Capítulo 4

Revisión de creencias

El campo de la revisión de creencias es lo suficientemente amplio (e interesante) como para dedicársele libros completos (como [1]) y toda una variedad de artículos que lo abordan desde ángulos diferentes. Es entonces imposible estudiar cada aspecto de él de manera correcta[1] y breve. Así, en este capítulo introduciremos al lector la teoría o modelo de revisión de creencias considerada por muchos como la más importante (como en [4, 2]): la teoría AGM. De hecho, expondremos sólo una parte de esta teoría y nos limitaremos a mencionar aquello que quede fuera de los alcances de este trabajo.

Ya en el capítulo anterior hemos hecho énfasis en la relación existente entre la lógica matemática y la filosófica. De manera análoga, la revisión de creencias, si bien cuenta con aplicaciones computacionales, parte de la filosofía[2], en particular de la *epistemología*[3]. Esto nos invita a abordar el tema de forma más *filosófica* que algunos de los artículos clásicos de la teoría de la revisión de creencias (como [19, 44]).

Cabe hacer una importante aclaración sobre el uso del término *revisión de creencias* en la literatura. Mientras que el término originalmente hace referencia a un tipo de operación de cambio de creencias, se le usa generalmente para referirse al área de estudio también llamada (*teoría del*) *cambio de creencias* o (*teoría de la*) *dinámica de creencias* [2]. Dicho uso es debido tal vez a la popularidad de este tipo de cambio de creencias. Así, cuando se habla de la *revisión de creencias* como un área de investigación, no hay que perder de vista que se está incluyendo mucho más que simplemente uno de los cambios de creencias que estudia.

[1] Entendiendo como correcto lo que es completo.
[2] De acuerdo al tratamiento que se da en [1].
[3] Que definida someramente es «el estudio del conocimiento y las creencias justificadas» [43].

4.1. Objeto de estudio de la revisión de creencias

Tal vez una de las mejores maneras de comprender en un principio los problemas estudiados por la teoría de la revisión de creencias sea a través de un ejemplo. Aquí parafraseamos uno que tomamos prestado de [1]:

> Oscar compró dos anillos de bodas de oro en una tienda en Casablanca. El vendedor afirmaba que los anillos eran de 24 kilates. Pese a ello, Oscar acudió con el vendedor de la tienda de al lado para verificar que en efecto fueran de oro.
>
> Tiempo después de la boda, Oscar estaba haciendo reparaciones y notó que un poco de ácido sulfúrico había manchado su anillo. En ese momento recordó que el agua regia es el único ácido que afecta al oro. Sorprendido, verificó que lo mismo pasaba con el anillo de su esposa.
>
> En vista de que sus creencias daban lugar a una contradicción, Oscar tuvo que *revisarlas*. Así que, como tenía más confianza en lo que había aprendido en sus clases de química que en su perspicacia como comprador, Oscar llegó a aceptar tristemente que los anillos no eran de oro.

Lo que ha sucedido en el ejemplo es que Oscar tuvo que transformar sus creencias en virtud de una pieza de información *contradictoria* que le ha indicado que parte de ellas no es correcta. Oscar aceptó la nueva información y rechazó algunas creencias que podían originar una contradicción con la nueva información. Esto es a lo que se conoce como *revisión de creencias* [1].

Esta situación es claramente diferente de lo que habría ocurrido si Oscar hubiera aprendido algún hecho *consistente* con sus creencias, como cuando aprendió algo sobre química por primera vez; es decir, un caso donde él no tiene una creencia que pudiera estar en contradicción con una nueva pieza de información. En tal situación él simplemente se limitaría a agregar la nueva pieza de información sin modificar sus creencias existentes. Este tipo de cambio en sus creencias es lo que se conoce como *expansión de creencias* [1].

Aparte de poder revisar sus creencias y expandirlas, Oscar podría preguntarse qué hubiera pasado si no hubiera comprado esos anillos en primer lugar. Para simular tal situación, él tendría que decidir, de entre sus creencias actuales, cuáles debería dejar de creer junto con la creencia de que compró los anillos y cuáles mantener.[4] Este último

[4] Al dejar de creer, generalmente se hace énfasis en dejar de creer también aquellas creencias que permiten deducir la creencia que se pretende retraer, aunque es posible pensar también en dejar de creer aquello cuyo único fundamento es la creencia que se pretende eliminar [1].

tipo de cambio es lo que se llama *contracción de creencias* [1].

En los tres casos que parten del ejemplo (revisión, expansión y contracción) se encuentran subyacentes dos objetos: *estados epistémicos* y *cambios* (o *transiciones*) entre estados epistémicos[5]. Según [1], en el área del cambio de creencias las teorías particulares (llamadas *teorías epistemológicas*) se encargan entonces de modelar tales estados epistémicos y su dinámica. Para lograr tal modelado, una teoría epistemológica debe contar con lo siguiente [1, 4][6]:

1. Una clase de modelos de *estados epistémicos*: modelos o representaciones de estados cognitivos de agentes racionales.

2. Una clasificación de las *actitudes epistémicas*: las posibles condiciones en las que se puede representar una creencia dentro de un estado epistémico.

3. Una clasificación de las *entradas epistémicas* que pueden llevar a cambios en los estados epistémicos.

4. Una clasificación de los *cambios epistémicos* o *cambios de creencias* que ocurren en respuesta a las entradas epistémicas.

5. Un conjunto de *criterios de racionalidad* que se encuentran en el meta-nivel de la teoría y sirven para evaluar las demás partes de ésta.

Con las primeras cuatro partes de la teoría epistemológica se forma luego un *sistema de creencias* [1, 4], que es el que finalmente modela ambos objetos de estudio.

Una característica importante de esta propuesta de [1] es que las teorías epistemológicas son *conceptuales* en el sentido de que no hacen asunción alguna respecto al *mundo exterior* a los modelos epistémicos: el modelado de los estados y su dinámica debe hacerse independientemente de las relaciones factuales entre las entradas epistémicas y el mundo exterior. Al hablar del modelo AGM (sección 4.3) haremos algunas observaciones al respecto.

4.2. Componentes de una teoría epistemológica

Antes de hablar sobre la teoría epistemológica (i.e. de dinámica de creencias) que nos interesa, es importante explicar algunos conceptos base y las partes de una teoría cualquiera.

[5] Donde revisión, expansión y contracción no son los únicos tipos de cambios posibles, como veremos en la subsección 4.2.5.3.

[6] Curiosamente, a pesar de basarse en [1], [4] omite mencionar el inciso 4.

Hay que aclarar que, debido al carácter filosófico de la revisión de creencias, sus conceptos tienen origen en filosofía y se toman de ahí para usarse en ambas facetas (si es que en algún momento pueden realmente disociarse completamente), la filosófica y la computacional. Aunque no es nuestra intención entrar aquí en profundidad sobre el significado filosófico preciso de algunas definiciones base (como la de *creencia*), consideramos ilustrativo introducirlas de manera breve, concisa y orientada a un contexto computacional para cumplir con el carácter didáctico de este documento.

4.2.1. *Creencias y conocimiento

Definir el concepto de *creencia* no es trivial.[7] Para intentar dar una noción, digamos que una *creencia* es una *proposición* que es considerada como verdadera por un agente (o sujeto), la cual puede ser o no verdadera en realidad [45].

El término *conocimiento* se encuentra muy relacionado con el de *creencia*. Considerando una vez más la existencia de un agente, puede definirse *conocimiento* como una *creencia verdadera justificada*: un agente X sabe p si y sólo si p es verdadero y X tiene una justificación para creer en p [43]. Es decir, un conocimiento es una creencia que es verdadera y que es creída no por mera suerte, sino porque existe una justificación para creerla.

Es común encontrar que mientras algunos trabajos utilizan la palabra *creencia* en términos compuestos que se refieren a sistemas de representación de creencias, otros recurren a la palabra *conocimiento* en su lugar para el mismo propósito. Tal diferencia puede apreciarse entre [5] y [46]. Sin embargo, en vista de que (como se dijo en la sección 4.1) las teorías epistémicas no deben atender a la relación entre las entradas y el mundo real, se entiende que el concepto de *verdad* es irrelevante o secundario para éstas [1]; entonces el término que debería usarse es el de *creencia*.

4.2.2. Representación de creencias

Si bien al presentar el ejemplo de la sección 4.1 se ha hablado del estado mental de Oscar como un *estado epistémico*, en general cuando se habla de *estados epistémicos* se está refiriendo a *idealizaciones* o *modelos* [1, 4]. Así, un *estado epistémico* no se limita solamente a entidades psicológicas, sino a cualquier representación idealizada (como puede ser un programa de computadora) [1].

[7] En [3] se muestra una definición de *creencia* desde el punto de vista de la inteligencia artificial, que es en esencia la misma que presentamos aquí, pero más detallada.

