

DEBATES

El problema de la representación en la filosofía cognitiva

The problem of representation in cognitive philosophy

JUAN C. VÉLEZ

Universidad del Valle (Cali, Colombia)

Recibido: 15-10-2007 Aprobado definitivamente: 15-01-2008

RESUMEN

La teoría representacional de la mente, basada en el concepto de *representación*, ha sido muy criticada, especialmente por recientes enfoques en la ciencia cognitiva, provenientes de la Biología y la Inteligencia Artificial. En este trabajo me centro especialmente en el punto de vista de Francisco Varela, quien sugiere la exclusión del término *representación* en la explicación de los sistemas cognitivos. Muestro que ello no es necesario, y que hay razones para hablar de representaciones en la relación que tenemos con el mundo en términos de conocimiento, por tanto, el *antirrepresentacionalismo* de Varela es inadecuado. En ese sentido me parece más afortunada la apreciación que hacen de la ciencia cognitiva y la filosofía de la mente Andy Clark y Pascual Martínez-Freire, y ésta es la postura que defenderé en contra de Varela.

PALABRAS CLAVE

MENTE, REPRESENTACIÓN, COGNITIVISMO, CONDUCTA, SISTEMA

ABSTRACT

The representational theory of the mind, based on the concept of *representation*, has been very criticized, specially by recent approaches to cognitive science, originated from Biology and Artificial Intelligence. In this work I focus my attention specially on the point of view of Francisco Varela, who suggests the exclusion of the term *representation* in the explanation of cognitive systems. I show that it is unnecessary, and that there are reasons to talk about representations in the relation that we have with the world in terms of knowledge, and therefore, Varela's *antirepresentationalism* is inadequate. In that connection the appreciation that Andy Clark and

Pascual Martínez-Freire do of cognitive science and the philosophy of the mind seems more fortunate to me, and this is the position that I will defend against Varela.

KEYWORDS

MIND, REPRESENTATION, COGNITIVISM, BEHAVIOR, SYSTEM

I. INTRODUCCIÓN

EN LA CORTA HISTORIA DE LAS CIENCIAS COGNITIVAS se han propuesto diversas teorías de la representación, para determinar, en términos de explicación, el tipo de procedimiento que subyace a la forma en que los sistemas, sean biológicos o artificiales, interactúan con un entorno. La conducta observada, los modelos computacionales, la investigación neurocientífica, ha generado un ámbito de discusión que ha arrancado desde sus cimientos los principales postulados de la teoría estándar, a saber, la teoría conductista del conocimiento. Esta característica general de los resultados obtenidos ha despertado una discusión interesante que gira en torno a la representación, término que arrastra e influye en la noción de racionalidad, lenguaje, conocimiento, pensamiento, inteligencia, y todo lo que constituye los estados mentales.¹

II. EL COGNITIVISMO

El tratamiento de la cognición que se hace desde este enfoque es en términos de un agente que resuelve problemas y toma decisiones (Guijarro 2000). Se defiende aquí una ontologización de las representaciones que configura y define nuestras mentes como una máquina sintáctica transformadora de símbolos, y la relevancia que tiene para la cognición semántica. Esta vía que la filosofía cognitiva ha tomado, a saber, el cognitivismo, ha sido desarrollada por los psicólogos cognitivos y filósofos de la mente Pylyshyn (1988); Newell (1987); Simon (1987); Fodor (1984, 1995); que desde los avances en la lógica aplicada a la Inteligencia Artificial (IA), ha dado lugar a la tesis del «computacionalismo», «representacionalismo» o «simbolismo», de acuerdo con la cual nuestra mente es un sistema que combina, procesa símbolos. La noción de representación en este programa de investigación queda expresada de la siguiente forma:

Los códigos o símbolos son clases de equivalencia de propiedades físicas que, por un lado, hacen que la conducta se produzca de la forma en que lo hace y que, por otro lado, son los portadores de interpretaciones semánticas que proporcionan el principio de alto nivel que se necesitan para su individuación, y

1 Para una introducción a la ciencia cognitiva ver Gardner (1988).

para establecer las generalizaciones. Como realistas acérrimos [...], concluimos que los códigos son psicológicamente reales, y que el cerebro es el sistema que los procesa, y que los códigos tienen de hecho un contenido semántico.²

La naturaleza de las computaciones de los símbolos se da de forma computacional, y por computacional se entiende que es posible formalizar tal contenido, siguiendo el modelo de la lógica formal. Donde la teoría se concentra es en la cognición semántica, puesto que el lenguaje es el modelo más representativo de nuestra conducta y nuestra inteligencia. La idea fundamental de esta posición es que la inteligencia, el razonamiento, el lenguaje, y gran parte de nuestros procesos mentales se rigen por estas reglas de carácter algorítmico, de lo que se concluye que nuestras mentes son máquinas sintácticas.

La noción de informívoro, en su sentido más tradicional, alude a una cosa que conoce o que tiene ciertas características que le hacen apta para conocer u obtener *información* del medio en que se encuentre, y esta información es recuperable mediante un proceso que implica *memoria*, la cual es definida como un conjunto de estructuras simbólicas. Es a ese nivel que nosotros los seres humanos, los computadores y gran parte de los mamíferos superiores somos denominados *sistemas de símbolos (SS)*.³ Luego, lo que tenemos en común con las computadoras es crear, manipular y procesar símbolos, y esos símbolos tienen una naturaleza abstracta. Hablar de ciencias cognitivas resulta más ventajoso, desde esa perspectiva, que de ciencias del conocimiento, ya que *percibir, creer, inferir, sentir*, son procesamientos de información. Un sistema de símbolos es un sistema adaptativo e inteligente que al enfrentarse a un medio debe ser capaz de producir conducta, en respuesta a una situación pero dependiente de su organización interna, tal como puso de relieve Newell. Por ello discernir lo que es bueno u óptimo para mantener un sistema, en su más amplio sentido, requiere de la discriminación de información. De ahí que sea usual definir el término de creencia como estado mental portador de información y sea materia de discusión el que un mamífero o un computador tengan creencias.

