

Información y pensamiento Perspectiva desde la filosofía de la inteligencia artificial

Information and thought.

View from the philosophy of artificial intelligence

Information et pensée

Une perspective depuis la philosophie de l'Intelligence Artificielle

Fecha de recibo: 05-07-10 - Fecha de aprobación: 06-11-10

JUAN CARLOS VÉLEZ

De la página 96 a la página 102

Resumen

El funcionalismo es la teoría filosófica que sirve de soporte a la inteligencia artificial. Esta ha modificado de manera revolucionaria la forma como nos representamos el mundo y a nosotros mismos desde el punto de vista del concepto de lo mental. Hace parte del paradigma cognitivo que rige la ciencia contemporánea y aporta desde la tecnología los fundamentos de lo que nos hace humanos, a saber, aspectos como la comunicación y la medicina; y dimensiones filosóficas como la mente, la subjetividad, etc. En este trabajo presentaré el contexto epistemológico en que nace la inteligencia artificial, es decir, sus supuestos metafísicos hasta las principales críticas que se le hizo en su momento en relación con la modelación de la mente computacional.

Palabras clave

Cognitivo, filosofía, funcionalismo, inteligencia artificial.

Abstract

The functionalism is the philosophical theory that serves as a support to the artificial intelligence (AI). This has changed the way we represent the world and ourselves in a revolutionary way, from the point of view of the concept of



mental. Part of the cognitive paradigm that governs contemporary science and from technology provides the foundations of what makes us human beings, such as communication and medicine; and philosophical dimensions as mind, subjectivity, etc. In this article I will present the epistemological context the (AI) was born, i.e. from their assumed metaphysical to the main criticism that was made in relation to the modeling of the computational mind.

Keywords

Cognitive, philosophy, functionalism, artificial intelligence.

Résumé

Le Fonctionnalisme est la théorie philosophique qui sert de support à l'Intelligence Artificielle (IA). Celle-ci a modifié d'une manière révolutionnaire la manière dont nous nous représentons le monde et nous mêmes du point de vue du concept du mental. Elle fait partie du paradigme cognitif qui régit la science contemporaine et apporte depuis la technologie les fondements qui nous font humains : des aspects comme la communication et la médecine ; et des dimensions philosophiques comme l'esprit, la subjectivité, etc. Je présenterai dans le travail suivant le contexte épistémologique de l'IA, c'est-à-dire, ses supposés métaphysiques jusqu'aux critiques principales qui lui ont été faites dans son moment en relation avec le modelage de l'esprit computationnel.

Mots clés

Cognitif, philosophie, fonctionnalisme, intelligence artificielle

1. *El funcionalismo*. Ubicar en su contexto la inteligencia artificial obliga a una excursión por los rasgos fundamentales de la filosofía de la mente y su principal problema: cuál es la naturaleza de los estados mentales. Se ha definido la mente como una colección de estados mentales (creencias y deseos) y el problema consiste en argumentar a favor o en contra de la mejor explicación que los caracterice. El funcionalismo es una de las teorías

más interesantes y prometedoras en la filosofía de la mente, al menos en toda la década de los ochenta y los noventa. Es, así mismo, una teoría causal de la mente que postula, a diferencia del conductismo (teoría que postula que los estados mentales son reducibles a patrones de conducta observable causados por estímulos del ambiente), estados internos, pero a diferencia del materialismo (teoría que postula que los estados mentales son reducibles

al lenguaje de la neurociencia), no identifica estos con estados cerebrales sino que los define como estados funcionales. De este modo un estado mental es definido como una entidad que tiene la propiedad de causar y ser causada por otros estados mentales (una definición circular que por ahora no importa), por ejemplo, mi *deseo* de que Uribe no sea reelegido es causado por mi *creencia* de que el impuesto de guerra no ha contribuido a mejorar



la economía. Y por funcionalismo basta decir que es la teoría que considera que un estado mental se define por su rol funcional, es decir, por su disposición a actuar de cierta manera. Siguiendo a Manuel García-Carpintero (1995), el funcionalismo es el desarrollo de algunas tesis del conductismo y la solución de los principales obstáculos que le presentaron en su momento; así como el conductismo es la reacción crítica frente a las dificultades epistemológicas de la teoría cartesiana de la mente.

El funcionalismo explica el vocabulario de lo mental, como el de las percepciones, los pensamientos o las acciones, desde el punto de vista de las *funciones que realizan*. Es una

teoría causal de los estados mentales porque a partir de cierto tipo de causas (estímulos) se modifica un estado interno preexistente como creencias, lo cual genera nuevas conductas o respuestas referidas al medio donde se originó la estimulación. El significado de los términos psicológicos se entiende de manera causal mediante elementos y subelementos que son funcionales.