Conceptualmente, los estados epistémicos son *estados de equilibrio*; más precisamente, se considera que los estados epistémicos deben estar en equilibrio interno, el cual es guiado por criterios de racionalidad [1, 4]. El agente (o ente) que posee al estado epistémico es quien se encarga de mantener tal equilibrio interno [1]. Tal requerimiento es, evidentemente, un requisito teórico ideal: considerando al estado de ejecución de una aplicación como un estado epistémico, no hay una fuerza natural que obligue a un programa de computadora a comportarse como las especificaciones de su diseño lo requieren.

4.2.2.1. Modelos justificacionistas y de coherencia

De acuerdo a [5], los modelos de estados epistémicos pueden clasificarse en dos grandes categorías: *modelos justificacionistas* y *modelos de coherencia*. También según [5], tales modelos pertenecen a dos teorías que pretenden responder a la pregunta: Â¿Al modelar un estado de creencias, deberían las justificaciones de las creencias ser parte del modelo? Ambas teorías son en un inicio *criterios de racionalidad* para llevar a cabo la operación de revisión de creencias[8], las cuales a su vez dan lugar a la necesidad de tener las dos categorías correspondientes de representaciones de creencias [3].

Los modelos justificacionistas, donde las justificaciones de las creencias son parte del modelo, pertenecen a la *teoría fundacionalista* [5]. Esta teoría postula el siguiente principio:

Principio de la incertidumbre negativa 4.1 ([3])**.** Una proposición para la cual ya no existe suficiente justificación no debe seguir siendo creída.

El principio anterior significa que (particularmente) una revisión de creencias consiste en eliminar primero toda creencia para la cual no existe ya una «justificación satisfactoria» y agregar luego aquellas creencias que se han justificado [5].

Ejemplos de modelos justificacionistas son: (i) Sistemas de mantenimiento de creencias, donde las creencias se representan mediante redes semánticas, siendo los nodos las creencias y las aristas las justificaciones para las creencias; (ii) Modelos bayesianos, donde las creencias se modelan como una función de probabilidad sobre un lenguaje, cuyo valor indica el grado de creencia que se tiene sobre una proposición particular. [1, 3]

Los modelos de coherencia, donde las justificaciones de las creencias no son incluidas en el modelo, pertenecen a la *teoría de coherencia*. Esta teoría postula como principio:

[8]Aunque no habría que limitar su motivación a la de ser un lineamiento para hacer revisión de creencias, sino cualquier tipo de cambio de creencias en general, pues de hecho tienen influencia en todo cambio de creencia que tenga lugar en sus modelos.

Principio de la economía informacional 4.2 ([3]). Cuando se incorporan nuevas creencias que contradicen a las anteriores, uno debe escoger mantener el conjunto de creencias que requiere cambio mínimo respecto al conjunto anterior, a la vez que se mantiene la consistencia.

Este principio también es conocido como *principio de conservación* [3]. Éste significa que una revisión de creencias debe de conservar tantas creencias, de las ya existentes, como sea posible, manteniendo al mismo tiempo la consistencia del sistema con respecto a la nueva creencia agregada [5].

Ejemplos de modelos de coherencia son: (i) Conjuntos de creencias (*belief sets*), donde las proposiciones que conforman las creencias son representadas mediante fórmulas de algún lenguaje formal (comúnmente el lenguaje $\mathcal{P}$ de la definición 2.29 en la página 20 como en [1, 19, 44]) y un conjunto de fórmulas cerrado bajo un operador de consecuencia lógica representa un estado epistémico; (ii) Modelos de mundos posibles, donde se trata de evitar representar proposiciones por medio de fórmulas particulares y se opta por relacionar una proposición a una o varias fórmulas considerando que una fórmula expresa cierta proposición si y sólo si la fórmula es verdadera en aquellos mundos posibles que constituyen el conjunto de mundos que representan a la proposición [1, 3]. Debe notarse que en ambos casos se pretende diferenciar entre fórmula y proposición, siendo la segunda el contenido esencial de la primera, como ya se ha explicado en la sección 3.1.

Existe un debate acerca de cual teoría debería ser preferida sobre la otra [5, 3]. Se ha argumentado a favor del enfoque de la teoría de coherencia en virtud de que se ha demostrado que la gente no lleva cuenta de la justificación de sus creencias [3]. Aunque por otro lado, el no llevar cuenta de la justificación para las creencias puede provocar que por ejemplo, si la contracción considera racional eliminar aquello que ya no tiene justificación, al realizar una contracción no se pueda saber qué creencias soportadas solamente por la que se quiere eliminar deberían retractarse también [1]. Sin embargo, una mejor respuesta a la pregunta «cuál tipo de representación es mejor», es aquella que dice que ambos enfoques son más bien complementarios [3].

4.2.3. Actitudes epistémicas

Las actitudes epistémicas son las condiciones o estados en los que se puede calificar a una creencia dentro de un estado epistémico [1, 4]. En general, la actitud epistémica de las creencias es descrita en términos de alguna *valuación* de los elementos del modelo epistémico [1]. Evidentemente las actitudes epistémicas varían en función del tipo de modelo que se elija para representar estados epistémicos. En el caso del modelo de conjuntos de creencias, la valuación de las actitudes epistémicas está dada por la relación de membresía, donde existen tres tipos de actitudes epistémicas dado un conjunto de creencias A y una fórmula x [1, 4]:

1. Si $x \in A$, x es aceptada en A.

2. Si $\neg x \in A$, x es rechazada en A.

3. Si $x \notin A$ y $\neg x \notin A$, x es indeterminada en A.

Otros sistemas de representación de creencias permiten representar actitudes epistémicas más complejas. Los basados en probabilidades, por ejemplo, pueden modelar grados de creencias, de tal forma que dos creencias pueden compararse para distinguir si alguna de ellas es más fuerte que otra. [1, 3]

4.2.4. Entradas epistémicas

Una entrada epistémica es una «fuerza externa que provoca un cambio epistémico» [1]. Si un estado epistémico está en equilibrio interno respecto a las fuerzas internas que lo gobiernan (i.e. los criterios de racionalidad impuestos por el tipo modelo de representación), las entradas epistémicas son estímulos externos (y las únicas fuerzas) que pueden llegar a provocar un *cambio de creencias* que lleve del estado de creencias actual (y en equilibrio) a otro estado de creencias (también en equilibrio) [1, 4]. En [1], donde se postulan las teorías epistemológicas, el concepto es bastante amplio, pues no se indica con exactitud en qué consiste una entrada epistémica; se dice que son *información externa* o *experiencias* y que su forma no es importante sino el *efecto* que tienen sobre los estados de creencias.

En vista de que lo importante de las entradas epistémicas es su efecto, éstas se *definen* a través de los cambios que causan en los estados epistémicos [1]. Al igual que con las actitudes epistémicas, dependiendo del tipo de modelo de representación elegido, se tendrá una determinada tipología de cambios epistémicos. Aunque [1] no parece decirlo explícitamente, la tipología de cambios epistémicos (y por consiguiente la de las entradas epistémicas) puede establecerse a diferentes niveles de detalle.

Si atendemos a los cambios de actitud epistémica que sufren las creencias individuales en un cambio epistémico dentro de un modelo de conjuntos de creencias, entonces se tienen dos tipos de entradas epistémicas: (i) las entradas que indican que una fórmula (i.e. una creencia) que antes no era aceptada, debe serlo ahora; (ii) las que indican que una creencia no debe ser creída más. Los tipos de cambios son, en el primer caso *adiciones* y en el segundo *derogaciones*. [1]

Otro nivel de detalle, o criterio, para crear una tipología de entradas epistémicas, es la que considera que los efectos que tienen pueden ser indicados como *restricciones* en los estados de creencias resultantes. La adición sería la restricción que dice que la nueva creencia debe ser aceptada[9] y la derogación sería la restricción de que la creencia sea indeterminada en el estado de creencias. [1]

Una última forma de definir las entradas epistémicas es identificarlas con *funciones* que toman estados de creencias y producen nuevos estados de creencia. Esto significa que se está relacionando las entradas epistémicas con la totalidad de los cambios que ocurren en un estado de creencias. [1]

4.2.5. Cambios epistémicos

Los cambios epistémicos son provocados por las entradas epistémicas y se llevan a cabo siguiendo reglas que determinan como debe cambiar el estado de creencias; tales reglas son llamadas *compromisos epistémicos*, *funciones de compromiso epistémico*, *funciones de compromiso*, *funciones de cambio*, entre otras formas [1]. En la literatura es difícil distinguir entre las funciones y los cambios mismos, pues las asunciones hechas provocan que ambos conceptos siempre *ocurran juntos*; nos gustaría hacer a continuación una distinción brevísima que hemos concluido a partir de la literatura y que permitirá dejar claro lo que intentamos decir por *ocurren juntos*.[10]

4.2.5.1. Cambios epistémicos y funciones de compromiso epistémico

La diferencia hacemos entre los dos conceptos es que el cambio epistémico es el efecto que tiene en un estado epistémico la aplicación de una función de compromiso epistémico.