Allen Newell, para definir la conducta de un proceso en el que hay una acción a distancia lo enuncia de la siguiente manera:

Designación: Una entidad X designa una entidad Y relativa a un proceso P, si, cuando P recibe X como entrada, su conducta depende de Y.⁴

Esta definición de designación es satisfecha por los SS en tanto hay símbolos que se comportan como la entrada al sistema que ponen en funcionamiento

2 Pylyshyn (1988), p. 67.

3 Allen Newell (1987), p. 58-63.

4 Allen Newell (1987), p. 75.

conjuntos de procesos que a su vez se comportan en función de lo que asignan los símbolos. Los procesos son de tipo simbólico y tienen que ver con lo que comúnmente se llama *programa*.

Este presupuesto es la base para la representación puesto que los procesamientos cognitivos se dan a partir de manipulación, transformación y combinación de símbolos que a su vez representan experiencias, percepciones y acciones.

Si, como insinuamos antes, las creencias son los estados mentales típicos portadores de información, ésta toma un matiz estrictamente mental. Recordemos que el mentalismo es una reacción frente al conductismo caracterizado éste por un desdén hacia los conceptos intencionales.⁵ En ese sentido más que de sistemas simbólicos de lo que se habla es de *sistemas creencias-deseos* debido a que los estados mentales básicos en la explicación de la conducta orientada a metas son deseos y creencias.

Las condiciones para la explicación de la conducta según la teoría computacional de la mente son:

- Que la conducta se presente mediante reglas formales que aludan a códigos o estados cognitivos diferentes; y que siendo un sistema formal no tiene interpretaciones.
- Pero considerado un sistema representacional, tales reglas deben respetar la interpretación semántica.

De ahí que se le conozca a este enfoque como cognición semántica. En consecuencia, la causación de la conducta es llevada a cabo por las propiedades físicas de las representaciones relacionadas con las pautas captadas en virtud de su referencia semántica. Según este enfoque, la computación y el conocimiento deben considerarse procesos regidos por *reglas* y estas reglas son lógicas, virtuales, abstractas. Por otro lado no puede haber computación sin representación puesto que los símbolos o códigos tienen que representar números, letras o palabras. Los argumentos que le dan soporte a la tesis de los sistemas de símbolos son la teoría de la decisión, la teoría de la percepción y la teoría del lenguaje del pensamiento. Revisemos el primero.

Según la teoría de la decisión la necesidad de explicar la conducta de un organismo en una situación compleja se apelan a las creencias que el organismo tiene, y por creencia se entiende todos los estados mentales portadores de información, una Teoría de la Decisión deberá consistir en:

5 Un ejemplo típico de ello lo encontramos en Quine. Ver el artículo de Guerrero, «Teoría de la evidencia y holismo moderado en Quine», de 1999.

- 1- Un agente se encuentra en una situación (S).
- 2- Estando en S, el agente cree que tiene diferentes opciones de conducta B1, B2, Bn que puede hacer.
- 3- El agente prevé las consecuencias resultantes a manera de hipótesis al realizar las opciones: «si se realiza B1 en S entonces probablemente se seguirá C1».
- 4- El agente prefiere unas consecuencias sobre otras estableciendo una jerarquía.
- 5- Así, de acuerdo con lo anterior, el agente elige una conducta.

La racionalidad, según su argumento, consiste en la contemplación y evaluación de las opciones de conducta, en términos de creencias-deseos, a fin de conseguir resultados. La racionalidad, en efecto, es el uso teórico, práctico y evaluativo de la razón, o en cómo adoptamos creencias, evaluamos hechos y tomamos decisiones.⁶

El argumento de Fodor consiste en que el modelo debe presuponer, para que sea válido, la utilización de un medio representacional por medio de los agentes. Es decir, el agente debe disponer de un medio para representar su situación actual, sus opiniones conductuales (creencias), lo que se imagina serán las consecuencias de sus acciones (deseos) y el orden de preferencia de tales consecuencias. Y nótese que aquí la representación es acerca de cosas que, obviamente, no han ocurrido aún. Además, estas representaciones son infinitas en el sentido de que el agente se puede enfrentar a nuevas estimulaciones sin haberlas antes experimentado. Hay aquí un uso de las representaciones al cual no habría que objetar porque cómo prescindir de estas al tener la imagen de algo desligado de su situación estimulativa actual.

III. CRITICAS A LA REPRESENTACIÓN

La crítica más importante al cognitivismo proviene de la Inteligencia Artificial y tuvo que ver los problemas que éste no pudo resolver. El conexionismo se inspira en la arquitectura del sistema nervioso central, *simulándola*,⁷ tal como lo ha expuesto Smolensky (1995). Para el conexionismo, de las redes nerviosas *emerge* una nueva conducta, a saber, los procesos cognitivos, por lo tanto, lo relevante es simular la actividad neuronal para dar cuenta, en tiempo real, por

6 No obstante es imposible hacer un seguimiento en torno a la discusión acerca de la racionalidad en este breve escrito sin ser injustos, por tanto remitimos al lector al artículo de F. Broncano (1995).

7 Para una reflexión sobre la noción de simulación véase Cuartas (2007).

ejemplo, de la percepción, el aprendizaje de un idioma, o el reconocimiento de rostros; el conexionismo no descarta la existencia de representaciones, salvo que éstas ya no se encuentran localizadas sino distribuidas. Me apresuro a decir que el programa conexionista inspiró no poco algunos de los postulados del nuevo giro cognitivo. Sin embargo, lo que hay que tener en cuenta es que en el enfoque conexionista el sentido no está localizado en símbolos particulares, sino que opera en función del estado global del sistema y está elaborado con el desempeño general en un área determinada. Como el estado global emerge de una red de unidades que son más densas que los símbolos, algunos investigadores se refieren al conexionismo como «paradigma subsimbólico».⁸

