

Método
Grupo Transdisciplinario de
Investigación en Ciencias Sociales
www.grupometodo.org

Borradores de Método

Área de Metodología
ISSN: 1692-9667

**El problema estructural de la I.A.
según Dennett**

Eduardo Salcedo Albarán

Documento 15
Agosto 27 de 2003

BORRADORES DE MÉTODO es un medio de difusión de las investigaciones del Grupo Método. Estos documentos son de carácter provisional, de responsabilidad exclusiva de sus autores y sus contenidos no comprometen a la institución.

Borradores de Método. No 15. Agosto 27 de 2003.
Editor Fundación Método
Colección Metodología

© Eduardo Salcedo Albarán
El problema estructural de la I.A. según Dennett

© Fundación Método. Carrera 8ª. No 37-10. ofi: 501. Telefax: (571) 4005765, Bogotá,
Colombia.

2003. Todos los derechos reservados.

Primera edición 2003.

Impreso en Colombia.

El problema estructural de la I.A. según Dennett

Eduardo Salcedo Albarán

esalcedo@grupometodo.org

Resumen

Este escrito tiene el propósito de reseñar y exponer algunos aspectos de la naturaleza del problema estructural de la I.A, que Daniel Dennett presenta en su artículo titulado *Las Ruedas del conocimiento: el problema estructural de la I.A.* Este escrito consta de tres partes: en la primera, se presenta un experimento para introducir la naturaleza del problema estructural; en la segunda, se presenta la metodología deductiva para enfrentar el problema estructural; en la tercera, se muestra su relevancia de la metodología deductiva en la investigación de I.A. y algunas de sus complicaciones.

Según Dennett, el problema estructural al que se enfrenta la I.A. puede entenderse como “un irritante impedimento técnico en robótica”¹; sin embargo, también puede entenderse como un problema epistemológico sobre cómo adquirimos conocimiento acerca del mundo, y cómo usamos este conocimiento para enfrentar los diferentes problemas de la vida cotidiana. En este documento reseño algunas consideraciones que Dennett presenta en su artículo titulado *Las Ruedas del conocimiento: el problema estructural de la I.A.* Doy total crédito a Daniel Dennett, de las ideas que se presentan en este documento, el cual sólo tiene el propósito de introducir el problema estructural de la I.A.

1. El problema estructural: robots que explotan

Para entender el problema estructural, Dennett apela a un experimento mental que a continuación pretendo reconstruir. Piénsese en un robot, llamado R1, al que se le da la instrucción de buscar su batería de repuesto en una habitación que está asegurada con llave; además, la batería está junto a una bomba próxima a detonar. R1 localiza la batería, que se encuentra en una carreta dentro de la habitación, y diseña un plan para retirarla jalando la carreta; sin embargo, cuando sacó la batería jalando la carreta, también sacó la bomba;

R1 sabía que la bomba estaba en la carreta dentro de la habitación, pero no se dio cuenta de que al jalar la carreta hacia fuera, la bomba saldría junto con la batería. ¡Pobre R1! Pasó por alto esta implicación obvia de su plan²

Luego, los diseñadores de R1, para enfrentar el problema y evitar la explosión, creyeron que era necesario que el robot no sólo considerara las implicaciones intencionadas de sus actos, sino también las implicaciones secundarias. Así, el siguiente robot diseñado, llamado R1D1, también encontró la batería y diseñó un plan de acción; sin embargo,

Acababa de deducir que jalar la carreta afuera de la habitación no cambiaría el color de las paredes de esta y estaba embarcándose en la comprobación de la siguiente implicación. [...] cuando explotó la bomba³

Con esta situación, los diseñadores consideraron necesario enseñarle al robot a diferenciar aquellas implicaciones que son pertinentes, y deben tenerse en cuenta al diseñar el plan de acción, de aquellas que no lo son; esto, con el propósito de que el robot no se que quedara congelado considerando todas y cada una de las implicaciones, relevantes e irrelevantes, para sus planes de acción. Llamaron a este tercer robot R2D1, el robot deductor, y le dieron la misma orden. Cuando el robot localizó la batería, se sentó afuera de la habitación. Los

¹ Ver Dennett, *Las Ruedas del conocimiento: el problema estructural de la I.A.*, pág. 168

² Ver Dennett, *Las Ruedas del conocimiento: el problema estructural de la I.A.*, pág. 167

³ Ver Dennett, *Las Ruedas del conocimiento: el problema estructural de la I.A.*, pág. 167

diseñadores, angustiados por ver que el robot se había quedado congelado de nuevo, le gritaron que hiciera algo y el robot respondió:

“lo estoy haciendo. [...] estoy ignorando diligentemente las miles de implicaciones que determine improcedentes. Tan pronto descubra una implicación improcedente, la sumo a la lista de las que debo ignorar”⁴

Así, mientras el robot se concentraba en omitir consideraciones innecesarias, explotó.

