

# EPISTEMOLOGÍA E HISTORIA DE LA CIENCIA

SELECCIÓN DE TRABAJOS DE LAS X JORNADAS

VOLUMEN 6 (2000), Nº 6

Pio García  
Sergio H. Menna  
Víctor Rodríguez  
Editores



ÁREA LÓGICO-EPISTEMOLÓGICA DE LA ESCUELA DE FILOSOFÍA  
CENTRO DE INVESTIGACIONES DE LA FACULTAD DE FILOSOFÍA Y HUMANIDADES  
UNIVERSIDAD NACIONAL DE CÓRDOBA



Esta obra está bajo una Licencia Creative Commons atribución NoComercial-SinDerivadas 2.5 Argentina



# Intencionalidad y cognición

## El rol de la representación en los sistemas biológicos y artificiales

Gustavo Agüero\*

### 1. Introducción

En los últimos años, se ha convertido en una actitud recurrente el ataque contra el “estilo clásico” de trabajo de la ciencia cognitiva contraponiéndolo con los “nuevos modos de trabajo” y las nuevas maneras de realizar las investigaciones. No obstante dentro, del llamado “estilo clásico” de la ciencia cognitiva se encuentran muchas estrategias y conceptos que los nuevos cognitivistas no están dispuestos a abandonar; entre los máspreciados recursos con los que se formó el cuerpo de las primeras investigaciones se encuentran dos reconocidas estrategias: *representar* y *computar*.

Pese a las críticas, muchos de los “nuevos enfoques”, cuyos líderes parecen ser modernos conexionistas, siguen asumiendo posiciones básicamente representacionistas, al estilo de los cognitivistas clásicos y quienes más fuertemente han cuestionado la estrategia de postular estados internos semánticamente caracterizados no tienen aún nada que ofrecer a cambio.<sup>1</sup>

Este trabajo no busca alimentar la bibliografía de las discusiones entre “clásicos” y “modernos” cognitivistas sino abordar un tema que ha recibido relativamente muy poco tratamiento siendo quizás uno de los más importantes del estudio de los sistemas cognitivos, me refiero a la propia estrategia *representacionista*. Téngase en cuenta que cuando hablamos de *la* estrategia no estamos haciendo referencia a la editorialmente importante discusión acerca de las arquitecturas de *representación*, ni tampoco habría que identificar la estrategia representacionista con alguna de sus variantes como lo son las estrategias simbólicas.

“Una explicación es representacionista, como afirma Clark (1999, pág. 197), si describe sistemas enteros de estados (locales o distribuidos) o procesos (secuencias temporales de estos estados) internos identificables, como entidades cuya función es contener tipos específicos de información acerca de estados de cosas corporales o externos. Por tanto la teorización representacionista se encuentra en los tramos superiores de un continuo de posibilidades cuyos límites inferiores, que no son representacionistas, incluyen meras correlaciones causales y casos muy simples de lo que suele denominarse «conexión adaptativa». Las conexiones que requieren suponer que los estados internos del sistema coordinan sus conductas (mediante la evolución, el diseño o el aprendizaje) con contingencias ambientales específicas.”

Otra manera interesante de plantear el rol de la representación en los sistemas cognitivos es considerar la pregunta de porqué, en una amplia variedad de casos, funcionan las explicaciones psicológicas ordinarias; me refiero no sólo a que resulta una útil estrategia predictiva, sino a que nos aporta mucha mayor comprensión de ciertos aspectos de nuestro

\* Universidad Nacional de Córdoba.

comportamiento en tanto seres humanos y del comportamiento de muchos otros organismos de lo que cualquier explicación a la mano podría aportar. Dejando de lado la *armonía pre-establecida*, el *paralelismo* u otro tipo de coincidencia milagrosa, la cuestión de cómo es posible que diferentes clases de organismos biológicos y artificiales puedan tener en común un cierto *respeto* por la psicología de creencias y deseos es el gran tema de la investigación sobre sistemas cognitivos.

De acuerdo a una influyente propuesta sobre la manera abordar el estudio de los sistemas que procesan información formulada entre otros,<sup>2</sup> por David Marr, la respuesta a esto seguramente tendrá que ser buscada en los mecanismos internos del sistema. La conocida estrategia explicativa fue presentada en los siguientes términos:

un extremo, el nivel superior, es la teoría abstracta del cálculo del mecanismo. En ella se caracteriza el funcionamiento de este como una proyección de un tipo de información en otra, se definen de modo preciso las propiedades abstractas de esta proyección y se demuestran su adecuación y conveniencia. En el centro aparece la elección de representación para la entrada y la salida así como el algoritmo a emplear para transformar una en otra, y en el extremo opuesto podemos observar los detalles acerca del modo en que se realizan físicamente el algoritmo y la representación – la arquitectura detallada del ordenador, por así decir.<sup>3</sup>

Por lo tanto el rol de la representación en la investigación cognitiva consiste en vincular propiedades psicológicas y propiedades computacionales, o dicho de otra manera, la teoría de la representación cumple la función de establecer la conexión entre el funcionamiento computacional y las explicaciones psicológicas (cfr. Cussins, 1990).

