

# 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



# Sistemas de cambio de creencias: una propuesta en el marco de la teoría de modelos

Eleonora Cresto\*

## 1. Introducción

Una teoría de cambio de creencias propone básicamente dos cosas: en primer lugar, propone algún modo de representar el estado epistémico de cierto agente, y, sobre todo, en segundo lugar, propone algún modo de representar el proceso por el cual el agente adquiere nuevas creencias o se deshace de las viejas. La empresa, recordemos, es concebida en un sentido normativo: se trata de capturar los rasgos esenciales que debe exhibir una dinámica de creencias para poder considerar que el agente en cuestión está dotado de racionalidad. La llamada teoría AGM (la teoría desarrollada por Alchourrón, Gärdenfors y Makinson<sup>1</sup>), es actualmente, en muchos sentidos, el punto de partida estándar para discutir estos temas.

Ahora bien, tanto AGM como sus variantes se ocupan de cambios racionales mínimos, y dejan explícitamente de lado las “revoluciones conceptuales” que puedan acaecer en el conjunto de creencias de un agente. Mi intención en este trabajo es mostrar que es posible construir una teoría de cambio racional de creencias que abarque también modificaciones no-estándar, o estructurales. Más aún, veremos que es posible construirla como una peculiar generalización de la teoría estándar AGM.

Por supuesto, antes de continuar es imprescindible aclarar qué se entiende por expresiones como “cambio no estándar” o “estructural”. Debo mencionar que *no* tengo en mente ninguna de las clásicas tesis kuhnianas sobre cambios de paradigma. En particular, no supondré que un cambio estructural entraña la incommensurabilidad del nuevo marco conceptual respecto del anterior. Y, por supuesto, no presupondré que dichos cambios sean *en principio* irracionales. (Si este fuera el caso no podrían ser incluidos en una teoría del cambio racional de creencias, desde luego). Por lo demás, lo que motiva el presente proyecto es la intuición de que AGM, y otras teorías afines, nos obligan a aceptar la siguiente situación indeseable:

Supongamos que concebimos los estados de creencias como conjuntos de proposiciones, representados ya sea por conjuntos de oraciones o por conjuntos de mundos posibles. Es claro que un agente puede tener tres actitudes básicas frente a una proposición: puede creerla, puede creer su negación, o puede permanecer en suspenso. Pero notemos que la categoría de permanecer en suspenso es extraordinariamente amplia (mejor dicho, *indeseablemente* amplia): comprende tanto proposiciones sobre las cuales el agente tiene genuinas dudas, así como proposiciones que éste podría no haber siquiera concebido como posibles objetos de creencia, porque no le resultan en absoluto significativas. Para ver esto, supongamos que un sujeto *X* es un hablante de castellano sin ningún tipo de instrucción científica. Las teorías estándar están forzadas a concluir, por ejemplo, que *X* ha suspendido el juicio acerca de si hay un agujero negro en el núcleo activo de la galaxia *xyz*. Sin embargo, esto no parece razonable. Por hipótesis, *X* no conoce el concepto de “núcleo galáctico activo” o “agujero negro”; no hay galaxias en su ontología, y, a su entender, “*xyz*” no nombra nada.

\* Columbia University

Sin embargo, en un modelo proposicional de cambio de creencias no hay manera de distinguir entre permanecer en suspenso sobre cierta (posible) creencia o ser completamente inconsciente de la mera posibilidad de formular la creencia en cuestión. La razón de esto es que el reino de las proposiciones debe ser considerado como dado desde el comienzo. Las proposiciones no pueden crearse ni destruirse: constituyen, de una vez por todas, el marco contra el cual la idea de creencia adquiere su significado.

Como se ve, el centro del problema lo constituye el contexto *proposicional* de teorías como AGM. Por su propia naturaleza, el medio proposicional no permite la posibilidad de modelar cambios en el propio marco. Y sin embargo, eso es lo que necesitaríamos hacer si, por poner un ejemplo un tanto burdo, en el caso anterior, *X* comenzara a estudiar astronomía. En lo que sigue propondré un proyecto para una generalización de la teoría AGM que de cuenta de estas situaciones.