El conexionismo fue el paso decisivo que preparó el terreno para el nuevo enfoque de la ciencia cognitiva. En 1991 E. Thompson, E. Rosch, y Francisco Varela publican *The Embodied Mind. Cognitive Science and Human Experience*. Allí se preguntan ¿cuál es el fundamento científico de la idea de que la mente es un procesador de información que selecciona rasgos que ya están dados en el mundo?, ¿se gana o se pierde algo si abandonamos la noción de representación?, ¿qué supuestos esconde la cognición tradicional? La denuncia que venía haciendo Varela desde su libro de 1990, *Conocer. Las ciencias cognitivas: Tendencias y perspectivas. Cartografía de las ideas actuales*, era que si suponemos que hay representaciones se postula un *realismo cognitivo*, es decir, que hay, en nuestra ontología, un sistema representacional, lo que significa que hay un sujeto que se representa el mundo. Para Varela «el mundo y quien lo percibe se definen recíprocamente». Para sustentar esta tesis debe eliminar la noción de representación, pues es a través de ésta que se plantea un realismo o un idealismo. Así, dice que:

Este énfasis de la mutua definición nos permite buscar una vía media entre el Escila de la cognición como recuperación de un mundo externo pre-dado (realismo) y el Caribdis de la cognición como proyección de un mundo interno pre-dado (idealismo). Ambos enfoques se basan en el concepto central de representación.⁹

En el realismo la representación se usa para recobrar lo externo, en el idealismo para proyectar lo interno. La cognición, según el giro que introduce el autor, se trata de disolver la dicotomía externo/interno entendiendo ésta como acción corporeizada.

De acuerdo con Varela la cognición como representación oculta los siguientes supuestos:

8 Ver la reconstrucción que hace Varela (1992), p. 127.

9 Varela (1992), p.202.

1. que habitamos un mundo con propiedades particulares del tipo que postula la física, como longitud, color, sonido, movimiento, etc;
2. que recobramos esas propiedades representándolas internamente;
3. que un nosotros subjetivo separado es quien hace estas cosas.

Pero ¿qué es lo que entiende Varela por representación? Hay un sentido no técnico de esta noción que se evidencia cuando lo contrastemos más adelante con el que propone Andy Clark. Varela propone dos sentidos de la noción de representación: un sentido débil que ejemplifica cuando frente a un mapa podemos decir que éste es acerca de algo, que lo *interpretamos* como que se refiere a un determinado lugar, al igual que una oración en el lenguaje representa un determinado estado de cosas. Pero hay un sentido fuerte que consiste en generalizar el anterior adquiriendo un compromiso epistemológico y ontológico de las representaciones como el que propone Fodor, es decir, cuando decimos que las representaciones son símbolos físicos (conceptos), instanciados en la arquitectura cerebral mediante lo que se denomina un sistema de representaciones.

III.1. COGNICIÓN ENACTIVA

La tesis principal de *The embodied Mind* es que en la enacción la cognición es la historia del acoplamiento corporal que hace *emerge* el mundo y que funciona a partir de redes y subredes sensorio-motrices interconectadas. Ello es lo que otorga significación y sentido al mundo o configura uno nuevo, según sea el caso. Esta tesis será reelaborada en el 2000, en su libro titulado *El Fenómeno de la vida*. Allí propone un programa de investigación en términos de lo que denomina Neurofenomenología, donde hace más explícita la noción de enacción. La preocupación fundamental es diferente: afrontar lo que Chalmers (1995) denominó «el problema duro de la conciencia».

«El enfoque enactivo (*enactive approach*) –como dice Martínez-Freire (2006)– se presenta como una alternativa tanto al cognitivismo como al conexionismo. Para éstos, el criterio de cognición continúa siendo una representación atinada de un mundo externo que está dado de antemano, pero para Francisco Varela las cuestiones relevantes que surgen en nuestras vidas no son predefinidas, sino cuestiones enactuadas, emergidas o alumbradas desde un trasfondo». En ese sentido el desplazamiento propuesto por Varela implica dejar de pensar en un sistema funcionalista constituido por inputs y outputs para darle paso a un sistema en el que los resultados de los procesos sean por sí mismos sus productos formando redes autónomas. Estas redes autónomas se definen por mecanismos de función internos de autoorganización, pero lo más importante es que no operan por representación sino por enacción. En este enfoque lo que interesa es la mente en relación con la acción, es la puesta en obra de un mundo y una mente a partir

de una historia de la variedad de acciones que un ser realiza en el mundo. Para la enacción la cognición no es la representación de un mundo pre-dado por una mente pre-dada. El mundo consiste, para el enfoque enactivo, en un conjunto de distinciones que son inseparables de la estructura cognitiva.

De lo que estamos hablando es de que hay muchas biología y por tanto hay muchas formas de ser en el mundo en la medida en que estas distinciones se realizan por diversas estructuras cognitivas (Vélez 2004). De esta manera el conocimiento depende del hecho de estar en un mundo del que nuestro cuerpo forma parte, al igual que el lenguaje y nuestra historia social. Nuestros cuerpos son estructuras físicas vividas, experimentadas, que nos permiten reflexionar sobre un mundo que no está hecho sino hallado, un mundo que paradójicamente parece estar allí antes que nosotros, pero que en últimas no está separado de nosotros.

III.2. BITTORIO

En «Structural coupling and the origin of meaning in a simple cellular automata» de 1988, Varela introduce el autómeta celular Bittorio para ejemplificar un sistema autónomo con la idea de acoplamiento estructural. Sin embargo, es en *The embodied Mind* que Bittorio es utilizado como argumento principal para alejarse de la hipótesis de la cognición como sistema de inputs y outputs. Un autómeta celular es considerado como «un espacio computacional constituido por muchas células discretas, cada célula puede estar en uno o varios estados y cambia o permanece en su estado según reglas determinadas (locales)».¹⁰

Bittorio es entonces un diseño matemático que se puede simular por computador.