De esta manera, según Dennett, se puede asegurar que los tres robots que explotaron padecieron el problema estructural. El primer robot no tenía un mecanismo que le permitiera considerar de manera completa las implicaciones de sus actos, mientras que el segundo y el tercero, aunque poseían esta capacidad, no podían hacerlo de manera eficiente. El segundo y el tercer robot se enfrentaron a una explosión computacional de cálculos.

2. Metodología deductiva para enfrentar el problema estructural de la I.A.

Se puede concluir que los robots que Dennett presenta y, en general, las máquinas de I.A., no están en capacidad de enfrentar eficientemente el problema estructural; mientras esto sea así, es muy difícil que una máquina pueda enfrentar de manera exitosa los problemas corrientes de la vida cotidiana.

Básicamente han sido dos los enfoques que han pretendido solucionar el problema estructural: un enfoque desde la deducción y otro desde la inducción. Desde la deducción se ha intentado mostrar que un correcto cálculo de predicados es suficiente para hacer todas las consideraciones necesarias de nuestros actos. Este enfoque fue inspirado por los métodos de resolución de teoremas y consistió formalizar y representar el conocimiento mediante axiomas y lógica común, para deducir los efectos de las acciones. Para permitir este cálculo, se utilizan unos axiomas antecedentes, llamados *axiomas estructurales*, que definen las condiciones y efectos de cada acto en un sistema particular. Con respecto a estos axiomas estructurales, Dennett dice que:

(...) a este conjunto de axiomas el sistema le aplicaría una acción -postulando la ocurrencia de una acción A en una situación S - y luego deduciría el efecto de A en S para obtener una descripción de la situación resultante S'.⁵

⁴ Ver Dennett, *Las Ruedas del conocimiento: el problema estructural de la I.A.*, pág. 168

⁵ Ver Dennett, *Las Ruedas del conocimiento: el problema estructural de la I.A.*, pág. 181

3. Contra la metodología deductiva

No obstante, se debe considerar el hecho de que nuestra actividad consciente no se asemeja, según lo que podemos percibir, a un constante cálculo de predicados; de esta manera, se puede proceder en la investigación sobre cualquiera de las siguientes hipótesis: o nosotros no pensamos mediante el cálculo de predicados, aunque un robot pueda hacerlo, o nosotros sí pensamos deductiva pero inconscientemente mediante el cálculo de predicados.

Dejando de lado la veracidad de cualquiera de las hipótesis sobre las cuales se sustenta el enfoque deductivo, Dennett dice que lo cierto es que este no ha mostrado funcionar de manera eficiente, salvo en casos triviales o casos de laboratorio. Una de las principales complicaciones del cálculo de predicados es que cualquier aspecto X de una determinada situación Y puede cambiar en una circunstancia Z , por lo cual se requiere introducir un axioma nuevo y distinto para cada una de las circunstancias Z que puede resultar de una modificación en el aspecto X dado en la situación Y .⁶

Esta consideración y formulación de axiomas se presenta como una tarea extremadamente difícil de realizar en la medida en que consistiría en la definición de un número muy elevado de axiomas, es decir, una explosión de axiomas. Esta complicación se intentó solucionar mediante la suposición de que nada cambia en una situación salvo aquello que se dice explícitamente que cambia; sin embargo, rápidamente los investigadores abandonaron esta suposición al percatarse de la inaplicabilidad de la cláusula *ceteris paribus* al mundo real. Además, la formalización de la cláusula *ceteris paribus* parece bastante complicada en la medida en que no implica una definición explícita de lo que significa:

¿Qué quiere decir que las de más cosas permanecen iguales? ¿Exactamente qué disposiciones de cuáles otras cosas se considera que permanecen iguales? Si tuviéramos que responder a esta pregunta sería inútil invocar la cláusula *ceteris paribus*, puesto que la utilizamos precisamente para evitar esa tarea. Si se pudiera responder esta pregunta no habría, en primer lugar, necesidad de invocar la cláusula⁷

