



Inteligencia Artificial. Revista Iberoamericana
de Inteligencia Artificial

ISSN: 1137-3601

revista@aepia.org

Asociación Española para la Inteligencia
Artificial
España

Luna, Carlos Daniel

Una generalización del modelo AGM de cambio de creencias

Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial, vol. 5, núm. 13, 2001, pp. 23-32

Asociación Española para la Inteligencia Artificial

Valencia, España

Disponible en: <http://www.redalyc.org/articulo.oa?id=92521303>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

Una Generalización del Modelo AGM de Cambio de Creencias

Carlos Daniel Luna

Instituto de Computación, Universidad de la República. Pedeciba Informática.
Casilla de Correo 16120, distrito 6, Montevideo, Uruguay.
cluna@fing.edu.uy

Resumen

La *teoría de cambio de creencias* busca caracterizar la dinámica del conocimiento, esto es, cómo debería ser la actitud epistémica de un agente racional frente a ciertas creencias, ante la llegada de nueva información. Este trabajo presenta un modelo de revisión de creencias alternativo al estándar AGM (Alchourrón, Gärdenfors y Makinson). El modelo propuesto, a diferencia de AGM, impide caracterizar un estado epistémico inconsistente y condiciona la aceptación de nueva información. Consecuentemente, resulta naturalmente adecuado, ya que la aceptación de creencias queda parametrizada y de esta manera el nuevo modelo permite un esquema cualitativo o cuantitativo de justificaciones para las creencias. El artículo incluye la formalización del *modelo paramétrico de revisiones consistentes* MPRC, el análisis de sus propiedades más relevantes, su relación formal con las operaciones de cambio AGM y las diferencias con el método de *semi-revisión*, concebido por Hansson.

Palabras clave: Lógica; Inteligencia Artificial; Teoría de Cambio de Creencias: modelo AGM, Semi-revisión.

1. Introducción

Dentro del campo de la Inteligencia Artificial existen diferentes áreas que tratan de modelar razonamiento humano. Para ello es necesario determinar un lenguaje en el cual se representará el conocimiento de un agente y definir reglas, que en base a tal conocimiento, puedan derivar nuevas conclusiones. La mayoría de estos sistemas parten de un lenguaje de primer orden (o un subconjunto del mismo) y realizan una extensión de forma tal que puedan tratar con información potencialmente inconsistente. El objetivo de estos sistemas es poder obtener nuevas conclusiones a partir de una base de conocimiento fija, imperturbable en el tiempo. Esto significa que tales sistemas se basan en la hipótesis de que el conocimiento básico de un agente no puede ser modificado. Sin embargo, existen otras áreas cuyo objetivo es caracterizar la dinámica del

conocimiento [1, 2, 3, 4, 5], esto es, cómo un agente racional actualiza su estado epistémico al recibir nueva información. Una de estas áreas se conoce como *teoría de cambio de creencias*, la cual mediante ciertas operaciones de cambio permite caracterizar la actitud epistémica de un agente racional después de incorporar nuevas creencias, o de rechazar parcial o totalmente el conocimiento disponible por él, en base a información externa recibida. La teoría de cambio de creencias irrumpe en la Lógica Filosófica y la Inteligencia Artificial en las dos últimas décadas. El paso inicial fue provisto por Levi [6] y Alchourrón, Gärdenfors y Makinson en [1] (comúnmente llamado el modelo AGM). Posteriormente, numerosos trabajos se han desarrollado en esta área y en torno a este modelo, que es, sin lugar a duda, el más trascendente de la teoría.

En el modelo AGM, la operación de revisión de creencias incorpora siempre la información que es propuesta. Es decir, privilegia a la nueva información con respecto a las creencias ya existentes en el conocimiento del agente. Esta suposición parece poco natural y puede causar resultados contrarios a la intuición, en ciertos casos. Además, obliga a considerar a una teoría inconsistente como un posible estado epistémico de un agente racional, siendo que en el modelo AGM una teoría inconsistente representa a todo el lenguaje y por ende carece de valor epistémico alguno.

Este trabajo tiene por objetivo analizar parcialmente la teoría AGM y presentar un modelo alternativo, en base a AGM, que impida caracterizar un estado inconsistente y condicione la aceptación de nueva información, siguiendo un enfoque justificacionista. La organización del trabajo es como sigue. Introducimos AGM en la sección 2 y en la 3 destacamos algunas falencias (limitaciones) del modelo, en base a las cuales introducimos el nuevo modelo paramétrico de revisiones consistentes MPRC, en la sección 4. Esta parte abarca la formalización axiomática de las operaciones, que describen el comportamiento deseado de las mismas, y el estudio de sus propiedades más relevantes. La sección 5 discute un importante trabajo relacionado, el método de *semi-revisión* [7]. Finalmente, exhibimos las conclusiones y las líneas futuras de investigación, en la última sección.

2. Texto General del Artículo

Existe una serie de problemas metodológicos y filosóficos involucrados en el proceso de cambio de creencias. Desde el punto de vista filosófico se presentan, entre otras, las siguientes cuestiones: ¿qué es hacer un cambio racional?, ¿cómo se pueden describir los estados de creencia de un agente? Preguntas de este estilo condujeron a Carlos Alchourrón, Peter Gärdenfors y David Makinson a adoptar un conjunto de postulados (modelo AGM) que, en su opinión, son los que corresponden al cambio de creencias de un agente racional [2, 3, 8, 9, 10]. Existen representaciones de estos postulados en términos de funciones de expansión, contracción y revisión de teorías, así como también en términos de modelos [11, 12].

2.1. Conjuntos de Creencias

Definición 1 (conjunto de creencias). Un conjunto de sentencias K es un conjunto de creencias,

abreviado CC, (una teoría lógica) ssi: Si $K \vdash \alpha$ entonces $\alpha \in K$, es decir, $K = \text{Cn}(K) \subseteq L$.¹

Cn es un operador de consecuencia lógica sobre un lenguaje L , que satisface: $A \subseteq \text{Cn}(A)$; Si $A \subseteq B$, entonces $\text{Cn}(A) \subseteq \text{Cn}(B)$; y, $\text{Cn}(A) = \text{Cn}(\text{Cn}(A))$ –para todo par de conjuntos de proposiciones A y B . Además, asumimos que Cn cumple las propiedades de compacidad, supraclasicidad y deducción [2]. En lo que sigue K denotará un CC. $K \vdash \alpha$ es una alternativa notación para $\alpha \in K$. K_\perp representa la teoría inconsistente², formada por todas las sentencias del lenguaje L y, $\text{Cn}(\emptyset)$ es el CC más pequeño, compuesto únicamente de las sentencias lógicamente válidas de L .

