UNIVERSIDAD NACIONAL DEL SUR

TESIS DE DOCTORADO EN FILOSOFÍA

MONOTONÍA, NO MONOTONÍA Y SEMIMONOTONÍA.

Una perspectiva filosófica y metalógica.

DAMIÁN OLIVAREZ STAGNARO

BAHÍA BLANCA ARGENTINA

2021

PREFACIO

Esta tesis se presenta como parte de los requisitos para optar al grado Académico de Doctorado en Filosofía, de la Universidad Nacional del Sur y no ha sido presentada previamente para la obtención de otro título en esta Universidad u otra. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en el ámbito del Departamento de Humanidades durante el período comprendido entre el 6 de mayo de 2014 y el 5 de marzo de 2021, bajo la dirección del Dr. Gustavo Bodanza y la codirección de la Dra. Cristina Behnisch (UNCo).

Damián Olivarez Stagnaro

UNIVERSIDAD NACIONAL DEL SUR

Secretaría General de Posgrado y Educación Continua

La presente tesis ha sido aprobada el …./…./…… , mereciendo la calificación de ……(……………………)

Dedicado a mi familia y a la memoria de Martín Stagnaro (QEPD).

Agradezco profundamente al Dr. Gustavo Bodanza por guiarme con sapiencia y generosidad, y la Dra. Cristina Behnisch por su acompañamiento y valiosos aportes. A los Dres. y Dras. que me instruyeron con rigor científico durante mi trayecto académico, les extiendo mi agradecimiento y gratitud.

Asimismo, agradezco a la Universidad Pública en su conjunto y al Estado de la República Argentina, y manifiesto mi compromiso con la defensa de los valores tradicionales vinculados a la educación pública, laica y gratuita. Particularmente, expreso mi gratitud con la UNS y su personal docente, no docente y estudiantes, por recibirme cálidamente y permitirme cursar mis estudios superiores en tan prestigiosa institución. Extiendo similar agradecimiento a la UNCo, universidad en la que me gradué y donde desenvuelvo mis actividades profesionales en la actualidad, en un ambiente de fraternal camaradería. Finalmente, pero no menos importante, manifiesto al CONICET, por confiar en mi capacidad y darme la oportunidad de acceder a una beca de posgrado, mi eterna gratitud y orgullo de haber pertenecido al programa de becarios.

RESUMEN

Una de las nociones fundamentales en el ámbito de la lógica no monótona (LNM) es precisamente el propio concepto de *no monotonía*, que constituye, según K. Schlechta, una *propiedad de sistemas* con plenos derechos. Por ello, resulta curioso que, en un campo teórico tan prolífico en formalismos y simbolismos matemáticos, no se disponga de un axioma, teorema o definición de la propiedad de monotonía en términos estrictamente formales, al modo de las propiedades metamatemáticas y metalógicas usuales de los sistemas de lógica simbólica (en cualquiera de sus enfoques, clásicos o no clásicos). Incluso, en textos de metalógica sobre LNM se han investigado una serie de propiedades, denominadas *centrales*, que sí tienen aquellas características. Un lógico del calibre de D. Gabbay, creador de tales propiedades, las consideraba como sustitutos de una definición estricta de no monotonía. Sin embargo, una por una, las propiedades centrales han sido puestas en cuestión o directamente derribadas mediante contraejemplos, en la bibliografía del área. Ante este panorama, el problema de una definición estricta de la propiedad de no monotonía adquiere una relevancia creciente, desde el punto de vista filosófico (de filosofía de la lógica). En esta tesis, se aborda ese problema siguiendo un camino que comienza por un análisis minucioso de la propiedad de *monotonía* de la lógica estándar. Luego, se avanza en el estudio de algunas descripciones y definiciones informales y semiformales de no monotonía, que resultan representativas de las que suelen encontrarse en la bibliografía de LNM. Asimismo, se abordan algunas de las objeciones a las propiedades centrales. A partir de este análisis, se realiza una propuesta de definición estrictamente formal de la propiedad de no monotonía, de carácter universal y metamatemático, con recurso a herramientas de lógica modal. La misma puede interpretarse semánticamente mediante teoría de modelos en la tesitura del sistema modal

T, como así también mediante teoría de modelos preferenciales. Posteriormente, se profundiza en el enfoque preferencial de LNM, a fin de mostrar de qué modo pueden ser restauradas las propiedades centrales, como *condiciones de coherencia* de los sistemas basados en modelos preferenciales. Tales condiciones parecen conducir a la búsqueda de un conjunto de propiedades que reflejen una *estabilidad relativa* de los agentes racionales. A partir de esta idea, se configura una nueva propuesta que parte de la definición estricta de no monotonía, y de una *conjetura del razonador semimonótono*, que desemboca en la formulación de una nueva propiedad metamatemática asociada a no monotonía: la propiedad de *semimonotonía*. Esta última condición, que pretende reflejar el comportamiento de un tipo particular de razonadores que normalmente no eliminan conclusiones previamente obtenidas, es interpretada en términos modelo-teoréticos con recurso a la estructura de *ideales*, dentro de la teoría de filtros. Con la labor analítica previa y la formulación de definiciones metamatemáticas de las propiedades monotonía y semimonotonía, el autor espera contribuir a mejorar los fundamentos metalógicos y filosóficos de la lógica no monótona.

ABSTRACT

One of the fundamental notions in the field of non-monotonic logic (LNM) is precisely the concept of non-monotony itself, which, according to K. Schlechta, constitutes a *property of systems*, with full rights. For this reason, it is curious that, in a theoretical field so prolific in mathematical formalisms and symbolisms, there is no axiom, theorem or definition of the property of non-monotony in strictly formal terms, in the way of the usual metamathematical and metalogical properties of symbolic logic systems (in any of

its approaches, classical or non-classical). Even in metalogical texts on LNM, a set of properties have been investigated, called *core properties*, which do have those characteristics. A logician of the height of D. Gabbay, creator of such properties, regarded them as substitutes for a strict definition of non-monotony. However, one by one, the core properties have been called into question in the area literature, or outright knocked down by counterexamples. At this background, the problem of a strict definition of the property of non-monotony acquires increasing relevance, from the philosophical point of view (in philosophy of logic). In this thesis, this problem is approached by following a path that begins with a careful analysis of the *monotony* property of standard logic. Then, it advances in the study of some informal and semi-formal descriptions and definitions of non-monotony which are representative of those typically found in the literature of NML. In addition, some of the objections to the core properties are commented. From this analysis, a proposal for a strictly formal definition of the property of non-monotony is made, in a universal and meta-mathematical fashion, with recourse to modal logic tools. It can be interpreted semantically through model theory in the *T* modal system tessitura, as well as through preferential model theory. Subsequently, the preferential approach of NML is analyzed with relative depth, to show how the core properties can be restored, in the way of *coherence conditions* of systems based on preferential models. Such conditions seem to lead to the search for a set of properties that reflect a *relative stability* of rational agents. From this idea, a new proposal is configured that starts from the strict definition of non-monotony, and from a conjecture of *the semi-monotonic reasoner*, which leads to the formulation of a new metamathematical property associated with non-monotony: the property of *semi-monotony*. This last condition, which aims to reflect the behavior of a particular type of reasoners who normally do not eliminate previously obtained conclusions, is interpreted in model-theoretical terms with recourse to the structure of *ideals*, within the theory of filters. With the previous analytical work and the formulation of meta-mathematical definitions of the monotony and semi-monotony properties, the author hopes to contribute to improve the metalogical and philosophical foundations of NML.

ÍNDICE

INTRODUCCIÓN Y OBJETIVOS

Ser no monótono es una propiedad, no el nombre de una lógica única (K. Schlechta)

En el último tercio del siglo XX ha surgido y se ha consolidado una nueva escuela de lógica, impulsada por las investigaciones en inteligencia artificial (IA). Esta familia de teorías denominada de conjunto como "lógica no monótona" (LNM) aborda tres tópicos fundamentales[1] que delimitan su campo disciplinario: 1) el estudio formal del razonamiento derrotable o *de sentido común*, que resulta en el desarrollo de sistemas o *marcos lógicos*; 2) el rechazo de la propiedad de monotonía de la lógica estándar (LE), y la investigación sobre propiedades metalógicas aplicables a sistemas de LNM; 3) el estudio de problemas relativos a la complejidad y computabilidad de los sistemas resultantes. De estos tópicos, claramente el aspecto estrictamente sistemático o *disciplinario* es el que ha recibido más atención y donde se han obtenido los mejores resultados, mediante el desarrollo de sistemas que modelan el razonamiento derrotable o de sentido común. Por otra parte, el aspecto metadisciplinario, particularmente el estudio

[1] Cfr. (Strasser & Antonelli, 2019).

de propiedades metalógicas satisfacibles por los diversos sistemas, ha recibido comparativamente menos atención que el primer tópico, y aún debe afrontar una serie de problemas metalógicos y filosóficos cuyos intentos de resolución no siempre han llevado a buen puerto. Algunos de estos *problemas de fundamento* serán atacados a lo largo de esta tesis, mientras que el tercer aspecto o tópico principal de las investigaciones en LNM, relativo a temas de computabilidad, no será tenido en cuenta.

Las dificultades de fundamento que afronta la LNM se originan en situaciones de diversa índole. En primer lugar, existen objeciones de carácter eminentemente práctico, en tanto algunas de las *propiedades centrales* que se han postulado para estas lógicas en la bibliografía del área, tales como *monotonía cauta* (MC) y *monotonía racional* (MR), han sido cuestionadas mediante contraejemplos (algunos de los cuales serán presentados en el **CAPÍTULO 2**). Este tipo de problemas refleja cierta tensión entre las propias características de los sistemas de LNM, de donde emergen tales contraejemplos, y los intentos de fundamentación metalógica de los mismos. Otras dificultades conciernen a las propias bases metalógicas y filosóficas de la LNM. Uno de los primeros en abordar este tipo de problemáticas fue D. Israel en su (1980). En términos generales, ambos tipos de problemática pueden resumirse en un señalamiento de cierta *debilidad de las bases filosóficas* (de filosofía de la lógica) de la LNM. Pero si sólo se realiza una declamación al respecto sin profundizar en problemáticas concretas, con el mero fin de abogar por el reemplazo de la LNM en favor de algún sistema de lógica (filosófica) no clásica, la crítica se convierte simplemente en un ataque ideológico en el marco de una disputa de paradigmas. Un ejemplo de esta actitud puede encontrarse en (Estrada & Garcia, 2009), donde se propone lisa y llanamente, reemplazar la LNM por la lógica paraconsistente. Posteriormente en el **CAPITULO 4** se retornará sobre las relaciones entre ambos tipos de lógica.

En esta tesis, se intentará realizar un aporte desde la metalógica y la filosofía de la lógica para la resolución de algunos de los problemas de fundamento de la LNM. Más precisamente, se buscará vincular ambos niveles de abordaje de la lógica, el nivel filosófico y el nivel metalógico, en relación con problemas específicos, y con el objetivo de contribuir a la fundamentación de la LNM. El punto de partida es el señalamiento de una dificultad que, según la evidencia disponible, ha sido pasada por alto por completo en la bibliografía de referencia (véase la sección **BIBLIOGRAFÍA** al final). Como se hará patente en los capítulos siguientes (particularmente en **2** y **3**), se han postulado numerosas propiedades metalógicas para la LNM en términos estrictamente formales, ya sea bajo la forma de reglas estructurales con recurso a relaciones de consecuencia no monótona, o bien en notación conjuntista, utilizando operadores no monótonos.

Tomando en cuenta que *no monotonía* es una propiedad (metalógica), tal como reza el epígrafe de esta **INTRODUCCIÓN**, y la más importante de todas en LNM, la que delimita un vasto y diverso conjunto de sistemas lógicos, resulta sorprendente que no se registre en la bibliografía disponible una definición estrictamente formal de esta propiedad, bajo la forma de un axioma, teorema o expresión metamatemática en la tesitura de la *tradición metalógica* (clásica y no clásica). En general, la mayoría de los autores considerados se contenta con afirmar que los sistemas de LNM no satisfacen monotonía, y unos pocos ofrecen algún tipo de definición semi-formal de no monotonía, como es el caso de D. Gabbay en su clásico (1985). Este último autor, además, postula un conjunto de propiedades metalógicas (agrupadas bajo su *definición 18*, posteriormente designadas como las *propiedades centrales* de la LNM) como candidatas a reflejar, de conjunto, el significado de no monotonía. Sin embargo, estas no solo constituyen un *intento indirecto* de definir el concepto, sino que además todas ellas han sido objetadas (véase **CAPÍTULO 2**). Al parecer, esta situación de carencia de fundamentación metalógica directa de la propiedad de no monotonía ha sido ignorada tanto por los defensores como por los detractores de la LNM.

Por otra parte, argüiblemente esta situación tiene una raíz filosófica (de filosofía de la lógica). Específicamente, sumado a la ausencia de una *fórmula* que exprese la propiedad de no monotonía, tampoco existe en la bibliografía disponible de LNM una profundización de la crítica de la mismísima propiedad de *monotonía* de LE cuyas características dieron origen al proceso teórico de creación de la LNM en primer lugar. Al no profundizarse en esa crítica, el propio concepto de no monotonía aparece envuelto en un halo de oscuridad, resultando en una noción que no parece requerir mayores aclaraciones, pero de la cual sólo se tienen intuiciones más o menos imprecisas. La hipótesis que guía la investigación presentada en el **CAPÍTULO 1** de esta tesis es que resulta inviable formular una crítica filosófica y metalógicamente sólida de la noción de monotonía, sin profundizar primero en su significado estricto, y en todos sus aspectos, incluyendo las principales formulaciones que ha recibido esta propiedad en los textos fundacionales de la década del '30 del siglo XX, como los de Tarski (1969a), (1969b) y Gentzen (1969).

Asimismo, al profundizar la indagación, se hará manifiesto que monotonía es una propiedad sumamente compleja, que se ramifica en una diversidad de temáticas filosóficas y metalógicas, desde las nociones de *consistencia*, *explosión lógica*, o el concepto mismo de *sistema lógico*, hasta el problema de la relación entre *inferencias e información*, entre otros tópicos. Sobre estos últimos temas sólo se realizarán, a lo largo del capítulo, digresiones y comentarios críticos (relativamente) breves, de manera subsidiaria, siempre y cuando resulte necesario para los fines de la comprensión completa y rigurosa del significado del concepto de monotonía. Sobre el debate relativo a la relación entre inferencias e información, en el **APÉNDICE 1** se presentará una discusión con mayor profundidad. Por otra parte, el tópico de la consistencia lógica exigirá un tratamiento recurrente, aunque accesorio, a través de todos los capítulos de esta tesis.

Sobre el *suelo firme* ganado en el **CAPÍTULO 1**, se procederá en el **CAPÍTULO 2** a *remontar el hilo de Ariadna* de la crítica de la propiedad de monotonía hasta sus aspectos más básicos, a fin de arrojar luz sobre la idea misma de no monotonía tanto desde un punto de vista filosófico como desde un punto de vista metalógico. El objetivo es generar una base teórica que permita explorar la posibilidad de formular esta propiedad en términos metamatemáticos. El autor espera que un formalismo de ese tipo pueda ayudar a fortalecer las bases filosóficas y metalógicas de la LNM, e incluso contribuir a clarificar y resolver otras problemáticas de fundamento que afectan a esta familia de lógicas. Particularmente, se presentarán las *propiedades centrales* tradicionales de la LNM utilizando la notación presente en (Makinson, 1994), para las cuales se consignarán algunos de los contraejemplos y objeciones más relevantes que pueden encontrarse en la bibliografía (y otras originales, como la crítica de la propiedad de *reflexividad/inclusión*). El resultado fundamental de este capítulo será la presentación de dos expresiones estrictamente formales de la propiedad de no monotonía, originales del autor (**2.3.1** y **2.3.2**), con recurso a herramientas de lógica modal. Quedará pendiente para el siguiente **CAPÍTULO 3** la reflexión respecto a las propiedades centrales: particularmente, se mostrará de qué modo pueden restaurarse las mismas, dentro de ciertos límites, con base en la tradición semántica de la LNM.