Una manera de enfrentar el problema estructural desde el enfoque deductivo, por adelantado, y sin tener que incluir nuevos axiomas para cada una de las aclaraciones necesarias con respecto a aquello que puede o no puede cambiar en una determinada situación, consiste en considerar la posibilidad de agrupar las variaciones de circunstancias por *estereotipos*. Así, tendría que formalizarse sólo algunos estereotipos específicos, sin necesidad de hacer explícita cada una de las variaciones. Estos estereotipos son escenarios típicos o paradigmáticos en los que, generalmente, se desarrollan nuestros actos; Minsky se refiere a ellos como *estructuras* y Schank como *guiones*. El uso de la metodología de guiones y estructuras no ha sido marginal en la I.A., por el contrario constituye un pilar de

⁶ Pareciera que en cierto tipo de ejercicio inductivo también se presenta una complicación similar, en la que es necesario repasar todas y cada una de las posibles variaciones que se pueden dar en todas y cada una de las posibles circunstancias posibles.

⁷ Ver Dennett, *Las Ruedas del conocimiento: el problema estructural de la I.A.*, pág. 180

esta investigación; de hecho, según John Searle, el argumento que usa para refutar la I.A. fuerte, conocido como *la habitación china*,⁸ se le ocurrió cuando leía acerca de los *programas de comprensión* de relatos de Roger Schank y Robert Abelson,⁹ diseñados en la Universidad de Yale. Como *input* se le puede proporcionar a uno de estos programas un relato del tipo *un hombre entra a un restaurante, pide una hamburguesa. Cuando le traen la hamburguesa, esta está quemada. El hombre sale enfurecido del restaurante sin pagar la hamburguesa*. Después, se le puede preguntar al programa *¿el hombre se comió la hamburguesa?* El *output*, esto es, la respuesta, será que no. En ninguna parte del relato se señala explícitamente que el hombre no comió la hamburguesa y, sin embargo, el programa es capaz de responder satisfactoriamente la pregunta. Esta capacidad se da gracias a que el programa cuenta con un *guión de restaurante* que indica cómo son los acontecimientos que generalmente suceden en un restaurante. En este guión se señala que, por lo general, cuando una persona se enoja al recibir la comida, y no paga la cuenta, entonces la persona no come lo que pidió. De esta manera, el programa aparea la pregunta con el guión de restaurante. Con estas respuestas satisfactorias acerca de relatos se argumenta que, como esta máquina satisface el test de Turing¹⁰, entonces *comprende* el relato tal como cualquier persona lo comprende. Con respecto a este enfoque, Dennett señala que:

Es importante advertir que en estos casos, el sistema no se rediseña a sí mismo (o aprende), sino que debe esperar a que un diseñador externo seleccione un diseño mejorado (...) ese comportamiento es profundamente irreal como modelo de la capacidad de reacción de los humanos a la vida cotidiana.¹¹

La anterior constituye una complicación del enfoque deductivo para enfrentar el problema estructural. Por este motivo, en los últimos desarrollos de I.A. se puede encontrar la simulación de sistemas de cognición infantil, que le permitan a la máquina aprender de su entorno.¹²

Referencias bibliográficas.

- Dennett, Daniel. "Las Ruedas del conocimiento: el problema estructural de la I.A.". En *Filosofía de la inteligencia Artificial*. Compiladora: Margarte Boden. México: fondo de cultura económica, 1990. 167 - 192
- Scassellati, Brian. "Foundations for a Theory of Mind for a Humanoid Robot", Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA, PhD Thesis, June 2001.

⁸ Cfr. Searle, *Mentes, cerebros y programas*.

⁹ Cfr. Schank y Abelson, *Guiónes, planes y entendimiento. Un estudio de las estructuras del conocimiento humano*.

¹⁰ Cfr. Turing, *La maquinaria de computación y la inteligencia*.

¹¹ Ver Dennett, *Las Ruedas del conocimiento: el problema estructural de la I.A.*, pág. 186

¹² Cfr. Scassellati, *Foundations for a Theory of Mind for a Humanoid Robot*.

- Schank, Roger y Abelson, Robert. *Guiones, planes, metas y entendimiento. Un estudio de las estructuras del conocimiento humano*. Barcelona: Paidós, 1987.
- Searle, John. "Mentes y cerebros sin programas". En *Filosofía de la mente y ciencia cognitiva*. Compilador: Eduardo Rabossi. Buenos Aires: Paidós, 1990. 413 – 444
- Turing, Alan. "La maquinaria de computación y la inteligencia". En *Filosofía de la inteligencia Artificial*. Compiladora: Margarte Boden. México: Fondo de cultura económica, 1990. pp. 53 – 80.