2.2. AGM: Una Teoría Epistemológica basada en Conjuntos de Creencias

Una Teoría Epistemológica provee un marco conceptual para modelar el dinamismo de las creencias y el conocimiento de un agente. A fin de lograr esto, una teoría provee también el modo para representar los elementos epistémicos y un conjunto de criterios de racionalidad para comprender los cambios epistémicos [2]. Los conjuntos de creencias son conjuntos de sentencias que pueden ser soportadas racionalmente por un individuo. El modelo AGM define una teoría epistemológica en base a conjuntos de creencias, como sigue.

Estados epistémicos (de creencia). Representados por *conjuntos de creencias* respecto de un lenguaje L . Llamaremos $\Sigma_K = \{K \mid K \text{ es un CC basado en } L\}$ al conjunto de las teorías lógicas para L .

Actitudes epistémicas. Para cada sentencia α del lenguaje, en cada estado epistémico K , puede ser: *aceptada* (si el agente cree en α , i.e. $\alpha \in K$), *rechazada* (cuando el agente acepta $\neg\alpha$, i.e. $\neg\alpha \in K$) o *indeterminada* (cuando un agente no tiene información acerca de α , es decir, ni acepta α ni la rechaza en su conocimiento K).

Cambios de creencia. El modelo provee tres operaciones de cambio: la *expansión* (+), la *contracción* (−) y la *revisión* (*) de teorías. Éstas se modelan mediante relaciones funcionales que toman un CC (estado epistémico) de Σ_K , una sentencia (creencia) de L y devuelven un nuevo CC (nuevo estado epistémico) de Σ_K .

¹ \vdash es una relación metalingüística. $\vdash : 2^L \rightarrow 2^L$ (caso infinito) ó $\vdash : 2^L \rightarrow L$ (caso finito).

² Una teoría lógica K es consistente si no existe una sentencia α del lenguaje tal que α y $\neg\alpha$ sean elementos de K .

Entradas epistémicas. Son aquellas dictadas por la experiencia y se consideran “fuerzas o razones externas”.

Criterios de racionalidad. A los inherentes a conjuntos de creencias se suma el criterio de “economía de la información”. El mismo consiste en retener la mayor cantidad de información posible cuando se produce un cambio de creencias. Un criterio implícitamente establecido por el modelo AGM indica que toda nueva información que provoca un cambio de creencias debe ser incorporada al estado epistémico del agente. Además, el estado de creencias de un agente, luego de un cambio, debe ser consistente siempre que esto sea factible³. Este último punto asegura que las sentencias aceptadas por un individuo comparten algún sentido de coherencia.

2.3. Operaciones de Cambio: Postulados

Si K es un CC consistente, entonces las expansiones transforman creencias que eran indeterminadas en aceptadas o rechazadas, las contracciones transforman creencias que eran aceptadas o rechazadas en indeterminadas, mientras que las revisiones transforman creencias aceptadas en rechazadas y viceversa. A continuación presentamos los postulados propuestos por el modelo AGM.

2.3.1 Expansión

La expansión de creencias se deriva de “aprender algo”. La expansión de K por ϕ se denota $K^+\phi$.

- (K⁺1) Si K es un CC, $K^+\phi$ es un CC
- (K⁺2) $\phi \in K^+\phi$
- (K⁺3) $K \subseteq K^+\phi$
- (K⁺4) Si $\phi \in K$, entonces $K^+\phi = K$
- (K⁺5) Si $K_1 \subseteq K_2$, entonces $K_1^+\phi \subseteq K_2^+\phi$
- (K⁺6) Para todo K y ϕ , $K^+\phi$ es el menor CC que satisface (K⁺1)-(K⁺5)

Existe una definición para el proceso de expansión en términos puramente lógicos y de la teoría de conjuntos.

Teorema 1. Una función de expansión satisface (K⁺1)-(K⁺6) sssi $K^+\phi = Cn(K \cup \{\phi\})$.

³ Al usar teorías lógicas y al asumir un operador de consecuencia supraclásico, implícitamente establecemos que las tautologías forman parte de todo estado epistémico. Luego, debido a la determinación de incorporar toda nueva información que causa una entrada epistémica, podemos llegar a obtener el estado inconsistente K_\perp .

2.3.2 Revisión

El siguiente tipo de cambio de creencias ocurre cuando una sentencia ϕ , que representa una entrada epistémica para el conjunto K , contradice las creencias ya existentes en K (si bien, en la práctica puede que ϕ no sea inconsistente con K , la motivación intuitiva es para el caso mencionado). Este tipo de cambio es “no monótono”; nuevas creencias se agregan, pero no todas las anteriores se conservan. Los postulados que se verán a continuación, si bien no determinan de manera única una función de revisión, enmarcan las posibles soluciones a los criterios de racionalidad de AGM. $K^*\phi$ denota la revisión de K por ϕ .

- (K^{*}1) Si K es un CC, $K^*\phi$ es un CC.
- (K^{*}2) $\phi \in K^*\phi$
- (K^{*}3) $K^*\phi \subseteq K^+\phi$
- (K^{*}4) Si $\neg\phi \notin K$, entonces $K^*\phi = K^+\phi$
- (K^{*}5) $K^*\phi = K_\perp$ sssi $\neg\phi \in Cn(\emptyset)$
- (K^{*}6) Si $\phi \leftrightarrow \psi \in Cn(\emptyset)$, entonces $K^*\phi = K^*\psi$
- (K^{*}7) $K^*(\phi \wedge \psi) \subseteq (K^*\phi)^+\psi$
- (K^{*}8) Si $\neg\psi \notin K^*\phi$, entonces $(K^*\phi)^+\psi \subseteq K^*(\phi \wedge \psi)$

(K^{*}1)-(K^{*}6) constituyen el conjunto de postulados básicos para la revisión, ya que no se refieren a revisiones con sentencias compuestas. (K^{*}7) y (K^{*}8) contemplan revisiones de sentencias compuestas (conjunciones).