## 2. Una propuesta en el marco de la teoría de modelos

Consideraré aquí que el estado epistémico total de un agente se representa por medio de un conjunto de oraciones interpretadas, tal que las creencias del agente son las oraciones que resultan verdaderas en la interpretación. Para esto, como es usual, supondremos un dominio no vacío de objetos, así como ciertas funciones de interpretación que mapeen constantes del lenguaje con objetos en el dominio, y predicados  $n$ -ádicos del lenguaje con conjuntos de  $n$ -tuplas ordenadas de objetos.

Así pues, consideraremos que el estado epistémico de un agente está representado por el par  $\langle L, I \rangle$ , donde  $L$  es un lenguaje de primer orden e  $I = \langle D, S \rangle$  es una interpretación para  $L$ , con un dominio no vacío  $D$  y un conjunto  $S$  de funciones de interpretación. La idea es que cada par  $\langle L, I \rangle$  determina exactamente el estado epistémico de un agente, en el sentido de que cada par fija lo que el agente cree, lo que descree (cree que es falso) y aquello sobre lo que suspende el juicio. Tales creencias e incertidumbres serán, respectivamente, las oraciones de  $L$  que resultan verdaderas, falsas, y sin interpretación en  $I$ .

Para obtener el efecto deseado,  $S$  no puede contener las funciones estándar de interpretación de la lógica clásica. La razón es que no podemos representar suspensiones de juicio a menos que permitamos huecos de valores de verdad en nuestras oraciones interpretadas. Tendremos que basarnos en asignaciones parciales, de manera que algunas oraciones de  $L$  resulten no ser ni verdaderas ni falsas en  $I$ . Para ello recurriré a lo que se conoce como lógica parcial. Una manera estándar de tratar las constantes en lógica parcial es permitiendo que un "no-objeto" (digamos, " $\otimes$ "), se agregue al dominio, de modo que la función denotación para nombres pueda asignar " $\otimes$ " a algunas constantes.<sup>2</sup> Por otra parte, trataré la interpretación de predicados de la siguiente manera. Supondré que en  $S$  encontramos (i) una función  $I_{p+}$  que ranquea sobre predicados de  $L$ , y asigna a cada predicado  $n$ -ádico un conjunto de  $n$ -tuplas ordenadas de objetos, tal que cuando aplicamos el predicado a cualquier  $n$ -tupla del conjunto producimos una oración verdadera. Nos referiremos a este conjunto como la extensión positiva del predicado. Encontramos, además, (ii) una función  $I_{p-}$  que relaciona, nuevamente, predicados con conjuntos de  $n$ -tuplas ordenadas, pero tal que la aplicación del predicado a sus correspondientes  $n$ -tuplas resulta en oraciones falsas. Este conjunto será la extensión negativa del predicado.

Adviértase que, cada vez que las extensiones negativa y positiva de un predicado dado efectúen una partición en el dominio, el predicado estará completamente definido. Pero no

es necesario que ello ocurra. Si ambas extensiones son mutuamente exclusivas pero no exhaustivas, habrá oraciones que no son ni verdaderas ni falsas. (Como dijimos, ellas representan las incertidumbres del agente). Esta será asimismo la situación si el predicado se aplica al “no objeto”; esto es, si es usado para predicar algo de un nombre sin denotación. Además, si hay al menos un predicado de  $L$  para el cual la intersección de las extensiones positiva y negativa no es nula, obtenemos el estado inconsistente de creencias.

Antes de continuar quiero hacer brevemente tres comentarios metodológicos:

(1) Notemos, primero, que aquí  $\langle L, I \rangle$  no determina únicamente las creencias de cierto agente, sino su estado epistémico total; esto es,  $\langle L, I \rangle$  no solamente incluye sus creencias, sino también sus incertidumbres o “suspensiones de juicio”. Las creencias positivas del agente no son sino un subconjunto  $L'$  de  $L$ , tal que la interpretación dada  $I$  constituye un modelo para  $L'$  (entendiendo aquí “modelo” con el sentido de “interpretación verdadera”).