Las posibilidades para la creación de funciones son casi ilimitadas (restringidas, una vez más, por las características del modelo de representación de creencias). Tal vez las funciones de cambio de creencias más comunes sean las tres presentadas en la sección 4.1 [5, 45]:

- *Expansión*: consiste en adicionar a un estado de creencias una nueva creencia, posiblemente junto con sus consecuencias, sin eliminar las creencias existentes

[9] Donde el *rechazar* una creencia es un caso particular.

[10] Luego de dicho comentario, al igual que en la literatura, no distinguiremos entre los cambios y las funciones al usar los términos, a menos que se indique lo contrario.

(i.e. sin ninguna consideración acerca de si el estado resultante es consistente o no).

- *Revisión*: consiste en adicionar a un estado de creencias una nueva creencia, posiblemente junto con sus consecuencias, asegurando que el nuevo estado de creencias sea consistente.

- *Contracción*: una contracción consiste en eliminar cierta creencia (y todo lo que la soporta) sin agregar nuevas creencias.

A cada una de las funciones le corresponde un cambio epistémico (los mismos que se describieron en la sección 4.1), el cual, estrictamente hablando, sólo tiene lugar bajo ciertas condiciones en el estado epistémico. Por ejemplo, si se fuera a aplicar una *función de compromiso epistémico de contracción* de manera arbitraria a un estado de creencias, ésta solamente eliminaría cierta creencia si existe en el estado de creencias, y por lo tanto el *cambio epistémico de contracción* sólo tendría lugar si se cumple esa condición. Sin embargo, en [1], se asume que ante una entrada epistémica, un *sistema de creencias*[11] aplica la función de compromiso epistémico adecuada, lo cual impide que dichas funciones se apliquen en *casos degenerados* (p. ej. aplicar la contracción de cierta creencia cuando ésta no es parte del estado de creencias) y por consiguiente todo cambio de creencias tiene lugar al aplicar su correspondiente función de compromiso epistémico: los cambios y las funciones *ocurren juntos*.

Ejemplifiquemos esto atendiendo al caso de los modelos de conjuntos de creencias. Existen seis tipos de cambios de actitud epistémica respecto a una creencia que pueden ocurrir durante un cambio epistémico (véase figura 4.1) [1, 4].

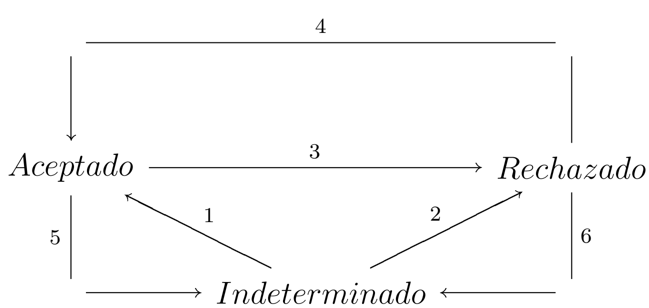

Figura 4.1: Posibles cambios de actitud epistémica para una creencia [1, 4].

En [1, 4], atendiendo a la actitud epistémica de la creencia contenida en la entrada

[11] Véase sección 4.1.

epistémica que estimula el cambio[12], se relacionan los puntos 1 y 2 de la figura 4.1 con la función de expansión, a los puntos 3 y 4 con la de revisión y a los puntos 5 y 6 con las contracciones. Sin embargo, desde nuestra forma de ver, tales relaciones entre cambios y funciones son solamente *idealizaciones teóricas*: dicho una vez más, no hay una fuerza natural que impida aplicar una contracción cuando la creencia que se pretende eliminar es desconocida.

4.2.5.2. Tipología de cambios de creencias y de entradas epistémicas

Siguiendo la línea de pensamiento de la subsección 4.2.4 para la creación de una tipología de entradas epistémicas dependiente de la tipología de cambios de creencias, las entradas epistémicas y los cambios epistémicos pueden verse relacionados a través de los compromisos epistémicos; por ejemplo, en un modelo de conjuntos de creencias, se pueden relacionar las adiciones con las expansiones y revisiones, considerando que la adición es el estímulo que indica que debe aceptarse una creencia y que una expansión o revisión es la respuesta del sistema de creencias a tal estímulo; en el caso de las derogaciones, éstas se corresponden directamente con las contracciones. Es posible, sin embargo, evitar asumir la existencia *a priori* de entradas epistémicas al considerar los estados epistémicos como entidad fundamental y luego definir las funciones de cambio, con las cuales se definen los cambios al igual que los tipos de entradas. [1]

4.2.5.3. Otros tipos de cambios de creencias

Para hacer énfasis en que los tres tipos de *funciones* de cambio de creencias presentados no son los únicos, mencionaremos brevemente y sin entrar en detalles técnicos una variante de la revisión de creencias: la actualización de creencias. Comúnmente, la diferencia entre la revisión y la actualización de creencias se intenta explicar como una oposición entre cambios de creencias sobre un mundo estático contra cambios de creencias sobre un mundo dinámico [47, 5, 3]. El ejemplo siguiente (una ampliación del presentado en [3]) ilustra lo que queremos decir:

> Imaginemos que actualmente creemos que en el cuarto de al lado, la ventana está abierta (x) o la puerta está abierta (y), es decir: $x \vee y$.
>
> Imaginemos que le pedimos a alguien que vaya al cuarto y nos diga si la ventana está abierta o cerrada, y que luego de hacerlo nos informa que la ventana estaba cerrada ($\neg x$). Si aplicamos la función de revisión a las

[12] Pues es posible (o más bien, es muy probable) que en un cambio epistémico cambie la actitud epistémica de otras creencias

creencias actuales, (en lógica clásica) esto nos lleva a creer (entre otras cosas) en $\neg x$ y en y, pues (como $\neg x$ no contradice ninguna creencia) sabemos que $x \vee y$, que $\neg x$ y además $\vdash (x \vee y) \wedge \neg x \rightarrow y$. Tal estado de creencias es evidentemente correcto.

Por otro lado, si pedimos que alguien vaya al cuarto y cierre la ventana, también creemos que la ventana no está abierta ($\neg x$); pero en este caso, si aplicamos la función de revisión, según ésta, al cerrar la ventana debemos creer que la puerta se encuentra abierta. Evidentemente tal estado de creencias no es correcto: la ventana pudo haber estado cerrada desde un principio, mientras que lo que se encontraba abierto era la puerta, por ejemplo.

Lo que sucede en el primer caso es que se ha recuperado nueva información sobre un mundo estático, mientras que en el segundo caso se ha realizado un cambio en el mundo (o tal vez no) y este cambio provoca que nuestras creencias dejen de reflejar el estado de dicho mundo. En los modelos de representación de creencias que no modelan el concepto de tiempo, la diferencia entre *estático* y *dinámico* no es en un inicio culpa de la función de revisión de creencias sino de la representación misma, como lo hace notar [47]: «lo que es esencial en la revisión de creencias no es que el mundo sea estático, sino que *el lenguaje utilizado para describir al mundo* es estático». Una solución al problema de la actualización de creencias, propuesta por Katsuno y Mendelzon, es la de asignar índices de tiempo a las proposiciones [2].

Sobre las actualizaciones de creencias pueden consultarse trabajos como [46, 47], donde [46] es uno de los primeros y más importantes (si no es que el primero y más importante) de los trabajos que pretenden estudiar este tipo de cambio (y función) de creencias.

4.2.6. Criterios de racionalidad

En este componente de las teorías epistemológicas se reúnen todos los criterios de racionalidad que rigen o permiten evaluar los otros tres elementos de la teoría [1]. Por ejemplo, si fuéramos a tener una teoría con un modelo de representación basado en conjuntos de creencias, tendríamos aquí los dos criterios de racionalidad que según [1] las acompañan:

- *Consistencia*: las sentencias de un conjunto de creencias no deben ser contradictorias.

- *Omnisciencia lógica*: el conjunto de creencias se asume cerrado bajo un operador de consecuencia lógica.

Otros ejemplos de criterios de racionalidad son algunos que se pueden (no necesariamente) aplicar a las funciones u operadores de cambio epistémico de revisión[13] [4], por ejemplo:

- *Consistencia*: el nuevo estado epistémico debe, si es posible, ser consistente.
- *Primacía de la nueva información*: la nueva información es siempre aceptada.
- *Economía informacional*: retener cuanto sea posible de las creencias preexistentes.