2.3.3 Contracción

Una contracción ocurre cuando una sentencia es refutada, pero ninguna sentencia es sumada. La contracción de K por una sentencia ϕ se denota $K^-\phi$. La contracción, al igual que la revisión, no está unívocamente definida en términos puramente lógicos y de la teoría de conjuntos, como ocurre en el caso de la expansión. Cuando realizamos la contracción de una teoría K por una creencia, algunas sentencias de K deben ser eliminadas a fin de que el sistema resultante, cerrado bajo consecuencia lógica, refleje el efecto de la actualización del conocimiento. Los siguientes postulados circunscriben el conjunto de soluciones para el caso de la contracción de teorías, siguiendo los criterios de racionalidad de AGM.

- (K⁻1) Si K es un CC, $K^-\phi$ es un CC
- (K⁻2) $K^-\phi \subseteq K$
- (K⁻3) Si $\phi \notin K$, entonces $K^-\phi = K$
- (K⁻4) Si $\phi \notin Cn(\emptyset)$, entonces $\phi \notin K^-\phi$
- (K⁻5) $K \subseteq (K^-\phi)^+\phi$
- (K⁻6) Si $\phi \leftrightarrow \psi \in Cn(\emptyset)$, entonces $K^-\phi = K^-\psi$

(K⁻7) $K^-\phi \cap K^-\psi \subseteq K^-(\phi \wedge \psi)$

(K⁻8) Si $\phi \notin K^-(\phi \wedge \psi)$, entonces $K^-(\phi \wedge \psi) \subseteq K^-\phi$

(K⁻1)-(K⁻6) son los postulados básicos para la contracción. (K⁻7) y (K⁻8) refieren a contracciones de conjunciones.

2.3.4 Relación entre la Contracción y la Revisión de Teorías

En las secciones previas la contracción y la revisión están dadas por dos conjuntos de postulados independientes, en el sentido de que los postulados de contracción no hacen referencia a los de revisión y viceversa. Sin embargo, es posible definir una operación en términos de la otra, tal cual lo ilustran las siguientes fórmulas de Levi [13] y Harper [14]: (Levi. Def *) $K^*\phi = (K^-\neg\phi)^+\phi$; (Harper. Def -) $K^-\phi = K \cap (K^*\neg\phi)$. Estas fórmulas se ven reforzadas por el siguiente teorema [3].

Teorema 2. Si la función de contracción ‘-’ satisface (K⁻1)-(K⁻8) y la expansión cumple (K⁺1)-(K⁺6), entonces la función ‘*’ obtenida por (Levi. Def *) verifica (K^{*}1)-(K^{*}8). Asimismo, si una función de revisión ‘*’ satisface (K^{*}1)-(K^{*}8) entonces la función ‘-’ obtenida por (Harper. Def-) verifica (K⁻1)-(K⁻8).

3. Observaciones a AGM

AGM es un modelo de *coherencia* y no *justificacionista* [15]. Toda entrada epistémica privilegia a la nueva información sobre el conocimiento disponible. En consecuencia, la consistencia de un estado no está siempre asegurada. Por otro lado, no se consideran grados de creencia, es decir, cada sentencia puede encontrarse sólo en tres estados posibles con respecto a una teoría lógica: *aceptada*, *rechazada* o *indeterminada*. A continuación presentamos un análisis crítico de AGM basándonos, fundamentalmente, en algunos de los aspectos sobresalientes sucintamente descritos en este párrafo y en la sección previa.⁴

3.1. Postulados de Éxito y sus Consecuencias

Una suposición adoptada por el modelo AGM es que una creencia que provoca un cambio en el estado epistémico de un agente se ve privilegiada respecto a las creencias concebidas previamente por él. Esto se evidencia en los postulados (K⁺2), (K⁻4) y (K^{*}2), conocidos como *postulados de éxito* para la

expansión, contracción y revisión, respectivamente. Esta suposición puede causar resultados contrarios a la intuición, en ciertos casos.

Ejemplo 1. Supongamos que un científico que estudia las propiedades de los elementos químicos tiene en su conjunto de creencias a ϕ : “todos los metales se dilatan al calentarse”. Asumamos además, que el científico llegó a tal conclusión luego de numerosos casos de prueba y análisis teórico-prácticos exhaustivos que justifican su creencia. Supongamos ahora que el científico observa mediante un experimento el siguiente hecho ψ : “cierto metal X no se dilata al calentarse”. Luego, ¿debe el científico abandonar su antigua creencia altamente justificada en función de este nuevo hecho, o debe, de alguna manera, contraponer ambos hechos de acuerdo a su “fuerza relativa”? Y en este último caso, ¿no esperaríamos intentar formalizar este análisis dentro de nuestro sistema de representación de conocimiento y de cambio de creencias? Ciertamente, cuando un científico hace una observación que entra en conflicto con su conocimiento previo, él no cambia inmediatamente sus creencias sino que realizará suficientes pruebas antes de abandonar su antigua creencia. Sería más natural (realista) pensar en un modelo en donde las creencias pudiesen incorporar su “importancia” o “justificación” como parte de las mismas, y en este contexto decidir aceptar o no una nueva creencia que contradice creencias previas. §

El postulado de éxito en una revisión AGM supone que la nueva creencia es más importante que cualquier otra del conocimiento disponible. Esto, traducido en nuestro ejemplo, significa que el científico debería realizar numerosas experiencias y análisis antes de revisar su conocimiento para incorporar ψ y rechazar ϕ . Parece natural pensar que cuando una creencia contradice creencias previas más importantes, entonces sea rechazada. Asimismo, en una expansión o contracción podríamos estimar que estas operaciones no tengan efecto si no se tiene una “explicación” adecuada de las sentencias que causan el cambio, aunque no existan problemas de consistencia en estos casos.