Este consiste en unidades o células interconectadas a manera de anillo. Una de estas células puede estar solamente en dos estados, uno o cero, activo o inactivo. Cuando el anillo es sumergido en un medio de unos y ceros aleatorios, una de sus células se topa con una de estas alternativas, llamadas perturbaciones. Una vez que ocurre esto el estado interno de la célula es sustituido por dicha perturbación, y dado que la célula está interconectada con otras células mediante su forma circular, todas ellas, como un sistema, modifican su configuración interna, es decir, se autoorganizan. Ahora bien, como el cambio en la configuración puede ser debido a una secuencia impar de perturbaciones, el sistema de ahí en adelante solo reconocerá este tipo de secuencias, más no las secuencias pares. En consecuencia, Bittorio se convierte en un «reconocedor de secuencias impares», sin necesidad de haber sido previamente programado para ello. El mundo de Bittorio es seleccionado a través de un mundo de azar

10 Martínez-Freire, 2006, p. 139.

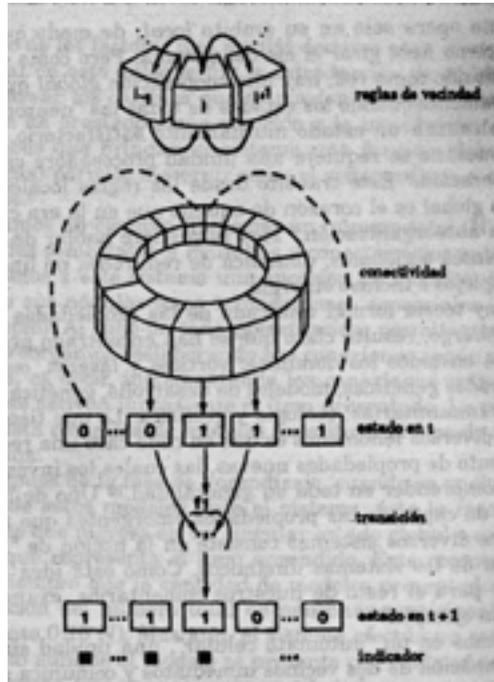


Imagen tomada de Varela, Thompson y Rosch, 1992, p.116.

mediante el acoplamiento estructural del sistema, cobrando entonces relevancia sólo en algunos de los aspectos, en este caso realizar distinciones de secuencias impares. El ejemplo debería hacer pensar en las posibilidades de la emergencia del sentido y la significación del mundo a partir de un sistema neuronal masivamente interconectado.

Como veremos, la crítica de Varela a la teoría representacional de la mente es injustificada en muchos aspectos en la medida en que desconoce la riqueza del sistema representacional cuando, por ejemplo, contempla casos de situaciones posibles, que no son necesariamente reales. La acción deliberada presupone decisiones entre resultados posibles. El error de Varela, como veremos, está en pensar que la cognición computacional se agota en las causas ambientales y desconoce afirmaciones del tipo: «algunas veces actuamos como actuamos porque nos parece la mejor manera de actuar teniendo en cuenta las opciones que consideramos. [...] Lo que queremos es considerarlas [las conductas deliberadas] como respuestas a resultados posibles».¹¹

11 Ver Fodor (1984) p. 53.

IV. COGNICIÓN ENCARNADA

En la cognición enactiva hay una fuerte preocupación por asignarle un papel preponderante al entorno que justamente influye en la conducta del sistema, en este caso biológico. Son el organismo y el entorno los que se van a determinar mutuamente en un acople que está guiado por la acción. Varela, defensor del enfoque antirrepresentacionista,¹² sostiene que hay que desterrar del vocabulario cognitivo la noción de representación puesto que lleva implícita la marca del realismo cognitivo, es decir, que hay unas propiedades de un mundo que hay que recuperar, previo uso de un sistema representacional pre-dado. La posibilidad de la actuación de un sistema dependería de lo que demanda el entorno, mientras que lo que propone Varela es que un sistema, como el caso del cerebro, al ser autoorganizativo no depende del medio externo. Hay que señalar que a Varela no le preocupa la semántica, ni el pensamiento, ni la psicología folk, ni el lenguaje, tan solo habla de sistemas en un sentido muy amplio, donde percepción y acción son indisolubles. Veremos cómo en el enfoque enactivo hay algunas inconsistencias, luego de mencionarlas nos permitirá llenar algunos vacíos, y salvar la teoría Representacional. En ello Clark nos servirá de punto de anclaje entre los modelos mencionados en los capítulos anteriores.

IV.1. LA NUEVA CIENCIA COGNITIVA

El objetivo de la ciencia cognitiva, según Andy Clark, consiste en cómo es materialmente posible el pensamiento corpóreo, dentro de la concepción de la mente biológica. En su libro *Being there. Putting brain, body and world together again*, de 1997, entiende por pensamiento o mente corpórea el estudio de la acción en tiempo real y situaciones reales como desplazamientos, agarre de objetos, la emulación virtual de la acción o la categorización de objetos. Esta propuesta rescata lo mejor de la cognición enactiva, pero sin rechazar radicalmente el computacionalismo. De hecho Clark reconoce lo mucho que le debe a las ideas de Varela, pero va más allá que éste. La hipótesis que va a defender es que en términos de la cognición, hay representaciones internas pero que no se agotan en explicaciones de estados internos, sino que hay que involucrar el entorno mismo y las relaciones dinámicas entre éste y el agente. Por entorno entiende el autor un nicho natural, que de alguna manera implica lo social.¹³

12 Varela es un caso de antirrepresentacionismo, como R. Brooks, T. Van Gelder, entre otros. Para una discusión sobre estas otras posturas ver el artículo de Martínez-Freire, «Putting representation, body and world together again», en *Contrastes. Cognición y Representación*, (2005).

13 Como señala el autor, esa es una diferencia básica entre su perspectiva y la de Heidegger, quien a través de la noción de trasfondo se ocupó del entorno estrictamente social en una

Uno de los principios defendidos por el autor es que la mente tiene un carácter de multiplicidad. Es decir, que su explicación no se puede abordar con el tipo de arquitectura funcionalista propia del cognitivismo, aquella que se basa en transformaciones seriales de unos símbolos en otros, en términos causales (García-Carpintero 1995), ni con la que expone el conexionismo, la que postula redes de nodos interconectados, donde la información se distribuye por el peso asignado. No hay un único sustrato sobre el que se desarrolle la actividad mental. Así, Clark propone que la cognición es ejecutada por una multiplicidad de máquinas virtuales adaptadas de acuerdo a las necesidades del sistema. Es importante empezar por resaltar la importancia que la noción de «conducta adaptativa» tiene para el modelo explicativo presente, pues esto es lo que justamente hay que explicar para dar cuenta del éxito en la acción.