La idea que subyace a estas investigaciones, como lo afirma Marr (1985, pág. 28), es la de que “debe existir un nivel de comprensión, en el que las tareas de procesamiento información (...) se analicen y se comprendan de modo independiente a los mecanismos y estructuras particulares que los implementan en nuestro cerebro.”

No voy a intentar definir aquí alguna noción clásica o nueva de “representación” ni caracterizar ningún uso en particular, este no pretende ser un trabajo conceptual, al menos en ese sentido; lo que pretende es considerar los supuestos teóricos y los rasgos más relevantes de la estrategia que más aportes ha hecho a las investigaciones sobre los procesos cognitivos: el *representacionalismo*.

## 2. La distinción entre “intención” y “cognición”

Si bien hay que reconocer que algunos filósofos hablan indistintamente de propiedades *intencionales* y propiedades *representacionales*, no seguiré aquí ese ejemplo, por el contrario, admitiré la posibilidad de realizar esta distinción lo cual constituye el fundamento mismo este trabajo, aunque también el de toda la empresa cognitiva en general. La razón más fuerte para aceptar la distinción entre *contenido intencional* y *contenido representacional* radica en que de lo contrario se pondrían sobre la teoría de representación restricciones demasiado fuertes y que no tienen que ver necesariamente con ofrecer una explicación correcta del modelo de la competencia del sistema (como veremos luego). Como afirma Cummins (1989, pág. 15), aunque la atribución de creencias resulta un caso legítimo de caracterización semántica, “no es una caracterización de alguna representación en el creyente.” No obstante, cuando descubrimos que cierta explicación psicológica o teoría de la competencia resulta *adecuada* para comprender el comportamiento de un sistema, “nos

esforzamos, como dice Dennett (1991, pág. 41), por interpretar algunos de sus estados o procesos internos como representaciones internas.”

Ahora bien, una característica distintiva del análisis estratificado de los sistemas es que el compromiso “realista” que se asuma con relación a los hechos y propiedades de un cierto nivel se manifiesta en el grado en que estos puedan ser “respaldados” por los hechos y propiedades que puedan encontrarse en los niveles de análisis más bajos. De esta manera, el “optimismo” propio de los realistas se expresa en la confianza que depositan en el hallazgo de mecanismos que *realicen o implementen* funcionalmente las propiedades psicológicas más abstractas que describen la competencia del sistema. Hay que tener en cuenta que de acuerdo a la estrategia explicativa aceptada, una mala comprensión o una visión descriptivamente inadecuada al nivel más abstracto provocará, sin lugar a dudas, confusiones y dejará preguntas sin responder en los niveles más bajos de descripción. El mismo Marr ha mostrado a través de sus investigaciones sobre la visión que hasta que no contemos con una comprensión clara y precisa a nivel computacional, que responda a preguntas acerca de *qué* capacidades desarrolla el sistema, no podemos dirigir, de forma adecuada, preguntas detalladas acerca *cómo* podría un programa estar diseñado para que tales capacidades sean efectivamente desarrolladas. De hecho, las propiedades abstractas del sistema definidas en el nivel superior cumplen la función de *sugerir* tipos adecuados de representación que puede emplear el sistema para llevar a cabo esas operaciones abstractas, considerando las características físicas del mecanismo que instancia el sistema.

Frente a esto, quienes comparten una actitud más declaradamente “pesimista” suponen que las cosas difícilmente resulten tan favorables al esquema estratificado. Junto a los nuevos conexionistas, filósofos influyentes como Dennett (1991, pág. 78) recuerdan a todo entusiasta el hecho de que “la descomposición del modelo de la competencia de alguien en partes, fases, estados, pasos o lo que sea, *no necesita* arrojar ninguna luz sobre la descomposición de las verdaderas partes mecánicas, fases, estados o pasos del sistema...”

