Agre está lleno de ideas relevantes, la más importante de las cuales es su exigencia de que la práctica técnica tenga conciencia de las metáforas filosóficas que permean la investigación en la implementación tecnológica.

Referencias

- Agre, P. (1994). Surveillance and Capture. Two Models of Privacy. *Information Society*, 10(2), 101-127.
- Agre, P. (1996). Computational Research on Interaction and Agency. En P. Agre & S. Rosenschein (eds.) *Computational Theories of Interaction and Agency*. Cambridge, MA/London: The MIT Press, 1-52.
- Agre, P. (1997). *Computation and Human Experience*. Cambridge/New York: Cambridge University Press.
- Agre, P. (1998). Beyond the Mirror World: Privacy and the Representational Practices of Computing. En P. Agre & M. Rotenberg (eds.) *Technology and Privacy: The New Landscape*. Cambridge, MA/London: The MIT Press, 29-61.
- Agre, P. (2002). The Practical Logic of Computer Work. En M. Scheutz (ed.) *Computationalism: New Directions*. Cambridge, MA/London: The MIT Press, pp. 129-142.
- Agre, P. (2005). The Soul Gained and Lost: Artificial Intelligence as a Philosophical Project. En S. Franchi & G. Güzeldere (eds.) *Mechanical Bodies, Computational Minds. Artificial Intelligence from Automata to Cyborgs*. Cambridge, MA/London: The MIT Press, pp.153-173.
- Agre, P. & Chapman, D. (1987). Pengi: An Implementation of a Theory of Activity. *Proceedings of the Sixth Annual Meeting of the American Association of Artificial Intelligence*. Seattle: Morgan Kaufmann, 268-272.
- Agre, P. & Chapman, D. (1988). What are plans for? MIT AI Laboratory Memo, 1050a.

- Agre, P. & Horswill, I. (1997). Lifeworld Analysis. *Journal of Artificial Intelligence Research*, 6(1), 111-145.
- Arbib, M. & Hesse, M. (1987). *The Construction of Reality*. Cambridge/New York: Cambridge University Press.
- Barwise, J. & Perry, J. (1983). *Situations and Attitudes*. Cambridge, MA/London: The MIT Press.
- Blumenberg, H. (2010). *Paradigms for a Metaphorology*. Trans. by R. Savage. New York: Cornell University Press.
- Brooks, R. (1999). *Cambrian Intelligence. The Early History of the New AI*. Cambridge, MA: The MIT Press.
- Carvin, A. (2009). The Mysterious Disappearance of Phil Agre. NPR. November 24, 2009. Recuperado de http://www.npr.org/blogs/alltechconsidered/2009/11/the_mysterious_disappearance_o.html
- Carvin, A. (2010) Missing Pioneer Phil Agre is Found Alive. NRP. Recuperado de [alltechconsidered/2010/01/missing_internet_pioneer_phil.html](http://www.npr.org/blogs/alltechconsidered/2010/01/missing_internet_pioneer_phil.html)
- Chalmers, D. (1996). *The Conscious Mind. In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- Dennett, D. (1997). Cog as a Thought Experiment. *Robotics and Autonomous Systems*, 20(2-4), 251-256.
- Dreyfus, H. (1992). *What Computers Still Can't Do*. Cambridge, MA: The MIT Press.
- Dreyfus, H. (2002a). Intelligence Without Representation: Merleau-Ponty's Critique of Mental Representation. *The Relevance of Phenomenology to Scientific Explanation. Phenomenology and the Cognitive Sciences*, 1(4), 367-383.
- Dreyfus, H. (2002b). Refocusing the Question: Can There Be Skillful Coping Without Propositional Representations or Brain

- Representations? *Phenomenology and the Cognitive Sciences*, 1(4), 413-425.
- Dreyfus, H. (2007). Why Heideggerian AI Failed and how Fixing it would Require making it more Heideggerian, *Philosophical Psychology*, 20(2), 247-268.
- Heidegger, M. (SZ). *Sein und Zeit*. Tübingen: Max Niemeyer Verlag, 1979.
- Heidegger, M. (GA 29/30). *Die Grundbegriffe der Metaphysik. Welt - Endlichkeit - Einsamkeit*. Gesamtausgabe Bd. 29/30. Ed. by F.-W. von Herrmann. Frankfurt am Main: Vittorio Klostermann, 1983.
- Heidegger, M. (Zoll). *Zollikoner Seminare. Protokolle-Zwiesgespräche-Briefe*. Ed. by M. Boss. Frankfurt am Main: Vittorio Klostermann, 1994.
- Husserl, E. (1970). *The Crisis of European Sciences and Transcendental Phenomenology. An Introduction to Phenomenological Philosophy*. Trans. by D. Carr. Evanston: Northwestern University Press.
- Kuhn, T. (1996). *The Structure of Scientific Revolutions*. Chicago: The Chicago University Press.
- Marres, R. (1989). *In Defense of Mentalism. A Critical Review of the Philosophy of Mind*. Amsterdam: Rodopi.
- Masís, J. (2020). A Sceptre of Wooden Iron: Heidegger and Cognitive Science. *The Philosopher*, 108(2), 70-74.
- Pescovitz, D. (2009). Missing: Phil Agre, Internet Scholar. *Boing Boing*, November 24, 2009. Recuperado de <http://boingboing.net/2009/11/24/missing-phil-agre-in.html>
- Preston, B. (1993). Heidegger and Artificial Intelligence. *Philosophy and Phenomenological Research*, LIII(1), 43-69.
- Rehberg, A. (2012). Heidegger and Cognitive Science. *Aporetic Reflections*. En J. Kiverstein & M. Wheeler (eds.) *Heidegger and Cognitive Science*. Basingstoke/New York: Palgrave Macmillan, 157-175.

- Rothman, D. (2009). Missing: Phil Agre, Online Pioneer, Vanishes from Net and Life. TeleRead, November 25, 2009. Recuperado de <http://www.teleread.com/ebooks/missing-phil-agre-online-pioneer-vanishes-from-net-and-life/>
- Sartre, J.-P. (1984). Being and Nothingness. A Phenomenological Essay on Ontology. Trans. by H. Barnes. New York/London: Washington Square Press.
- Sengers, P. (1995). From the Belly of the Devil: Critical Theory in Scientific Practice. *Parallax*. 2(1), 151-159.
- Smith, B. C. (1996). On the Origin of Objects. Cambridge, MA/London: The MIT Press.
- Travers, M. (2009). Phil Agre, an Appreciation. *Omniorthogonal*, November 21, 2009. Recuperado de <http://omniorthogonal.blogspot.de/2009/11/phil-agre-appreciation.html>
- Winograd, T. (1995). Heidegger and The Design of Computer Systems. En Feenberg, A. & Hannay, A. (ed.) *Technology and the Politics of Knowledge*. Bloomington: Indiana University Press, pp. 108-127.
- Young, J. (2009). Friends and Colleagues Mount a Search for a Missing Scholar, Philip Agre. *The Chronicle of Higher Education*, November 19, 2009. Recuperado de <http://chronicle.com/article/FriendsColleagues-Search/49222/>