En resumen, si hablamos acerca de la importancia o justificación en el lenguaje en el que se representa tanto el conocimiento de un agente como las creencias que desencadenan los cambios, entonces el postulado de éxito para las revisiones deja de ser razonable. Por otro lado, si consideramos al postulado de éxito como razonable, cabe preguntarnos: ¿qué significa hacer una revisión de nuestras creencias por una contradicción?, ¿tiene sentido que una contradicción forme parte de algún estado epistémico de un agente racional?; en este

⁴ Un análisis más completo del modelo AGM puede consultarse en [21].

sentido también el postulado (K^*5) resulta intuitivamente cuestionable cuando afirma: $K^*\perp=K\perp$. Pero, ¿puede una contradicción ser aprendida?, ¿no debería ser $K^*\perp=K$? y, en este último caso violar el postulado de éxito, o equivalentemente: ¿las contradicciones deberían formar parte del lenguaje que representa a las creencias que disparan los cambios en una teoría lógica? Quizás una solución más racional en este caso sea considerar que una revisión es una función parcial, en vez de total. En este punto, podríamos tener lenguajes distintos para representar las creencias de un agente en un estado epistémico y para representar las creencias por las cuales el agente provoca un cambio.

Una alternativa para eliminar el privilegio de la nueva información sobre las creencias que posee un agente sería: primero, agregar la nueva creencia por la cual revisamos nuestro conocimiento al conjunto de creencias y luego restaurar la consistencia eliminando las creencias que sean necesarias de acuerdo al modelo estándar. Esto podría consistir en abandonar parte del viejo conocimiento o descartar la creencia que provoca el cambio, siempre en un plano de igualdad de condiciones entre las viejas y la nueva creencia. Claro que deberíamos asumir una representación finita de un estado epistémico (bases de creencias o bases de conocimiento [4, 16]) y no un conjunto clausurado que represente a todo el lenguaje, en el caso de una inconsistencia. Esta alternativa la expone Hansson en su modelo *semi-revisión* [7], que será analizado en la sección 5. La sección 4 presenta un nuevo modelo para la incorporación de creencias –que difiere de AGM y *semi-revisión*–, donde el éxito de la operación está condicionado y la consistencia asegurada.

3.2. Actitudes Epistémicas y Concepción de las Operaciones de Cambio

El hecho de que una sentencia sólo pueda encontrarse en tres estados posibles con respecto a una teoría lógica, nos obliga a realizar una pobre representación de la realidad. Si pudiésemos hablar de los distintos grados de creencia de las sentencias en el lenguaje, lo cual resulta racionalmente apropiado, ¿debería valer, por ejemplo, P: “si K es consistente y $\phi \in K$ entonces $K^*\phi=K$ ”? Claramente, este hecho, ¿no podría estar aumentando la importancia relativa de ϕ ? y en este caso, ¿no debería reflejarse esto en el nuevo estado epistémico? Cabe aclarar que P se deduce de los postulados en el modelo AGM y que nuestros estados epistémicos estarían constituidos no sólo por fórmulas, sino también por la “importancia epistémica” de las mismas. La importancia o “fuerza relativa” podría ser expresada, por ejemplo,

mediante el número de veces que una sentencia es observada, una probabilidad u otro criterio cuantitativo, o de acuerdo a una “explicación o justificación”, o por medio de la relación de consecuencia lógica, o de algún otro criterio cualitativo. Además, “la importancia” de una creencia debería poder variar, tal como ocurre en la realidad con los hechos que uno cree o, al menos la importancia comparativa entre las creencias debería poder modificarse.

Ejemplo 2. Sea $K=Cn(\{\phi \wedge \psi\})$ y supongamos que realizamos $K^*\neg(\phi \wedge \psi)$. Entonces, si en algún momento creemos que ϕ es más importante que ψ , sería razonable que $K^*\neg(\phi \wedge \psi)=Cn(\{\phi\})$, pero si en otro instante consideramos que ψ es más importante que ϕ , lo razonable sería: $K^*\neg(\phi \wedge \psi)=Cn(\{\psi\})$. §

Luego, parece inapropiado que la operación de revisión sea una función si el lenguaje de representación no es lo suficientemente rico para expresar los detalles significativos de los estados epistémicos. AGM soluciona este problema considerando que la fuerza relativa de las creencias está implícita en la elección de la función de revisión ‘*’. Lo que no queda claro es si luego de un cambio en las creencias ‘*’ debería ser la misma función, ya que al aprender una nueva fórmula sería razonable pensar que la importancia epistémica de las creencias pudiera modificarse.

Ejemplo 3. Sea K un estado epistémico que contiene la creencia ϕ de un científico, ϕ : “Los metales tienen la propiedad P” y luego aprende, mediante reiteradas observaciones, que X cantidad de metales tienen esa propiedad. ¿No sería natural que ϕ aumentara su importancia epistémica?. §

Este ejemplo nos muestra que tres actitudes epistémicas son insuficientes para reflejar nuestro conocimiento y el grado de creencia sobre ciertos hechos. Una concepción más realista debería contemplar grados de creencia sobre el conocimiento. Por lo tanto, una concepción más apropiada de la operación de revisión –y de las operaciones de cambio en general– sería una función de un estado epistémico por un fórmula en un nuevo estado epistémico, pero donde los estados no fueran sólo conjuntos de creencias, sino que la importancia epistémica (fuerza relativa) de las mismas estuviera incluida en la representación. Esto nos permitiría modelar la noción de cambio de creencias con mayor detalle, ya que podría pensarse como una generalización del modelo AGM.

4. MPRC: Un Modelo Paramétrico de Revisiones Consistentes

La idea de aceptar a K_{\perp} como una teoría lógica que eventualmente represente el estado epistémico de un agente racional, luego de un cambio en sus creencias, parece naturalmente inadecuada –tal como se analizó previamente en 3.1. El origen de tal problema surge, obviamente, al incorporar creencias al estado epistémico de un agente, ya que las contracciones no generan inconsistencias sobre teorías consistentes. En una expansión, la incorporación imprudente de creencias causa que si una sentencia $\neg\phi$ está en un estado epistémico K , $K^+\phi$ constituya el estado de creencias inconsistente del agente, representado por $K_{\perp}=\perp$. Luego, resulta inadecuado aceptar a la expansión como una operación *coherente* de incorporación de nuevas creencias. Por otro lado, en las revisiones, por ser Cn un operador de consecuencia supraclásico (toda teoría lógica incluye a las tautologías), y a causa de la hipótesis subyacente en el modelo AGM por la cual toda nueva creencia debe ser aceptada, estamos obligados a considerar a K_{\perp} como un posible estado epistémico de un agente racional (ver postulados K^*2 y K^*5). En particular, K^*5 ($K^*\phi=K_{\perp}$ sssi $\neg\phi \in Cn(\emptyset)$) es un postulado demasiado fuerte; no sólo asegura que la consistencia de la información está asegurada, siempre que esto sea lógicamente posible, sino también que ella es restaurada cuando la nueva creencia no es inconsistente. Esto determina que si $\neg\phi \notin Cn(\emptyset)$ entonces $(K_{\perp})^*\phi \neq K_{\perp}$!!.