La teoría de la mente corpórea, defendida por el autor, postula que la actividad de un organismo en su entorno es parte integrante de sus estados mentales. Es decir, *la interacción cuerpo-mundo constituye la mente*, por tanto, la mente no se encuentra desconectada del cuerpo y sus correspondientes actividades, sino que la mente se encuentra extendida por el entorno. Este proceso de interacción o retroalimentación está mediatizado por *trayectorias representacionales*, esto es, información descargada por agentes humanos cuando constituyen elementos culturales, y en general, proporcionada por el mundo. La idea es identificar estas trayectorias que se encuentran tanto en la arquitectura del sistema como en el entorno.

IV.2. AGENTES AUTÓNOMOS

El argumento que presenta Clark para defender su tesis ha sido extraído de la nueva robótica, a saber la teoría de los *agentes autónomos*. Entendemos por *agente autónomo* un sistema encarnado (embebido), diseñado para satisfacer objetivos internos o externos mediante sus propias acciones en interacción continua, y a largo plazo con el entorno en el que está situado. La clase de los *agentes autónomos* englobará, de esta manera, a todos los animales y a los robots autónomos (Sánchez 2001). La cuestión, por tanto, estará en abstraer los detalles particulares de su implementación (por ejemplo, células nerviosas vs. máquinas; músculos vs. motores) para entender lo fundamental de esta clase de sistemas: la acción como principal problema para un agente encarnado, y la reflexión como optimizador de su conducta. Esto es importante porque se podría

suerte de metafísica. Véase el apartado 8.8 «Raíces», p. 224, de su *Estar ahí*, donde afirma que «mi versión de “estar ahí” es considerablemente más amplia e incluye todos los casos donde el cuerpo y el entorno local aparecen como elementos en una actividad extendida de resolución de problemas».

confundir este modelo como el del conductismo clásico, es decir, no se trata de que el agente solo reaccione a estímulos mediante respuestas; más bien, significa que un *agente autónomo* tiene que combinar, cuando lo requiera, las respuestas en tiempo real, con sus metas a largo plazo, y estas metas se supone que ya son información lista para ser utilizada por cualquier componente del sistema.

Los robots móviles programados parcialmente para realizar funciones en entornos reales como, por ejemplo, oficinas llenas de objetos, son considerados agentes autónomos. En estos robots se simplifica la inteligencia humana, en términos de procesos formales e inferenciales, a favor de sus elementos biológicos básicos como la integración de funciones sensoriales y motrices para afrontar contingencias de su entorno. Estos robots tienen la posibilidad de autorrepararse y autoabastecerse porque su diseño es diferente: están hechos de varios subsistemas o capas que producen diversas actividades que se sustituyen entre sí, independientemente de un significador o razonador central y de una memoria o archivador. Así, una capa que tiene como función detectar objetos e inmovilizar al robot, si se encuentra inactiva, otra capa puede generar movimientos en otra dirección aleatoria, otorgando una meta más concreta a otra capa llevándole a otro escenario. La idea importante con esta arquitectura, llamada de subsunción,¹⁴ es que al no haber una sede de control central, las conductas de cada capa compiten entre sí por las entradas del entorno. Algo similar al objetivo de la arquitectura conexionista pero con la diferencia de que ésta sigue tratando problemas como el aprendizaje de fonemas y el reconocimiento de rostros a partir de la simulación del entorno, mientras que en este nuevo enfoque lo que se simula son los dispositivos periféricos sensoriales y motrices, pero el entorno es «real». Las conductas solo se pasan entre ellas una señal simple, así la distinción entre cognición y percepción se diluye porque no hay un lugar donde las entradas se traduzcan o trasladen a un código central. Es decir, lo que hay es cooperación y competición entre conjuntos distribuidos de esquemas, y he aquí el punto, no emplean representaciones en el sentido clásico del término sino en el sentido de *representaciones parciales donde cada una capta un aspecto en particular de las interacciones del sistema*. Por representaciones parciales debe entenderse lo opuesto a representaciones exhaustivas.

La característica de este tipo de arquitectura consiste en utilizar una menor capacidad de cálculo, por tanto no actúa utilizando una interpretación interna del mundo, puesto que el mundo es dividido en situaciones mutuamente excluyentes y cada situación dispara diferentes acciones donde cada situación está definida por uno o más sensores. Lo novedoso de esta propuesta es que no se trata de

14 La arquitectura de subsunción fue desarrollada por Rodney Brooks en los 80 y consiste en un modelo basado en el comportamiento reactivo. La idea es descomponer las tareas del robot en conductas simples.

simular las capacidades cognitivas de alto nivel como razonamiento, aprendizaje y conocimiento sino acciones más cotidianas como esquivar obstáculos o recoger objetos particulares.

IV.3. OTRA MANERA DE ENTENDER LAS REPRESENTACIONES

La mente, al no tener ahora el tremendo gasto computacional que implica representar exhaustivamente una situación, utiliza parte de energía en funciones de acoples con su entorno mediante cierta actividad interna que bien podemos llamar representaciones internas. De esta actividad emerge una pauta de organización del sistema sensorio-motriz con capacidad de acomodarse a los cambios en el entorno. Sin embargo, como ha señalado Clark, el entorno no siempre está presente en tiempo real.