En algunas ocasiones el último es equivalentemente enunciado como en [1]:

- *Mínimo cambio*: realizar siempre el cambio mínimo para la aceptación de la nueva creencia.
- *Pérdida mínima*: la perdida de información debe ser mínima con el objeto de no eliminar creencias innecesariamente.

Sobre el principio de la *economía informacional* se ha argumentado a favor y en contra. Algunos hacen notar que la noción de *minimalidad* que implica no es apropiada para todas las clases de modelos de estados epistémicos [1]. Otros argumentan que en la clase de modelos de conjuntos de creencias[14] existen casos donde el criterio no es del todo racional[15] [1]. GÃ€rdenfors argumenta a favor de tal criterio diciendo que no debe ser tomado literalmente, sino que debe de seguirse en la medida de lo posible; opuesto es el argumento de Levi que dice que la economía informacional interpretada en términos de *cantidad de creencias* no es correcta, sino que habría que considerarla respecto a algún tipo de *valor informacional* [48].

4.3. El modelo AGM

La teoría dominante de cambio de creencias en las ciencias computacionales es la que propone el modelo AGM, llamado así en honor a sus autores, Alchourrón, GÃ€rdenfors y Makinson [2]. La mayoría de las características de una teoría epistemológica que se

[13] Aunque algunos de ellos también pueden aplicarse a la expansión y/o contracción.

[14] Que es donde se aplica tradicionalmente este criterio de racionalidad a los operadores.

[15] Por ejemplo, al contraer una fórmula x, es necesario contraer lo que la permite deducir, pero por el criterio de economía informacional no es necesario contraer aquellas creencias y que se derivan de lo que se debe contraer, aún si lo único que permite que permite deducir a y sean tales creencias.

han mencionado en la sección anterior, corresponden a las del modelo AGM; dicho modelo está basado en conjuntos de creencias y modela los tres tipos de (funciones de) cambio de creencias que hemos venido analizando en el capítulo[16].

4.3.1. Breve historia de la revisión de creencias

El estudio de la revisión de creencias comienza en la década de los 70's con trabajos de William Harper e Issac Levi [4]. Tiempo después, dos corrientes se vieron unidas por la necesidad subyacente de modelar dos cambios de creencias interdependientes: el trabajo de Alchourrón y Makinson que buscaba la definición de un operador de contracción y el de GÃ€rdenfors que buscaba un operador de revisión [49].

Carlos Alchourrón, quien originalmente hubiera estudiado leyes, se vio enfrentado con el problema de la derogación dentro de códigos legislativos. En el contexto de la filosofía de las leyes, la derogación no se veía como algo problemático: para derogar uno simplemente elimina la ley en cuestión. Sin embargo, en vista de que en realidad el proceso de derogación es todo menos simple y unívoco, Alchourrón se vio preocupado por encontrar una solución. Fue entonces que Alchourrón, junto con su colega cercano Eugenio Bulygin, formuló y publicó su problema en 1975[17]. Él acudió entonces a David Makinson en busca de algún tipo de análisis formal. De esa manera comenzaron a trabajar ambos en el problema de la derogación. Para el final de los 70's empezaron a vislumbrar como realizar el proceso de derogación. Eventualmente llegaron a ver que el problema podía enunciarse de forma más general que como lo habían venido haciendo, si se consideraba que la representación consiste en algún tipo de lenguaje formal. [49]

Por su parte, Peter GÃ€rdenfors estaba interesado en encontrar una semántica para los condicionales que estuviera libre de los compromisos ontológicos que suponía la ya existente. Él propuso una solución, entre 1978 y 1982, definiendo reglas o postulados que permitieran evaluar los condicionales, las cuales modelaban las operaciones de revisión y contracción. En 1982, Alchourrón y Makinson vieron que los postulados de GÃ€rdenfors se veían satisfechos por un caso particular de las construcciones de derogación (i.e. contracción) que habían formulado. Ese mismo año, al publicar ese trabajo en la revista *Theoria*, editada por GÃ€rdenfors, entraron en contacto con él. Así fue como se formó finalmente el trío AGM. [49, 4]

En 1985 el grupo AGM publicó [19], el cual es el trabajo seminal donde se presenta el modelo de *contracción de elección parcial* (*partial meet contraction*) que da origen a muchos trabajos subsecuentes: *contracción segura* (*safe contraction*), *contracción de nú-*

[16]Expansión, revisión y contracción.
[17]Aunque en realidad Cornides ya lo había notado en 1969 [49].

cleo (*kernel contraction*, una generalización de la contracción segura), funciones basadas en *firmeza epistémica* (*epistemic entrenchment*) y *sistemas de esferas* [4].

4.3.2. Características del modelo

Las partes de la teoría epistemológica AGM son [3, 4][18]:

- Los estados epistémicos se representan con conjuntos de creencias donde las proposiciones son representadas mediante fórmulas del lenguaje $\mathcal{P}$ [19] [20], el operador de consecuencia lógica Cn asumido es *finitario*[21] y tal que $Cn_{Cl}(\varnothing) \subseteq Cn(\varnothing)$, por lo tanto los conjuntos de creencias son *cn-teorías*[22] [23] *supraclásicas*.
- Las actitudes epistémicas son las que hemos mencionado en la subsección 4.2.3.
- Las entradas epistémicas y los cambios epistémicos[24] quedan definidos al especificar las funciones de compromiso epistémico de expansión, revisión y contracción.
- Los criterios de racionalidad que gobiernan al modelo son los mencionados en la subsección 4.2.6.

Debe notarse que, gracias a la supraclasicidad del operador de consecuencia, en el modelo existe solamente un conjunto inconsistente, i.e. el conjunto de todas las fórmulas del lenguaje (en nuestro caso, $\mathcal{P}$) [1, 4].

La parte de esta teoría que recibe más atención en la literatura ([1, 4, 3, 5, 19, 44]) es la descripción de las funciones de compromiso epistémico, en particular las de revisión y contracción[25], con lo cual quedan completas las partes de la teoría. Una asunción fundamental del modelo AGM es que para cada tipo de cambio de creencias, por cada conjunto de creencias y cada creencia de entrada existe un único resultado[26] [1]; puesto

[18] Aunque en [5, 1] se trata el modelo AGM también, es en las referencias citadas en las que se indica explícitamente que las partes mencionadas pertenecen en efecto al modelo AGM.

[19] Véase definición 2.29 en la página 20.

[20] Aunque en realidad se asume un lenguaje *por lo menos* proposicional, como $\mathcal{P}$, con lo que se deja abierta la posibilidad a utilizar un lenguaje que permita, por ejemplo, un cálculo de primer orden (véase subsección 3.1.2).

[21] Véase definición 2.60 en la página 26.

[22] Véase definición 2.49 en la página 23.

[23] Entonces, por la notación 2.50 en la página 23, se utilizará A para representar conjuntos de creencias.

[24] De los cuales, el primero no es mencionado en la literatura (excepto por [1]) y el segundo es mencionado con una fuerte relación con las funciones de compromiso epistémico que les corresponden.

[25] Pues definir la operación de expansión es una tarea trivial.

[26] Asunción que parece no ser indicada como requisito de las teorías epistémicas en general en [1], aunque el término *función de compromiso epistémico* ya sugiere que se está hablando de relaciones funcionales y no de simples relaciones. Pese a la teoría sobre las teorías epistémicas de [1], en [5] se menciona una teoría de cambio de creencias donde la revisión no está unívocamente determinada.

en términos formales, cada operador de cambio de creencias $\diamond \in \{+, \dot{+}, \dot{-}\}$ (expansión, revisión y contracción, respectivamente) es una $\diamond : Cn\,(\wp\,(\mathcal{P})) \times \mathcal{P} \to Cn\,(\wp\,(\mathcal{P}))$.

Notación 4.3. En adelante, utilizamos la notación $A \diamond x$ en lugar de $\diamond(A, x)$ para todo operador de cambio de creencias, a menos que se indique lo contrario.

Como hemos venido mencionando[27], la expansión trata de modelar (idealmente) el cambio de creencias que ocurre cuando se tiene información para creer cierta proposición y no se posee información previa. Entonces se tiene que:

Definición 4.4. La función de compromiso epistémico de *expansión* del modelo AGM, $+ : Cn\,(\wp\,(\mathcal{P})) \times \mathcal{P} \to Cn\,(\wp\,(\mathcal{P}))$, es tal que $+\,(A, x) \overset{\text{def}}{=} Cn\,(A \cup \{x\})$.

Notación 4.5. En adelante, evitamos utilizar la notación $+\,(A, x)$ y preferimos utilizar $Cn\,(A \cup \{x\})$ para indicar una expansión de la cn-teoría A por x. Este estilo es el mismo que se sigue en [19, 44].