4.1. Una Nueva Teoría Epistemológica

En función del análisis previo presentaremos un modelo de incorporación de nuevas creencias –en base a AGM– que satisfaga los criterios de racionalidad enunciados en la sección 2.2, restringidos a justificaciones y a la condición: “toda teoría lógica debe ser consistente”. Proponemos utilizar sólo una operación de aceptación de nuevas creencias: la *revisión*, y en relación a AGM, las nuevas creencias no serán privilegiadas respecto a las existentes en la teoría.

Utilizaremos en la formalización una función *justifica*, que dada una teoría lógica K y una creencia ϕ , retorna uno de tres resultados posibles: $\{\phi\}$, si el agente justifica ϕ en relación a su estado epistémico K ; $\{\neg\phi\}$, si el agente en K justifica $\neg\phi$; y finalmente, \emptyset si el agente no puede justificar ϕ ni $\neg\phi$ en relación a su estado de creencias K . El modelo MPRC se define entonces como una teoría epistemológica de la siguiente manera:

Estados epistémicos. Representados por *conjuntos de creencias consistentes* respecto de un lenguaje L . Llamaremos $\Sigma_c=\{K \mid K=Cn(K) \subseteq L, \text{ con } \perp \notin K\}$ al conjunto de todos los estados de creencias consistentes para L .

Actitudes epistémicas. Para cada sentencia ϕ del lenguaje y cada conjunto de creencias consistente K , existen 7 actitudes epistémicas diferentes, a saber:

- AE1: $\text{justifica}(\phi, K)=\{\phi\}$, si $\phi \in K$ y $\neg\phi \notin K$
- AE2: $\text{justifica}(\phi, K)=\{\phi\}$, si $\phi \notin K$ y $\neg\phi \notin K$
- AE3: $\text{justifica}(\phi, K)=\{\neg\phi\}$, si $\phi \notin K$ y $\neg\phi \in K$
- AE4: $\text{justifica}(\phi, K)=\{\neg\phi\}$, si $\phi \notin K$ y $\neg\phi \notin K$
- AE5: $\text{justifica}(\phi, K)=\emptyset$, si $\phi \in K$ y $\neg\phi \in K$
- AE6: $\text{justifica}(\phi, K)=\emptyset$, si $\phi \in K$ y $\neg\phi \notin K$
- AE7: $\text{justifica}(\phi, K)=\emptyset$, si $\phi \notin K$ y $\neg\phi \in K$

Las siguientes no constituyen actitudes epistémicas válidas del modelo, ya que se excluyen sentencias cuyas negaciones estén justificadas:

$\text{justifica}(\phi, K)=\{\phi\}$ y $\neg\phi \in K$;

$\text{justifica}(\phi, K)=\{\neg\phi\}$ y $\phi \in K$.

Cambios de creencia. El modelo provee solamente dos operaciones de cambio: la *revisión* (\times) y la *contracción* (\sim) de teorías consistentes. Éstas se modelan mediante relaciones funcionales que toman un CC (estado epistémico) de Σ_c , una sentencia (creencia) de L y devuelven un nuevo CC (nuevo estado epistémico) de Σ_c .

Entradas epistémicas. Definidas igual que en el modelo AGM.

Criterios de racionalidad. A los inherentes al modelo AGM se antepone la aceptación condicionada por justificaciones al privilegio de la nueva información. Asimismo, se suman dos criterios: “toda teoría lógica debe ser consistente” y “ninguna creencia en un estado epistémico puede tener como justificación a su negación”.

4.2. Revisión en Σ_c

La axiomatización propuesta para una función de revisión (\times) de Σ_c por L en Σ_c , es una generalización de la propuesta AGM, que consta de 9 axiomas. El estado epistémico de un agente debe ser consistente, es decir,

(Rev.1) $K^{\times}\phi \in \Sigma_c$

El segundo postulado expresa que una sentencia ϕ será incorporada en el estado de creencias K de un

agente si él justifica ϕ en K , o al menos no justifica $\neg\phi$.

(Rev.2) Si $\text{justifica}(\phi, K) \neq \{\neg\phi\}$, $\phi \in K^\times\phi$

Sólo aquellas creencias cuyas negaciones no estén justificadas pueden pertenecer a un estado epistémico. Esto es lo que afirma el siguiente postulado,

(Rev.3) Si $\phi \in K$, $\text{justifica}(\phi, K) \neq \{\neg\phi\}$

El axioma anterior supone un proceso de "descreencia" luego de una actualización y podría formularse como una restricción de la función *justifica*, ya que no refiere directamente a la operación de revisión. Al igual que en el modelo AGM, la revisión es un subcaso de la expansión clásica de creencias, esto es,

(Rev.4) Si $\text{justifica}(\phi, K) \neq \{\neg\phi\}$, $K^\times\phi \subseteq Cn(K \cup \{\phi\})$

El quinto axioma establece cuándo la revisión de creencias coincide con la expansión clásica: "si el agente aprueba la sentencia o en caso contrario, si no la reprueba y su negación no es parte de su estado epistémico", es decir,

(Rev.5) Si $[\text{justifica}(\phi, K) = \emptyset \text{ y } \neg\phi \notin K]$ ó $\text{justifica}(\phi, K) = \{\phi\}$, $Cn(K \cup \{\phi\}) \subseteq K^\times\phi$

El siguiente postulado especifica que el estado de creencias de un agente permanece inalterado al intentar incorporar una creencia cuya negación está justificada.

(Rev.6) Si $\text{justifica}(\phi, K) = \{\neg\phi\}$, $K^\times\phi = K$

Similarmente a la axiomatización AGM, la revisión debe ser independiente de la sintaxis de las sentencias a incorporar. Formalmente,

(Rev.7) Si $Cn(\phi) = Cn(\psi)$, $K^\times\phi = K^\times\psi$

Los dos postulados subsiguientes extienden la axiomatización previa a conjunciones de sentencias, de manera similar a la propuesta en el modelo AGM para la operación (*).