La diferencia con el conductismo es que para éste un estado mental se define en términos de estímulo ambiental y respuesta en forma de conducta, mientras que el funcionalista refiere un estado mental a una gran variedad de estados mentales conectados causalmente,

lo que excluye el reduccionismo estímulo-respuesta conductista.

El funcionalismo es también un método que crea modelos abstractos de los procesos mentales constituidos por mecanismos que pueden incluso buscarse en el sistema nervioso a partir de modelos constituidos por redes y subredes neuronales, y así mismo, utilizarse en modelos de inteligencia artificial creando unidades y subunidades. Este modelo, también considerado un sistema, está soportado en una descripción funcional, y ésta es, a su vez, la descripción de un proceso causal, es decir, las relaciones entre inputs y outputs.

Así cuando un sistema (u objeto) satisface una descripción funcio-



nal se dice entonces que el objeto la implementa o la realiza; una descripción funcional es la descripción de un suceso causal. Un ejemplo de descripción funcional tomado de García-Carpintero consiste en una máquina expendedora de café, a la cual se le pueden insertar, a modo de inputs, dos tipos de monedas: de doscientos y de quinientos; tiene dos posibles estados S1 y S2 que son los estados internos del sistema. Si está en S1 se le introduce una moneda de quinientos, devuelve un café (output) y sigue en S1. Si luego se le introduce una moneda de doscientos pasa a S2; estando en S2 se introduce una moneda de quinientos, devuelve un café y dos monedas, una de cien y otra de doscientos. *La descripción funcional es posible por la estructura causal.* Un programa de computador tiene una descripción funcional muy compleja pero que incluye estados internos especificados por la descripción, los cuales se pueden ordenar en secuencias causales.

El punto interesante es que para el funcionalismo *no hay creencias sin representaciones*. Es decir, los estados internos son representaciones que un hablante tiene en la cabeza, pero el tipo de representación se considera un estado funcional en la medida en que equipara una creencia precisamente con información que el sistema tiene y que puede actualizar en el mismo sentido como la máquina expendedora de café la tiene.

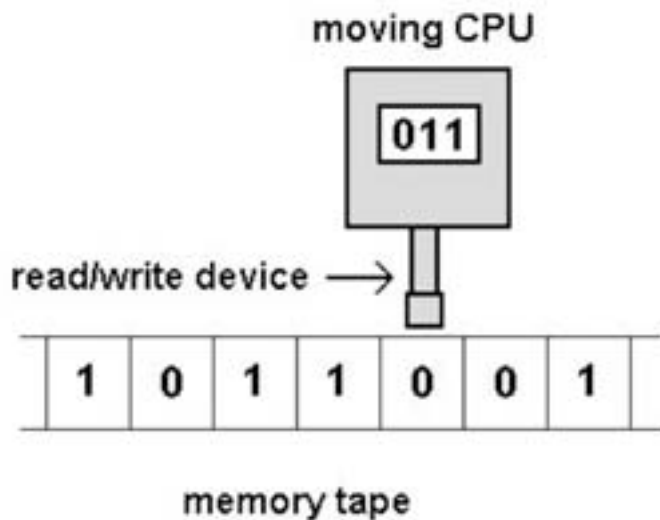
Pero si los estados mentales son definidos en términos de descripciones funcionales, ¿puede esta descripción funcional, por sofisticada que sea, dar cuenta de la competencia lingüística, es decir, su capacidad para usar el lenguaje? Para que la dimensión simbólica del lenguaje permita expresar contenidos, tanto generales como específicos, el lenguaje consta de expresiones elementales y reglas para combinarlos. La especificidad de estas reglas es ser algorítmicas.

Las propiedades del lenguaje son entonces la productividad y la sistematicidad; la primera tiene que ver con la capacidad infinita de construir y entender oraciones, la sistematicidad expone la idea de que las relaciones conceptuales no son arbitrarias, es decir que el hablante conoce las categorías gramaticales, los significados de las palabras y las posibles combinaciones. La tesis funcionalista plantea que estas propiedades del lenguaje son compartidas por el pensamiento y para ello recurre a la tesis del *lenguaje del pensamiento*.

Dada la caracterización anterior del análisis funcionalista en términos de las interacciones causales entre los estados internos, en el caso anterior entre creencias y deseos, hay otras versiones del funcionalismo que mantienen dicha estructura, a saber, el funcionalismo de tabla de máquina y el funcionalismo computacional.