Si esto es así, entonces el proyecto cognitivista debería tratar de obtener el diseño de un mecanismo o modelo computacional cuyo comportamiento pueda, en circunstancias y condiciones *adecuadas*, describirse mediante la atribución de estados intencionales. Entonces, el diseño propuesto por el cognitivista debe *crear la ilusión* de que el sistema funciona porque tiene como estados *internos* aquellos que le son *atribuidos* por el método intencional. Aceptar esto no significa afirmar que el sistema *no deba* o no tenga que tener estados internos que den cuenta del éxito de la explicación psicológica, significa que es suficiente con que el sistema  *Cree la ilusión* de que posee dichos estados, respondiendo a las expectativas de la estrategia intencional.

Hay dos maneras en las cuales el nivel 2 (de procesamiento) podría relacionarse estructuralmente con el nivel más abstracto de explicación: la manera más sencilla es la que propone el cognitivismo clásico, según la cual podría implicar representaciones simbólicas de reglas que son leídas y seguidas por el sistema; la manera más compleja podría implicar conocimiento tácito de reglas que son seguidas por el sistema.

Desde un punto de vista conexionista, la relación entre la estructura de la teoría de la competencia y el nivel de procesamiento es totalmente diferente del que propone la visión clásica del cognitivismo. El nivel 2 conexionista no refleja en ningún sentido fuerte los aspectos estructurales de la teoría de la competencia. La relación, según el ilustrativo ejemplo propuesto Rumelhart y Mc Clelland (1986) es en algún sentido análoga a la que existe

entre la mecánica clásica y la física cuántica: “los principios newtonianos describen y predicen el comportamiento de los sistemas físicos en una amplia variedad de casos” no obstante el hecho de que el universo físico, no sigue las leyes clásicas, aunque hay que reconocer, como señala Clark (1990), que bajo ciertas condiciones especificables, “se comporta como si lo fuera.”

En esto radica precisamente la importancia de distinguir entre *intención* y *representación*, como señaláramos anteriormente, no necesitamos poner tan fuertes restricciones sobre los mecanismos representacionales como las que se derivarían por ejemplo de la exigencia de que para cada contenido intencional haya un estado interno que lo soporte.

Ahora bien, definir la cuestión acerca de si el sistema posee estados internos que “respaldan” o “legítiman” en algún sentido los estados psicológicos intencionales o si solo posee un mecanismo que *crea la ilusión* de que hay tales estados internos, es algo que tiene que ser resuelto, al menos en parte, en base a la investigación empírica.

### 3. Representación e indeterminación

Es una norma aceptada entre los investigadores de sistemas cognitivos aquella que expresa que, cuando se descubre *algo* cuyo comportamiento puede comprenderse en base a alguna estrategia de explicación psicológica se intenta reconocerlo como un *sistema* o un dispositivo de cierta complejidad con una organización tal que le permite acceder, de manera *relevante*, a diferentes estados y procesos, y se busca interpretar algunos de dichos estados y procesos como representaciones internas de su entorno, en virtud de las cuales el sistema regula su comportamiento.

Muchos de los filósofos que ven al realismo intencional como una posición desmesurada tienen motivos que se inspiran fuertemente en la tesis quinenana de la indeterminación de la traducción. Esta es básicamente la tesis interpretacionista según la cual encontrar sentido a las preferencias de un hablante de una lengua diferente a la propia consiste en reconocer, en su conducta verbal, la existencia de un modelo o patrón. Es decir, la traducción consiste en encontrar significado a las expresiones de alguien que no comparte nuestro lenguaje, o bien podría decirse consiste en reconocer las expresiones mismas. Hay que considerar que esta tesis interpretacionista depende, en parte al menos, de asumir que “el significado no es una existencia psíquica; es primariamente una propiedad de la conducta,” tal como lo expresó el propio Quine (1969) citando a Dewey.

Una de las posibilidades que abre esta indeterminación es la de que puedan existir más de un modelo de interpretación “correcta” respecto del *significado* de las expresiones del hablante o del *contenido* de las intenciones que le son atribuidas. Resulta posible el caso en que bajo cierta interpretación, se atribuya a un sujeto la creencia *A* y bajo cierta otra la creencia *B*, resultando que tanto *A* como *B* explican adecuadamente el comportamiento del sujeto. Esto es lo que podemos obtener en base a toda la evidencia disponible, no nos ayuda aquí el más profundo conocimiento de los hechos neuronales o de cualquier otro tipo.

Es precisamente el hecho de que puedan existir interpretaciones alternativas igualmente “adecuadas” lo que constituye para algunos el argumento más fuerte en contra del realismo intencional.