(Rev.8) $K^\times(\phi \wedge \psi) \subseteq Cn(K^\times\phi \cup \{\psi\})$

(Rev.9) Si $[\text{justifica}(\psi, K^\times\phi) = \emptyset \text{ y } \neg\psi \notin K^\times\phi]$ ó $\text{justifica}(\psi, K^\times\phi) = \{\psi\}$, $Cn(K^\times\phi \cup \{\psi\}) \subseteq K^\times(\phi \wedge \psi)$

Restricciones Adicionales. La función *justifica* debe verificar las siguientes restricciones.

(J.1) Para toda creencia ϕ y toda teoría consistente K , $\text{justifica}(\phi, K) \in \{\{\phi\}, \{\neg\phi\}, \emptyset\}$.

(J.2) Si $\phi \in Cn(\emptyset)$, para todo conjunto de creencias consistente K , $\text{justifica}(\phi, K) = \{\phi\}$. Esto es, las

tautologías forman parte de toda teoría consistente y son justificadas por cualquier agente.

(J.3) Para toda creencia ϕ y toda teoría consistente K , $\text{justifica}(\phi, K) = \text{justifica}(\neg\phi, K)$. Es decir, en relación a cada creencia y su negación, el agente en un estado epistémico ó bien justifica la creencia, ó su negación, ó no posee una justificación definida.

(J.4) Si $Cn(\phi) = Cn(\psi)$,
 $\text{justifica}(\phi, K) = \{\phi\}$ y $\text{justifica}(\psi, K) = \{\psi\}$, ó
 $\text{justifica}(\phi, K) = \{\neg\phi\}$ y $\text{justifica}(\psi, K) = \{\neg\psi\}$, ó
 $\text{justifica}(\phi, K) = \emptyset$ y $\text{justifica}(\psi, K) = \emptyset$

Es decir, sentencias lógicamente equivalentes poseen justificaciones "equivalentes".

(J.5) El operador de consecuencia debe extender conservativamente el criterio de justificación para sentencias derivadas de creencias cuyas negaciones no están justificadas. Sea X es un conjunto de sentencias, no necesariamente clausurado, debe verificarse que: Si $\forall \phi \in X$. $\text{justifica}(\phi, X) \neq \{\neg\phi\}$, entonces $\forall \phi \in Cn(X)$. $\text{justifica}(\phi, X) \neq \{\neg\phi\}$

(J.6) Las implícitas en los axiomas Rev.1 y Rev.3.

4.3. Propiedades de MPRC

A continuación presentamos algunas de las propiedades relevantes del modelo MPRC⁵. Implícitamente asumimos a todas las variables, de fórmulas y conjuntos, cuantificadas universalmente.

(P.1) El estado de creencias de un agente racional será representado siempre por un conjunto de creencias consistente. Esto es, $\perp \notin K$ (para cualquier K) y obviamente, $\text{justifica}(K, \perp) = \{\neg\perp\} = \{T\}$, es decir, $\perp \notin K^\times\perp$.

(P.2) La aceptación de creencias es condicionada –queda parametrizada– a través de la función *justifica*, que podría adaptarse a un esquema de justificaciones para las creencias (ver, por ejemplo, [17] y [18]). Ninguna creencia perteneciente al estado epistémico de un agente puede poseer una justificación para su negación. Sin embargo, no toda creencia justificada o de justificación indefinida es parte de su estado epistémico hasta que no decida incorporarla.

(P.3) El postulado de éxito para la revisión ($\phi \in K^\times\phi$) no se satisface. Si $\text{justifica}(\phi, K) = \{\neg\phi\}$, $K^\times\phi = K$. Esto ocurre cuando $\neg\phi$ es una tautología o simplemente porque el agente justifica $\neg\phi$ en función de su estado epistémico.

⁵ Las demostraciones completas son omitidas por razones de espacio.

(P.4) Es importante destacar que la función *justifica* puede variar para cada estado epistémico K y por cada creencia ϕ . Así, aún cuando $\phi \in K^\times \phi$, no necesariamente sucede que *justifica*(ϕ, K) coincide con *justifica*($\phi, K^\times \phi$). Un agente cauto, o poco crédulo, podría considerar que aunque *justifica*(ϕ, K)= $\{\phi\}$, *justifica*($\phi, K^\times \phi$)= \emptyset , si $\phi \notin K$, dejando abierta la posibilidad a incorporar eventualmente $\neg\phi$ y descartar ϕ . Esta propiedad intenta salvar parcialmente las limitaciones de la revisión AGM descritas en 3.2.

(P.5) *justifica*(ϕ, K) $\neq \{\neg\phi\}$ sssi $\phi \in K^\times \phi$. La propiedad se deriva a partir de Rev.2, Rev.3 y Rev.6. Luego, el éxito en la incorporación de nuevas creencias queda restringido exclusivamente a sentencias que no poseen una justificación para su negación.

(P.6) Si $\phi \in K$ y *justifica*(ϕ, K)= \emptyset , $\neg\phi \in K^\times \neg\phi$ (por postulado Rev.2) y $\phi \notin K^\times \neg\phi$ (por postulado Rev.1). Esto es, ante la indeterminación en la justificación de una sentencia perteneciente al estado epistémico de un agente, se prefiere la nueva información relacionada con la creencia ("privilegio" de AGM restringido).

(P.7) $K^\times \phi \subseteq Cn(K \cup \{\phi\})$. Se deduce a partir de Rev.4, Rev.6 y del hecho que $K \subseteq Cn(K \cup \{\psi\})$, para cualquier sentencia ψ .

(P.8) Si $\phi \in K$ entonces $K^\times \phi = K$. Revisar por una sentencia perteneciente al estado epistémico de un agente mantiene invariable su conocimiento. La justificación se obtiene a partir de Rev.3, Rev.4, del hecho que si $\phi \in K$ entonces $Cn(K \cup \{\phi\}) = K$, y por los axiomas Rev.5 y Rev.1.

(P.9) Si $\phi \in K$ ó $\phi \notin K^\times \phi$, entonces *justifica*(ψ, K)=*justifica*($\psi, K^\times \phi$). Si la creencia a incorporar pertenece al estado epistémico de un agente ó posee una justificación contraria, el agente mantiene la justificación de las creencias en el estado epistémico anterior. La prueba se sigue de Rev.6, P.5, P.8 y de la funcionalidad de *justifica*.