En el caso de la revisión y la contracción surge el grave problema de que no es posible enunciar una definición explícita (como la de la expansión) usando solamente términos lógicos y de teoría de conjuntos [5]. Esto sucede porque el modelo de conjuntos de creencias no es lo suficientemente rico para modelar la información que por lo menos permita determinar qué creencias deben ser aceptadas y cuáles eliminadas [1]; por ejemplo, de una cn-teoría tal que $x \in A$, $y \in A$ y $x \wedge y \to z \in A$ (y sus consecuencias), no existen razones puramente lógicas que indiquen cuales de las fórmulas es mejor eliminar al querer contraer a A respecto a z [5]. Es por ello que en [1, 5] se señala que existen dos maneras de especificar las funciones *formalmente*: formular *postulados de racionalidad* que deben ser satisfechos por las funciones (i.e. *el qué*)[28] y *modelos constructivos* de las funciones (i.e. *el cómo*). Los postulados están inspirados principalmente en el criterio de la *economía informacional*, mientras que algunos modelos constructivos toman en cuenta además el grado de *firmeza epistémica*[29] [1]. Además, para la expansión y revisión (en los postulados y las construcciones) se asume el criterio de la *primacía de la información*; para la contracción se asume que se debe contraer la creencia que se pide siempre que sea posible.

Una característica fundamental de este modelo es que no se encarga de discriminar la aceptación o no de la información contenida en las entradas epistémicas. Si bien no se

[27] Véase sección 4.1 y subsección 4.2.5.

[28] Lo mismo puede hacerse para la operación de expansión, aunque es un caso raramente mencionado: [1] ha sido la única obra revisada donde se han encontrado postulados para la expansión.

[29] Criterio que dice que «algunas proposiciones son más útiles que otras en cuestionamientos y toma de decisiones» y son las creencias con menor firmeza epistémica las que deben retractarse cuando alguna creencia deba ser eliminada [1, 5].

dice lo contrario en [1] para el caso general de cualquier teoría epistémica, sí se afirma que (como ya mencionamos en la sección 4.1) a las teorías epistémicas no les concierne la relación entre el mundo real y los sistemas de creencias que soportan. La decisión de actuar como consecuencia de una entrada epistémica u omitirla no está modelada en la teoría AGM[30]. Esta tarea le corresponde a otros tipos de teorías, como la de Galliers mencionada en [5] o el trabajo de [11][31].

4.3.3. Postulados

Ésta es la parte de la teoría AGM en la que se centra nuestra investigación. En nuestra opinión, los postulados son más importantes que los modelos constructivos porque no comprometen a la teoría con un solo modelo de sus operadores. Además, si se pretende obtener una teoría de cambio de creencias en lógicas no clásicas, son el objeto de estudio indicado para descubrir si la justificación detrás de ella es válida bajo dichas lógicas en primer lugar. Una vez logrado esto, sería posible proceder a desarrollar modelos constructivos y luego aplicaciones concretas.

Aunque los postulados de la teoría no son en sí parte de los criterios de racionalidad de la teoría epistémica AGM, se inspiran directamente en los que ya hemos mencionado en la subsección 4.3.2 y la subsección 4.2.6. Además, estos se encuentran divididos en un conjunto de postulados *básicos* y un conjunto de postulados *suplementarios*, como veremos en breve.

Ya que los postulados se basan en consideraciones puramente lógicas, pues son enunciados usando el lenguaje objeto de la clase de modelos de representación de creencias asumida, estos solamente logran circunscribir un conjunto de implementaciones racionales [1, 5]. En ese sentido fallan en cumplir con la asunción de que los cambios de creencias se encuentran definidos unívocamente.

Antes de introducir los postulados, cabe aclarar que la asunción de que el operador de consecuencia (o la relación de consecuencia) subyacente en el modelo AGM es *supraclásica* da lugar a diferentes versiones de los postulados que son equivalentes bajo lógica clásica, y cuyas variaciones son relativamente *inofensivas*. En algunos casos, los postulados son escritos de manera diferente con notaciones equivalentes (bajo lógica clásica); por ejemplo, en [5, 1] se usa $\vdash x \leftrightarrow y$ en lugar de $Cn(\{x\}) = Cn(\{y\})$, como

[30]Aunque sí lo está, como un supuesto que ya mencionamos en la subsección 4.2.5.1, la elección de la función adecuada dependiendo del contenido de la entrada epistémica.

[31]En este último, interesantemente se consideran varias formas de aceptar la información externa que proviene de varias fuentes, como son aceptar algo que todas las fuentes consideran cierto o combinar información de varias fuentes en casos donde no todas afirman lo mismo.

en [4, 19, 44].[32] En otros casos, existen por lo menos tres versiones de los postulados de AGM que son, estrictamente hablando, diferentes. Las versiones, incluso entre artículos y obras de los mismos autores, a veces contienen postulados que en otras publicaciones no se incluyen[33]. Haremos algunos comentarios sobre los casos donde la variación es de contenido y no de notación.

4.3.3.1. Postulados AGM de contracción

A la luz de las diferentes versiones de postulados hemos considerado pertinente construir nuestra versión particular de los postulados de contracción seleccionado diferentes elementos de entre la literatura, procurando que el conjunto sea lo más apegado a la versión del trabajo seminal [19]. Así, tenemos que:

Definición 4.6. Los postulados *básicos* de racionalidad para la contracción en el modelo AGM son:

$$A \dot{-} x = Cn\left(A \dot{-} x\right) \tag{$\dot{-}1$}$$

$$A \dot{-} x \subseteq A \tag{$\dot{-}2$}$$

$$\text{Si } x \notin A \text{, entonces } A \dot{-} x = A\,. \tag{$\dot{-}3$}$$

$$\text{Si } x \notin Cn\left(\varnothing\right), \text{ entonces } x \notin A \dot{-} x\,. \tag{$\dot{-}4$}$$

$$A \subseteq Cn\left(\left(A \dot{-} x\right) \cup \{x\}\right) \tag{$\dot{-}5$}$$

$$\text{Si } Cn\left(\{x\}\right) = Cn\left(\{y\}\right), \text{ entonces } A \dot{-} x = A \dot{-} y\,. \tag{$\dot{-}6$}$$

Ha de notarse que ya que todos los estados epistémicos se representan con cn-teorías, los postulados sólo deben aplicarse a cn-teorías, por lo que se asume que cada postulado debe leerse como «si $A = Cn(A)$, entonces ...»[34] [35].

En el caso de $(\dot{-}1)$, hemos tomado la versión de [4]; su motivación es clara, asegurar que el resultado de una contracción mantiene las propiedades del modelo de representación [4]. Como $A \dot{-} x$ se forma a partir de A al eliminar algunas creencias, se considera

[32] Es fácil ver que ambas son equivalentes en lógica clásica. En este capítulo damos preferencia a la segunda de ellas en la formulación de los postulados por ser más elemental que la primera y porque ambas no son equivalentes en toda lógica; cuando en la literatura se usa $\vdash x \leftrightarrow y$ se pretende decir que $Cn\left(\{x\}\right) = Cn\left(\{y\}\right)$. Reservamos la notación $\vdash x \leftrightarrow y$ para las demostraciones del capítulo siguiente, donde hacemos la aclaración de porque esto no afecta la generalidad de las pruebas.

[33] O excluyen postulados que otros si enuncian.

[34] Curiosamente en algunas ocasiones pueden verse postulados para los que se hace muy explícita tal condición y postulados para los cuales se omite completamente; por ejemplo, en [19] $(\dot{-}2)$ no posee condición alguna, pero $(\dot{-}3)$ sí.

[35] Al hacer incapié en esta asunción se evita formular condiciones como la de $(\dot{-}3)$ en [19], donde a diferencia de aquí, se dice «si $x \notin Cn\left(A\right)$, entonces $A \dot{-} x = A$».

racional no introducir nuevas creencias en $A \dot{-} x$; esa es la motivación de $(\dot{-}2)$ [1, 5]. Cuando se hace una contracción, el criterio de la *economía informacional* requiere que nada sea eliminado de un estado A cuando $x \notin A$; esta es la utilidad de $(\dot{-}3)$ [1, 5]. El objetivo de la contracción es que sea *exitosa* siempre que sea posible[36], lo cual motiva a $(\dot{-}4)$; si $x \in Cn(\varnothing)$ significa que por $(\dot{-}1)$ x no puede ser eliminado [1, 5]. La justificación para $(\dot{-}5)$ es que «el criterio de la economía informacional requiere que $A \dot{-} x$ sea un subconjunto 'grande' de A. (...) postulamos que *todas* las creencias en A sean *recuperadas* luego de contraer y expandir respecto a la misma creencia» [1]. En [1, 5, 3] $(\dot{-}5)$ se muestra con la condición de que debe ser el caso que $x \in A$, pues la inclusión se conserva en el otro caso: por $(\dot{-}3)$ cuando $x \notin A$ se tiene que $A \subseteq Cn((A \dot{-} x) \cup \{x\}) = Cn(A \cup \{x\})$. Pese a ser redundante al abarcar ambos casos, hemos elegido la versión de [4][37] por ser la utilizada en las primeras etapas de nuestra investigación. Finalmente, $(\dot{-}6)$ representa el criterio de que la operación de contracción no debe (y de hecho ninguna otra) estar guiada por el modelo de representación de creencias, sino por las creencias mismas [4].