Para ese fin, se ofrecerá una exposición sintetizada de los fundamentos del enfoque preferencial de la LNM, particularmente en las versiones de (Kraus et al., 1990) y (Schlechta, 2004). Este **CAPÍTULO 3** resultará, sin duda, el más abstruso debido a la complejidad matemática de la propuesta de los autores, donde los límites entre lógica, teoría de modelos y álgebra se vuelven difusos (al estilo de la vieja escuela metamatemática de Tarski). Asimismo, en términos filosóficos se afrontarán importantes y complicadas aporías relativas a la interpretación del símbolo de *relación de consecuencia no monótona* y la distinción entre lenguaje objeto y metalenguaje en la propuesta de los autores citados. Estas últimas resultan temáticas subsidiarias a los fines de esta tesis, pero no siempre será posible darles un tratamiento secundario o accesorio debido al alto impacto que pueden tener en el eje más general del capítulo: la exploración de vías para la restauración de las propiedades centrales de la LNM (para una ampliación de la digresión sobre las mencionadas aporías véase **APENDICE 2**). Adicionalmente, la exposición de los fundamentos de la propuesta de K. Schlechta (más bien, de un recorte temático sintetizado de la misma, según los fines de esta tesis) resultará de utilidad para la propuesta, en el **CAPITULO 4**, de una variante original de la propiedad de no monotonía bajo la *conjetura del razonador semimonótono*, que tomará forma en la novedosa expresión **4.2.1** (propiedad de semimonotonía), fundada semánticamente mediante *teoría de filtros*.

En suma, con las expresiones **2.3.1** y **2.3.2** el autor cree ofrecer una alternativa clara para contribuir a la resolución del problema de la ausencia de una definición metamatemática de la propiedad de no monotonía. Con **4.2.1**, se ofrece una nueva perspectiva para enfocar un tipo particular de agentes racionales: aquellos que, aun siendo no monótonos, normalmente no eliminan conclusiones previamente obtenidas. En el camino, muchos problemas y cuestiones quedan planteados rumbo a futuras indagaciones, en tanto la tarea de fundamentación metalógico-filosófica de la LNM se presenta como descomunal, teniendo en cuenta la enorme complejidad y diversidad de teorías y sistemas dentro del área. El objetivo planteado aquí es mucho más humilde: aportar a la resolución de *algunos problemas de fundamento* de la LNM. Con la labor analítica presentada a lo largo de esta tesis y las definiciones mencionadas, el autor espera haber logrado su cometido.

CAPÍTULO 1 – MONOTONÍA

INTRODUCCIÓN

Dado el panorama actual de la LNM descripto en **INTRODUCCIÓN Y OBJETIVOS**, cabe preguntarse con la mayor seriedad y profundidad posibles: ¿qué significa *monotonía*? (en LE). La motivación de este capítulo es la búsqueda de una respuesta a esa pregunta.

Para ese objetivo, en primer lugar, se procederá a presentar algunas de las definiciones informales de "monotonía". Las que se consignan aquí resultan lo suficientemente representativas de las descripciones informales de tal propiedad en el ámbito de la lógica, y constituyen un buen medio para aproximarse al asunto. Posteriormente, se procederá a abordar las representaciones formales de monotonía. Existen dos modos principales de representación formal de este concepto, que podrían etiquetarse como "representación estilo Gentzen" y "representación estilo Tarski". La inmensa mayoría de las exposiciones formales de monotonía que se encuentran en artículos y manuales al uso se basan en alguna de los estilos de presentación mencionados. Aquí se expondrán ambas formulaciones, la tarskiana y la gentzeniana, junto a una breve descripción del marco teórico en el que se enrolan, como así también de conceptos asociados, precisiones teóricas y consideraciones históricas que de un modo u otro aportan a la comprensión de la noción de monotonía.

La formulación tarskiana del concepto en cuestión es anterior a la gentzeniana, en tanto fue presentada por primera vez en el año 1928, en ocasión de un mitin de la Sociedad de Matemáticos Polacos. La disertación de Tarski en esa convención fue publicada dos años más tarde, en 1930, bajo el título *Sobre algunos conceptos fundamentales de las metamatemáticas* (Tarski, 1969b). Por otra parte, la formulación de Gentzen se remonta al artículo del año 1934 titulado *Investigaciones sobre la deducción lógica* (Gentzen, 1969), en el cual el autor publica los resultados de su tesis doctoral de 1933. Es destacable que ambos autores al momento de presentar los resultados de sus investigaciones no habían superado los 30 años: en 1928 Tarski tenía sólo 26 años, mientras que Gentzen en 1933 apenas había cumplido 24 años. Esto resulta sorprendente teniendo en cuenta el profundo impacto que han generado en la lógica contemporánea las contribuciones de juventud de ambos autores. Aunque sus investigaciones fueron contemporáneas, no existen certezas respecto a si cada autor estaba al tanto de los resultados de su colega, o bien si fueron completamente independientes.

Volviendo a la descripción del capítulo, luego de la presentación de las expresiones formales de monotonía se pondrán en juego estas últimas con las definiciones informales del mismo concepto. En primer lugar, se analizará la relación entre monotonía, razonamientos y sistemas. En segundo lugar, se abordará el complejo asunto de la relación entre monotonía e información, tópico que se enmarca en la problemática más general de la relación entre inferencias e información. Por último, en base a los resultados de las secciones anteriores, se ofrecerá una definición informal alternativa del concepto de monotonía, que se espera que sirva a los fines de responder, al menos de modo aproximado, la pregunta que motiva el capítulo.

1.1 DEFINICIONES INFORMALES

En una aproximación inicial, se presentarán algunas definiciones informales de la propiedad de monotonía. En la bibliografía considerada, estas no aparecen acompañadas de otros formalismos asociados, por lo que los autores al parecer las consideran suficientes y adecuadas para explayar el significado definido:

1.1.1 "Dado un argumento deductivo, si se agrega nueva información o nuevos hechos (a las premisas iniciales), no se pierden conclusiones" (Copi, 2014:26).

1.1.2 "Esta propiedad dice que las consecuencias lógicas de un conjunto de hipótesis crece monotónicamente en relación con las hipótesis, en otras palabras, que agregar nueva información A a las hipótesis Γ no puede resultar en la retractación de ninguna consecuencia B de Γ" (Thomason, 2011:782).

1.1.3 "Las lógicas deductivas tradicionales son siempre monótonas: la adición de nuevas premisas (axiomas) nunca puede invalidar viejas conclusiones (teoremas)" (Lukaszewicz, 1990:77).

Estas definiciones, en un primer análisis, plantean la existencia de una relación crucial entre la propiedad de monotonía y la idea de información nueva[2]. Un segundo aspecto importante que se desprende de las mismas es que en algunas se afirma que monotonía es una propiedad de sistemas lógicos (**1.1.2**, **1.1.3**), y en otra de razonamientos (**1.1.1**). Posteriormente se abordará este problema, quedando abierto por el momento. Así,

[2] También hechos, axiomas, hipótesis o premisas nuevas.

parafraseando las citas anteriores, es posible obtener un esquema de definición informal que funciona como aproximación inicial al concepto de monotonía:

1.1.4: *La propiedad de monotonía [de un sistema/de un razonamiento] es una condición que garantiza que, dado un conjunto de premisas, ante el agregado de información nueva a las mismas no se pierden conclusiones previamente obtenidas.*

1.2 MONOTONÍA SEGÚN TARSKI

A continuación, se realizará una exposición sobre la formulación tarskiana del concepto de monotonía, mencionando a su vez otros conceptos asociados que contribuyen a profundizar la comprensión de este última. El punto de partida de Tarski en los textos de 1930 es la noción de "sistema deductivo", la cual se aborda desde un punto de vista metadisciplinario, en el sentido de una reflexión que toma como objeto un *sistema de signos* (que constituye a su vez una disciplina teórica sobre otro objeto de estudio). Esta reflexión metadisciplinaria es denominada por Tarski como "metamatemática" con preferencia sobre otras denominaciones que en el marco de su pensamiento son sinónimas: "metodología de las ciencias deductivas" y "metalógica". La idea básica es que los resultados de la metamatemática valen universalmente para cualquier sistema deductivo, y por lo tanto incluyen a la lógica y la matemática. Esto tiene sentido en el marco de la propuesta global de Tarski, ya que en el corpus teórico que este autor ha creado a lo largo de los años, las fronteras entre lógica, semántica, teoría de modelos, álgebra, geometría y teoría de conjuntos resultan artificiales[3].

Cabe destacar que la metamatemática no sólo *estudia* teorías formales deductivas, sino que también *construye* su objeto de estudio[4]. En otras palabras, se trata de una metadisciplina a la vez descriptiva y normativa. Ahora bien, en cuanto al aspecto normativo, cabe preguntarse, ¿cómo se construye una disciplina deductiva? La respuesta es simple: mediante el *método deductivo*. Idealmente, una disciplina científica debería permitir la definición de cualquier expresión o término perteneciente a la misma, y la justificación de todas sus aserciones. Sin embargo, esto no es posible en la práctica, ya que la definición de cualquier expresión se efectiviza mediante la utilización de otras expresiones, y estas a su vez requieren nuevas definiciones, lo cual conduce a una *regresión ad infinitum* hacia el significado. Por ello, en toda disciplina deductiva debe existir un pequeño grupo de términos que no requieran definición, por ser inmediatamente comprensibles, los cuales se denominan *términos primitivos*. Estos permiten definir, en última instancia, todas las otras expresiones del sistema, evitando así la regresión al infinito.

En cuanto a la justificación, se procede de un modo similar: se toman algunas de las sentencias del sistema como axiomas, aquellas que no requieren otros fundamentos por tener la característica de ser autoevidentes. Estos axiomas permiten establecer la verdad de cualquier sentencia verdadera del sistema, por medio de la demostración y con el auxilio de definiciones y de otras sentencias justificadas previamente. Ahora bien, en virtud de la construcción de una disciplina deductiva según la receta que ofrece la metamatemática tarskiana, ocurre que tal disciplina adquiere ciertas características

[3] Cfr. (K. Simmons, 2009: 511).
[4] Cfr. (K. Simmons, 2009:512).

especiales, las cuales son descriptas por la propia metamatemática. Estas investigaciones pueden darse en dos niveles: un nivel que se enfoca en *metamatemáticas especiales* (metadisciplinas relativas a una ciencia en particular), y un nivel de máxima abstracción y generalidad, donde los resultados de la metamatemática valen para toda disciplina deductiva[5]. Los escritos de Tarski que serán tomados como referencia a continuación, (Tarski, 1969a) y (1969b), se ubican en el nivel de nivel de máxima generalidad en el primer caso, y en el nivel de una metadisciplina cuyo objeto son los sistemas de LE, en el segundo. El concepto de monotonía es formulado en ambos niveles, por lo cual a los fines de dilucidar el significado de este no se requiere establecer mayores distinciones entre los mencionados niveles de abstracción metadisciplinaria.

La exposición subsiguiente se diferencia de los textos de referencia de A. Tarski en términos de notación, en tanto en estos últimos se utiliza una notación de teoría de conjuntos basada en (Fréchet, 1924), pero aquí se utilizará la notación moderna que puede encontrarse, por ejemplo, en (Rosen, 2011). Junto a la notación moderna, se utilizarán los siguientes símbolos: partiendo del conjunto universal de sentencias S, para sentencias cualesquiera de S se utilizará la variable x; y para subconjuntos de S las letras $\Delta, \Gamma, \Theta, \Phi, \Omega,\ \Psi$. Asimismo, se han reemplazado las apariciones de la frase "si…, entonces…" por la implicación clásica "→". Si bien en los textos citados sólo se utilizan símbolos de teoría de conjuntos (junto a lenguaje natural), en el contexto de este trabajo la opción de hacer uso del condicional material en la formulación de los axiomas y teoremas no genera ninguna dificultad. No obstante, en el **CAPÍTULO 2** se hará patente que este tipo de implicación puede resultar problemática al momento de intentar definir formalmente el concepto de "no monotonía".

[5] Cfr. (K. Simmons, 2009).

Dicho esto, lo primero a tener en cuenta es que existen en la perspectiva tarskiana dos nociones básicas asociadas a la idea de *sistema*: las nociones de *sentencia* y *consecuencia*. Las sentencias, más precisamente las sentencias significativas, constituyen un concepto primitivo (no definido) en el marco de los estudios metamatemáticos, cuyo significado concreto debe ser especificado por cada disciplina formalizada. Dado S, el conjunto de todas las sentencias significativas[6], si se aplican reglas de inferencia (más bien, si se aplica el operador Cn que representa el conjunto de todas las reglas de inferencia de un sistema deductivo) a subconjuntos de S, se obtienen conjuntos de consecuencias[7]. Es decir, para un conjunto cualquiera de sentencias $\Delta \subseteq S$, al aplicar el operador Cn al mismo se obtiene otro conjunto de sentencias, $Cn(\Delta) \subseteq S$, que es el conjunto de todas las consecuencias del conjunto Δ. Este conjunto $Cn(\Delta)$, denominado también *conjunto consecuencia* o *clausura deductiva* se define como "la intersección de todos los conjuntos que contienen al conjunto Δ y que son cerrados bajo las reglas de inferencia consideradas" (Tarski, 1969a:63). Un conjunto cerrado bajo las operaciones consideradas es aquel en el cual la aplicación de las operaciones (cada una de ellas) a elementos del conjunto, siempre devuelve un elemento del conjunto. En otras palabras, el conjunto $Cn(\Delta)$ es la intersección de los puntos fijos del operador Cn cuando se aplica a los superconjuntos de Δ. Esta última definición podría expresarse formalmente como sigue[8]:

1.2.1 Conjunto consecuencia: $Cn(\Delta) = \bigcap \Gamma_i : (S \supseteq \Gamma_i \supseteq \Delta) \wedge (Cn(\Gamma_i) = \Gamma_i)$

[6] Como el concepto de "sentencia significativa" no está definido, tampoco lo está el conjunto S. Sólo es posible aproximarse a sus significados, mediante el estudio de los axiomas o principios de los sistemas deductivos.

[7] Una *consecuencia* es también una sentencia de S. En el marco de la propuesta de Tarski se trata de un concepto primitivo, no definido

[8] Tanto esta definición como su expresión formal (**1.2.1**) no aparecen en los textos de Tarski, son originales del autor de esta tesis.

Cabe destacar que la clausura deductiva de un conjunto inconsistente será igual al conjunto S, fenómeno lógico conocido como *explosión lógica*. Volviendo a la presentación de Tarski, los siguientes cuatro axiomas pueden entenderse como principios o propiedades de los sistemas deductivos, los cuales a su vez aportan a una aproximación genérica a los conceptos primitivos de sentencia y consecuencia. En principio los axiomas (y sus teoremas) se presentan con referencia a un conjunto de sentencias universal S, pero según el teorema de relativización[9] se cumplen de igual modo para dominios de sentencias (subconjuntos de S).

1.2.2 Cardinalidad del conjunto S: $|S| \leq \aleph_0$

1.2.3 Inclusión: $(\Delta \subseteq S) \rightarrow (\Delta \subseteq Cn(\Delta)) \subseteq S)$

1.2.4 Idempotencia: $(\Delta \subseteq S) \rightarrow (Cn(Cn(\Delta)) = Cn(\Delta))$

1.2.5 Compacidad: $(\Delta \subseteq S) \rightarrow (Cn(\Delta) = \bigcup Cn(\Phi_i)) \colon \Phi_i \subseteq \Delta \wedge |\Phi_i| < \aleph_0$

El axioma **1.2.2** expresa que el conjunto de todas las sentencias es finito o bien infinito numerable. **1.2.3** afirma que las sentencias de Δ están contenidas en las consecuencias de Δ, y que las consecuencias de Δ son sentencias. El axioma **1.2.4** afirma que el conjunto $Cn(\Delta)$ es un punto fijo del operador Cn cuando se aplica al conjunto $Cn(\Delta)$. Finalmente, el axioma **1.2.5** significa, en palabras de Tarski, que "toda consecuencia de un conjunto de sentencias es una consecuencia de un subconjunto finito de ese conjunto y viceversa" (Tarski, 1969a:64). La importancia de este último axioma para las ciencias deductivas queda de manifiesto en el hecho de que del mismo se deriva el teorema presentado abajo que es la expresión formal del concepto de monotonía en la formulación tarskiana[10]:

[9] Cfr. Th. 6 en (Tarski, 1969a:67).
[10] Cfr. Th. 1 en (Tarski, 1969c:32).