2. *El funcionalismo de tabla de máquina.* Esta versión del funcionalismo fue desarrollada a partir de los pioneros trabajos del matemático inglés Alan Turing, alrededor de los treinta. El objetivo de Turing, considerado brevemente, era el de ofrecer una descripción física de la mente. Pero no cualquier propiedad de la mente, sino específicamente lo que ésta hace al calcular. Turing compara al hombre y a una máquina precisamente en el nivel de las operaciones abstractas que ejecutarían al computar números. Como los cálculos son lo suficientemente simples y mecánicos, la distinción entre hombre y máquina no es relevante para caracterizarlos y describirlos. Turing lo hace imaginando un hipotético sistema hecho de una cinta que contiene casillas o cuadrados vacíos, donde es posible escribir un símbolo que generalmente es un número. El sistema se puede desplazar a la izquierda o a la derecha, gracias a que consta de una unidad de ejecución. La unidad de ejecución es un “dispositivo” que siempre presenta un estado interno, y que al posarse sobre una casilla hace que el sistema pase a otro estado. Esto se debe a que ejecuta ciertas operaciones, que a su vez están determinadas por las reglas condicionales. Entonces estas reglas son las que especifican el estado interno del sistema, concretamente de la unidad de ejecución.

La acción de la máquina consiste en escribir un símbolo en la



casilla en blanco y cambiar el estado interno de la unidad de ejecución. Borrar e imprimir un símbolo o dejar el que está, de acuerdo con unas instrucciones, es básicamente lo que esperaríamos al calcular. La actividad de cálculo consiste entonces en identificar una condición con una función que le corresponde a cada número o símbolo. Se puede hablar así de conocimiento en la máquina puesto que el instante en el que se encuentra la máquina es un estado de memoria que identifica un objeto abstracto (símbolo) con cada estado de memoria. Calcular significa pasar de una configuración interna a otra según una tabla de instrucciones. Un punto importante que habría que resaltar es que la máquina debe ser capaz de determinar el output que producirá a partir de la introducción de cierto input, siempre y cuando se siga un

número finito de pasos. De nuevo tenemos aquí que la interacción causal de elementos internos es la impronta del funcionalismo.

3. Inteligencia artificial. Por otro lado, otra versión del funcionalismo computacional es la que ha desarrollado la IA. Esta consiste en equiparar los símbolos y las reglas para manipularlos, con las operaciones mentales. Aquí la analogía entre computadores y mentes humanas se hace mucho más fuerte. Incluso el mismo Turing proporcionó, en su famoso artículo del cincuenta: “Maquinaria, computadora e inteligencia”, el criterio de comparación en su famoso test de Turing. Este consiste en que la equivalencia entre hombre y máquina se pone de manifiesto al exhibir ambos el mismo output o conducta a partir del mismo input (que generalmente son preguntas); lo que habría que

resaltar es que la obtención de estos mismos resultados se logra implementando distintas secuencias de pasos. Lo que le interesa al teórico de la inteligencia artificial es desarrollar máquinas que realicen la misma secuencia de pasos que un ser humano al ejecutar las mismas funciones cognitivas.

4. *Críticas a la inteligencia artificial.* El funcionalismo computacional se compromete con lo que Searle denominó IA fuerte, es decir, el programa de investigación desarrollado en torno a la búsqueda del algoritmo que tenga la capacidad de emular las tareas cognitivas humanas, implementándolo en una arquitectura funcional que subyace al manejo de la información. El funcionalismo computacional hace equivaler los procesos mentales con los procesos algorítmicos determinantes en la producción de conducta. Hay que decir que a pesar de las duras críticas al funcionalismo, como principalmente el que debido a su liberalidad adscribe mentalidad a diestra y siniestra, o que deja de lado el aspecto cualitativo de la experiencia mental, sigue siendo uno de los enfoques en filosofía de la mente que actualmente se renueva una y otra vez, así es el caso del funcionalismo teleológico, el cual considera que los estados mentales son el resultado de presiones selectivas, por tanto, son útiles para el organismo, pero este no es el lugar para hacer una exposición de sus tesis.

El filósofo más reconocido por los ataques a la inteligencia artificial ha sido John Searle. En su libro “El misterio de la conciencia” afirma que el cerebro es una máquina con la capacidad de pensar, luego, lo que problematiza no es la noción máquina aquí implicada sino el término pensamiento pues lo que resulta inadmisible es que un computador pueda llegar a pensar porque para hacerlo es necesario conocer los significados de las palabras (semántica), y lo que hace un computador es solo aplicar reglas de cálculo para manipular símbolos (sintaxis). Un computador nunca podrá pensar y ser consciente aunque se comporte de manera inteligente, por ejemplo al jugar al ajedrez.