De entre los postulados básicos de contracción, $(\dot{-}5)$, conocido como *postulado de recuperación*, es particularmente controversial [5]. A pesar de que es fácil ejemplificar escenarios donde carece de racionalidad (véase [2]), existen argumentos a favor, como el de GÃ€rdenfors presentado en [48].

Aparte del conjunto de postulados básicos, se tiene un conjunto *suplementario* que caracteriza contracciones respecto a conjunciones, con lo que se pretende atacar el caso de cambios de creencias (en este caso contracciones) *iterados* o *compuestos* [1, 4, 5].

Definición 4.7 ([5])**.** Los postulados *suplementarios* de racionalidad para la contracción en el modelo AGM son:

$$(A \dot{-} x) \cap (A \dot{-} y) \subseteq A \dot{-} (x \wedge y) \qquad (\dot{-}7)$$

$$\text{Si } x \notin A \dot{-} (x \wedge y)\text{, entonces } A \dot{-} (x \wedge y) \subseteq A \dot{-} x\,. \qquad (\dot{-}8)$$

Si una creencia no se suprime al contraer A por x ni al contraerla por y, no debería ser eliminada al contraer A por la conjunción $x \wedge y$; $(\dot{-}7)$ modela esto [1, 4]. Por otro lado, parece racional pedir que al contraer una conjunción $x \wedge y$ de un conjunto de creencias A, se debe dejar de creer en x, en y o en ambos[38]; tal es la motivación para $(\dot{-}8)$ [1, 4].

[36] Principio análogo al de *primacía de la nueva información.*

[37] Que es la de [19] pero sin la condición redundante «(...) siempre que $A = Cn(A)$».

[38] Curiosamente, en [1, 4] se utiliza la frase «se debe dejar de creer, o bien en x o bien en y», cuya contraparte literal en inglés denota una disyunción exclusiva, pero como se ve en [1, 5] es posible que ambos x y y sean eliminados.

4.3.3.2. Postulados AGM de revisión

Al igual que en el caso de los postulados de contracción, hemos construido nuestra propia versión de los postulados siguiendo la de [19] en la medida de lo posible. También, como en los de contracción, los postulados de revisión solamente se refieren a cn-teorías y deben leerse con el antecedente «si $A = Cn(A)$, entonces ...».

Definición 4.8. Los postulados *básicos* de racionalidad para la revisión en el modelo AGM son:

$$A\dot{+}x = Cn\left(A\dot{+}x\right) \tag{$\dot{+}1$}$$

$$x \in A\dot{+}x \tag{$\dot{+}2$}$$

$$A\dot{+}x \subseteq Cn\left(A \cup \{x\}\right) \tag{$\dot{+}3$}$$

$$\text{Si } \neg x \notin A\text{, entonces } Cn\left(A \cup \{x\}\right) \subseteq A\dot{+}x\,. \tag{$\dot{+}4$}$$

$$\text{Si } \neg x \notin Cn\left(\varnothing\right)\text{, entonces } A\dot{+}x \text{ es consistente.} \tag{$\dot{+}5$}$$

$$\text{Si } Cn\left(\{x\}\right) = Cn\left(\{y\}\right)\text{, entonces } A\dot{+}x = A\dot{+}y\,. \tag{$\dot{+}6$}$$

Hemos tomado ($\dot{+}1$) de [4] y su justificación es análoga a la de ($\dot{-}1$). El criterio de *primacía de la nueva información* es modelado por ($\dot{+}2$) [4]. Los postulados ($\dot{+}3$) y ($\dot{+}4$) pretenden decir que la expansión es un caso especial de la revisión ($A\dot{+}x = Cn\left(A \cup \{x\}\right)$ si $\neg x \notin A$); en particular ($\dot{+}3$) dice que en una revisión no debe incluirse más información de la que sería agregada por una expansión [1, 3]. En los primeros trabajos, como [19, 48], ($\dot{+}3$) y ($\dot{+}4$) son presentados como un solo postulado «si $\neg x \notin A$, entonces $Cn\left(A \cup \{x\}\right) = A\dot{+}x$». Originalmente, entre los postulados de revisión se encontraba la *identidad de Harper*, que eventualmente fue eliminada y presentada por separado (y de la cual hablamos en la siguiente subsección) [4]; el enunciar a ($\dot{+}3$) y ($\dot{+}4$) como dos postulados separados parece obedecer a esto. Con esta separación se tienen dos postulados que ilustran muy bien el hecho de que $A\dot{+}x \subseteq Cn\left(A \cup \{x\}\right)$ sin importar si $\neg x \in A$ o $\neg x \notin A$, y que $A\dot{+}x = Cn\left(A \cup \{x\}\right)$ si $\neg x \notin A$; pero es fácil ver que «si $\neg x \notin A$, entonces $Cn\left(A \cup \{x\}\right) = A\dot{+}x$» es una mejor forma de codificar ($\dot{+}3$) y ($\dot{+}4$), pues con tal postulado ($\dot{+}3$) es en redundante bajo lógica clásica: si $\neg x \in A$, por la definición de expansión (definición 4.4), $A\dot{+}x \subseteq Cn\left(A \cup \{x\}\right) = \mathcal{P}$. Pese a todo ello, en vista de que la explosividad[39] no es una característica de toda lógica, encontramos conveniente la formulación de ($\dot{+}3$) y ($\dot{+}4$).[40] En [1, 5] ($\dot{+}5$) aparece como $\neg x \in Cn\left(\varnothing\right)$ sii $A\dot{+}x = \mathcal{P}$,

[39]Como en la definición 3.25 en la página 48.

[40]En [5, 1] se dice que la separación responde a razones técnicas, pero parece que nunca se dice claramente cuáles.

cuya motivación es evitar a toda costa los conjuntos de creencias inconsistentes. Es posible ver que $(\dotplus 5)$ es redundante bajo lógica clásica en el caso de $\neg x \in Cn(\varnothing)$, pues por $(\dotplus 1)$ y $(\dotplus 2)$ se tiene que $A \dotplus x = \mathcal{P}$; por esta redundancia y en vista de la irracionalidad de identificar la inconsistencia de un conjunto de creencias A con la igualdad $A = \mathcal{P}$ en lógicas no clásicas, nuestra versión de $(\dotplus 5)$ es como la que aparece en [19] bajo el nombre de $(\dotplus 4)$. Finalmente, $(\dotplus 6)$ tiene la misma motivación que $(\dot{-}6)$ [4].

Definición 4.9 ([5])**.** Los postulados *suplementarios* de racionalidad para la revisión en el modelo AGM son:

$$A \dotplus (x \wedge y) \subseteq Cn((A \dotplus x) \cup \{y\}) \tag{$\dotplus 7$}$$

$$\text{Si } \neg y \notin A \dotplus x, \text{ entonces } Cn((A \dotplus x) \cup \{y\}) \subseteq A \dotplus (x \wedge y)\ . \tag{$\dotplus 8$}$$

Según [1], los postulados suplementarios de revisión son una generalización de $(\dotplus 3)$ y $(\dotplus 4)$. Su propósito, al igual que su contraparte para contracción, es indicar cómo realizar revisiones de creencias *iteradas* o *compuestas* [1, 5]. La idea es que si $A \dotplus x$ es una revisión de A y se desea cambiar $A \dotplus x$ agregando nuevas creencias, tal cambio debería hacerse usando expansiones de $A \dotplus x$ siempre que sea posible [1, 5].