1.2.6 Monotonía (a): $(\Delta \subseteq \Gamma \subseteq S) \rightarrow (Cn(\Delta) \subseteq Cn(\Gamma))$

Para profundizar el concepto, se presenta a continuación un teorema que es consecuencia directa de **1.2.3**, **1.2.4** y **1.2.6**, y que es una expresión análoga a **1.2.6**[11]:

1.2.7 Monotonía (b): $((\Delta \cup \Theta) \subseteq S) \rightarrow ((\Delta \subseteq Cn(\Theta)) \leftrightarrow (Cn(\Delta) \subseteq Cn(\Theta)))$

Este teorema afirma que todas las sentencias de un conjunto Δ pertenecen al conjunto consecuencia de otro conjunto Θ syss todas las consecuencias del conjunto Δ pertenecen al conjunto consecuencia de Θ. A continuación, se presentan otros dos teoremas equivalentes a **1.2.6**[12]:

1.2.8 Monotonía (c): $(\Delta_i \subseteq S) \rightarrow (\bigcup Cn(\Delta_i) \subseteq Cn(\bigcup \Delta_i))$

1.2.8 significa que la unión de la colección de conjuntos de consecuencias de una colección de conjuntos Δ_i es un subconjunto del conjunto de consecuencias de la unión de tal colección de conjuntos Δ_i.

1.2.9 Monotonía (d): $(\Delta_i \subseteq S) \rightarrow (Cn(\bigcap \Delta_i) \subseteq \bigcap Cn(\Delta_i))$

En **1.2.9** se afirma que el conjunto de consecuencias de la intersección de una colección de conjuntos Δ_i es un subconjunto de la intersección de la colección de conjuntos de consecuencias de una colección de conjuntos Δ_i. Con el objetivo de precisar aún más el significado de **1.2.6**, se formularán a continuación dos corolarios de tal teorema[13]:

1.2.10 Cor. (a): $(\Delta \subseteq \Gamma \subseteq Cn(\Delta) \subseteq S) \rightarrow (Cn(\Delta) = Cn(\Gamma))$

---

[11] Cfr. (Tarski, 1969a:64).
[12] Cfr. (Tarski, 1969a:65).
[13] Estos corolarios no aparecen en los textos de Tarski, son originales del autor de esta tesis.

Prueba: 1) Sea $\Delta \subseteq \Gamma \subseteq Cn(\Delta)$ [supuesto de prueba directa]; 2) $\Gamma \subseteq Cn(\Delta)$ [se sigue de 1)]; 3) $Cn(\Gamma) \subseteq Cn(Cn(\Delta))$ [por teorema de monotonía]; 4) $Cn(\Gamma) \subseteq Cn(\Delta)$ [por idempotencia en 3)]; 5) $\Delta \subseteq \Gamma$ [se sigue de 1)]; 6) $Cn(\Delta) \subseteq Cn(\Gamma)$ [por teorema de monotonía]; 7) $Cn(\Delta) = Cn(\Gamma)$ [se sigue de pasos 4) y 6) por definición de igualdad de conjuntos].

Coloquialmente, cuando se agregan (a un conjunto inicial Δ) sentencias que pertenecen al conjunto formado por la intersección de todos los conjuntos de consecuencias del conjunto inicial de sentencias, en el conjunto de consecuencias del superconjunto de Δ se conservan todas las consecuencias que estaban presentes en el conjunto de consecuencias del conjunto inicial Δ, y no se obtienen consecuencias que no estuvieran ya presentes en el conjunto de consecuencias de Δ. El corolario (a) representa, por decirlo de algún modo, el *núcleo interior* del concepto de monotonía, en tanto expresa la estabilidad de una base de conocimiento (consistente) incrementada mediante *sus propias* consecuencias deductivas[14].

1.2.11 Cor. (b): $((\Delta \subset \Gamma \subseteq S) \wedge ((\Gamma - \Delta) \cap Cn(\Delta) = \emptyset)) \rightarrow (Cn(\Delta) \subset Cn(\Gamma))$

Prueba: 1) Sea $\Delta \subset \Gamma$ y $(\Gamma - \Delta) \cap Cn(\Delta) = \emptyset$ [supuesto de prueba directa]; 2) $Cn(\Delta) \subseteq Cn(\Gamma)$ [por teorema de monotonía aplicado al primer miembro de la conjunción en 1)]; 3) Dado $a \in (\Gamma - \Delta)$ [donde a es una sentencia cualquiera de S]; 4) $a \notin Cn(\Delta)$ [se sigue de 3) y del segundo miembro de la conjunción de 1)]; 5) $a \in \Gamma$ [se sigue de 3)]; 6) $a \in Cn(\Gamma)$ [por inclusión en 5)]; 7) $a \notin Cn(\Delta) \wedge a \in Cn(\Gamma)$ [se sigue de pasos 4) y 6)]; 8) $Cn(\Delta) \subset Cn(\Gamma)$ [se sigue de los pasos 2) y 7) por definición de subconjunto propio].

[14] Este corolario se asemeja a la propiedad de "monotonía cauta", atribuida a sistemas de LNM. La diferencia principal entre el corolario (a) y monotonía cauta es que esta última se asocia a un operador C que representa un cálculo no monótono. Véase **CAPÍTULO 2**.

Esta expresión afirma, en primer lugar, que cuando se agregan (a un conjunto inicial Δ) sentencias que no pertenecen al conjunto formado por la intersección de todos los conjuntos de consecuencias del conjunto inicial de sentencias, en el conjunto de consecuencias del superconjunto de Δ no se pierden consecuencias respecto a las que estaban presentes en el conjunto de consecuencias de Δ. En segundo lugar, que se obtienen consecuencias que no estaban presentes en el conjunto de consecuencias del conjunto inicial Δ, pues el conjunto de consecuencias del superconjunto propio de Δ es un superconjunto propio del conjunto de consecuencias de Δ. El corolario (b) representa el *núcleo exterior* de la propiedad de monotonía, en tanto expresa la estabilidad de una base de conocimiento (consistente) incrementada mediante sentencias *externas* relativamente a las consecuencias deductivas de la base de conocimiento inicial. Por supuesto, tanto **1.2.10** como **1.2.11** se cumplen también si la base de conocimiento es inconsistente, pero en este caso debido a la explosión lógica se difumina cualquier frontera metafórica entre un núcleo interior y uno exterior de la propiedad de monotonía.

De manera resumida, y a modo de recapitulación, el teorema de monotonía **1.2.6**, teoremas equivalentes y corolarios significan que "la operación Cn en el dominio de conjuntos de sentencias es monótona" (Tarski, 1969a:65). En otras palabras, que el conjunto consecuencia o clausura deductiva de cualquier superconjunto de un conjunto dado Δ contiene todas las consecuencias de Δ. Por último, cabe resaltar que en (Tarski, 1969b) se agrega un grupo de axiomas especiales que valen para sistemas deductivos clásicos entre los cuales se encuentra el siguiente axioma conocido como *ley de no contradicción* o de explosión lógica, según el cual *todo se sigue de una contradicción* (ex contradictione quodlibet)[15]. O también, como se mencionó antes, la clausura deductiva de

[15] Cfr. (Tarski, 1969b:32).

un conjunto inconsistente es igual al conjunto S. Este fenómeno lógico descubierto por Duns Scotto en el siglo XIII, origina una de las principales críticas a la LE en tanto este tipo de lógica carece de las herramientas adecuadas para manipular contradicciones. De esa crítica han surgido algunas de las lógicas no clásicas más importantes, como la lógica intuicionista, la lógica paraconsistente y en parte también la propia LNM. Por otra parte, la ley de no contradicción permite construir el *principio de consistencia* de la LE[16], según el cual (en un sistema deductivo) si un conjunto de sentencias es consistente, entonces su conjunto de consecuencias es consistente. La conversa también se cumple: si un conjunto de sentencias es inconsistente, su clausura deductiva también lo es. En otras palabras, *un sistema deductivo no genera inconsistencias*, ya que es imposible derivar contradicciones a partir de un conjunto de sentencias consistente, y, por otra parte, si un conjunto de consecuencias es consistente, necesariamente su conjunto de premisas lo es. Esta es una característica extremadamente importante de los sistemas deductivos. En LNM, existen sistemas que pueden generar inconsistencias a partir de conjuntos de sentencias consistentes, como algunos sistemas de lógicas default[17] y sistemas preferenciales[18]. Posteriormente se retornará, de manera subsidiaria, sobre el tópico de consistencia en las secciones y capítulos subsiguientes. A continuación, se presentarán formulaciones propias de la ley de explosión lógica y el principio de consistencia que si bien no coinciden exactamente el original de (Tarski, 1969b), sí respetan su espíritu y permiten desplegar algunas consecuencias que parecen haber sido pasadas por alto por este último autor, en relación a la consistencia del conjunto S. Dado un conjunto inconsistente $\perp = \{x, \ldots, \neg x\}$ con $x, \neg x \in S$, y $\Delta \subseteq S$ se cumplen:

[16] Cfr. (Tarski, 1969b:34).
[17] Cfr. (Schlechta, 2004:47).
[18] Cfr. (Makinson, 1994:81).

1.2.12 Ley de explosión lógica: $Cn(\bot) = S$

1.2.13 Principio de consistencia: $(\Delta \not\supseteq \bot) \leftrightarrow (Cn(\Delta) \neq S)$

La expresión **1.2.12** no requiere mayores precisiones, establece simplemente que la clausura deductiva de un conjunto de sentencias inconsistente contendrá todas las sentencias de S. Por otra parte, la expresión **1.2.13** admite la situación $\Delta = S$, que genera un escenario interesante: dado $(Cn(\Delta) = S) \rightarrow (\Delta \supseteq \bot)$, que es una consecuencia directa de **1.2.13** (utilizando sólo reglas de LE), al introducir aquella igualdad se obtiene la expresión condicional $(Cn(S) = S) \rightarrow (S \supseteq \bot)$. De modo que, si fuera verdadero $Cn(S) = S$ esto implicaría $S \supseteq \bot$, es decir, que el conjunto S de todas las sentencias significativas sería inconsistente. En efecto, por definición **1.4.2** (véase en la sección **4** de este capítulo) es verdadero $Cn(S) = S$ y es un *sistema deductivo* en sentido estricto. Esto implica que S es inconsistente. Este resultado no es mencionado explícitamente por Tarski en su (1969a) ni en (1969b). Intuitivamente, resulta razonable que el conjunto de todas las sentencias significativas contenga para cada sentencia su negación, no obstante, dadas ciertas características de la propuesta de Tarski que serán abordadas posteriormente, el sistema $Cn(S) = S$ puede generar algunas dificultades.

1.3 MONOTONÍA SEGÚN GENTZEN

En el texto de referencia (Gentzen, 1969) el autor presenta dos tipos de sistemas lógicos, uno de los cuales consiste en un *cálculo de deducción natural*[19], mientras que el segundo

[19] Posiblemente Gentzen se inspiró para construir este cálculo en algunos desarrollos de A. Heyting, tal como se menciona en (von Plato, 2014).

consiste en un *cálculo de secuentes*[20]. Asimismo, para cada tipo de cálculo, se establecen dos variantes, una intuicionista y otra clásica. El cálculo de deducción natural se conforma mediante un conjunto de reglas de introducción y eliminación de constantes lógicas, conjunto que ha permanecido estable hasta la actualidad conformando el *núcleo sintáctico* de la LE[21]. Por otra parte, el cálculo de secuentes integra aspectos y reglas del cálculo de deducción natural, junto a un tipo de reglas especiales denominadas *estructurales*, que prescinden por completo del uso de constantes lógicas, y que resultan de gran importancia para el estudio de propiedades de los sistemas lógicos.

La motivación general que impulsó a Gentzen en sus investigaciones, parte de la convicción de que la representación de las inferencias deductivas realizada en sistemas como los de Frege, Russell y Hilbert (desde la segunda mitad del siglo XIX), se aleja demasiado de las formas de deducción utilizadas en la práctica concreta de los matemáticos[22]. Estos autores convirtieron a la lógica en una disciplina axiomática alejada de sus orígenes intuitivos como *ciencia del razonamiento*. Particularmente, a Gentzen lo guiaba la idea de que la deducción es un modo de razonamiento hipotético o suposicional, basado en supuestos temporarios, tal como había sido la norma en lógica desde la época de Aristóteles[23]. En otras palabras, buscaba capturar el modo en el que razonan *efectivamente* los matemáticos cuando realizan pruebas de teoremas[24]. El cálculo de deducción natural se acerca bastante a ese objetivo, al ofrecer un instrumental más intuitivo para la demostración de teoremas en comparación con los sistemas axiomáticos,

[20] Aunque el mismo fue desarrollado fundamentalmente por Gentzen, existen antecedentes en la obra de P. Hertz, según lo consignado en (Roetti, 2016:140).
[21] Cfr. (von Plato, 2014:1167).
[22] Cfr. (Gentzen, 1969:68).
[23] Cfr. (von Plato, 2014:1168).
[24] Cfr. (Bimbó, 2015:5).

y cuenta con la gran ventaja de ofrecer herramientas para la formalización del razonamiento suposicional de un modo simple y directo.

Por su parte, el cálculo de secuentes va un paso más allá respecto al cálculo de deducción natural, en tanto se enfoca en la *relación de consecuencia*, ofreciendo un enorme potencial para analizar pruebas y para investigar sobre las características de los sistemas lógicos[25]. Es que las reglas estructurales tomadas por separado ofrecen un verdadero "marco metalógico" (Legris, 1999:237) para el estudio de las relaciones de consecuencia lógica y sus propiedades, como así también para la resolución de problemas tales como la consistencia de la lógica de predicados, el problema de decisión, y la no derivabilidad de la ley del *tertium non datur* en lógica intuicionista. En lo que sigue, la exposición se centrará exclusivamente en el cálculo de secuentes, en lo que concierne a las propiedades de la relación de consecuencia lógica (particularmente monotonía y propiedades asociadas), omitiendo la distinción entre cálculos intuicionistas y clásicos (en tanto particularmente la propiedad de monotonía vale para ambas variantes del sistema gentzeniano).

Un secuente tiene la forma $A_1, \dots, A_n : B_1, \dots, B_n$ donde $A_1, \dots, A_n$ y $B_1, \dots, B_n$ son secuencias ordenadas de fórmulas de un lenguaje de primer orden, y ":" representa una relación de derivabilidad. Las fórmulas que están a la izquierda de ":" se denominan *antecedentes*, mientras que las fórmulas de la derecha se denominan *sucedentes*. Intuitivamente, el esquema de secuente significa que de los antecedentes se infiere deductivamente al menos uno de los sucedentes. En ese sentido, las fórmulas del

[25] Ídem.

sucedente expresan cada una de las conclusiones que pueden obtenerse alternativamente a partir de un conjunto de premisas[26].

En este punto, conviene abrir un paréntesis respecto de la relación entre el signo ":" (derivabilidad[27]) de los sistemas de cálculo de secuentes, con el signo "⊢" (deducibilidad[28]) de los sistemas de LE. El símbolo de "⊢", también conocido como "torniquete", fue utilizado por primera vez por Frege en su Conceptografía (Frege, 1967), donde se distingue entre un mero juicio o proposición (antecedido por "─") y la *aserción* de ese juicio, la afirmación de que ese juicio es verdadero. En este último caso, el juicio es antecedido por "⊢"[29]. En otras palabras, en el sistema de Frege todo lo que sigue al torniquete, es un *teorema*[30]. Posteriormente, el signo "⊢" sería considerado como símbolo metalingüístico (relativamente a un lenguaje objeto de LE) para la noción de consecuencia sintáctica[31].

En el marco del cálculo de secuentes gentzeniano, no es correcto en líneas generales identificar ":" con "⊢", pero existe un caso donde es posible *interpretar*, con reservas, derivabilidad como deducibilidad: en el caso de que el sucedente de ":" no sea vacío ni contenga más de una fórmula bien formada (Roetti, 2016:145). Aun así, hay que tener en cuenta la diferencia de status de ambos signos, en tanto ":" es un *signo auxiliar* del cálculo de secuentes, mientras que "⊢" es un *signo metalingüístico* en LE. En el resto de los casos, la identificación es errónea en tanto el signo de derivabilidad posee una mayor

[26] Cfr. (Legris, 1999: 236)
[27] En el original (Gentzen, 1969), el signo de derivabilidad es "→". Aquí se adopta la convención presente en (Legris, 1999) de simbolizar derivabilidad como ":" para evitar la confusión con el signo de implicación material de LE.
[28] Cfr. (Roetti, 2016:145). También se lo denomina como "demostrabilidad".
[29] La interpretación del uso de este símbolo en el propio sistema de Frege ha generado intensos debates, los cuales son referidos en (Lazzer, 2006).
[30] Cfr. (Martin-Löf, 1996) y (Bimbó, 2015:15).
[31] Cfr. (Alchourrón, 1995).

generalidad que el signo de deducibilidad, puesto que: a) “tanto el antecedente como el sucedente del signo “:” pueden ser vacíos, tanto el primero como el segundo, o ambos”; y b) “el sucedente de “:” puede contener más de una expresión” (Roetti, 2016:146)[32]. Hecha esta aclaración, y volviendo a las propiedades de los sistemas lógicos, lo primero a tener en cuenta es la condición mínima que debe cumplir una relación de consecuencia lógica, indicada por el denominado “secuente inicial” que expresa la propiedad de *reflexividad*:

1.3.1 Reflexividad $A : A$

Este secuente no es un axioma en sentido estricto, ya que no es una tautología y más aún, no es una fórmula, sino una representación del *tránsito* de una deducción[33]. Su significado remite al *principio de identidad*, y se vincula con la propiedad de inclusión tarskiana (**1.2.3**). Junto a **1.3.1**, las siguientes reglas estructurales (entre otras) expresan propiedades de la relación de consecuencia lógica:

1.3.2 Debilitamiento izquierdo $\dfrac{\Gamma : \Theta}{A, \Gamma : \Theta}$

1.3.3 Debilitamiento derecho $\dfrac{\Gamma : \Theta}{\Gamma : \Theta, A}$

1.3.4 Corte $\dfrac{\Gamma : \Theta, C \qquad C, \Delta : \Lambda}{\Gamma, \Delta : \Theta, \Lambda}$

[32] Por otra parte, en (Marraud, 1998:89) se afirma que no se debe confundir “>” (aquí, “:”) con “⊢” por otros motivos: según este autor, el signo “>” permite representar estructuras generales de los sistemas lógicos mientras que “⊢” define la estructura básica de los objetos que manipula el sistema (encadenamientos deductivos de fórmulas).