(2) En segundo lugar, no es necesario concebir las oraciones en cuestión como el contenido de las creencias (o de las “indecisiones”) del agente; en realidad, hay buenas razones para rechazar esta posición.<sup>3</sup> En lugar de ello, puede entenderse que dichas oraciones representan directamente las actitudes epistémicas del agente. Ellas son creencias o incertidumbres – no su contenido.

(3) Por último, quiero destacar que el par  $\langle L, I \rangle$  debe entenderse como una representación de estados epistémicos. Bajo ningún punto de vista intento decir, por ejemplo, que el lenguaje  $L$  sea el vehículo sintáctico real usado por los agentes para comunicarse; ni tampoco intento decir que las creencias o suspensiones de juicio sean realmente el resultado de interpretar una determinada base sintáctica mediante un dominio de objetos y un conjunto de funciones de interpretación como el que estoy proponiendo.

### 3. AGM en el marco de la nueva propuesta

Es fácil ver que en este marco, revisiones, contracciones y expansiones en el estilo de AGM se pueden interpretar como peculiares modificaciones de  $S$  (esto es, cambios en las funciones de interpretación) para un  $L$  y  $D$  fijos. Así, cambios de creencia ordinarios (estilo AGM) consisten en el pasaje de un  $\langle L, \langle D, S \rangle \rangle$ , a un nuevo par  $\langle L, \langle D, S' \rangle \rangle$ . Por ejemplo, consideremos el caso en el cual, en términos de AGM, ni  $A$  ni  $\neg A$  representan proposiciones creídas por el sujeto. Traducir esta situación equivale a asegurar que ni  $A$  ni  $\neg A$ , de un  $L$  dado, resultan verdaderas en  $\langle D, S \rangle$ , lo cual significa que necesitamos que las funciones de interpretación en  $S$  sean tales que no asignen a  $A$  ningún valor de verdad. Ahora, imaginemos que, nuevamente en terminología de AGM, el sujeto expande de un conjunto de creencias  $K$  a  $K^+_A$ . En nuestro esquema, esto es equivalente a decir que  $S$  fue modificado de manera tal que  $A$  ahora resulta verdadero en  $\langle D, S^+_A \rangle$  (donde “ $S^+_A$ ” es el nuevo conjunto de funciones de interpretación). Un razonamiento similar se aplica a las revisiones y contracciones. Puede mostrarse que todos los axiomas de AGM pueden traducirse sin demasiado trabajo a nuestra propuesta. Aquí no puedo extenderme demasiado sobre los detalles técnicos, pero sí quisiera ilustrar brevemente el procedimiento con los postulados de expansión:

*Postulados AGM para la expansión de un conjunto de creencias:*<sup>4</sup>

(<sup>+</sup>1)  $K^+_A$  es un conjunto de creencias.

(<sup>+</sup>2)  $A \in K^+_A$

(\*3)  $K \subseteq K^+_A$

(\*4) Si  $A \in K$ , entonces  $K^+_A = K$

(\*5) Si  $K \subseteq H$ , entonces  $K^+_A \subseteq H^+_A$

(\*6) Para todo  $K$  y todo  $A$ ,  $K^+_A$  es el conjunto más pequeño que satisface (\*1) – (\*5)

### Traducción:

Representaremos el estado epistémico  $\langle L, \langle D, S \rangle \rangle$  expandido con  $A_L$  (donde  $A_L$  es una fórmula de  $L$ ) como:  $\langle L, \langle D, S^+_A \rangle \rangle$ . Sean  $[I_{p^+}]^+_A$ ,  $[I_{p^-}]^+_A$  y  $[I_C]^+_A$  las nuevas funciones en  $S^+_A$ . Entonces:

(+1)  $\langle L, \langle D, S^+_A \rangle \rangle$  representa un estado epistémico.