(P.10) La expansión de creencias está también parametrizada (condicionada).

Si *justifica*(ϕ, K)= $\{\neg\phi\}$ y $\neg\phi \notin K$, entonces $\neg\phi \notin K^\times \phi$ y $\phi \notin K^\times \phi$. Se sigue de Rev.6 y de Rev.2.

(P.11) La revisión en el modelo MPRC es una extensión, aunque no conservativa, de la revisión AGM. Si bien esta última intuitivamente resulta de la primera considerando la restricción $\forall \phi, K. \text{justifica}(\phi, K) = \emptyset$, la misma viola la condición J.2 de la función *justifica*, que conjuntamente con la restricción J.3 eliminan la posibilidad de introducir información contradictoria en un estado epistémico.

(P.12) Una relación formal entre la operación de revisión de AGM (notada $*$) y la revisión consistente introducida (notada \times) se sigue de los resultados previos y de J.5:

Si *justifica*(ϕ, K)= $\{\neg\phi\}$ ó $\phi \in K$ entonces $K^\times \phi = K$ sino, si $\neg\phi \notin K$ entonces $K^\times \phi = Cn(K \cup \{\phi\})$ sino, $K^\times \phi = Cn(\{\psi \in K^* \phi \mid \text{justifica}(\psi, K^* \phi) \neq \{\neg\psi\}\})$.

4.4. Contracción en Σ_c

El objetivo ahora es definir a la operación de contracción en el modelo formalizado a partir de la revisión. Una relación en base a la revisión ' \times ' y la contracción AGM ' \sim ' se establece a partir de la identidad de Levi y la propiedad P.12. Sin embargo, la identidad anterior no define explícitamente una nueva operación de contracción.

La manera más natural de definir a la operación de contracción es a través de la fórmula de Harper, presentada en 2.3.4: $K^\sim \phi = K \cap (K^\times \neg\phi)$, con ' \times ' la operación de revisión en MPRC y ' \sim ' la contracción. Esta definición equivale, para las 7 actitudes epistémicas en consideración, a la siguiente definición

$K^\sim \phi = K^\sim \phi$, si *justifica*(ϕ, K)= \emptyset y $\phi \in K$

$K^\sim \phi = K$, en otro caso

(' \sim ' denota la contracción AGM)

Observemos que existen dos procesos de eliminación de información: un proceso de sustracción por contracción explícita con ' \sim ' y un proceso de "descreencia" implícito que se refleja en la restricción de la función *justifica*, plasmada en el postulado Rev.3. La contracción implícita sanciona la eliminación del siguiente conjunto del estado epistémico del agente, luego una operación de actualización 'op' (' \sim ' ó ' \times '), sin tener en cuenta el postulado Rev.3 en la definición de 'op': $\{\psi \mid \psi \in K^{\text{op}} \phi \text{ y } \text{justifica}(\psi, K^{\text{op}} \phi) = \{\neg\psi\}\}$. Este conjunto representa las sentencias que perdieron su justificación debido al proceso de actualización. Asimismo, notemos que la contracción por sentencias justificadas no produce cambios en el conocimiento. Es decir, el "éxito" de ' \sim ' está también restringido.

5. Un Trabajo Relacionado

Existen algunos trabajos recientes que caracterizan modelos de cambio de creencias donde no se satisface el postulado de éxito para las revisiones. Éstos generalmente se enmarcan en un contexto *justificacionista* y utilizan "argumentos" o "explicaciones" para las creencias. Las revisiones se

realizan contraponiendo argumentos ante situaciones conflictivas. La necesidad de utilizar argumentos en un proceso de revisión surge porque un agente inteligente antes de incorporar nueva información a su conocimiento, normalmente espera una justificación para la misma, por lo que la noción de “argumento, explicación o justificación” puede ser completamente aplicable en estos casos. El modelo MPRC introducido en este artículo presenta una formalización unificadora y más abstracta de estos sistemas, a través de una extensión paramétrica del principal modelo aceptado en la teoría, AGM.

En [7], Hansson plantea un modelo alternativo a AGM, llamado *semi-revisión*. Este difiere respecto del modelo standard –al igual que el propuesto en este trabajo– en que la sentencia que provoca una revisión no siempre es aceptada. Hansson presenta, además, un método constructivo para la *semi-revisión*, el cual requiere un tratamiento eficiente de inconsistencias locales, que resulta factible en conjuntos de creencias no clausurados (bases de creencias) y no en conjuntos clausurados. El método de *semi-revisión* de creencias es una forma de cambio de creencias que no asigna ningún valor o prioridad especial a la nueva información. De acuerdo a esto, la *semi-revisión* por una sentencia es el proceso de decidir o analizar no solamente la sentencia propuesta, sino también la negación de ésta. El resultado de la *semi-revisión* será que el sistema de creencias incorpore la sentencia propuesta, o incorpore la negación de la sentencia (este caso no vale en MPRC) o bien que permanezca inmune al cambio propuesto, no incorporando a ninguna de ellas. El proceso de *semi-revisión* no especifica en base a qué criterios o con qué parámetros tomar las decisiones mencionadas, es decir ofrece una caracterización sintáctica pero carece de una contraparte semántica que pueda guiar el proceso central: el de efectuar una decisión por una sentencia, por su negación o por ninguna de ellas [19, 20]. En este trabajo dicho proceso está parametrizado, pudiendo instanciarse con distintos esquemas cualitativos o cuantitativos de justificación, como los citados en la sección 3.2.

Semi-revisión se caracteriza para conjuntos de creencias en base a una colección de postulados. De éstos, algunos son compatibles con la operación de revisión introducida en este artículo. Los postulados de *semi-revisión* que se satisfacen en MPRC son: *closure*, *stability*, *internal exchange*, *inclusion*, *negation-neutral inclusion*, *vacuity* 2, *strong consistency*, *consistency retainment* y *extensionality*. Los postulados que no se verifican son: *weak success*, *negation-neutrality* y *vacuity*. Asimismo, existen propiedades que el modelo MPRC cumple y

semi-revisión no, muchas de las cuales fueron enunciadas previamente en 4.3.⁶ En particular, *semi-revisión* no asegura siempre la consistencia de un estado epistémico actualizado, criterio que es central en MPRC.