[33] Con las reservas del caso, un secuente inicial con antecedente vacío y con una y sólo una fórmula en el sucedente puede considerarse como un axioma, cfr. (Roetti, 2016:144-145).

Como interpretación de **1.3.2**, vale lo siguiente: dada una secuencia cualquiera de fórmulas en el antecedente de un secuente, de las cuales se deriva una secuencia de fórmulas en el sucedente, la introducción de fórmulas arbitrarias a la secuencia del antecedente no modifica la secuencia derivada en el sucedente. Intuitivamente, esta regla garantiza que al agregar fórmulas *en conjunción* con el conjunto inicial de premisas no se pierden consecuencias obtenidas previamente. Por su parte, **1.3.3** puede entenderse como la aserción de que, dada una secuencia cualquiera de fórmulas en el antecedente de un secuente de las cuales se deriva una secuencia de fórmulas en el sucedente, la anexión de fórmulas arbitrarias a la secuencia del sucedente no modifica los resultados de la derivación previa. Intuitivamente, se trata de conclusiones alternativas o *en disyunción* con las conclusiones obtenidas anteriormente. Cabe resaltar que tanto en **1.3.2** como en **1.3.3** no se requiere una dependencia de hecho entre las secuencias iniciales del sucedente y del antecedente, respecto a las fórmulas arbitrarias anexadas[34]. Ambas reglas, tomadas de conjunto, constituyen la versión gentzeniana de la propiedad de monotonía. Como es patente, la noción de monotonía es más amplia en Gentzen que en la propuesta de Tarski, en tanto la versión tarskiana **1.2.6** (y expresiones equivalentes) corresponde sólo a debilitamiento izquierdo. En general, cuando se debate sobre la propiedad de monotonía, no se suele tomar en cuenta el debilitamiento derecho gentzeniano (**1.3.3**). Cabe destacar que para algunos sistemas de LNM se postula una propiedad con el mismo nombre, pero diferentes características, por lo que no debe confundirse con **1.3.3** (véase **CAPÍTULO 3**, propiedades **3.1.4** y **3.2.16**).

Por otra parte, la regla **1.3.4** adquiere relevancia debido a su rol en el *hauptsatz*, el teorema fundamental del cálculo de secuentes. Esta propiedad de *corte* o *transitividad* puede entenderse intuitivamente como la justificación de la acción de *saltarse pasos* en una

[34] Cfr. (Legris, 1999:237).

demostración, a fin de simplificar las operaciones. Por su parte, el hauptsatz establece que toda derivación puede ser transformada en una derivación con el mismo secuente final, y en la cual la regla de corte no es utilizada[35]. Esto a su vez permite demostrar que el sistema tiene la *propiedad de la subfórmula*: en el tránsito de una derivación libre de corte, toda fórmula de la demostración es una subfórmula del antecedente del secuente inicial, es decir, que "no desaparecen fórmulas a lo largo de las derivaciones en el sistema" (Legris, 1999:237). Mediante el recurso al cálculo de secuentes y las propiedades de corte, subfórmula y el hauptsatz, Gentzen pudo arribar (por otros medios y con otras herramientas) a resultados similares a los establecidos por Tarski en el principio de consistencia (**1.2.13**), pues le permitieron construir una prueba de la consistencia de los sistemas deductivos al demostrar que el secuente $: A \wedge \neg A$ no es derivable. Intuitivamente, la prueba consiste en demostrar que de un antecedente vacío no se sigue una contradicción[36]. Por último, es importante destacar que las reglas estructurales de reflexividad, debilitamiento y corte permiten representar el significado abstracto de la noción de *consecuencia lógica*, y en ese sentido cumplen una función metalógica análoga a los cuatro axiomas del sistema tarskiano[37].

1.4 COMPARACIÓN ENTRE DEFINICIONES INFORMALES Y FORMALES DE MONOTONÍA

[35] Cfr. (Gentzen, 1969:87).
[36] Para profundizar la demostración de consistencia (Gentzen, 1969), (von Plato, 2014) y (Roetti, 2016).
[37] Cfr. (Alchourrón, 1995:39).

Volviendo a las definiciones informales de monotonía, según el esquema **1.1.4** se plantea en primer lugar el problema respecto a si monotonía es una propiedad de razonamientos o de sistemas. Para abordar este asunto, se tomará en cuenta la presentación formal tarskiana y posteriormente la gentzeniana.

Como se indicó previamente, el tema de los escritos de Tarski de 1930 es el concepto de *sistema* (deductivo). Clásicamente, la idea de un sistema lógico remite a dos componentes básicos: un lenguaje simbólico artificial y un cálculo lógico. Este paradigma tiene sus orígenes en el pensamiento de Leibniz, particularmente en sus propuestas de una *characterica universalis* o sistema universal del conocimiento, y de un *calculus ratiocinator*[38]. Un ejemplo paradigmático de la concepción clásica es la conceptografía fregeana (Frege, 1967). El sistema de Frege se compone de un lenguaje de fórmulas que aspira a representar un compendio de las verdades científicas de las ciencias formales e idealmente de otras ciencias, y un cálculo lógico de primer orden. En el caso de Tarski, en términos generales el autor se enmarca en el paradigma clásico, pero describe el concepto de sistema en términos conjuntistas, como un conjunto especial de sentencias, del siguiente modo:

1.4.1 Sistema deductivo: *Un sistema es un conjunto de sentencias que contiene todas sus consecuencias*[39].

Esta definición refiere en última instancia a un lenguaje artificial de sentencias y a un cálculo lógico. Aunque el cálculo no es explícitamente mencionado, se trata de un componente esencial en tanto lo que define las consecuencias de un conjunto de sentencias es, justamente, un operador que representa un conjunto de reglas de inferencia.

[38] Para un estudio en profundidad de estos tópicos, véase (van Heijenoort, 1967) y (Peckhaus, 2004).
[39] Cfr. (Tarski, 1969a:69-70). Allí se establece que "sistema deductivo", "sistema cerrado" y "sistema" son sinónimos.

De modo que la teoría de los sistemas de Tarski puede enrolarse en el paradigma clásico, aunque en principio no se orienta a la descripción de una *characterica universalis*, sino de múltiples sistemas. No obstante, el autor no rechaza por completo la posibilidad de un *sistema universal* de ese tipo, al postular una virtual clase $\mathcal{L}$ entendida como la *clase de los sistemas deductivos*[40] (la colección de todos los sistemas). Según este planteo, para cualquier conjunto de sentencias Θ, se cumple lo siguiente:

1.4.2 Clase de los sistemas deductivos: $(\Theta \in \mathcal{L}) \leftrightarrow ((\Theta \subseteq S) \wedge (Cn(\Theta) = \Theta))$

En otras palabras, un conjunto de sentencias es un sistema deductivo syss es un subconjunto de S y es un punto fijo del operador Cn. Ahora bien, por definición **1.4.2** no se excluye la posibilidad de un sistema $\Theta = S$ tal que $Cn(S) = S$. Es decir, un sistema compuesto por todas las sentencias significativas y sus consecuencias deductivas. Por otra parte, Tarski formuló un teorema que afirma que la suma de todos los sistemas es también un sistema[41]. Es decir, que la clase $\mathcal{L}$ podría ser entendida como un conjunto (el conjunto de todos los sistemas) y sería también, un sistema. La pregunta que surge inmediatamente es la siguiente: ¿es este *sistema universal* $\mathcal{L}$ una *characterica universalis* (según la propuesta clásica de Leibniz)? ¿Se trata del *compendio de todo el conocimiento humano*? En este caso, emergería un enorme problema filosófico (de filosofía de la lógica), ya que tal *characterica universalis versión tarskiana* sería inconsistente, dado que contiene $Cn(S) = S$ (que es un sistema inconsistente, como se consignó previamente en la sección **1.2**). Ahora bien, incluso suponiendo que el sistema $\mathcal{L}$ no es una *characterica universalis*, ni un compendio de todo el conocimiento humano, esto no resta importancia al problema, ya que de todos modos se trata de la *clase de todos los sistemas*, el *sistema de los sistemas deductivos*, a fin de cuentas, el *sistema universal*. El hecho de que este sistema $\mathcal{L}$ sea

[40] Cfr. (Tarski, 1969a:70).
[41] Cfr. Th. 12 en (Tarski, 1969a:71).

inconsistente resulta cuanto menos extraño, teniendo en cuenta la clásica *exigencia de consistencia* de la LE. No podrá profundizarse aquí sobre el impacto de esta situación para la propuesta global de Tarski, quedando el problema abierto a futuras investigaciones.

Volviendo al concepto de *sistema* en relación con la propiedad de monotonía, en los escritos de 1930 de Tarski los axiomas y teoremas metamatemáticos (algunos de los cuales se reflejan en la sección **1.2**) pueden entenderse, en su unidad, como una *exposición del concepto de sistema lógico (deductivo)*, ya que los mismos describen características fundamentales de tales sistemas. De modo que en el marco de la propuesta tarskiana sería más adecuado definir monotonía como una propiedad de sistemas, y no de razonamientos. Sin embargo, es importante destacar que tanto monotonía como el resto de las propiedades de los sistemas deductivos remiten directamente al operador Cn, el cual, si bien puede ser interpretado genéricamente como una función, a fin de cuentas, representa un *conjunto de reglas de inferencia*, es decir, un cálculo lógico. Cuando Tarski sostiene que Cn es monótono en el dominio de conjuntos de sentencias, está afirmando que el cálculo lo es[42]. De modo que puede afirmarse legítimamente que en la propuesta tarskiana *un sistema es monótono en virtud del cálculo* lógico que lo constituye, y por lo tanto sería más preciso afirmar que concibe la propiedad de monotonía como una propiedad del cálculo deductivo.

En el caso de Gentzen, también enrolado en líneas generales con la concepción clásica de *sistema*, en sus "Investigaciones…" se refiere a las reglas estructurales (entre ellas la regla de debilitamiento) como "propiedades específicas" del cálculo natural (Gentzen, 1969:68). En el texto citado el autor utiliza los términos "sistema" y "cálculo" como

[42] Cfr. (Tarski, 1969a:65).

sinónimos (con clara preferencia del segundo término por sobre el primero), aunque no ofrece mayores precisiones al respecto. Nuevamente, queda claro que debilitamiento/monotonía no es una propiedad de razonamientos, sino del cálculo. Volviendo a las definiciones informales de monotonía, es patente que predicar monotonía de razonamientos es erróneo, mientras que predicar esa propiedad respecto a sistemas es apropiado en general, aunque en particular puede precisarse aún más la definición. Lo más adecuado al pensamiento de Tarski y Gentzen es predicar monotonía relativamente al cálculo (deductivo).

1.4.3: *Monotonía es una propiedad del cálculo deductivo.*

Algo similar puede afirmarse de gran parte de las propiedades metalógicas estudiadas por estos autores, lo cual pone al cálculo (conjunto de reglas de inferencia) en primer plano como *negocio de la lógica*. Esta idea no es novedosa, tal como destaca J. Hintikka en su (Hintikka & Sandu, 2007:15) cuando afirma que "a las reglas de inferencia se las considera a menudo el alfa y el omega de la lógica". Establecido **1.4.3**, restaría abordar el problema de la relación entre información y monotonía. Esta cuestión puede inscribirse dentro una problemática de mayor generalidad, que resulta ser un tópico de gran importancia en filosofía de la lógica: la relación entre inferencias e información. Al respecto, lo primero que hay que mencionar es que el concepto de información está presente de las definiciones informales de monotonía, donde se establece una conexión entre esta propiedad y la noción de *información nueva*, pero no ocurre lo mismo con las expresiones formales de monotonía. En la bibliografía utilizada para reconstruir perspectiva tarskiana de monotonía, particularmente en (Tarski, 1969a), (Tarski, 1969b) y (K. Simmons, 2009), no existe ningún tipo de referencia a los conceptos de información e información nueva. En el teorema de monotonía **1.2.6** y formulaciones equivalentes no existe restricción alguna relativamente a las sentencias que pueden anexarse a un conjunto dado de sentencias, al cual se aplica el operador Cn. Del mismo modo, en la presentación gentzeniana no caben dudas respecto a la relación del concepto de monotonía y la noción de información nueva: no existe en los artículos de referencia (Gentzen, 1969; Legris, 1999; Roetti, 2016; von Plato, 2014) ningún tipo de mención sobre el concepto de información, menos aún respecto a su *novedad*. Las reglas estructurales de debilitamiento remiten a la introducción de *fórmulas arbitrarias* en el antecedente o en el sucedente de un secuente.

En este punto, es inevitable abrir una digresión relativamente extensa en torno a algunos de los debates relativos a la problemática general de la relación entre inferencias e información, a fin de continuar profundizando sobre el concepto de monotonía. El primer paso obligado es reconstruir los planteos de la *teoría clásica de la información* (TCI), que es la teoría filosófica tradicional sobre la relación inferencias/información y que se encuentra en el centro del debate actual. La TCI cuenta con representantes modernos y contemporáneos tales como C. S. Peirce y la escuela neopositivista, pero aquí se optará por la formulación peirceana (Peirce, 1994: CP 2) debido a su simplicidad. El punto de partida de la versión peirceana de la TCI es la consideración de que toda proposición es un *símbolo informacional*, y a su vez las proposiciones componen argumentos (al modo habitual, algunas proposiciones funcionan como premisas mientras que otras son conclusiones de esas premisas). Existen dos clases de argumentos o razonamientos, denominados *razonamientos explicativos* (analíticos, deductivos); y *razonamientos ampliativos* (sintéticos, no deductivos). En el marco de la propuesta peirceana los argumentos explicativos son (todos ellos) argumentos válidos[43] (y se da por supuesto que

[43] Peirce toma en cuenta, fundamentalmente, los diferentes tipos de silogismos aristotélicos, que son explicativos siempre y cuando se cumpla la garantía de que "los hechos presentados en las premisas no pueden bajo ninguna circunstancia imaginable ser verdaderas sin involucrar la verdad de la conclusión, la cual es en consecuencia aceptada con modalidad necesaria" (Peirce, 1994: CP 2.778).

sus premisas son consistentes). Asimismo, se sostiene que todas las demostraciones matemáticas son argumentos explicativos[44]. La característica principal de los argumentos explicativos es que los hechos o la información respecto a hechos, contenida en la conclusión de los mismos, se encuentra ya indicada en las premisas. Es decir, que en la conclusión no hay información que no esté de algún modo disponible en las premisas, y en todo caso la conclusión presenta esa información de un modo diferente, o resumido, o resaltando aspectos que en las premisas pueden ser pasados por alto[45].

Por otra parte, los razonamientos ampliativos pueden ser, según Peirce, de tres subtipos: inductivos, abductivos y analógicos. A diferencia de los argumentos explicativos, los ampliativos no son válidos, y la verdad de las premisas no obliga a aceptar la verdad de la conclusión. De hecho, las conclusiones de estos argumentos sólo son aceptadas provisionalmente, aunque pueden serlo indefinidamente, hasta que la experiencia indique que esa conclusión ya no está justificada[46]. Este tipo de argumentos tiene a su vez la característica de que los hechos o la información contenida en la conclusión no se encuentra contenida en las premisas. De modo que lo afirmado en la conclusión de un argumento ampliativo *va más allá* de lo afirmado en sus premisas. Se trata de hechos diferentes en cada caso. Por ello, Peirce afirma que los razonamientos ampliativos son los únicos que incrementan el conocimiento. En los siguientes postulados se sintetiza la TCI:

1.4.4: *Los razonamientos explicativos son no informativos, en el sentido de que sus conclusiones no aportan información nueva relativamente a la información contenida en sus premisas.*

[44] Cfr. (Peirce, 1994: CP 2.680).
[45] Ídem.
[46] Cfr. (Peirce, 1994: CP 6.40).