(Notemos que esto determina un cierto conjunto de creencias positivas.  $K^+_A$ , en el sentido de AGM, no es sino un subconjunto de  $L$ , tal que, como ya dijimos, la interpretación dada  $\langle D, S^+_A \rangle$  constituye un modelo para  $K^+_A$ ).

(+2)  $A$  es verdadero en  $\langle D, S^+_A \rangle$

(+3) Para todo predicado  $F \in L$ , y dado el dominio  $D$ ,

$$I_{p^+}(F) \subseteq [I_{p^+}(F)]^+_A, \text{ y } I_{p^-}(F) \subseteq [I_{p^-}(F)]^+_A;$$

y para todo nombre  $c \in L$ , si  $I_C(c) = a$  y  $a \neq \emptyset$ , entonces  $[I_C(c)]^+_A = a$ ,

(esto es, los nombres que tienen una denotación "real" en  $S$ , mantienen la misma denotación en  $S^+_A$ ).

(+4) Si  $A$  es verdadero en  $\langle D, S \rangle$ , entonces  $S = S^+_A$

(+5) Para todo  $F \in L$ , y para un  $D$  dado, si  $I_{p^+}(F) \subseteq J_{p^+}(F)$ , y  $I_{p^-}(F) \subseteq J_{p^-}(F)$ ,

$$\text{entonces } [I_{p^+}(F)]^+_A \subseteq [J_{p^+}(F)]^+_A, \text{ y } [I_{p^-}(F)]^+_A \subseteq [J_{p^-}(F)]^+_A;$$

y para todo nombre  $c$ , si es cierto que  $[I_C(c) = a \text{ y } a \neq \emptyset]$ , entonces  $J_C(c) = a$  entonces, si  $[I_C(c)]^+_A = a$  y  $a \neq \emptyset$ , entonces  $[J_C(c)]^+_A = a$ .

(+6) Para todo  $\langle L, \langle D, S \rangle \rangle$  y toda  $A_L$ , el conjunto  $S^+_A$  en  $\langle L, \langle D, S^+_A \rangle \rangle$  sólo puede diferir de  $S$  en lo que es estrictamente necesario para satisfacer (+1)-(+5).

(En realidad, debemos notar que todos los postulados de AGM siguen vigentes, pero relativizados a un estado epistémico  $\langle L, \langle D, S \rangle \rangle$ . La idea de "traducirlos" al nuevo análisis es que, de esta manera, más adelante podremos incluir las instrucciones para cambios no estándar en el mismo único conjunto de postulados. Además, cabe mencionar que puede aceptarse en el contexto de la nueva propuesta, si así lo deseáramos, la estrategia típica de AGM para construir una función de contracción – la llamada "partial meet contraction").

## 4. Cambios no estándar

Ahora bien, la parte interesante de este análisis es que permite mayor libertad para efectuar cambios que son imposibles de analizar en teorías estándar. Básicamente, la idea es que ahora podemos estudiar cambios que tienen lugar cuando modificamos también  $L$  y  $D$ , además de  $S$ .  $L$  y  $D$  pueden expandirse o contraerse, y pueden variar simultánea o independientemente uno del otro. Veamos aquí algunos de los muchos ejemplos posibles de cambios no estándar:

(a) Cambios de  $\langle L, \langle D, S \rangle \rangle$  a  $\langle L', \langle D', S' \rangle \rangle$ , donde  $L \subset L'$  y  $D \subset D'$ . Estos constituyen posiblemente los casos más atractivos. Recordemos el ejemplo con el cual comenzamos este trabajo. Es razonable pensar que para expandir el cuerpo de creencias de  $X$  con una nueva afirmación del tipo "hay un agujero negro en el núcleo galáctico activo de la galaxia  $xyz$ ", nuevos predicados y nombres deben ser incorporados al lenguaje, mientras que la ontología de  $X$  debe ampliarse con la introducción de nuevos objetos. Sólo entonces el agente está capacitado para suspender el juicio, creer o descreer la afirmación en cuestión.