En la naturaleza hay correlaciones causales del tipo luz solar/movimiento de las hojas, o secreciones/estímulos, luna/mareas, y no parece que se pueda argumentar que hayan implicadas representaciones. En ese caso, dada la simplicidad de los sistemas (o agentes), no se autoriza el empleo de la explicación representacionista. Pero existen claramente casos en donde la complejidad de las relaciones entre los agentes no son ya de acople sino de desacople. Este desacople permite justamente sustituir las señales del entorno por simulaciones «virtuales» de éste, en casos donde se hace necesario hacer predicciones creando modelos de la realidad. Particularmente pienso en entornos humanos, que además de que están hechos de obstáculos físicos, son las palabras o «artefactos lingüísticos» los que lo pueblan. El aspecto semántico de la representación no se puede excluir de una teoría cognitiva de la mente. Clark en ese sentido apunta que si se entiende las relaciones entre agente/entorno como complejas y dinámicas un sistema explotará «un conjunto entero de correlaciones por sus contenidos semánticos».¹⁵

La compatibilidad entre ítems específicos de información y la necesidad de postular grados de representaciones es evidente en situaciones como la de razonar sobre objetos inexistentes, fuera del alcance del tiempo y la distancia, es decir, no es posible adoptar la actitud antirrepresentacionista de Varela puesto que con su ejemplo de un sistema autoorganizado, «Bittorio», solo es posible explicar algunos casos de la cognición humana. Negar que haya representación y computación es ir en contra de los logros conseguidos por la ciencia y psicología cognitiva, e incluso concederle una oportunidad al conductismo, como ha señalado Martínez-Freire.

La información en forma de señales que sustituyen el entorno en caso de que éste sea inestable, o no exista, permite más plasticidad en la conducta, y

15 Clark (1999), p. 196.

sobre todo mayor capacidad de predicción de situaciones contingentes pero igualmente reales. Calcular trayectorias estelares o la bolsa de valores, hacen parte de la inestabilidad de un entorno en el que hay agentes humanos, haciendo interpretaciones de acontecimientos, clasificándolos o imaginándolos, para luego sacarles provecho. Ya Dennett e incluso Vygotsky se habían fijado en la función que tenían los apoyos externos de la conducta o la cognición. Así como los niños se apoyan en otros durante el desarrollo de la marcha bípeda, asimismo nos apoyamos en el entorno marcándolo con etiquetas lingüísticas principalmente. Pero para ello es necesario disponer de recursos computacionales que permitan abstraer los rasgos que son comunes a objetos que a simple vista se nos presentan como distintos entre sí, como ocurre con los colores o logotipos distintivos de los diferentes tipos de música.

De lo que se está hablando es de clases, universales, o especies. Estamos entonces frente a un sistema de representaciones internas que tiene la capacidad de uniformar entradas sensoriales provenientes de objetos disímiles, o de distinguir estas codificaciones en objetos similares.

Hay representación cuando hay entidades que contienen tipos específicos de información acerca de estados de cosas externas, una perspectiva que aún cuando es un proceso dinámico corpóreo que se deja ver en la adaptación de la conducta, supone actividad computacional.

Lo que ha demostrado la neurociencia es que el cerebro humano integra diversas fuentes de información. Incluso Rodolfo Llinás está de acuerdo cuando explica cómo los organismos segmentados, para generar movimiento integrado, necesitan de un mecanismo de abstracción que resuelva la locomoción. En acciones motrices, por ejemplo al agarrar un objeto, necesitamos de un proceso interno que realiza o emule una copia de dicha acción y devuelva una señal como si llegara de las entradas sensoriales, así el sistema predice cómo debería ser la acción propioceptiva a partir de la retroalimentación virtual. Lo importante es que este emulador interno no se debe considerar como un modelo detallado de la realidad, sino como el resultado de un tipo de arquitectura ajustable y potenciado por el entorno. Clark señala que,

En vez de un argumento claro en contra del computacionalismo y el representacionalismo *en general*, nos encontramos con un conjunto de evidencias que indican que no descubriremos explicaciones computacionales y representacionales *correctas* si no damos la importancia que merece al papel del cuerpo y del entorno local: un papel que incluye definir problemas y, en ocasiones, resolverlos.¹⁶

16 *Ibid.* p. 205.

Sólo es viable hablar de representación cuando los estados internos se coordinan en un entorno de situaciones posibles, los cuales pueden ser leídos por subsistemas del agente y así obtener información sobre el entorno. Estamos autorizados a hablar de representaciones cuando la conexión entre estados internos y el entorno se vuelve muy compleja. Pero lo más importante es que un sistema es denominado complejo cuando *puede utilizar sus códigos internos para generar actividad intrínseca independiente de la estimulación sensorial para generar procesos denominados imaginación o reflexión*.

Clark propone, muy acertadamente, hablar de representaciones personalizadas o locales, en lugar de hacerlo en términos de representaciones exhaustivas, es decir, aquéllas que dependen de la acción, y no de la predeterminación simbólica del mundo.

V. CONCLUSIONES

Ahora pasaré a las críticas. En relación con el término «representación» hemos visto que para algunos «no hay representación sin lenguaje», así sea un lenguaje del pensamiento, para otros cualquier «registro mental» aunque sea simple cuenta como representación, o bien se interpreta como mapas o modelos del entorno parciales o con muchos detalles, e incluso tienen un rol funcional como el de imaginar la conducta futura de los demás, para actuar en función de ello. Examinemos críticamente el tratamiento que se ha hecho de la representación mental con otra luz.

En *De cuerpo presente* Varela dice que se pueden distinguir dos tipos de representación, 1) Representaciones en sentido débil con las que meramente interpretamos el lenguaje, los mapas, etc, y 2) Representaciones en sentido fuerte como símbolos físicos, como el mentalés. Varela acepta el primer sentido pero rechaza el segundo ya que éste tiene un fuerte compromiso con el realismo. Como ha señalado Martínez-Freire en su artículo «El enfoque enactivo en las ciencias cognitivas» (2006), ambos sentidos están estrechamente correlacionados, son codependientes. En efecto, Varela acepta que realizamos *interpretaciones* del mundo suponiendo que éste es de tal o cual manera, pero el problema es que las interpretaciones de las que habla son representaciones de segundo orden, dentro de una jerarquía de representaciones, donde las interpretaciones son representaciones de representaciones. Las segundas no podrían existir sin las primeras. Así las representaciones mentales primarias son internas (o externas) al sistema. Sin duda que toda representación necesita ser interpretada, lo cual requiere de un intérprete.