1.4.5: *Los razonamientos ampliativos son informativos, en el sentido de que sus conclusiones aportan información nueva relativamente a la información contenida en sus premisas.*

Como se mencionó previamente, el postulado **1.4.4** da por supuesto que los razonamientos explicativos son razonamientos deductivos (válidos) con premisas consistentes. En el caso contrario las inferencias deductivas pueden ser informativas, puesto que (en el marco de LE) dada la ley de no contradicción **1.2.12** ocurre que partiendo de un conjunto inconsistente de premisas pueden obtenerse todas las sentencias del lenguaje. El postulado **1.4.5** no ha generado polémicas significativas, pero **1.4.4** ha sido reiteradamente criticado en el ámbito de la filosofía de la lógica, debido a que genera consecuencias que se consideran anti-intuitivas o paradójicas. A continuación conviene introducir, como marco de referencia actual, algunos elementos de la *teoría de la información semántica* (TIS), que es un conjunto de teorías lógico-filosóficas que toman como punto de partida los planteos de la *teoría matemática de la comunicación* (TMC) creada por C. Shannon (Shannon, 1948). Esquemáticamente, la TMC sería la versión sintáctica de una *teoría general de la información*, mientras que la TIS constituye la formulación semántica de la misma teoría. Tanto la TMC como la TIS son relevantes para el debate actual sobre el problema de la relación entre inferencias e información, particularmente por el rol que juega en ese debate el denominado *principio de relación inversa* (PRI) de Carnap y Bar-Hillel. Estos autores, en su (Carnap & Bar-Hillel, 1952), sentaron las bases de la TIS y a su vez formularon el PRI entre la probabilidad de una sentencia s (que describe un evento, situación o mundo posible) y la cantidad de información semántica transportada por la misma[47].

1.4.6: *La información semántica transportada por una sentencia es inversamente proporcional a la probabilidad de los eventos descriptos en tal sentencia.*

Este criterio toma la forma de una medida cuantitativa de la información semántica, que suele ser calculada en términos de *bits* de información[48]. Entre las formulaciones para el cálculo cuantitativo de la información semántica, se encuentra el siguiente axioma que sistematiza el PRI[49]:

1.4.7: $I(s) = -\log_2 P(s)$

Intuitivamente, a menor probabilidad P de que efectivamente ocurra lo descripto en una sentencia s, mayor cantidad de información semántica I transporta s y viceversa. Tomando a **1.4.7** como axioma de la TIS, y en tanto la probabilidad de una tautología es igual a 1 (en términos modales, tautología como una sentencia que es verdadera en todos los mundos posibles), se sigue inmediatamente teorema **1.4.8** presentado a continuación. Del mismo modo, dado que la probabilidad de una contradicción es igual a 0 (es falsa en todos los mundos posibles), se sigue el teorema **1.4.9**. Cabe aclarar que los teoremas presentados aquí constituyen una formulación metamatemática alternativa, propia del autor de esta tesis, de los teoremas T7-8 b) y c) en (Carnap & Bar-Hillel, 1952: 20,21).

1.4.8: $I(\top) = -\log_2 1 = 0$ bits

---

[47] Se suele afirmar que K. Popper formuló previamente ideas filosóficas similares al PRI, sin embargo cualquier formulación sistemática del mismo presupone la teoría de Shannon, tal como consigna L. Floridi en su (2009).

[48] También pueden utilizarse otras unidades de medida. El bit (binary digit) es la medida convencional de la TMC.

[49] Cfr. (Carnap & Bar-Hillel, 1952), y (Floridi, 2009).

1.4.9: $I(\perp) = -\left(\lim_{x \to 0^+} \log_2 x\right) = \infty \text{ bits}$

El teorema **1.4.8** puede entenderse como un correlato informacional (en términos de PRI) para el postulado **1.4.4** de la TCI. Partiendo de un conjunto de premisas consistentes y construyendo una demostración, al aplicar el metateorema de la deducción a la misma se obtiene una tautología, las cuales son no informativas (transportan 0 bits). Tanto el postulado **1.4.4** como el teorema **1.4.8** se vinculan, a su vez, con la denominada "paradoja de Cohen – Nagel[50]" (también, *escándalo de la deducción*[51]), que cuestiona la utilidad de las inferencias deductivas por ser no informativas. Por otra parte, el teorema **1.4.9** establece, en términos metamatemáticos, una condición de la lógica deductiva que recibe el nombre de *paradoja de Carnap – Bar-Hillel*, según la cual la cantidad de información semántica contenida por una contradicción es infinita, y por lo tanto una sentencia de este tipo es "demasiado informativa para ser verdadera" (Carnap & Bar-Hillel, 1952: 8). En otras palabras, se trata del correlato informacional del principio *ex contradictione quodlibet* (ley explosión lógica **1.2.12**).

Hasta la fecha, en filosofía de la lógica se han ofrecido diferentes alternativas para la solución de ambas paradojas, aun así, no está claro que se haya arribado a soluciones satisfactorias[52]. Lo que sí resulta evidente es que en la búsqueda de soluciones a tales resultados antiintuitivos de la TCI, se ha generalizado el rechazo del postulado **1.4.4**, a tal punto que ese rechazo ha encarnado en el sentido común. En efecto, la no informatividad de las inferencias explicativas puede resultar chocante, pero es un hecho establecido sobre sólidas bases metalógicas y metamatemáticas. Para cualquier

[50] Según la denominación que le asigna a este problema D'agostino en su (D'Agostino, 2014). Por su parte, los propios Cohen y Nagel lo denominan como "paradoja de la inferencia". Véase (Cohen & Nagel, 1934: 173).
[51] Cfr. (Hintikka, 1973).
[52] Para un recuento de las mismas, ver (D'agostino & Floridi, 2009).

demostración (en términos clásicos) se puede probar su *no informatividad* aplicando herramientas metalógicas y de teoría de la información, del siguiente modo: 1) aplicando el metateorema de la deducción para obtener una tautología; 2) aplicando el cálculo de informatividad según la fórmula del axioma **1.4.7** (cuyo resultado está contemplado en el teorema **1.4.8**). De modo que una crítica exitosa del postulado **1.4.4** debería derribar el metateorema de deducción, el teorema **1.4.8** de la TIS, o al menos presentar ejemplos de razonamientos explicativos informativos. Pero el metateorema de deducción goza de buena salud para sistemas de LE[53]. Lo mismo ocurre con el teorema **1.4.8** en el marco de la TIS. Por último, la existencia de ejemplos de razonamientos explicativos informativos (es decir, razonamientos deductivos con premisas consistentes que produzcan información no contenida en las premisas) no parece haber sido probada fehacientemente, aunque se ha teorizado en extenso sobre la esa posibilidad en textos como (Hintikka, 1970a), (Hintikka, 1970b), (Hintikka, 1973), (Popper, 1976), (Hintikka & Sandu, 2007), (D'agostino & Floridi, 2009), (Frederick, 2011), (Frederick, 2014) y (D'Agostino, 2014) entre otros. En el **APÉNDICE 1** al final de esta tesis se analizarán algunas de estas propuestas.

En la actualidad, el postulado **1.4.4** de la TCI sigue siendo adecuado para describir las propiedades informacionales de los razonamientos deductivos sistematizados en el marco de LE, aun aceptando los resultados antiintuitivos o paradójicos que genera ese postulado, como problemas no resueltos[54]. Por supuesto, la LE, que se originó primordialmente para

---

[53] De ello da cuenta J. Porte, en su (Porte, 1981). El autor afirma que el metateorema sólo es falso en el marco de sistemas no estándar que admiten derivaciones que involucran predicaciones con variables libres.

[54] Cabe mencionar que la supuesta "paradoja de la inferencia" o escándalo de la deducción fue etiquetada como "pseudoparadoja" por Cohen y Nagel en 1934, aunque autores actuales que abordan este problema (como D'agostino) parecen ignorar por completo la respuesta presente en "Chapter IX" de (Cohen & Nagel, 1934). Véase **Apéndice 1** al respecto. Por otra parte, la paradoja de Carnap – Bar-Hillel parece no tener solución en el marco de la LE.

modelar un tipo determinado de demostraciones matemáticas, no es el único sistema lógico disponible y posiblemente no tenga sentido defender la TCI para sistemas no estándar. Pero eso es harina de otro costal. En lo que concierne a LE, se trata de un postulado importante, el cual aún es posible refinar, reformulando algunos términos y aprovechando el instrumental metalógico y de teoría de la información disponible para ofrecer una presentación formal y sistemática del mismo. Lo primero a tener en cuenta al observar el postulado **1.4.4** son los conceptos de *información* y *novedad*, los cuales deberían estar estrictamente definidos, a fin de evitar ambigüedades. El concepto de información queda bien establecido con ayuda del axioma **1.4.7** de la TIS, y no se requieren más precisiones al respecto. Para el concepto de novedad la situación es más complicada, pues este involucra una fuerte carga de subjetividad en tanto *lo novedoso* siempre es novedoso para algún sujeto.

Afortunadamente, Cohen y Nagel se han ocupado de este tema en su *Introducción a la lógica…*, donde establecen una distinción entre novedad lógica y novedad psicológica. Descartando la segunda por carente de interés, se define la primera como "la independencia lógica de lo expresado en la conclusión respecto a las premisas" (Cohen & Nagel, 1934:175). En un razonamiento deductivo (con premisas consistentes) la información que acarrea la conclusión carece de novedad respecto a la información acarreada por las premisas, pues la conclusión no es lógicamente independiente de sus premisas. Esto puede observarse simplemente construyendo una demostración que modele tal razonamiento, y luego aplicando el metateorema de la deducción y el teorema **1.4.8**. Por contraposición, en un razonamiento ampliativo la conclusión es lógicamente independiente (no se sigue) de las premisas, y por lo tanto la información que contiene es lógicamente novedosa respecto a la información contenida en las premisas. En "términos tarskianos", la distinción de Cohen – Nagel permite construir dos conjuntos especiales relativos a cualquier conjunto consistente de sentencias Δ de un sistema deductivo estándar, Ω_Δ^1 y Ω_Δ^2 los cuales permiten diferenciar entre dos tipos de sentencias en función de su novedad lógica respecto a Δ:

1.4.11 Sentencias tipo 1: $\Omega_\Delta^1 = \{x : x \in \{S \cap Cn(\Delta)\}\}$

Se trata del conjunto de todas las consecuencias deductivas de Δ. Estas sentencias, que en el caso más trivial pueden ser idénticas a sentencias de Δ, representan información no novedosa relativamente a Δ.

1.4.12 Sentencias tipo 2: $\Omega_\Delta^2 = \{x : x \in \{S - Cn(\Delta)\}\}$

Se trata de todas las sentencias de S que no son consecuencias deductivas de Δ. Representan información lógicamente novedosa respecto a Δ. De manera que sólo un cálculo ampliativo que permita producir consecuencias Ω_Δ^2 a partir de un conjunto Δ consistente estaría produciendo, legítimamente, información novedosa. Así, tomando en cuenta la clasificación entre sentencias de tipo 1 y sentencias de tipo 2 relativamente a un conjunto Δ consistente, es posible reformular los corolarios **1.2.10** y **1.2.11** del teorema de monotonía tomando en cuenta la novedad lógica de las sentencias. Sea Δ consistente, se propone la siguiente variante de **1.2.10**:

1.4.13: $((\Delta \cup \Psi) \subseteq S) \wedge (\Psi \in P(\Omega_\Delta^1)) \rightarrow (Cn(\Delta) = Cn(\Delta \cup \Psi))$

Esta expresión significa que en un sistema deductivo no se pierden consecuencias de un conjunto inicial de sentencias Δ consistente si se agrega al mismo un conjunto de sentencias Ψ que pertenece al conjunto potencia de Ω_Δ^1, en otras palabras, cuando las sentencias que se adicionan a Δ son idénticas a sentencias existentes previamente en Δ, o bien cuando las sentencias que se adicionan a Δ son consecuencias deductivas de Δ no idénticas a sentencias de Δ. En segundo lugar, se propone la siguiente variante del corolario **1.2.11** (con Δ consistente):

1.4.14: $((\Delta \subset (\Delta \cup \Psi) \wedge (\Psi \in P(\Omega_{\Delta}^{2})) \ \rightarrow (Cn(\Delta) \subset Cn(\Delta \cup \Psi))$

Esta expresión significa que en un sistema deductivo se conservan las consecuencias de un conjunto inicial de sentencias consistente Δ si se agrega al mismo un conjunto de sentencias Ψ que pertenece al conjunto potencia de Ω_{Δ}^{2} (es decir, información lógicamente novedosa), dicho de otro modo, cuando las sentencias que se adicionan a Δ ni son idénticas a sentencias existentes previamente en Δ, ni son consecuencias deductivas de Δ no idénticas a sentencias de Δ. Con **1.4.13** y **1.4.14** queda claro que la propiedad de monotonía se cumple independientemente del tipo de información contenida en las premisas adicionales.

1.5 DEFINICIÓN ALTERNATIVA DE MONOTONÍA

A modo de conclusión del capítulo, a continuación, se presentará una definición alternativa de monotonía, que sintetiza el análisis previo. La misma funciona como descripción informal del mencionado concepto, tanto si se toma en cuenta la expresión formal tarskiana o la gentzeniana (debilitamiento izquierdo), presentadas previamente. En ambos casos, vale lo siguiente:

1.5.1 *La propiedad de monotonía es una condición del cálculo de LE que garantiza la conservación de las consecuencias de un conjunto cualquiera de sentencias cuando se amplía ese conjunto, independientemente del tipo información que representen las* *sentencias adicionales: ya sea que ésta sea lógicamente novedosa o lógicamente no novedosa relativamente a la información representada en el conjunto inicial de sentencias.*

CAPÍTULO 2 – NO MONOTONÍA

INTRODUCCIÓN

La propiedad de monotonía de LE había permanecido libre de controversias hasta el surgimiento, a mediados del siglo XX, de una nueva disciplina dentro de las ciencias computacionales, denominada "Inteligencia Artificial" (IA). Uno de sus fundadores, J. McCarthy, proponía en su (1959) la utilización de LE para la construcción de programas con *sentido común*. El autor partía de considerar que no existía un modo de formalizar los procesos mediante los cuales los humanos aprenden de la experiencia. Para construir un programa que pudiera ejecutar este tipo de aprendizaje es necesario un lenguaje que permita representar hechos acerca de situaciones, objetivos y efectos de acciones, como así también obtener conclusiones a partir de esos hechos para orientar la acción. La expectativa de McCarthy era que la LE resultara adecuada para la formalización de la información y los procesos de razonamiento involucrados en el sentido común. Aun así, el autor advertía que esa teoría lógica había sido desarrollada originalmente para los fines de las matemáticas, y que sólo se habían realizado unos pocos progresos en la formalización de las ciencias empíricas. Más tarde quedaría claro que las características propias de la LE en tanto teoría que modela el razonamiento deductivo dificultan o directamente impiden su aplicación a otros tipos de inferencia. M. Minsky, otro de los fundadores de la IA, fue el primero en poner de manifiesto en su (1974) que la propiedad de monotonía es la fuente del problema, y en ese sentido puede ser considerado el *padre* de las lógicas no monótonas (LNM). Según este autor, en un sistema de LE, cada nuevo axioma permite nuevos teoremas, ninguno de los cuales puede desaparecer. No existe un modo de agregar información al sistema para decirle que existen ciertas conclusiones que no deben ser obtenidas. Además, el autor criticaba la exigencia de consistencia de la LE, expresada en el principio de consistencia **1.2.13** y que resulta de enorme importancia para la lógica matemática, pero inconveniente para el modelado del conocimiento ordinario. Las personas al razonar deben lidiar con inconsistencias, y por ello es necesario crear sistemas que permitan manipularlas, evitando los resultados indeseados que se generan en LE por motivo de la ley de explosión lógica (**1.2.12**).