(b) Cambios de  $\langle L, \langle D, S \rangle \rangle$  a  $\langle L', \langle D, S' \rangle \rangle$ , donde  $D$  permanece igual, pero  $L \subset L'$ . Estos cambios pueden verse como representaciones de estipulaciones lingüísticas. Por ejemplo, pueden ser apropiados para representar el caso en que bautizamos con un nuevo nombre algo que ya estaba en nuestro dominio, o tal vez para el caso en que acuñamos un predicado unificador para varios objetos que ya eran parte de nuestra ontología.

(c) Cambios de  $\langle L, \langle D, S \rangle \rangle$  a  $\langle L, \langle D', S' \rangle \rangle$ , en los cuales  $D'$  es más chico que el dominio inicial  $D$ . Estos cambios representan casos en los cuales por alguna razón decidimos restringir nuestros compromisos ontológicos. Parecen adecuados para representar situaciones en las cuales un agente concluye que cierto aparente "objeto" no era más que una alucinación, o tal vez que ciertos conceptos (tales como "flogisto" o "éter") no se corresponden con nada en el mundo.

Por supuesto, todos estos cambios no estándar deben ser regulados con precisión. Aquí, por razones de tiempo, voy a dar solamente algunas pautas de cómo estas indicaciones podrían proceder. El caso más simple probablemente sea el de las *expansiones* de  $L$  o  $D$ . Supongamos, por ejemplo, que modificamos nuestro par inicial  $\langle L, \langle D, S \rangle \rangle$  con la adición de una oración  $A$ , formulada en algún lenguaje  $L_A$  y verdadera para alguna interpretación  $\langle D_A, S_A \rangle$ . Supongamos que  $A$  no pertenece al lenguaje  $L$  de nuestro par inicial, y que  $D_A$  (la denotación de las constantes de  $A$ ) no está incluida en  $D$ . El primer paso consiste en expandir  $L$  con todos (y solamente) los nombres y predicados que aparezcan en  $A$  pero no estén en  $L$ . El segundo paso deberá ser expandir el dominio con todos (y solamente) los objetos  $o_i$  que no estén en  $D$  (el dominio del par inicial), y que hayan sido asignados (por una función denotación en  $S_A$ ) a alguna constante  $c$  en  $A$ .

Regimentar las contracciones de  $D$  o  $L$  es un poco más complejo. En primer lugar, debemos reconocer que ningún input posible (esto es, ninguna posible nueva información sobre una oración  $A$  cualquiera) puede proporcionar, por sí mismo, una razón para contraer el lenguaje o el dominio de nuestro estado de creencias original. La razón es clara. Del hecho de que  $A$  contenga ciertos nombres y predicados en lugar de otros, todo lo que podemos decir es que ahora  $L$  deberá contener esos mismos nombres y predicados, si es que no los tenía ya. Más allá de esto, nada se sigue respecto de cómo debe ser el lenguaje en el par final (en el estado de creencias final). En particular, nada se sigue sobre el vocabulario que *no* apareció en  $A$ . Pese a todo, entiendo que es posible regimentar contracciones no estándar. Para ello la estrategia será la de usar la navaja de Occam *luego* de haber llevado a cabo algún otro cambio en el estado de creencias en cuestión. Supongamos, por ejemplo, que luego de una contracción estándar, en el estilo AGM, la extensión positiva y negativa de cierto predicado se identifican con la clase vacía. En ese caso, podemos sentirnos justificados en remover dicho predicado del lenguaje de nuestro par final. En el mismo espíritu, si, luego de efectuar una modificación estándar en cierto estado epistémico, un objeto de

nuestro dominio aparece solamente en las extensiones negativas de cada predicado del lenguaje, bien podemos pensar en eliminarlo como un elemento genuino del dominio final.

## 5. Algunas consecuencias

La teoría que acabo de delinear ha sido concebida como una peculiar generalización de AGM. Es decir que estoy dando los postulados de AGM por sentado – si bien habrá que hacer algunos ajustes menores a la hora de incorporar los nuevos postulados para cambios no estándar. Es claro que nadie negó nunca que AGM trate acerca de cambios de creencias *relativas* a un marco proposicional dado. Lo que aquí intenté fue una estrategia para pensar cómo ese marco puede, él mismo, variar, y de manera *racional*.