6. Conclusiones y Trabajos Futuros

El principal resultado del modelo AGM es haber destacado puntos importantes de la “dinámica de la información”. Básicamente establece requerimientos mínimos para cambios racionales de creencias y abstrae de detalles de bajo nivel de formalismos específicos. Esto fue hecho principalmente investigando las relaciones entre dos estados consecutivos de creencias y la información causante de la transición. Sin embargo, la teoría establece una condición de privilegio de la nueva información causante de un cambio de creencias y no permite concebir un modelo justificacionista, consecuencia que lleva, además, a aceptar a una teoría inconsistente como posible estado epistémico de un agente racional.

El modelo de actualización de conocimiento propuesto en este trabajo impide caracterizar un estado inconsistente y condiciona la aceptación de nueva información. Del estudio desarrollado se desprende que MPRC resulta naturalmente adecuado, ya que la aceptación de creencias queda parametrizada, pudiendo adaptarse el modelo a un esquema de justificaciones para las creencias. El artículo presentó, además, las principales propiedades de MPRC, su relación formal con las operaciones de cambio AGM y las diferencias con el método de *semi-revisión*.

Un trabajo futuro consiste en desarrollar caracterizaciones constructivas de las operaciones del modelo MPRC, cuyas implementaciones impliquen costos computacionales razonables. La idea es usar bases de creencias (bases de conocimiento), en lugar de conjuntos de creencias, como modelo de estado epistémico y definir caracterizaciones constructivas en relación a las de AGM (*partial meet* [1], *epistemic entrenchment* [3] y *safe contraction* [10]), en virtud de las correspondencias establecidas con las operaciones del modelo MPRC, en P.12 y la sección 4.4. Luego, resulta necesario redefinir los postulados Rev.1-Rev.9 para bases de creencias, ya que en este contexto admitimos $K \neq Cn(K)$. Un trabajo relacionado es [7]. Trabajar con bases de creencias permite diferenciar creencias básicas y derivadas de

⁶ Las demostraciones son omitidas por razones de espacio.

un agente racional, como así también hablar de implementaciones computacionales que tengan costos razonables [4, 5, 7, 9, 16].

Otro trabajo pendiente es investigar particulares instanciaciones de la función de justificación. Por ejemplo, en torno al paradigma argumentativo [17, 18], u otro criterio cualitativo ó cuantitativo. Finalmente, analizar un modelo en el cual el postulado Rev.3 no sea exigido, tanto para conjuntos clausurados (conjuntos de creencias) como para no clausurados (bases de creencias). Esta alternativa simplificaría la operación de revisión, ya que se eliminaría el proceso de "descreencia" implícito (contracción implícita), aunque sentencias no justificadas podrían pertenecer al estado epistémico de un agente, luego de un cambio de creencias. Consecuentemente, la implementación de la operación de revisión para bases de conocimiento resultaría de un costo menor que al considerar Rev.3 –y su adaptación a conjuntos no clausurados.

Referencias

1. C. Alchourrón, P. Gärdenfors and D. Makinson. *On The Logic of Theory Change: Partial Meet Contraction and Revision Functions*. Journal of Symbolic Logic, 50:510-530, 1985.
2. P. Gärdenfors. *Knowledge in flux: Modeling the Dynamics of Epistemic States*. Cambridge, MA: The MIT Press, Bradford Books, 1988.
3. P. Gärdenfors and D. Makinson. *Revision of Knowledge System Using Epistemic Entrenchment*. Second Conference on Theoretical Aspect of Reasoning About Knowledge: 83-95, 1988.
4. S. O. Hansson. *A Textbook of Belief Dynamics: Teory Change and Database Updating*. First Draft, Upsala University, Sweden, 1993.
5. S. O. Hansson. *A Dyadic Representation of Belief*, pp. 89-121 in Gärdenfors (de.), *Belief Revision*, Cambridge University Press, Cambridge, 1992.
6. I. Levi. *The Enterprise of Knowledge*. MIT Press, Cambridge, 1980.
7. S. O. Hansson. *Semi Revision*. Journal of Applied Non-Classical Logic, in press.
8. C. Alchourrón and D. Makinson. *Hierarchies of Regulations and Their Logic*. New Studies in Deontic Logic, R. Hilpinen, ed. Dordrecht: Reidel, 123-148. 1980.
9. C. Alchourrón and D. Makinson. *On The Logic of Theory Change: Contraction Function and their Associated Revision Functions*. Theoria 48:14-37, 1982.
10. C. Alchourrón and D. Makinson. *On the Logic of Theory Change: Safe Contraction*. Studia Logica 44:405-422, 1985.
11. A. Grove. *Two Modelling for Theory Change*. Auckland Philosophy, 1986.
12. Alberto Mendelzon and Hirofumi Katsumo. *Propositional Knowledge Base Revision and Minimal Change*. University of Toronto, 1990.
13. I. Levi. *Direct Inference*. The Journal of Philosophy, 74:5-29, 1977.
14. W. L. Harper. *Rational Conceptual Change*. PSA 1976, 2:462-494, 1977.
15. P. Gärdenfors. *The Dynamics of Belief Systems: Foundations vs. Coherence Theories*. Revue Internat de Philosophie, 44:24-46, 1990.
16. A. Fuhrman. *Theory Contraction through Base Contraction*. J. of Philosophical Logic 20:175-203, 1991.
17. M. Falappa y G. Simari. *Revisión de Bases de Conocimiento mediante Argumentos*. CIMAFA'95, Cuba, Enero de 1995.
18. M. Falappa y G. Simari. *Propiedades del Operador de Revisión mediante Argumentos*. CACIC'95, Bahía Blanca, Argentina, 1995.
19. G. Parra y G. R. Simari. *Semi-Revisión Plausible en Bases de Creencias*. Workshop ATIA'97, UNLP, Argentina, 1997.
20. G. Parra y G. R. Simari. *Una operación de Consolidación basada en Plausibilidad*. Workshop ATIA'98, U.N. del Comahue, Argentina, 1997.
21. C. Luna. *Revisión de Creencias: Observaciones y Propuestas*. ICIE'98, Bs. As., Argentina, 1998.