Por otra parte, uno de los objetivos fundamentales de la IA es la construcción de un *agente racional artificial*. Para ese fin, se propone modelar y emular computacionalmente el razonamiento humano de manera global, en todos sus aspectos y manifestaciones, sin limitarse al razonamiento estrictamente demostrativo (deductivo) tal como ocurre en LE. En virtud del principio teórico vigente en IA que establece que *razonamiento es computación* (Poole, 1998:1), cualquier tipo de razonamiento debería poder ser representado formalmente y emulado computacionalmente. Por este motivo, desde sus orígenes la IA se ha enfocado en la formalización de ciertos tipos de razonamiento de uso cotidiano que caen fuera de los alcances de la LE por carecer de corrección lógica[55], pero que juegan un rol importante en el establecimiento de conclusiones. Por supuesto, la indagación teórica sobre este tipo de razonamientos se remonta a los orígenes de la lógica,

[55] Cfr. (Russell & Norvig, 1994).

en la obra de Aristóteles, pero había recibido menos atención por parte de los lógicos entre fines del s. XIX y el segundo tercio del siglo XX[56]. Estos razonamientos se engloban bajo la categoría de *razonamientos ampliativos* en el marco de la TCI (tal como se consignó en el **CAPÍTULO 1**), aunque esta denominación no es habitual en LNM[57]. Ahora bien, es importante tener en cuenta que los procesos de razonamiento de sentido común no excluyen necesariamente a la deducción, en tanto y en cuanto los sujetos realizan deducciones más o menos simples en su vida cotidiana. Sin embargo, el interés de la LNM se centra primordialmente en el razonamiento ampliativo, por ejemplo, las diferentes formas de razonamiento inductivo y probabilístico (cuyo estudio formal precede históricamente a la LNM), pero también otros esquemas o tipos de razonamiento cuyo abordaje formal es (relativamente) novedoso, tales como el razonamiento *prima facie*, por *default*, y por *circunscripción*, entre otros. Si bien las inferencias ampliativas

[56] En esa época, el objetivo primordial de los lógicos (principalmente europeos) era orientar sus investigaciones a los fines de la fundamentación de las matemáticas. Por ello su interés se centraba primordialmente en la representación formal del razonamiento matemático, incluyendo principalmente los diferentes tipos de razonamiento deductivo, pero también algunos tipos de razonamiento ampliativo como la inducción y el razonamiento probabilístico. Así, el estudio de otros modos de inferencia no deductiva de uso cotidiano quedó confinado a la lógica filosófica. Durante esta etapa, los avances en el programa de fundamentación de las matemáticas llevaron a la resolución de importantes problemas metamatemáticos, al establecimiento de la LE como teoría lógica paradigmática y a la generalización del ideal leibniziano de formalizar todo el conocimiento humano mediante LE (en el sentido de una *characterística universalis*). Sin embargo, en el último tercio del siglo XX, más precisamente a partir de la década del ‘70, comenzó a crecer entre los lógicos un movimiento crítico hacia la LE, dentro del cual se pueden enmarcar la LNM (originada en EE. UU.) y por otra parte las teorías de lógica informal (LI) originadas en Canadá. Por supuesto, no son las únicas corrientes que critican aspectos de la LE, pero tienen en común el señalamiento de la inconveniencia de esa teoría para dar cuenta de los modos de razonamiento utilizados cotidianamente en el despliegue del sujeto en el mundo, y por ello tanto la LNM como la LI toman a este tipo de razonamiento como su objeto de estudio. Eventualmente, ambas *escuelas lógicas* han llegado a criticar la concepción heredada de la LE según la cual los razonamientos deben ser considerados como *productos*, y se comenzó a enfocarlos como *procesos*, en términos de diálogos o juegos dialógicos entre dos jugadores (enfoque dialéctico). No obstante, existen diferencias fundamentales entre ambas corrientes en lo que concierne a la construcción de sus objetos de estudio, las metodologías de representación y análisis, y las motivaciones. La diferencia más notable, de hecho, pasa por el rechazo de la LI de la utilización de lenguajes formales, mientras que en LNM esa es una condición necesaria para cualquier sistema, necesidad que se explica en función de los objetivos de la IA. De todos modos, por lo dicho anteriormente la LNM puede enmarcarse históricamente junto a la LI como parte de la “reacción norteamericana” a la lógica matemática europea de principios del siglo XX.

[57] En (Makinson, 1989) se refiere a este tipo de razonamientos como “no deductivos”, otros autores del área los etiquetan como razonamientos *no monótonos* (por ejemplo, en (Brewka, 1991), (Ginsberg, 1994)), *de sentido común* ((Winograd, 1980), (Lukaszewicz, 1990)), o *derrotables* ((Pollock, 1987), (Simari & Loui, 1992)).

no poseen corrección lógica en términos de validez, en la práctica con frecuencia ocurre que resultan aceptables y las conclusiones que producen tienen un grado de justificación suficiente como para ser *creídas* por los agentes. En contraste, para un razonador que evalúa sus razonamientos recurriendo únicamente a LE, las conclusiones de las inferencias ampliativas no son aceptables.

Estas dificultades muestran el límite de la pretensión de extender el dominio de aplicación de la LE a otros ámbitos por fuera del razonamiento puramente demostrativo de las matemáticas, particularmente en lo que respecta a la representación formal y evaluación del razonamiento práctico. Por otra parte, la LNM se enfoca en un fenómeno lógico que no ha sido tenido en cuenta por la LE, como es la *eliminación de conclusiones* previamente obtenidas. Este fenómeno también ha sido descripto como *revisión de creencias*[58], y como una condición de *derrotabilidad*[59] (de los procesos de razonamiento de sentido común). Los sujetos razonan espontáneamente obteniendo conclusiones, luego vuelven a razonar y tanto pueden obtener nuevas conclusiones, como eliminar otras anteriormente inferidas, y así sucesivamente, según necesidad práctica. En el desafío de modelar este proceso, se hace patente otro de los límites de la LE, ya que los sistemas que satisfacen la propiedad de monotonía conservan obligatoriamente todas las conclusiones obtenidas previamente.

De modo que, en términos generales, los objetivos de la LNM pueden resumirse en los siguientes: a) formalización de los procesos de razonamiento de sentido común, primordialmente en lo que concierne a la representación formal de inferencias ampliativas; y b) modelización del fenómeno de eliminación de conclusiones. Respecto a este último fenómeno, el mismo es asociado a una característica metalógica especial,

---

58 Cfr. (Mc Dermott & Doyle, 1980).
59 Cfr. (Pollock, 1987).

denominada usualmente como *no monotonía*, característica que constituye en el correlato formal del fenómeno práctico considerado. Pero, más allá de estos objetivos generales, la LNM es un campo teórico de gran complejidad que comprende un cúmulo de teorías con diversos enfoques y metodologías[60]. Según mencionaba G. Brewka hace casi dos décadas, este campo de estudio permanecía *en estado de flujo* (Brewka, 1991:ix), donde constantemente se publican nuevas teorías y se refinan las existentes. En la actualidad, el panorama sigue siendo complejo, aunque se realizan esfuerzos encaminados a lograr una mayor cohesión teórica y metodológica del campo[61]. En ese marco, la noción de no monotonía constituye un eje teórico transversal a todos los sistemas de LNM (quizá el único), y por eso constituye un tema de interés para la filosofía de la lógica. No obstante, tal como se indicó en **INTRODUCCIÓN Y OBJETIVOS**, la bibliografía del área carece de una expresión puramente formal, metamatemática, de no monotonía, en el sentido de una fórmula que exprese una propiedad satisfacible por todos los sistemas considerados no monótonos. Esto resulta extraño, ya que se trata, por decirlo de algún modo, de *la propiedad de las propiedades* en LNM. Habiendo profundizado previamente en el **CAPÍTULO 1** sobre la propiedad de monotonía, se impone un análisis similar sobre no monotonía, que comenzará por un abordaje filosófico que puede reducirse a la búsqueda de una respuesta para la siguiente pregunta: *¿qué significa no monotonía?* Una vez desplegado tal análisis, el autor espera que quede allanado el camino para ensayar una formulación metamatemática de la mencionada noción.

---

[60] Entre las teorías más difundidas de LNM pueden consignarse las siguientes (según el nombre que reciben en su idioma original, el inglés): *negation as failure*, *circumscription*, *default logic*, *relevance logic*, *autoepistemic logic*, *defeasible reasoning*, *argumentation system*. Esta familia de teorías además ha recibido los nombres comunes de *exotic logics* (Brewka, 1991), *non-monotonic reasoning* y *non-monotonic formal systems* (Dov Gabbay, 1994), pero también *logics for defeasible reasoning* y *argumentation systems* (Prakken & Vreeswijk, 2001), entre otras denominaciones.

[61] Por ejemplo, en (D. Gabbay & Schlechta, 2016).

2.1 DEFINICIONES INFORMALES DE NO MONOTONÍA

En la bibliografía del área, es habitual encontrar aseveraciones simples respecto a que los sistemas de LNM lisa y llanamente no satisfacen la propiedad de monotonía de LE. Esta es la lectura más inmediata que puede hacerse del concepto de no monotonía. Un ejemplo es la siguiente definición que puede caracterizarse como *minimalista*:

2.1.1 "Los sistemas de lógica no monótona quebrantan la propiedad de monotonía" (Lukaszewicz, 1990:77):

Tal lectura minimalista e inmediata del concepto puede ser reflejada en el siguiente postulado:

2.1.2 No *monotonía consiste en la mera negación de la propiedad de monotonía.*

Por otra parte, pueden encontrarse otro tipo de definiciones informales, con características más complejas. A continuación, se presentará un grupo de tales definiciones, que resulta representativo de este segundo tipo de abordaje del concepto:

2.1.3 "La inferencia es no monótona, ante el aprendizaje de un nuevo hecho (...) podemos vernos forzados a retractar nuestra conclusión" (Ginsberg, 1994:3).

2.1.4 "Los sistemas de lógica no monótona son lógicas en las que la introducción de nuevos axiomas puede invalidar viejos teoremas" (Mc Dermott & Doyle, 1980: 41).

2.1.5 "El razonamiento de sentido común frecuentemente es no monótono. En muchas situaciones sacamos conclusiones que son abandonadas a la luz de nueva información" (Brewka, 1991: 2).

El núcleo teórico de este grupo de definiciones informales es la caracterización de la idea de no monotonía como *posibilidad de eliminación de conclusiones* previamente obtenidas de un conjunto de premisas, ante la ampliación de ese conjunto mediante la adición de proposiciones que representan información nueva. Esta característica común de las mencionadas definiciones puede tomarse como uno de los postulados esenciales de la LNM, del siguiente modo:

2.1.6 *No monotonía consiste en la posibilidad de eliminación de consecuencias previamente inferidas*.

El fenómeno de eliminación de consecuencias se concreta por medio del propio cálculo del sistema, mediante reglas especiales. Es decir, los cálculos de LNM permiten tanto obtener como eliminar conclusiones, y en esto difieren radicalmente del cálculo de LE, donde sólo es posible lo primero. A continuación, se ejemplificará la eliminación de conclusiones en un sistema de *lógica default* (LD). El ejemplo es una versión ampliada del clásico argumento de Tweety[62]:

(1) $Ave(Tweety)$ [Axioma]

(1a) $Desconocido(\neg Vuela(Tweety))$ [*Oráculo*, regla de inferencia no deductiva que puede entenderse como un algoritmo que mapea la base de conocimiento[63] en busca de un determinado valor semántico]

(1b) $Desconocido\big(\neg Vuela(x)\big) \wedge Ave(x) \Rightarrow Tentativamente(Vuela(x))$

[*Default*, condicional por defecto]

[62] Cfr. (Perlis, 1987).

[63] En IA, el concepto de base de conocimiento remite a un conjunto de sentencias que o bien son axiomas del sistema, o bien son teoremas. Las mismas representan el conjunto del conocimiento del agente, es decir la totalidad de sus premisas y las consecuencias obtenidas de estas.

(1c) $Tentativamente(Vuela(x)) \Rightarrow Vuela(x)$ [Regla de inferencia no deductiva denominada *salto a la conclusión*]

(2) $Vuela(Tweety)$ [Consecuencia no deductiva obtenida a partir de todas las líneas anteriores]

(3) $Avestruz(Tweety)$ [Nuevo axioma]

(3a) $Avestruz(x) \longrightarrow \neg Vuela(x)$ [Axioma]

(4) $\neg Vuela(Tweety)$ [Consecuencia deductiva de (3), (3a)]

(4a) $Vuela(x) \wedge \neg Vuela(x) \longrightarrow Suprimir(Vuela(x)) \wedge Atípico(x)$

[Regla de eliminación]

(4b) $Atípico(Tweety)$ [Consecuencia deductiva de (2), (4), (4a)]

(4c) $Suprimir(Vuela(Tweety))$ [Consecuencia deductiva de (2), (4), (4a)]

Como se observa en este caso, la eliminación de conclusiones ocurre como un subproducto de la emergencia de una inconsistencia (entre las líneas (2) y (4)). Como se mencionaba en la **INTRODUCCIÓN** de este capítulo, la gestión de inconsistencias es uno de los principales *negocios* de la LNM, negocio que se remonta a los orígenes de la disciplina con la propuesta de M. Minsky. Volviendo a las definiciones informales **2.1.3-5**, resulta interesante mencionar que en este grupo de definiciones se observa nuevamente, al igual que en **1.1.1-3** para el caso del concepto de monotonía, la dicotomía entre razonamientos y sistemas, y la recurrencia del concepto de *información novedosa*. Al respecto, sería razonable suponer que para no monotonía son aplicables consideraciones similares a las establecidas para la propiedad de monotonía en el **CAPÍTULO 1**, en tanto referente negativo del concepto de no monotonía. Por un lado, no monotonía debería entenderse como una característica estructural del cálculo especificada en reglas o procedimientos de eliminación, como es patente en la línea (4a) del ejemplo anterior. Del mismo modo, en cuanto a la relación entre inferencias e información, en tanto la propiedad de monotonía se cumple independientemente del tipo de información adicional relativamente a un conjunto dado de sentencias, sería razonable esperar que la condición de no monotonía sea independiente del tipo de información adicional. Pero esto conllevaría una posibilidad extraña: en un sistema no monótono con base de conocimiento consistente debería existir tanto la posibilidad de eliminación de conclusiones ante la adición de sentencias que representan información novedosa (no son consecuencias deductivas de la base de conocimiento), tal como el axioma de la línea (3); pero también ante la adición de sentencias que son consecuencias deductivas de la propia base de conocimiento, es decir, ante la emergencia de información no novedosa. Posteriormente en la sección **2.3** se volverá sobre este asunto.

2.2 SIGNIFICADO DE NO MONOTONÍA: NEGACIÓN DE MONOTONÍA, DEFINICIÓN SEMI-FORMAL Y APROXIMACIÓN INDIRECTA

Una expresión puramente formal, metamatemática, de la propiedad en cuestión (preferentemente en términos “tarskianos”) facilitaría el análisis y la contrastación de las definiciones informales, en virtud de la potencia analítica del abordaje de las propiedades metalógicas en términos conjuntistas. Lamentablemente, no hay evidencias de la existencia de una expresión ese tipo en la bibliografía metalógica sobre LNM. Tomando como referente negativo de la noción de no monotonía el propio concepto de monotonía, expresado en el teorema **1.2.6**, cabe preguntarse ¿es factible representar formalmente no monotonía mediante la mera negación del teorema de monotonía? A simple vista, este movimiento vendría a reflejar en el plano metalógico la lectura minimalista monotonía expresada en el postulado **2.1.2**. Sin embargo, como se explayará a continuación, esa opción genera resultados indeseados, los cuales posiblemente expliquen la ausencia de una definición metamatemática que exprese de un modo directo la idea de *negación de monotonía*. Se intentará expresar formalmente tal idea tomando como referente negativo el teorema **1.2.6**. Dado que esa fórmula aplica a un operador Cn de LE, se deberá reemplazar ese operador por uno que represente un cálculo de LNM. En (Makinson, 1994) se propone la utilización del operador "C", que simboliza un cálculo no monótono genérico.