Un intérprete, siguiendo el artículo «Representación e interpretación» del filósofo español, es «algo o alguien que da sentido a una representación». El tipo

de interpretación puede considerarse sólo como la representación del objeto a secas, como cumpliendo una función, o como referente en el entorno lingüístico del intérprete. Interpretar es pensar, y pensar es una actividad representacional, una actividad en la que se mantiene una relación con otra cosa. Otro sentido de las interpretaciones se da cuando el objeto es representado en un tiempo que ya ha transcurrido, o que aún no ha ocurrido. La consideración del pasado y el futuro posible, son también la marca de la verdadera cognición.

En efecto, son las representaciones secundarias las que permiten que se diluya el acople entre sistema cognitivo y sistema ecológico a favor de lo que ha denominado Martínez-Freire *desacople*. Lo significativo, al menos en el caso humano es el desdoblamiento de la realidad producido por las representaciones sobre las representaciones. El problema con Varela es que estaría considerando acoples permitidos solo mediante «representaciones sensorio-motrices». Lo particular de su argumento es que al admitir representaciones en el sentido débil está hablando de representaciones de segundo orden, o interpretaciones que necesitan de la ontologización de las representaciones.

No hay que olvidar que es precisamente el lenguaje (y por tanto la teoría) el que genera el desajuste respecto a las propiedades de las cosas; hay variedad representacional que permite la construcción de situaciones hipotéticas, ausentes o fantásticas que permiten a quien usa y comprende el lenguaje «vivir en otra realidad».

A propósito de las oraciones, el filósofo, Norwood Russell Hanson en su libro *Observación y explicación. Patrones de descubrimiento*, de 1958, ya había advertido que estas «pueden ser usadas para hacer aseveraciones, descripciones, suministrar narraciones, informes, etc., ninguno de los cuales depende de la posesión de algunas propiedades en común con aquello a lo que se refiere la oración».¹⁷ De hecho, lo que hacemos con el lenguaje es desajustar el corsé de la realidad perceptible y limitante que encontramos, por ejemplo, en las representaciones pictóricas o en las grabaciones. La descripción de experiencias distintas sensorialmente simultáneas sólo son posibles en el lenguaje.

La causa de este malentendido es que se piensa en la cognición como acciones, movimientos de agarre, percepciones, desplazamientos en entornos físicos, pero se olvida que esta excesiva simplificación de la cognición deja de lado, por ejemplo, la metarrepresentación, una habilidad que tenemos los humanos para representarnos los estados mentales de otros organismos. Así este nuevo nivel de representación requiere del contraste entre lo que ocurrió con lo que debió de haber ocurrido, para aplicarlo en un entorno cognitivo. No hay que olvidar que nuestro entorno no es ya tan biológico, está plagado de *trayectorias representacionales*, que permiten el aprovechamiento de las diferentes versiones del mundo y de las mentes de otras personas.

17 N. R. Hanson (1977), p. 109.

Varela en su afán de desactivar el realismo no considera que las interpretaciones estén montadas sobre las representaciones que ocurren en un sistema cognitivo del intérprete causadas por el mundo. Por el contrario, o el mundo se confunde con la cognición, o lo alumbra o hace emerger, como resultado de la acción efectiva. Nuestro autor no acepta el acceso directo o indirecto al mundo; el mundo es enactuado, y al igual que para Clark, es el resultado de acoplamientos. Sin embargo, ¿si se postula que hay representaciones primarias, estas lo son de qué? ¿Qué es lo que representan? Creo que lo mejor es postular que existe la Realidad, algo allá afuera que lleva mucho más tiempo que yo, incluso que nuestra especie, y más aún anterior a la vida misma; pero esta realidad es interpretada en tanto mundo percibido, por el particular proceso evolutivo de la estructura cognitiva de cada especie, incluyendo por supuesto, la nuestra. Es por esto que los seres humanos tenemos versiones del mundo que concuerdan en algunos aspectos y no lo hacen en otros: porque tenemos similitudes biológicas y cognitivas debido a las mismas presiones selectivas como especie. En consecuencia, existe un mundo ontológicamente independiente de nuestras mentes que contienen diferentes puntos de vista (históricos, culturales, intelectuales) sobre éste. Algunas de estas versiones estarán mejor adecuadas para ciertos propósitos que otras. El antirrealismo de Varela solo sería posible si se exagera las diferencias epistemológicas entre perceptores, sobre todo humanos, negando incluso el conocimiento objetivo que pueda conciliarse mediante el lenguaje (Martínez-Freire 2007). Pero se deriva del error de estar atrapado en la vieja idea de que «las cosas representan otras cosas en virtud de la semejanza entre ellas».

Este criterio supone un realismo ingenuo que consiste en que la mente hace copias de los objetos externos, es decir, si se piensa en una esfera roja es por que de hecho debe haber algo en el mundo como una esfera roja. Pero semejanza no es lo mismo que representación, como ha puesto de relieve el psicólogo J. Perner (1994), si así fuera no podríamos explicar las representaciones erróneas, por ejemplo cuando la imagen de un objeto está distorsionada por problemas en el sistema visual, o en el entorno, pues no se supone que haya algo en la realidad de ese modo. Así mismo el criterio de semejanza no da cuenta de la representación de entes no existentes, imaginados o hipotéticos como en el caso de «Pegaso». La semejanza sea quizás aplicable a representaciones primarias donde hay más fidelidad entre el reflejo y el mundo.

Afirmo entonces que no es posible negar con los argumentos que propone Varela la existencia de un mundo exterior a nuestra mente representacional, que es captado indirectamente por las experiencias mentales; como ha defendido Martínez-Freire el mundo existe para las mentes «en el sentido en que la única y posible experiencia de realidades se da en las representaciones mentales». Las representaciones son el rasgo básico de nuestra mente, sea ésta artificial o biológica.