4.3.3.3. Relación entre contracción y revisión

Luego de introducir los postulados AGM, es fácil ver que ambos conjuntos (de contracción y revisión) son independientes el uno del otro en el sentido de no se hace referencia a la revisión en los postulados de contracción y viceversa. Queda entonces la pregunta de si existe alguna relación entre ambas operaciones, y la respuesta es *sí.* [5, 1]

En 1977 Isaac Levi argumentó que las únicas dos formas legítimas de cambio son las expansiones y las contracciones, y que además las revisiones pueden analizarse como una secuencia compuesta por contracciones y expansiones [1]. Formalmente su tesis es que:

Definición 4.10. La *identidad de Levi* es:

$$A \dotplus x \stackrel{\text{def}}{=} Cn(A \dot{-} \neg x \cup \{x\}) \tag*{\textit{Levi}}$$

Interesantemente, con ella es posible demostrar que:

Teorema 4.11 ([1, 5])**.** *Sea un operador de contracción $\dot{-}$ que satisface los postulados AGM* $(\dot{-}1)$ *a* $(\dot{-}4)$ *y* $(\dot{-}6)$*, entonces la función de revisión $\dotplus$ obtenida utilizando Levi satisface* $(\dotplus 1)$ *a* $(\dotplus 6)$*.*

Demostración. Tomamos la demostración de [1], la cual se realiza sin apelar a propiedades de la lógica subyacente. Los postulados $(\dotplus 1)$ y $(\dotplus 2)$ se siguen del uso de Levi. $(\dotplus 3)$ es consecuencia de $(\div 2)$ y teoría de conjuntos. Para demostrar $(\dotplus 4)$ se asume que $\neg x \notin A$, entonces por $(\div 3)$ se sabe que $A \div \neg x = A$ y luego $Cn\,(A \cup \{x\}) = Cn\,(A \div \neg x \cup \{x\}) = A \dotplus x$. $(\dotplus 5)$ se sigue básicamente de $(\div 4)$, pues si $\neg x \notin Cn\,(\varnothing)$, entonces $\neg x \notin A \div \neg x$. $(\dotplus 6)$ se sigue directamente de $(\div 6)$.

□

Debe notarse que no es necesario utilizar $(\div 5)$ para que un operador de contracción de lugar a uno de revisión que cumpla todos los postulados básicos de revisión [1]. Una relación similar en el caso de los postulados suplementarios es:

Teorema 4.12. *Sea un operador de contracción* $\div$ *que satisface los postulados AGM* $(\div 1)$ *a* $(\div 6)$*. Entonces (i) si* $(\div 7)$ *es satisfecho por el operador de contracción,* $(\dotplus 7)$ *es satisfecho por la operación de revisión, y (ii) si* $(\div 8)$ *es satisfecho por la contracción,* $(\dotplus 8)$ *es satisfecho por la revisión.*

Demostración. La demostración del teorema 5.31 y la del teorema 5.37 corresponden a los incisos (i) y (ii) respectivamente; ambas son válidas bajo lógica clásica.

□

Muy curiosamente, en [1, 5] se enuncia el teorema anterior con las mismas condiciones que el teorema 4.11; sin embargo, en la misma demostración de [1] se utiliza $(\div 5)$. Este *error* queda evidenciado al ver las demostraciones de [19] y la versión del mismo teorema encontrado en [4].

En el mismo año que Levi argumentara la identidad que lleva su nombre, Harper propuso que $y \in A \div x$ si y sólo si $y \in A$ y $y \in A \dotplus \neg x$ [1]. La idea es que, como $A \dotplus \neg x$ es el mínimo cambio de A necesario para aceptar $\neg x$, contiene tanto de A que no implique a x como sea posible [1]. Formalmente esto significa:

Definición 4.13 (*). La *identidad de Harper* es:

$$A \div x \stackrel{\text{def}}{=} A \cap A \dotplus \neg x \qquad \textit{Harper}$$

Y de manera similar al caso de la identidad de Levi, se tienen dos teoremas que relacionan a la revisión con la contracción:

Teorema 4.14 ([1, 5])**.** *Sea un operador de revisión* $\dotplus$ *que satisface los postulados AGM* $(\dotplus 1)$ *a* $(\dotplus 6)$*, entonces la función de contracción* $\div$ *obtenida utilizando Harper satisface* $(\div 1)$ *a* $(\div 6)$*.*

Demostración. Es fácil demostrar por teoría de conjuntos y propiedades de un operador de consecuencia abstracto[41] que $(\dot{-}1)$ se cumple por Harper y $(\dot{+}1)$. Por Harper y teoría de conjuntos, $(\dot{-}2)$ se cumple. Asumiendo que $x \notin A$ y que en la lógica subyacente $\vdash \neg\neg x \rightarrow x$, por $(\dot{+}3)$, $(\dot{+}4)$ y Harper, se deduce $(\dot{-}3)$. Asumiendo que $x \notin Cn(\varnothing)$ y que en la lógica subyacente $\vdash \neg\neg x \rightarrow x$, por $(\dot{+}2)$, $(\dot{+}5)$ y Harper, se deduce $(\dot{-}4)$. Por $(\dot{-}3)$, $(\dot{-}5)$ se satisface cuando $x \notin A$; asumiendo $x \in A$, por Harper, teoría de conjuntos, definición de expansión y $(\dot{+}2)$, se cumple $(\dot{-}5)$. Por Harper y $(\dot{+}6)$, $(\dot{-}6)$ se satisface.

□

Teorema 4.15. *Sean las mismas condiciones que en el teorema 4.14. Entonces (i) si* $(\dot{+}7)$ *es satisfecho por el operador de revisión,* $(\dot{-}7)$ *es satisfecho por la operación de contracción, y (ii) si* $(\dot{+}8)$ *es satisfecho por la revisión,* $(\dot{-}8)$ *es satisfecho por la contracción.*

Demostración. La demostración se omite aquí; ésta puede encontrarse en [1].

□

4.3.3.4. Relación de la equivalencia clásica de los postulados y la investigación

En vista de que la equivalencia entre postulados básicos depende de mínimas condiciones por parte de la lógica subyacente, en esta investigación nos interesa descubrir las condiciones que permiten asegurar la equivalencia de los postulados suplementarios en lógicas no clásicas. De tal forma se avanza el conocimiento que permita obtener en un futuro operadores de cambio de creencias bajo lógicas no clásicas, pues como señala [5], la equivalencia de ambos conjuntos completos de postulados significa que un modelo constructivo para un operador proporcionaría automáticamente una construcción para el otro.

Cabe señalar que pese a que $\nvdash \neg\neg x \rightarrow x$ bajo Int y G_3, y la demostración del teorema 4.14 lo requiere, ello no constituye un teorema de imposibilidad para obtener la equivalencia del conjunto básico de postulados; no nos hemos concentrado en el proyecto en investigar tal equivalencia, y más aún, en el capítulo siguiente queda claro que no es necesaria para perseguir nuestros objetivos.

Finalmente, nos gustaría señalar que nuestra investigación hace un fuerte uso de la identidad de Levi y omite por completo a la de Harper. Entonces solamente hemos estudiado la validez del teorema 4.12 y de una variación del teorema 4.15 presentada en [19], donde se hacen las mismas asunciones que en el teorema 4.12. Esto pone en

[41] Véase definición 2.54 en la página 24.

claro el hecho de que el trabajo que hemos comenzado con este proyecto no es trivial en el sentido de que existen varias formas de considerar equivalentes ambos conjuntos de postulados.

4.3.3.5. Aptitud de los postulados AGM

Una pregunta que se intenta responder en [1] es si el conjunto de postulados es adecuado para modelar la revisión de creencias. Ahí, se presentan una serie de consecuencias que se desprenden de ellos[42] y se comparan con algunos postulados que parecen racionales a simple vista, pero que luego de eso son evidentemente contra-intuitivos. Se argumenta que debido a que los postulados AGM excluyen otras propuestas, a que son equivalentes entre ellos (como se dijo en la sección previa) y a que se demuestra (como veremos en la sección siguiente) que son equivalentes a ciertos modelos constructivos, estos constituyen un conjunto adecuado [1]. Sin embargo, parece que el problema queda abierto en la medida en que lo racional es en gran parte subjetivo. El cuestionar la racionalidad de los postulados es la base para algunas otras teorías epistémológicas, como es la semántica de mundos posibles [4].

4.3.4. Construcciones

A la par con la creación de los postulados, surgen modelos constructivos que pretenden satisfacerlos. Si bien existe una variedad de construcciones, aquí sólo mencionaremos brevemente tres propuestas de los integrantes del trío AGM para la contracción, las cuales se basan directamente en el principio de la *economía informacional* [1, 5, 4].

4.3.4.1. Maxichoice

La idea de la que se parte es que, al formar una contracción $A \dot{-} x$, el principio de la economía informacional requiere que la contracción contenga la mayor cantidad de creencias de A como sea posible de tal modo que no se derive x. Ello lleva a pensar inevitablemente en algún tipo de solución que tenga que ver con un *subconjunto maximal* de A que no implica a x [1]. Se dice que T *es un subconjunto maximal de* A *que falla en implicar* x si y sólo si (i) $T \subseteq A$, (ii) $x \notin Cn(T)$ y (iii) para todo S tal que $T \subset S \subseteq A$, $x \in Cn(S)$; la última condición implica que si T fuera a ser expandido con alguna fórmula de $A \setminus T$, tal expansión contendría x [5]. El problema con tal idea es que puede existir más de un subconjunto de este tipo [4].