2.2.1 $\neg((\Delta \subseteq \Gamma \subseteq S) \rightarrow (C(\Delta) \subseteq C(\Gamma)))$

La expresión **2.2.1** puede ser interpretada como la siguiente expresión de segundo orden:

2.2.2 $\neg\forall\Delta\forall\Gamma((\Delta \subseteq \Gamma) \rightarrow (C(\Delta) \subseteq C(\Gamma)))$

A su vez, esto equivale a:

2.2.3 $\exists\Delta\exists\Gamma((\Delta \subseteq \Gamma) \wedge (C(\Delta) \nsubseteq C(\Gamma)))$

En otras palabras, definir no monotonía como la negación de monotonía implica que *debe* existir efectivamente un Δ cuya ampliación resigne conclusiones. Esta condición establece una exigencia demasiado fuerte, y choca con la idea de no monotonía como *posibilidad* de eliminación de conclusiones (postulado **2.1.6**). Quizá por este motivo, los intentos de representación formal del concepto de no monotonía suelen seguir un camino que podría denominarse "indirecto", en tanto comienzan con una caracterización informal o semi-formal no conjuntista, y continúan con la formulación de propiedades formales *asociadas*, denominadas a menudo “propiedades centrales” de la LNM, las cuales se espera que sirvan para clarificar el significado del concepto. Este enfoque ha sido iniciado por D. Gabbay en su (1985), trabajo pionero sobre temas metalógicos de LNM, y continuado por D. Makinson y otros autores, llegando a constituirse en una referencia ineludible. Gabbay comienza por definir “monotonía” como una propiedad según la cual si $P \vdash A$, entonces $P \cup P' \vdash A$, mientras que para “no monotonía” propone la siguiente definición semi-formal donde se introduce por primera vez el símbolo de *relación de consecuencia no monótona* “$|\sim$” o *torniquete no monótono* (TNM):

2.2.4 En sistemas no monótonos, en ocasiones ocurre que $P \mid\sim A$ pero no ocurre que $P \cup P' \mid\sim A$ (Gabbay, 1985: 439-440).

Esta descripción viene a expresar una intuición similar a la contenida en el postulado **2.1.6**, en tanto expresa la *posibilidad* de eliminación de conclusiones ante la ampliación del conjunto de premisas. La relación de consecuencia no monótona o TNM remite a un cálculo donde existen reglas o procedimientos tanto para inferir como para eliminar conclusiones. El conjunto particular de reglas/procedimientos definirá las características específicas de cada sistema de LNM, pero independientemente de la diversidad de cálculos posibles entre esas reglas es obligatorio que existan reglas o procedimientos de eliminación. Este sería uno de los requisitos imprescindibles para un sistema para ser considerado no monótono. Posteriormente, el autor postula un conjunto de propiedades asociadas a su definición semi-formal, agrupadas bajo su *definición 18*, las cuales tomadas de conjunto expresarían el significado del propio concepto de no monotonía. Cabe destacar que existen dos posturas respecto al status metalógico de estas propiedades: la postura fuerte del propio Gabbay según la cual las mismas describen las características esenciales y universales de los sistemas no monótonos[64]; y por otra parte una postura débil o prudente, propuesta por D. Makinson, según la cual tales propiedades sólo describen ciertas *características deseables* o incluso *interesantes* de los sistemas no monótonos. Mientras que Gabbay utiliza una notación inspirada en Gentzen, presentando las propiedades a modo de reglas estructurales, Makinson (1994) reformula las mismas en términos conjuntistas, notación que se utilizará a continuación. La primera propiedad es *inclusión*[65]. Dado un operador C no monótono, se postula la condición de inclusión (análoga a **1.2.3** para un operador puramente deductivo Cn), que afirma que todas las sentencias de una base de conocimiento pertenecen a la clausura no monótona de tal base:

2.2.5 $\Delta \subseteq C(\Delta)$

Junto a inclusión se postula la propiedad de *cumulatividad,* que puede entenderse como una condición de estabilidad del sistema (agente)[66], pues al incorporar a un conjunto de sentencias Δ (base de conocimiento) las propias consecuencias de Δ ocurre que: 1) no se pierden consecuencias respecto a $C(\Delta)$; 2) no se obtienen nuevas consecuencias relativamente a las que ya estaban presentes en $C(\Delta)$:

2.2.6 $(\Delta \subseteq \Gamma \subseteq C(\Delta)) \to (C(\Delta) = C(\Gamma))$

La propiedad de cumulatividad es simplemente un modo de representar conjuntamente otras dos propiedades: *monotonía cauta* y *corte*[67]. Monotonía cauta expresa el comportamiento del sistema ante la incorporación de sentencias de $C(\Delta)$ a Δ (lo que genera un conjunto ampliado Γ), que garantiza la condición expresada en la cláusula 1) de **2.2.6**. Esta propiedad establece una *garantía de conservación de consecuencias con*

---

[64] En tanto la definición 18 en (Gabbay, 1985) responde a la pregunta *¿qué es una lógica no monótona?*
[65] En (Gabbay, 1985): *reflexividad*.
[66] Cfr. (Makinson, 1994)).
[67] En (Dov Gabbay, 1985): *monotonía restringida* y *transitividad*, respectivamente.

restricciones, que operaría sólo en casos especiales, cuando el agente *acumula sus conclusiones en las premisas*. Se supone que el proceso de acumulación de conclusiones en las premisas sin pérdida de conclusiones previamente obtenidas es una característica común del razonamiento práctico. En (Makinson, 1994) se afirma que monotonía cauta puede ser entendida como una especie de *irreversibilidad* en la obtención de conclusiones de un agente. Es decir que una vez inferida, una proposición puede ser retenida independientemente de qué otras proposiciones inferidas por el agente sean agregadas al "stock de información útil" (Makinson, 1994:44).

2.2.7 $(\Delta \subseteq \Gamma \subseteq C(\Delta)) \rightarrow (C(\Delta) \subseteq C(\Gamma))$

Corte puede entenderse como la condición del sistema que expresa el hecho descripto en la cláusula 2) de **2.2.6** para el caso cuando se agregan a un conjunto Δ sólo sentencias que son consecuencias de Δ:

2.2.8 $(\Delta \subseteq \Gamma \subseteq C(\Delta)) \rightarrow (C(\Gamma) \subseteq C(\Delta))$

Asimismo, se formula la propiedad de *supraclasicalidad* en (Makinson 1994:45) para sistemas no monótonos cerrados bajo conectivas veritativo-funcionales. Gabbay afirmaba en (1985:446) que tal propiedad (a la que denomina "compatibilidad de $|\!\sim$ con $\vdash$") es común a todos los sistemas no monótonos. Esta propiedad hace redundante la propiedad de inclusión:

2.2.9 $Cn(\Delta) \subseteq C(\Delta)$

Finalmente, se presentará la propiedad de *conservación de la consistencia*. Como se mencionó en la **INTRODUCCIÓN** a este capítulo, una de las críticas de M. Minsky a la LE, que dio impulso al surgimiento de la LNM, se dirige contra el principio de consistencia (**1.2.13**), que junto a la ley de explosión lógica (**1.2.12**) expresan la imposibilidad de manipular contradicciones en LE. No obstante, en LNM no se renuncia por completo a la idea general de un principio de consistencia. La clave es que, si bien no puede cumplirse un principio tan fuerte como el de LE, al menos puede garantizarse que, en tanto y en cuanto un subconjunto de las creencias de un agente sea clásicamente consistente, sus consecuencias no monótonas también lo serán. En (Gabbay, 1985:446) se agrupa esta propiedad junto con la de supraclasicalidad (**2.2.9**) bajo el nombre de "consistencia de la información nueva con la información vieja": si Δ es consistente, $\Delta \cup C(\Delta)$ también lo será. En (Makinson, 1994:51) esta condición se presenta formalmente del siguiente modo (donde $\mathcal{L}$ es un lenguaje de LNM):

2.2.10 $(Cn(\Delta) \neq \mathcal{L}) \rightarrow (C(\Delta) \neq \mathcal{L})$

Por principio de consistencia de LE, si $Cn(\Delta)$ es consistente, Δ también lo es, de modo que **2.2.10** equivale a aseverar que en un sistema de LNM no puede ocurrir que Δ sea consistente y $C(\Delta) = \mathcal{L}$. En otras palabras, que el cálculo no debería generar inconsistencias, a menos que la base de conocimiento sea inconsistente. Por otra parte, como es esperable, no está garantizado que, si el conjunto de consecuencias es consistente, también lo sea la base de conocimiento (como ocurre en **1.2.13**), en tanto la LNM cuenta con reglas de eliminación que permiten suprimir inconsistencias. Con esta propiedad se completa el cuadro de *propiedades centrales* o asociadas, que (siguiendo a Gabbay) perfilarían un significado general de *no monotonía* mediante una especie de "aproximación indirecta"[68]. Ahora bien, cabe preguntarse por el status de estas propiedades centrales. Como se mencionó antes, en (Gabbay, 1985) sin dudas existe una pretensión de universalidad en el sentido de que se supone que tales propiedades son

[68] Por supuesto, existen otras propiedades y condiciones que aplican a sistemas de LNM, muchas de las cuales se presentan en (Makinson, 1994).

satisfacibles por todos los sistemas de LNM, de hecho, para el autor las mismas definen el propio concepto de no monotonía.

Por otra parte, en (Makinson, 1994) se adopta la perspectiva de Gabbay en relación a enfocar las LNM's mediante una relación de consecuencia no monótona (TNM) y un conjunto de propiedades. Sin embargo, el autor se refiere a estas últimas como propiedades meramente "deseables" o incluso "interesantes" (Makinson, 1994:39). Quizá consciente de la existencia de objeciones y contraejemplos a las propiedades centrales, Makinson realiza estas salvedades y termina por advertir que posiblemente no existe una única relación de consecuencia no monótona ni un único conjunto de propiedades asociadas. Puede interpretarse la postura de Makinson como enrolada en la tradición de Gabbay, pero sin adherir al universalismo de este último, el cual es reemplazado por una postura mucho más prudente que denomina *pluralista*. De cualquier modo, la principal diferencia con relación al status de las propiedades centrales parece ser *ad hoc*, o al menos, sólo de matiz o énfasis: para Gabbay son *propiedades universales*, mientras que para Makinson son *propiedades interesantes*; es decir, no se descarta que sistemas de LNM no satisfagan alguna de ellas, pero aun así conforman un núcleo teórico central para el área. Esto se hace patente a lo largo del artículo de referencia, y en ese sentido, puede afirmarse que la *pretensión de universalidad* de Gabbay se convierte en *pretensión de generalidad* en Makinson.

Lo que genera este debilitamiento del status de las propiedades centrales es la existencia en la bibliografía del área de un conjunto de problemas teóricos, objeciones y contraejemplos a las propiedades centrales, algunos de los cuales serán consignados a continuación. En primer lugar, se debe tomar en consideración el *factum* de que ningún agente racional es completamente consistente. Este fenómeno es descripto crudamente por D. Perlis cuando afirma, como un aspecto insoslayable de la vida cotidiana, que "la inconsistencia es el estado de cosas normal en razonamiento de sentido común" (Perlis, 1987: 180). Esta afirmación fáctica se fundamenta teóricamente en una paradoja descubierta por D. Israel en su (1980):

2.2.11 Paradoja de Israel: *Cualquier sujeto racional debe estar dispuesto a admitir que al menos una de sus creencias es falsa. Por lo tanto, el sujeto posee un conjunto de creencias inconsistente, dado que no hay interpretación posible bajo la cual todas sus creencias sean verdaderas.*

Por definición, un agente racional asigna el valor *verdadero* a sus creencias, en tanto (se supone que) no es razonable creer en falsedades. Pero si el agente acepta que al menos una de sus creencias (sentencias a las que atribuye el valor verdadero) es falsa, se ve obligado a admitir que su base de conocimiento es inconsistente. Teniendo en cuenta que uno de los objetivos principales de la LNM es la modelización del razonamiento práctico, tal como es efectuado por agentes racionales concretos en la vida cotidiana (a fin de emularlo computacionalmente), la paradoja de Israel genera un impacto concreto en la propiedad de *conservación de la consistencia*, una de las más importantes en metalógica de LNM. Por un lado, si un sistema no monótono trabaja con bases de conocimiento inconsistentes, **2.2.10** se cumple trivialmente y pierde relevancia. Por otra parte, si el sistema no admite bases inconsistentes, mediante la introducción de algún tipo de cláusula ad hoc[69], la propiedad se cumpliría de manera no trivial y ganaría en importancia, pero tal sistema se estaría alejando irremediablemente del comportamiento de los razonadores de sentido común, lo cual resultaría al menos extraño dados los objetivos de la LNM. Pero no sólo la *inconsistencia generalizada* de las bases de conocimiento plantea un problema para esta propiedad **2.2.10**, pues no todos los sistemas de LNM la satisfacen.

[69] Como al parecer ocurre en sistemas basados en *lógica modal generalizada*. Véase **CAPÍTULO 3** sección **2**.

Mientras que para Gabbay es una propiedad universal, el propio Makinson resalta que algunos sistemas preferenciales no la cumplen[70]. Según lo mencionado en el capítulo anterior, existen sistemas de LNM que literalmente introducen contradicciones allí donde las había, pues, independientemente de las características específicas de tales sistemas, a priori el hecho de que no satisfagan **2.2.10** implica que aun siendo $Cn(\Delta) \neq \mathcal{L}$ ocurre $C(\Delta) = \mathcal{L}$.

Ahora bien, ante este cuadro de situación, se impone una breve digresión. En primer lugar, sobre los posibles vínculos entre la aporía de la inconsistencia del sistema $\mathcal{L}$, la paradoja de Israel y la generación de inconsistencias en sistemas de LNM, sólo puede especularse filosóficamente. Pero resulta interesante desde el punto de vista metadisciplinario, que tanto en el campo de la lógica puramente deductiva como en el ámbito de la LNM se haya arribado, por vías diferentes, a resultados similares en relación con el tópico clásico de *exigencia de consistencia* del conocimiento como parámetro normativo, a saber: tal exigencia (en cualquiera de sus formas) debería abandonarse. En la propuesta clásica tarskiana, la alternativa es *consistencia o explosión lógica*, pero la inconsistencia ha logrado *colarse* en el sistema deductivo más importante de todos: el *sistema de todos los sistemas*. En LNM, mientras algunos autores realizan esfuerzos por formular variantes más débiles del principio de consistencia **2.2.13**, otros descubren la generación de inconsistencias a partir de bases consistentes en ciertos sistemas (derribando cualquier intento de ese tipo), y además se descubre el fenómeno concreto de inconsistencia generalizada de los razonadores de sentido común. Quizá estas situaciones reflejan, con diferentes niveles de abstracción, un fenómeno mucho más profundo que el de meras aporías lógico-filosóficas. Quizá el asunto que está en juego es la posibilidad cierta de

[70] Cfr. (Makinson, 1994:81). Sin embargo, según el autor, los sistemas preferenciales que no satisfacen la propiedad de conservación de la consistencia aún pueden aspirar a satisfacer una *versión debilitada* de la misma, denominada *preservación relativa de consistencia* (Makinson, 1994:51).

una *inconsistencia global del conocimiento humano* como estado normal de cosas. [Fin de digresión]

Volviendo a la paradoja **2.2.11**, en LNM existen dos enfoques para lidiar con esa situación. Una posibilidad es la *restauración de la consistencia*, pero también puede optarse por *tolerar* las inconsistencias, sin eliminarlas (Befenhart et al., 1997:18). En el primer caso, ocurre que de un Δ inconsistente se obtiene $C(\Delta)$ consistente, tal como se ejemplificará a continuación. Para el siguiente ejemplo, se considera una relación de consecuencia no monótona *libre* basada en subconjuntos máximos consistentes (SMC) de una base de conocimiento[71]. La idea intuitiva es que es las consecuencias libres de una base de conocimiento son las sentencias que pertenecen a la intersección de los SMC de la base considerada.

2.2.12 Sea $\Delta = \{p \wedge q, \neg p \wedge q, r\}$; donde $SMC_1 = \{p \wedge q, r\}$ y $SMC_2 = \{\neg p \wedge q, r\}$. En consecuencia, el conjunto $Libre(\Delta) = SMC_1 \cap SMC_2$ y por lo tanto $C(\Delta) = \{r\}$. Si posteriormente se adiciona la sentencia $\neg r$ a la base de conocimiento, se obtiene $\Gamma = \Delta \cup \{\neg r\}$ por lo cual $C(\Gamma) = \emptyset$.

En **2.2.12**, que es una versión con diferencias de notación de un ejemplo presentado en (Strasser & Antonelli, 2019:3), se observa cómo se restaura la consistencia, pues ocurre que Δ es inconsistente, mientras que $C(\Delta)$ es consistente. Asimismo, el operador C hace fallar monotonía cuando se amplía la base de conocimiento mediante la adición de $\neg r$. Pero lo más llamativo de este ejemplo es un aspecto que no ha sido tenido en cuenta por los autores citados, y es que el mismo muestra que las relaciones de consecuencia no monótonas que restauran la consistencia también pueden hacer fallar la propiedad de

[71] Cfr. (Rescher & Manor, 1970).

inclusión (en Gabbay, *reflexividad*), en tanto en el ejemplo $\Delta \supset C(\Delta)$ y $\Gamma \supset C(\Gamma)$. Al fallar inclusión/reflexividad, tampoco se verifica supraclasicalidad, pues como se ejemplifica en **2.2.12** para alguna sentencia x se cumple que $x \in \Delta$, y dada la propiedad de inclusión para sistemas deductivos $\Delta \subseteq Cn(\Delta)$ ocurre que $x \in Cn(\Delta) \wedge x \notin C(\Delta)$. En definitiva, inclusión/reflexividad y supraclasicalidad parecen resultar aceptables sólo para sistemas que no restauran la consistencia. Esto puede considerarse una verdadera aporía metalógica para la LNM, ya que al parecer el precio a pagar por manipular contradicciones de este modo (restaurando consistencia) es bastante alto: la *no satisfacción* de dos propiedades muy importantes. Vale destacar que si bien el ejemplo **2.2.12** no es original del autor de esta tesis, sí lo es el análisis que lo identifica como un contraejemplo de inclusión/reflexividad y supraclasicalidad.