REFERENCIAS BIBLIOGRÁFICAS

- BRONCANO, F. (1995): «El control racional de la conducta», *La mente humana*, Edición de F. Broncano, Madrid, Editorial Trotta, S.A.
- CLARK, A. (1999): *Estar ahí. Cerebro, cuerpo y mundo en la nueva ciencia cognitiva*. Barcelona, Ediciones Paidós Ibérica, S.A.
- CLARK, A. (2002): «That special something: Dennett on the making of minds and selves», in A. Brook and D. Ross (Eds) *Daniel Dennett*, Nueva York, Cambridge University Press.
- CUARTAS, J. M., (2007): «Modelar el cerebro: un sobrevuelo al conexionismo», en *Los rumbos de la mente. Ensayos sobre el yo, lo mental natural y la Inteligencia Artificial*, Editorial San Pablo, Universidad Pedagógica Nacional.
- DENNETT, D. (1998): «Los verdaderos creyentes: La estrategia intencional y por qué funciona», *La actitud intencional*, Barcelona, Editorial Gedisa, S.A.
- DENNETT, D. (2000): «La creación del pensamiento», *Tipos de mentes*, Madrid, Editorial Debate, S.A.
- FODOR, J. (1984): *El lenguaje del pensamiento*, Madrid, Alianza Editorial, S.A.
- FODOR, J. (1995): «Las actitudes proposicionales», en *Filosofía de la mente y ciencia cognitiva*, Eduardo Rabossi (compilador), Barcelona, Ediciones Paidós Ibérica, S.A.
- GARCÍA-CARPINTERO, M. (1995): «El Funcionalismo», *La mente humana*, Edición de F. Broncano, Madrid, Editorial Trotta, S.A.
- GARDNER, H. (1988): «Las ciencias cognitivas. Perspectiva histórica», *La nueva ciencia de la mente. Historia de la revolución cognitiva*, Barcelona, Ed. Paidós.
- GUERRERO, G. (1999): «Teoría de la evidencia y holismo moderado en W. V. Quine», *Praxis filosófica. Filosofía antigua*, N° 8/9, Departamento de Filosofía, Universidad del Valle.
- GUIJARRO, J. L. (2000): «Cognición», en *Compendio de Epistemología*, Jacobo Muñoz & Julián Valverde (compiladores), Madrid, Trotta.
- HANSON, N. R. (1977), *Observación y explicación: guía de la filosofía de la ciencia. Patrones de Descubrimiento. Investigación de las bases conceptuales de la ciencia*, Madrid, Alianza Editorial.
- MARTÍNEZ-FREIRE, P. (1995): *La nueva filosofía de la mente*, Barcelona, Editorial Gedisa, S.A.
- MARTÍNEZ-FREIRE, P. (2005): «Putting representation, body and world together again», en *Contrastes. Cognición y Representación*. Edición de P. Martínez-Freire, Suplemento 10, Málaga.
- MARTÍNEZ-FREIRE, P. (2006): «El Enfoque Enactivo en las Ciencias Cognitivas», *Ludus Vitalis*, Vol. XIV, num. 26.
- MARTÍNEZ-FREIRE, P. (2007): «Representación e Interpretación», en *La importancia del conocimiento. Filosofía y Ciencias Cognitivas*, 2ª edición, La Coruña, Netbiblo.
- NEWELL, A.: «Sistemas de símbolos físicos», en *Perspectivas de la ciencia cognitiva*, D. Norman (compilador), Barcelona, Paidós.

- NORMAN, D. (1987): «¿Qué es la ciencia cognitiva?», en *Perspectivas de la ciencia cognitiva*, D. Norman (compilador), Barcelona, Paidós.
- OJEDA, C. (2001): «Francisco Varela y las ciencias cognitivas», en *Revista Chilena de neuro-psiquiatría*, 39(4): 286-295.
- PERNER, J. (1994): «El concepto de representación», en *Comprender la mente representacional*, Barcelona, Ediciones Paidós.
- PYLYSHYN, Z. (1988): *Computación y conocimiento*, Madrid, Editorial Debate.
- RAMOS, J. (2000): «Simbolismo vs. Conexionismo: la estructura de las representaciones», en *Mentes reales. La ciencia cognitiva y la naturalización de la mente*, Juan José Botero, Jaime Ramos y Alejandro Rosas (compiladores), Siglo del hombre Editores, Universidad Nacional de Colombia, Departamento de Filosofía.
- SÁNCHEZ Chacón, R. (2001): *Cognición situada: una aproximación robótica*. Tesis de Maestría en Inteligencia Artificial, Universidad Veracruzana.
- SIMON, H. (1987): «Ciencia cognitiva: la más nueva ciencia de lo artificial», en *Perspectivas de la ciencia cognitiva*, D. Norman (comp.), Barcelona, Paidós Ibérica, S.A.
- SMOLENSKY, P. (1995): «La estructura constitutiva de los estados mentales conexionistas: una respuesta a Fodor y Pylyshyn», en *Filosofía de la mente y ciencia cognitiva*, E. Rabossi (comp.), Barcelona, Ed. Paidós.
- VARELA, F. (1990): *Conocer. Las ciencias cognitivas: Tendencias y perspectivas. Cartografía de las ideas actuales*, Barcelona, Gedisa.
- VARELA, F., Thompson, E., Rosch, E., (1991): *The Embodied Mind. Cognitive Science and Human Experience*, Cambridge (MA), MIT Press. Versión castellana: *De cuerpo presente. Las ciencias cognitivas y la experiencia humana* (trad. Carlos Gardini), Barcelona, Gedisa, 1992.
- VÉLEZ, J. C. (2004): «Relativismo perceptivo, realismo y enacción», *Praxis Filosófica*, Nueva serie, N. 19, Julio- Diciembre, Departamento de Filosofía, Universidad del Valle.

JUAN CARLOS VÉLEZ es licenciado en Filosofía y Magíster en Filosofía por la Universidad del Valle, Cali, Colombia. Profesor Facultad de Educación, Universidad Católica *Lumen Gentium*. Profesor Facultad de Humanidades, Universidad del Valle. Investigador del grupo *Episteme. Filosofía y ciencia*. Departamento de Filosofía, Universidad del Valle.

Publicaciones recientes:

«Relativismo perceptivo, realismo y enacción», en revista *Praxis Filosófica*, Nueva serie, N. 19, Julio- Diciembre, 2004, Departamento de Filosofía, Universidad del Valle. En preparación: «Teoría de la mente y psicología intencional».

Línea de investigación:

Filosofía de la mente y de las ciencias cognitivas.

Dirección electrónica: karlegion@yahoo.es