[42] Algunas de las cuales también se encuentran en [5].

Para mitigar el problema de la multiplicidad de resultados, un primer intento es usar una *función de selección* ζ para elegir alguno de los conjuntos maximales de A que falla en implicar x [5]; si se identifica tal conjunto como $A \perp x$, éste cumple con: (i) $x \in Cn(\varnothing)$ si y sólo si $A \perp x = \varnothing$, y (ii) $x \notin A$ si y sólo si $A \perp x = \{A\}$ [4]. La función de selección elige a un elemento $\zeta(A \perp x)$ de $A \perp x$ siempre que $A \perp x \neq \varnothing$. Así, puede definirse lo que en [19] se llamó *maxichoice contraction function* (*contracción de elección individual*, de acuerdo a [4]) [5]:

$$A \dot{-} x \stackrel{\text{def}}{=} \begin{cases} \zeta(A \perp x) & \text{si } x \notin Cn(\varnothing) \\ A & \text{en cualquier otro caso} \end{cases} \qquad \textit{Maxichoice}$$

Es posible demostrar que cualquier contracción maxichoice es equivalente a el conjunto de postulados básicos de contracción y una condición adicional [5]. El problema es que las contracciones producidas por este tipo de función son tales que, para cualquier cn-teoría A y cualquier fórmula x, para toda $y \in \mathcal{P}$, o bien $(x \vee y) \in A \dot{-} x$ o $(x \vee \neg y) \in A \dot{-} x$; esto significa que, utilizando Levi, las funciones de revisión definidas a partir de dicho operador de contracción son *completas*[43][5, 48].

4.3.4.2. Full meet contraction

Otro intento para definir una función de contracción podría ser considerar el otro *extremo*, es decir, considerar que la contracción es la intersección de todos los elementos de $A \perp x$ [4, 5]:

$$A \dot{-} x \stackrel{\text{def}}{=} \begin{cases} \cap(A \perp x) & \text{si } x \notin Cn(\varnothing) \\ A & \text{en cualquier otro caso} \end{cases} \qquad \textit{Meet}$$

Tal tipo de función fue llamada *full meet contraction function* en [19] (*contracción de elección total*, de acuerdo a [4]). En contraste con las funciones maxichoice, full meet contraction produce resultados que son *demasiado pequeños*; mientras que de manera similar, toda función de contracción definida como una función full meet contraction es equivalente con los postulados básicos de contracción y una condición adicional [4, 5].

4.3.4.3. Partial meet contraction

Como un tercer intento, podría pensarse en intersectar solamente *algunos* de los elementos de $A \perp x$ [5]. En este caso, una *función de selección* ζ podría utilizarse para elegir un *subconjunto* $\zeta(A \perp x)$ no vacío de $A \perp x$, si $A \perp x$ no está vacío, o $\zeta(A \perp x) = A$

[43] Es decir, tales que para todo $y \in \mathcal{P}$, $y \in A \dotplus x$.

en caso de que $A \perp x = \varnothing$ [5]. Entonces la función de contracción conocida como *partial meet contraction function* (*contracción de elección parcial*, según [4]) queda como [5]:

$$A \dot{-} x \stackrel{\text{def}}{=} \cap\zeta\left(A \perp x\right) \qquad \textit{Partial meet}$$

Esta función es equivalente con el conjunto de postulados básicos [5]. Más aún, si la función de selección se construye utilizando una relación de orden *transitiva* y *reflexiva* sobre la unión de la familia de todos los conjuntos $A \perp x$ tales que $x \in A$ y $x \notin Cn\left(\varnothing\right)$, entonces se obtiene un tipo de función llamado *contracción de elección parcial transitivamente relacional*; tal función es equivalente con el conjunto completo de postulados de contracción.

4.4. Otros modelos

En esta sección mencionamos dos modelos o teorías epistemológicas diferentes a la AGM para proveer una visión más general del área de estudio.

4.4.1. Revisión de creencias para bases de creencias

Evidentemente, el modelar estados de creencias usando conjuntos de creencias (i.e. conjuntos de fórmulas de algún lenguaje cerradas bajo un operador de consecuencia lógico) es posible sólo de manera teórica: es imposible físicamente representar la cantidad infinita de creencias que existen en un conjunto de creencias. Más aún, es racional pensar que cuando un humano realiza revisiones o contracciones nunca lo hace sobre un conjunto de creencias A, sino sobre un subconjunto finito[44] T'_A de A tal que $Cn\left(T'_A\right) = A$, al cual se le llama *base para el conjunto de creencias A* [5].

El modelo AGM es interesante pese a estas limitaciones: de él se parte para crear variantes que permitan manejar bases de creencias en lugar de conjuntos de creencias [5]. Sin embargo, el considerar bases de creencias trae consigo su propia clase de problemas: es posible tener dos bases diferentes T' y S' tales que $Cn\left(T'\right) = Cn\left(S'\right)$; todavía más complicado es el hecho de que el el resultado de aplicar una función de cambio de creencias $*$ a cada base puede hacer que impliquen conjuntos de creencias diferentes, es decir $Cn\left(T' * x\right) \neq Cn\left(S' * x\right)$. Para responder a las necesidades propias de las bases, el enfoque utilizado por algunos trabajos es el de proponer una serie de postulados de racionalidad sobre los operadores de cambio de creencias [5]; en [3] pueden encontrarse tales postulados.

[44]Recordemos de la subsección 4.3.2 que el operador de consecuencia lógica se asume *finitario*.

4.4.2. Modelo de mundos posibles o esferas

En este modelo se utiliza una representación de creencias *sintáctica* [4]. Este modelo fue propuesto por Grove en 1986 [1]. En él se considera la existencia de un conjunto M_L de conjuntos maximales consistentes del lenguaje[45] [4]. Cada elemento de M_L es un *mundo posible* y cada creencia es representada por un conjunto de mundos posibles, o lo que es lo mismo, cada creencia es representada por un subconjunto de M_L [2]. Los conjuntos de creencias A se representan también por subconjuntos de mundos posibles $\parallel A \parallel = \{M \in M_L \mid A \subseteq M\}$ [4]. Usando tal representación de creencias, se define lo que es un *sistema de esferas centrado en* $\parallel A \parallel$, con el cual se pueden especificar funciones de contracción y revisión que son *completas* y *correctas* respecto al conjunto completo de postulados AGM correspondiente a cada una [1, 4].

4.5. Aplicaciones de la revisión de creencias

Para finalizar este capítulo, de entre las posibles aplicaciones de la teoría de revisión de creencias, mencionamos dos que son particularmente interesantes.

4.5.1. Semántica para proposiciones condicionales

Una de las aplicaciones originales que motivaron la creación de la teoría AGM es poder crear una *semántica* para los *condicionales contrafácticos* y los *abiertos* [4, 1]. Un condicional contrafáctico puede tener la forma «si x fuera el caso, entonces y sería el caso», mientras que uno abierto podría ser de la forma «si x es el caso, entonces y será el caso» [5, 1]. Según [5, 1, 4], el objetivo es utilizar un operador de revisión de creencias para descubrir si una proposición condicional es *aceptada* en cierto estado epistémico o no, ya que se supone que existe una fuerte relación entre las proposiciones condicionales y la revisión, la cual es:

$$x > y \in A \; sii \; y \in A \dotplus x$$

Tal equivalencia es una reformulación del *test de Ramsey para las proposiciones condicionales* [5]. Hay que notar que se asume que existe en el lenguaje objeto un símbolo $>$ y que no se trata de enunciar *condiciones de verdad* para la proposición $x > y$, sino condiciones de *aceptabilidad* en el estado de creencias representado por A. En [1] se dedica un capítulo completo a esta aplicación.

[45] Es decir, lo que comúnmente se conoce en cursos básicos de lógica como *teorías completas*.

4.5.2. Razonamiento no monótono

Una aplicación interesante de las teorías de revisión de creencias es la creación de *semánticas de razonamiento no monótono*. Las lógicas que hemos presentado en el capítulo 3 son *finitarias*, lo que significa, entre otras cosas, que cumplen con la condición de *monotonía* [46]. Evidentemente, si fuera a construirse una estructura matemática que incluyera un operador de revisión y/o uno de contracción y con algún tipo de sistema deductivo, estaríamos frente a una lógica *no monótona*. Otro acercamiento es el de [50] que, por ejemplo y entre otras cosas, indica la relación existente entre la revisión de creencias y algunas teorías de razonamiento no monótono, en particular con la *semántica preferencial*; se dice que es posible definir una relación de consecuencia lógica *preferencial* utilizando un operador de revisión definido para un modelo de coherencia de representación de creencias.

[46] Véase (2.2) en la página 24.