Quisiera mencionar ahora algunas consecuencias de este enfoque. Notemos, antes que nada, que da respuestas al problema que planteamos al comienzo: efectivamente, nos permite discriminar casos en los que las teorías tradicionales de cambio de creencias estipularían, engañosamente, que el agente suspende el juicio.

Pero este análisis entraña otras consecuencias filosóficas. En primer lugar, recordemos que en todo este trabajo hemos convenido en que las oraciones interpretadas a las que nos hemos remitido *no* constituyen el *objeto* de las actitudes epistémicas del agente, sino que ellas *son* tales actitudes epistémicas, y, en conjunto, conforman su estado total de creencias e incertidumbres. De esta manera, evitamos tratar con el enojoso problema del “contenido”, y, en caso de que la idea de “proposición” nos resulte teóricamente sospechosa, es bueno saber que ya no necesitamos comprometernos con su existencia.

En segundo lugar, hemos enriquecido el concepto de racionalidad presupuesto. Llevar a cabo un cambio racional en un cuerpo de creencias e incertidumbres es ahora un proceso complejo, porque tales cambios pueden tener lugar de maneras radicalmente diversas. Esto conlleva algunos resultados interesantes. Por ejemplo, se ha argumentado muchas veces que una meta importante de toda investigación racional (aunque por cierto no la única meta) consiste en reducir el número de cuestiones no decididas a un mínimo; en otros términos, deberíamos tratar de expandir nuestro conjunto de creencias hasta llegar a un conjunto máximamente consistente. Si este fuera el caso, teorías en el estilo de AGM nos proveen las restricciones formales que debemos obedecer para cumplir esta meta. Pero nuestro nuevo escenario entraña que ninguna meta de este estilo puede ser inteligiblemente formulada. El ideal de revisar y expandir hasta llegar a un conjunto máximamente consistente de creencias está esencialmente mal concebido. La razón de esto es, por supuesto, que conjuntos máximamente consistentes sólo pueden ser definidos para un lenguaje dado. Pero una vez que permitimos que  $L$  y  $D$  cambien junto con nuestras funciones de interpretación, no hay un límite prefijado que podamos alcanzar. En otras palabras, mientras que la expansión racional para un  $L$  y  $D$  dados tiene un límite, en nuestra propuesta dicho límite se ha transformado en un blanco móvil. En síntesis, nuestra nueva noción de racionalidad impide que formulemos la mera posibilidad de obtener alguna vez la representación de una completa “historia del mundo”, y enfatiza la esencial relatividad respecto de un lenguaje en la que se encuentra cualquier historia parcial que podamos concebir. Espero que estos resultados permitan hacer justicia por lo menos a algunas importantes intuiciones pre-teóricas sobre los conceptos de racionalidad y cambio de creencias.

## Notas

<sup>1</sup> Véase por ejemplo Alchourrón, Gärdenfors y Makinson (1985), o Gärdenfors (1988).

<sup>2</sup> Para referencias sobre lógica parcial puede verse Blamey (1986).

<sup>3</sup> Véase Levi (1991), pp. 21 y ss.

<sup>4</sup> Aquí sigo la presentación de Gärdenfors (1988), cap.3.

## Referencias

- Alchourrón, C., Gärdenfors, P. and Makinson, D. (1985), "On the logic of theory change: Partial meet functions for contraction and revision", *Journal of Symbolic Logic* 50, pp. 510-530.
- Blamey, Stephen (1986), "Partial Logic", in Gabbay, D., and Guenther, F. (eds.), *Handbook of Philosophical Logic*, vol. III, D. Reidel Publishing Company.
- Gärdenfors, Peter (1988), *Knowledge in Flux -- Modeling the Dynamics of Epistemic States*, Cambridge, Mass, The MIT Press.
- Levi, Isaac, (1991), *The Fixation of Belief and Its Undoing -- Changing Beliefs Through Inquiry*, Cambridge, Cambridge University Press.