En estos términos, la discusión se propone como un enfrentamiento que coloca a una tesis ontológica – la indeterminación de la traducción – frente a ciertas hipótesis explicativas (de fuerte tono realista). La cuestión es entonces, si puede objetarse el valor de una hipóte-

sis explicativa realista o no, sobre la base de una tesis ontológica de carácter pragmático o no. Tyler Burge (1986) ha enfocado la situación con gran claridad al afirmar que "es simplemente ilusorio pensar que las preconcepciones metafísicas o aún cognoscitivas provén un estándar para juzgar las ontologías o los esfuerzos explicativos de las ciencias particulares" y agrega "las cuestiones acerca de lo que existe, de cómo se individualúan las cosas y qué se reduce a qué, surgen en referencia a las prácticas descriptivas y explicativas corrientes. Las respuestas que se proponen a estos problemas no pueden utilizarse, por sí mismas, para criticar un modo, por otra parte exitoso de descripción y explicación."

El éxito o el fracaso de una hipótesis empírica probablemente dará elementos que ayuden a definir el carácter de la tesis ontológica acerca de la naturaleza representacional de los estados internos; pero una vez más, una tesis ontológica tiene una relevancia filosófica de la que los prejuicios metafísicos carecen. Por supuesto que esto no pretende restar valor a la tesis de Quine, sino que busca, por otro lado, señalar que no tiene en sí misma la fuerza que algunos le atribuyen como para desechar una hipótesis explicativa sobre la sistematicidad del pensamiento.

#### 4. Representación y cognición

La estrategia *representacionista*, es decir, la estrategia de postular estados internos del sistema que poseen contenido semántico tiene como objetivo principal contribuir a dar una explicación del éxito de la estrategia predictiva de atribución intencional, mostrando cómo es posible para un sistema físico tener estados mentales.

Por lo tanto, las preguntas fundamentales en torno a esta estrategia son cuestiones tales como ¿cuándo puede decirse que un sistema posee estados internos o representaciones? ¿cómo reconocer un mecanismo representacional? o ¿qué sistemas posee tales mecanismos?

En primer término, así como no aceptamos como un hecho bruto no analizable que algo pueda representar otra cosa, de la misma manera, tampoco aceptamos como incuestionable que los sistemas biológicos y en particular los seres humanos o su cerebro, incluyan como estados internos, representaciones de su mundo exterior.

En algún sentido, que intentaré precisar, la pregunta acerca de si un sistema posee representaciones internas guarda cierta analogía con la pregunta acerca de si el sistema posee estados psicológicos como creencias y deseos, es decir, si tiene estados mentales. Si bien esto podría concederse cautelosamente, sugeriré algo bastante más difícil de conceder, a saber, que también existe cierta analogía en las respuestas a estas preguntas. Lo que pretendo señalar aquí con esta analogía es, por un lado, la cuestión nada obvia de que existen representaciones internas y por otro la pertenencia de estas problemáticas a los niveles interpretativos superiores o más abstractos que explican el comportamiento de los sistemas complejos. Es cierto que hay muchas maneras de representar algo, pero si de algo depende que un objeto sea una representación de algo, no es en absoluto de una propiedad intrínseca o individualista, sino de una propiedad funcional; es decir, del papel que el objeto desempeña dentro de un sistema. Esto significa, particularmente, que la capacidad de algo para representar es un hecho que puede establecerse con total objetividad.

La adopción de la estrategia representacionista de base computacional constituye la manera más tradicional de tratar la cognición, aunque si bien, en los últimos años han surgido algunas estrategias alternativas al representacionismo, resulta difícil igualar los mé-

ritos alcanzados por el par representación y computación. La idea básica del representacionista es la de que ciertos procesos y estados internos del sistema poseen la capacidad de covariar de acuerdo a ciertos parámetros con las modificaciones del entorno o del propio sistema. Bajo tales condiciones decimos que un estado o proceso del sistema representa información relevante acerca de su entorno o de sí mismo y que esa información le permite regular su comportamiento. Estas son características generales de la estrategia representacional más allá de que se considere una noción más o menos robusta de representación o de un tipo local o distribuido.

## 5. Las implicancias del representacionalismo

Pues las teorías representacionistas constituyen la mayor parte de los intentos hasta hoy realizados en la investigación cognitivista, pero ¿qué implica el representacionalismo?

Implica básicamente aceptar una explicación del funcionamiento de un sistema sobre la base de su comportamiento; es decir, de sus relaciones con su entorno próximo. Una explicación de base conductual, implica a su vez, la atribución de estados y procesos que harían comprensible y racional el comportamiento del sistema; no obstante, no se supone que al utilizar esta estrategia de atribución de intenciones se estén formulando hipótesis sobre posibles estados y procesos internos del sistema, todo lo cual podría ser sometido a algún tipo de verificación.