Sin embargo, no todos los filósofos piensan igual. Daniel Dennett cree que la perspectiva de Searle adolece del prejuicio antropocentrista, es decir, que el pensamiento quizás sea más amplio de lo que suponemos desde nuestra peculiaridad humana. Además, uno podría decir que los computadores son solo una extensión de nuestras mentes, y como tal, los computadores serían una consecuencia de la teoría de la evolución, un proceso de continuidad lógico, incluso algorítmico. El punto de quiebre señalado por Searle es que los símbolos abstractos utilizados por un programa no son suficientes para lograr la candidatura a pensamiento, ya que carecen de las capacidades causales



para producirlo, porque estas solo son posibles en un medio de ejecución, a saber, el cerebro. El cerebro tiene entonces unas propiedades causales exclusivas que no tendría otro medio.

Por esto Searle esperanzadamente advierte que cuando se

conozcan las propiedades causales del cerebro para producir la conciencia, estas podrán reproducirse en máquinas artificiales y crearla, pero las operaciones de cómputo por sí solas no bastan.

A manera de punto final, el principal problema de las teorías funcionalistas es la incapacidad explicativa de dar cuenta del problema de la naturaleza de la conciencia: es decir, de la propiedad básica de la mente que nos permite experimentar el mundo desde la perspectiva de la primera persona. Tener un punto de vista acerca de la realidad tal cual se nos presenta o la construimos, es aquello que compartimos quizás con muchas especies de animales; ese es el problema grueso, lo que sin embargo no le quita fuerza pues no hay actualmente una teoría que lo haga satisfactoriamente. Como dice García-Carpintero en su artículo ya citado, el funcionalismo es una propuesta conceptualmente atractiva capaz de arrojar luz sobre muchos problemas en la filosofía de la mente.

El funcionalismo descrito es una de las mejores alternativas al conductismo en la medida en que define los estados mentales como estados funcionalmente descritos: cuando alguien está en cierto estado mental es asumir que la descripción interna de ese estado depende de las complejas correlaciones entre inputs y outputs. Es por esto que esta postura nos sirve de plataforma

para contar la historia de la ciencia cognitiva o lo que se ha denominado la revolución cognitiva, donde el actor principal es la noción de conocimiento entendido como cognición. ¿Por qué?, en primer lugar en las ciencias cognitivas los distintos casos de conocimiento comienzan a entenderse como procesamientos de información: sentir, percibir, creer, inferir, recordar, querer que algo sea el caso, son procesamientos de información. En segundo lugar, estos casos de conocimiento no son exclusivamente humanos sino que los exhiben algunos animales (vertebrados) y algunas máquinas (computadores). En consecuencia, establecemos y pertenecemos, junto con los computadores y los animales, a una nueva categoría de entidades: la de los informívoros.

Bibliografía

- Dennett, D. (1987): “Los verdaderos creyentes: La estrategia intencional y por qué funciona” (cap.) *La actitud intencional*, Editorial Gedisa, S.A., 1998, Barcelona.
- Dennett, D. (1991): “Versiones múltiples frente a teatro cartesiano”, (cap., 5) en *La conciencia explicada*, 1995, Ediciones Paidós Ibérica, S.A.
- Dennett, D. (1996): “La creación del pensamiento”, (cap.5) en *Tipos de mentes*, de la versión castellana, Editorial Debate, S.A., Madrid, 2000.
- García-Carpintero, M. (1995): “El Funcionalismo”, *La mente humana*, Edición de F. Broncano, Editorial Trotta, S.A., 1995.
- Martínez-Freire, P. (1995): *La nueva filosofía de la mente*, Editorial Gedisa, S.A. 1995, Barcelona.
- Martínez Freire, P. (2002): “La revolución cognitiva”, P. Web: freire@uma.es
- Searle, J. (1989): “Mentes y cerebros sin programas” en *Filosofía de la mente y ciencia cognitiva*, Rabossi (compilador), Ediciones Paidós Ibérica, S.A., 1995.
- Searle, J. (1992): *El redescubrimiento de la mente*, CRÍTICA (Grijalbo Mondadori, S.A.), 1996.
- Simon, H. (1981): “Ciencia cognitiva: la más nueva ciencia de lo artificial”, en *Perspectivas de la ciencia cognitiva*, Norman (comp.), Paidós Ibérica, S.A., 1987.
- Turing, A., (1950): “Máquina, computadora e inteligencia”, en *Controversia sobre mentes y máquinas*, 1986, Biblioteca de divulgación científica, *Muy interesante*.