Es decir, la estrategia representacionista en ciencias cognitivas equivale a aceptar la estrategia explicativa de análisis por niveles como lo sugirieron Marr y otros; esto es, comenzar por la comprensión adecuada del comportamiento y los propósitos del sistema y sólo entonces iniciar la búsqueda sobre los mecanismos subyacentes que “justifiquen” la conducta manifiesta. Esto parece contradecir, de alguna manera, aquello que afirmaba Dennett acerca de que la descomposición del modelo de la competencia no *necesita* arrojar ninguna luz sobre la descomposición de las verdaderas partes mecánicas del sistema. Podría aceptarse que no exista necesariamente un estricto “isomorfismo” entre los niveles de la competencia y de la ejecución, no obstante afirmar que el análisis de un nivel “no necesita arrojar ninguna luz” sobre el otro equivale a rechazar el análisis por niveles que propone Marr y que Dennett mismo acepta y recomienda explícitamente. De hecho el beneficio principal de esta estrategia de análisis consiste precisamente en que las buenas explicaciones a un nivel deben allanar el camino para los niveles inferiores, el propio Dennett (1991, pág. 77) se refiere a esta cuestión señalando que “el punto de vista estratégico más revelador de Marr es que si uno tiene una visión totalmente equivocada acerca de cuál es la descripción a nivel computacional de su sistema (...) sus tentativas de teorizar a niveles más bajos se verán confundidas por acertijos artefacticos espurios.” Esto señala la principal ventaja de la estrategia explicativa por niveles, sin embargo Dennett (1991, pág. 77) considera que “lo que Marr subestima es el punto hasta el cual las descripciones a nivel computacional pueden también confundir al teórico que olvida cuán idealizadas están.”

La respuesta a este injustificado escepticismo parece ser bastante sencilla, o bien se acepta la estrategia explicativa con lo que ello implica o bien se la rechaza; pero aceptarla suponiendo que la descripción de un nivel no tiene porqué favorecer la descripción de otro equivale a aceptarla rechazando la principal ventaja que la estrategia ofrece.

La cuestión de si las descripciones de nivel 1 o psicológicas arrojan alguna luz sobre el nivel *algorítmico* no puede ser una cuestión empírica tal como se plantea, aceptar la estra-

tegia de análisis de Marr implica aceptar que la relación entre los niveles tiene decididamente un fuerte carácter normativo; esto significa decir que las descripciones del nivel 1 están pensadas, en el esquema explicativo, especialmente para arrojar luz sobre la teoría de la ejecución; no disponemos de ninguna guía alternativa hacia los niveles más bajos de análisis del sistema.

## 6. Conclusiones

Para finalizar diremos que, efectivamente, lo que hace representacional a un estado o proceso internos en un sistema es la función que desempeña en la regulación de su comportamiento (ver Dennett, 1991, pág. 41). Esto es otra manera de decir que existe una íntima conexión entre los niveles explicativos 1 y 2, o entre explicaciones psicológicas ordinarias y los modelos de procesamiento subyacentes. En síntesis, la estrategia de postular representaciones internas para explicar el comportamiento de un sistema biológico o artificial es quizás la manera más acabada de aplicación de los métodos hermenéuticos en la explicación de sistemas físicos complejos.

Difícilmente un abordaje no representacionista pueda aportar la comprensión que ofrece esta estrategia al vincular explicaciones intencionales con explicaciones causales a través de mecanismos del sistema que pueden manipular información.

## Notas

<sup>1</sup> Una interesante presentación de estas discusiones se halla en Clark (1999)

<sup>2</sup> Entre otros como Chomsky, Newell y Simon.

<sup>3</sup> Marr, D. (1985, pág. 33).

## Bibliografía

Clark, A. (1999) *Estar ahí*, Ed. Paidós.

Boden, M. (Ed.) (1994) *Filosofía de la Inteligencia Artificial*, Ed. Fondo de Cultura Económica.

Cussins, A. (1990) La construcción conexionista de conceptos, *Filosofía de la Inteligencia Artificial*, (Ed.) Boden, M., Ed. Fondo de Cultura Económica.

Marr, D. (1985) *La visión*, Ed. Alianza.

Dennett, D. (1991) *La actitud intencional*, Ed. Gedisa.

McClelland, J. - Rumelhart, D. - Hinton, G. (1986) Las representaciones distribuidas; Boden (Ed.) 1994.

Quine, W. (1969) *La relatividad ontológica y otros ensayos*, Ed. Tecnos.

Burge, T. (1986) Individualism and psychology; *The Philosophical Review*, XLV, nº 1.