Continuando con el análisis de las propiedades centrales, hay que agregar que la propiedad de *monotonía cauta* ha sido puesta en cuestión reiteradamente. En (Prakken & Vreeswijk, 2001) se presentan varios contraejemplos a esta propiedad, uno de los cuales especifica el comportamiento de sistemas probabilísticos en una situación donde no hay inconsistencias en la base de conocimiento:

2.2.13 Dada una base de conocimiento $\Delta = \{p, s\}$; un cálculo que contiene reglas probabilísticas y gestiona las inconsistencias mediante un criterio de comparación de la fuerza relativa de los enlaces inferenciales; y los siguientes esquemas de argumentación, donde "$\longmapsto$" representa pasos de inferencia no especificados:

(1) $p \longmapsto q \longmapsto r$

(2) $s \longmapsto \neg r$

Ante inconsistencias, el sistema compara las cadenas de argumentación que soportan a cada conclusión contradictoria, en busca del enlace inferencial más débil. La fuerza de los enlaces es cuantificada en términos probabilísticos, fluctuando entre 0 y 1. En el ejemplo, en la cadena (1) la fuerza de la derivación de q desde p es 0.7 y de r desde q es 0.85; mientras que en (2) la fuerza de la derivación de $\neg r$ desde s es 0.8. Ante la inconsistencia, el sistema optará por eliminar el término contradictorio cuya derivación contenga el enlace inferencial con menos fuerza, de modo que (2) derrota a (1) porque su enlace más débil es 0.8 mientras que el enlace más débil de (1) es 0.7. Esta situación no presenta dificultades en relación a la propiedad de monotonía cauta. Pero el asunto cambia si en vez de (1) se tiene:

(1a) $p \longmapsto q$

En este caso, q y $\neg r$ están justificados debido a que no caen en contradicción con ninguna otra sentencia. Por lo tanto, pueden ser agregados a la base de conocimiento y de ese modo $\Gamma = \{p, s, q, \neg r\}$. Ahora bien, a partir de q puede construirse un nuevo argumento:

(1b) $q \longmapsto r$

En este caso, ante la inconsistencia el análisis arroja que el enlace más débil de (1b) es 0.85, por lo que el sistema optará por eliminar la conclusión de (2). Así, falla monotonía cauta en tanto se cumple $\Delta \subseteq \Gamma \subseteq C(\Delta)$ pero $\exists x(x \in C(\Delta) \wedge x \notin C(\Gamma))$. En otras palabras, se pierden conclusiones previamente obtenidas al ampliar la base de conocimiento mediante la adición de sentencias que son consecuencias de esa misma base. [Fin del contraejemplo]

Del ejemplo **2.2.13** se sigue trivialmente $(\Delta \subseteq \Gamma \subseteq C(\Delta)) \wedge (C(\Delta) \neq C(\Gamma))$, con lo cual también queda refutada la propiedad compuesta de cumulatividad, aunque esto no es suficiente para refutar la propiedad de corte. Para esta última propiedad, existen contraejemplos también basados en razonamiento probabilístico. La clave de estos contraejemplos reside en un sistema probabilístico atribuye un grado de apoyo de las premisas a la conclusión que disminuye a medida que la extensión de la prueba crece. Intuitivamente, cuando la probabilidad condicional $p(\beta|\alpha)$ es mayor que $p(\gamma|\alpha \wedge \beta)$ podría debilitarse demasiado la inferencia $\alpha|\!\sim\gamma$ al aplicar corte. En (Antonelli, 2005:5) se ofrece un contraejemplo de este tipo. Por último, resulta interesante mencionar el caso de otra propiedad que ha sido propuesta para sistemas preferenciales, pero que ha sido refutada posteriormente, tal es el caso de *monotonía racional*[72]:

2.2.14 $(\neg\alpha \notin C(\Delta) \wedge\ \beta \in C(\Delta)) \rightarrow (\beta \in C(\Delta \cup \{\alpha\}))$

Esta propiedad establece que no se pierden conclusiones de una base de conocimiento Δ si se le adicionan sentencias cuya negación no es una consecuencia de la base de conocimiento considerada. En otras palabras, si las sentencias adicionales no son inconsistentes con las consecuencias de la base de conocimiento, no deberían forzar una revisión de creencias. A simple vista esta condición resultaría aceptable para sistemas de LNM, pero le ha sido opuesta una objeción en la forma del siguiente contraejemplo[73]:

2.2.15 Considere tres compositores: Verdi, Bizet y Satie. La información disponible inicialmente indica que Verdi es italiano, mientras que Bizet y Satie son franceses. Así, para la base de conocimiento $\Delta = \{It(v), Fr(b), Fr(s)\}$ se sigue no monotónicamente $C(\Delta)\ = \{It(v), Fr(b), Fr(s)\}$. Posteriormente una fuente confiable afirma que Verdi y Bizet son compatriotas. Esta proposición adicional impide inferir tanto $It(v)$ (debido a que Verdi podría ser francés), como $Fr(b)$ (pues Bizet podría ser italiano). Sin embargo,

[72] Cfr. (Lehmann & Magidor, 1992).
[73] Cfr. (Antonelli, 2005: 8).

es posible seguir sosteniendo que $Fr(s)$. De este modo, $C(\Delta \cup \{Co(v,b)\}) = \{Fr(s)\}$. Ahora bien, si posteriormente se agrega una proposición que afirma $Co(v,s)$, ya no será posible seguir sosteniendo $Fr(s)$. De ese modo, finalmente $C(\Delta \cup \{Co(v,b)\} \cup \{Co(v,s)\}) = \{Co(v,b), Co(v,s)\}$, lo cual significa que Verdi, Bizet y Satie son todos de la misma nacionalidad. Esta situación refuta la propiedad de monotonía racional, en tanto $\neg Co(v,s) \notin C(\Delta \cup \{Co(v,b)\})$, $Fr(s) \in C(\Delta \cup \{Co(v,b)\})$ y $Fr(s) \notin C(\Delta \cup \{Co(v,b)\} \cup \{Co(v,s)\})$. [Fin de contraejemplo]

Los contraejemplos mencionados hasta aquí no sólo ponen de manifiesto las dificultades existentes para el establecimiento de propiedades de LNM, sino que además afectan, en términos generales, el fundamento metamatemático de estas lógicas, pues bloquean cualquier intento de definir la propiedad de no monotonía mediante algún tipo de *aproximación indirecta*, tal como proponía Gabbay con sus propiedades centrales.

2.3 SIGNIFICADO DE NO MONOTONÍA: PROPUESTA DE DEFINICIÓN FORMAL

Del análisis realizado en este capítulo se desprenden algunas conclusiones en relación con la clarificación del significado del concepto de no monotonía. En primer lugar, se presentaron algunas definiciones informales, que permitieron delinear dos postulados básicos respecto a las acepciones más difundidas en la bibliografía: **2.1.2** y **2.1.6**. Asimismo, se descartó la versión más simplista reflejada en el postulado **2.1.2** en tanto su expresión formal en términos conjuntistas (la negación del teorema **1.2.6**) genera resultados indeseados. Posteriormente, tomando como referencia la definición semi-formal **2.2.4** de Gabbay, quedó claro que el asunto es un poco más complejo que la simple negación de la propiedad de monotonía, y que la noción fundamental que debería ser reflejada en cualquier definición formal de no monotonía es la de *posibilidad de eliminación de consecuencias* previamente obtenidas (en línea con el postulado **2.1.6**). Este resultado podría traducirse, según la distinción entre contextos intensionales/extensionales en (GAMUT, 2009), como la detección de serias dificultades para una definición de la propiedad en cuestión en una perspectiva extensional, y la intuición de una especificación que remite a un contexto intensional, para ese fin.

Por otra parte, se han puesto de relieve las dificultades de una *aproximación indirecta* al concepto mediante la formulación de las propiedades centrales, las cuales han sido objetadas poniéndose en tela de juicio su status metateórico. Debido a esto, es difícil sostener actualmente, como lo hacía Gabbay en su (1985), que éstas expresan el significado estricto de no monotonía, y que son *propiedades universales* dentro del ámbito de la LNM. El enfoque de Makinson respecto a las propiedades centrales logra salvar algunas de las dificultades derivadas de la pretensión de universalidad originaria de la propuesta de Gabbay, pero paga el precio de renunciar a definir metamatemáticamente no monotonía (mediante una aproximación indirecta). Una formulación directa y sin rodeos de una expresión puramente simbólica que permita dar cuenta de esta última propiedad en el terreno metalógico, reflejando la idea de *posibilidad de eliminación de consecuencias*, continúa siendo una cuenta pendiente para la metalógica orientada a LNM.

A continuación, se realizará una propuesta para la construcción de una expresión de este tipo, según el principio de *ausencia de garantía de conservación de consecuencias*. Teniendo en cuenta el doble carácter descriptivo/prescriptivo de la metamatemática y la metalógica[74] (véase **CAPÍTULO 1** sección **2**), puede afirmarse que una expresión de ese tipo no sólo estaría *describiendo* el comportamiento de los sistemas de LNM, sino también *prescribiendo* una característica que se considera primordial para que un sistema sea considerado una LNM. La siguiente formulación de la propiedad de no monotonía toma como punto de partida la definición semi-formal **2.2.4**, ampliamente aceptada en el ámbito de la LNM. Además, extrapola el operador modal "□" al terreno de la metalógica. Una interpretación conveniente del operador es análoga a la que recibe en el sistema modal T (sistema K + axioma de necesitación[75]).

2.3.1 $\neg\Box((\Delta \subseteq \Gamma \subseteq S) \rightarrow (C(\Delta) \subseteq C(\Gamma)))$

Suponiendo que la expresión **2.3.1** es falsa, se sigue necesariamente que si $\Delta \subseteq \Gamma$, entonces no hay ningún mundo posible donde $\exists x(x \in C(\Delta) \wedge x \notin C(\Gamma))$. Pero de **2.2.4** se interpreta que eso sí es posible. En tanto esta última definición de Gabbay es ampliamente aceptada en LNM, puede afirmarse que es inadmisible tal contradicción, por lo que la propiedad **2.3.1** se cumple para sistemas de LNM. Por otra parte, en una interpretación que extrapole los axiomas y teoremas necesarios del sistema T hacia la metalógica, **2.3.1** equivale a:

2.3.2 $\Diamond\neg((\Delta \subseteq \Gamma \subseteq S) \rightarrow (C(\Delta) \subseteq C(\Gamma)))$

En ese contexto, una especificación semántica admisible es la siguiente: dado un modelo M con un conjunto de mundos posibles W, una relación de accesibilidad R y una función de verdad V, **2.3.2** será válido si para todo mundo $w \in W$ existe al menos un $w' \in W$ tal que wRw' y $(\Delta \subseteq \Gamma) \wedge (C(\Delta) \nsubseteq C(\Gamma))$ es verdadero en w'. Intuitivamente, la expresión indica que es posible que falle monotonía en el modelo. Otra interpretación viable de

[74] E incluso de la *lógica* misma como disciplina global, idea que se defiende en (Antonelli, 2005:ix).
[75] Cfr. (Hughes & Cresswel, 1996).

2.3.1 consiste en considerar el operador modal como remitiendo a mundos posibles en la tradición de la teoría de modelos preferenciales, la cual será profundizada en el **CAPÍTULO 3**. La idea básica detrás de esta interpretación consiste en considerar un marco $< M, \prec >$, donde $\prec$ indica un orden de preferencias entre los mundos de M. Entonces, **2.3.1** es válida en el marco porque para algún mundo $m \in M$ existe un $m' \in M$ tal que $m' \prec m$ y $(\Delta \subseteq \Gamma \subseteq S) \rightarrow (C(\Delta) \subseteq C(\Gamma))$ es falso en m'.

Volviendo a la expresión **2.2.1**, nótese que la misma refleja la simple negación del teorema de monotonía **1.2.6**, y en ese sentido se trata de la negación de una implicación material, lo cual genera resultados indeseados debido a las particulares condiciones en las cuales se verifica la negación de un condicional de ese tipo. Por su parte, **2.3.1** puede entenderse como la negación de un condicional estricto[76], donde se ha aplicado teorema de necesitación $A \Rightarrow \Box A$ (si A es un teorema, entonces "es necesario A" es un teorema[77]) al teorema **1.2.6**. Esto parece indicar que cualquier definición formal de la propiedad de no monotonía exige recursos de lógica modal. En los **CAPÍTULOS 3** y **4** se hará patente esta estrecha relación entre lógica modal y LNM. De cualquier modo, lo más importante aquí es que la propiedad **2.3.1** no sólo evita las dificultades de la simple negación de monotonía, sino que refleja efectivamente la intuición del concepto de no monotonía como *ausencia de garantía de conservación de consecuencias*. Un sistema que satisface **2.3.1-2** tiene una estructura tal que admite al menos un mundo posible en el cual no se cumple monotonía. Así, el sistema se comporta monotónicamente al menos en algunos mundos. Dependiendo de las características del sistema considerado y de la base de conocimiento, quizá en la mayoría (pero no en todos). Una intuición similar a esta última será presentada en el **CAPÍTULO 4**, bajo la conjetura del razonador semimonótono.

[76] Cfr. (Hughes & Cresswel, 1996) cap. 11.
[77] Cfr. (Hughes & Cresswel, 1996:25).

Nótese también que la expresión **2.3.1** no establece ninguna restricción en el aspecto informacional, lo cual parece abrir una extraña posibilidad. Para sistemas de LNM con bases de conocimiento consistentes, se cumple de manera no trivial la propiedad **2.3.3**, equivalente a **2.3.1**:

2.3.3 $\neg\Box\Big(\big((\Delta \cup \Psi) \subseteq S\big) \wedge (\Psi \in P(\Omega_{\Delta}^{1} \cup \Omega_{\Delta}^{2})) \rightarrow \big(C(\Delta) \subseteq C(\Delta \cup \Psi)\big)\Big)$

La posibilidad (no probada, aunque *ausencia de prueba* no significa *prueba de ausencia*) de que sentencias de tipo 1 (relativamente a una base consistente Δ) puedan forzar eliminación de consecuencias chocaría con la idea de que sólo la información novedosa puede hacer fallar monotonía. Pero la inconsistencia generalizada de las bases de conocimiento de los agentes racionales (paradoja de Israel) resta importancia a este problema, el cual queda abierto por el momento. Finalmente, se propone la siguiente definición informal que complementa las definiciones formales **2.3.1-2**:

2.3.4 *La propiedad de no monotonía es una condición del cálculo de LNM según la cual no está garantizada la conservación de consecuencias de una base de conocimiento de un agente racional, cuando se amplía esa base.*

Lo interesante de definir el concepto de no monotonía en términos de **2.3.1-2** y **2.3.4** es que se aborda de manera directa su significado en términos metamatemáticos y metalógicos, sin necesidad de recurrir a propiedades asociadas. Además, se establecen bases metalógicas abstractas que permiten tanto describir como prescribir las condiciones que cumple/debe cumplir un sistema para ser considerado una LNM. Estas ventajas resultan importantes si se toma en cuenta el carácter plural, diverso y caótico[78] de las numerosas teorías y sistemas que conforman el área, donde resulta muy complejo

[78] Cfr. (Makinson, 1994).

establecer un hilo conductor que permita delimitar metalógicamente el campo disciplinario en cuestión. En ese sentido, estas definiciones aportan a la fundamentación metamatemática de la LNM (en el sentido amplio tarskiano de una propiedad abstracta satisfacible por todo sistema de LNM). Ahora bien, una vez ganado un fundamento abstracto universal para las LNM's bajo la forma de una propiedad que debe satisfacer todo sistema para considerarse no monótono, cabe preguntarse ¿qué ocurre con las propiedades centrales? ¿los contraejemplos y objeciones resultan suficientes para descartarlas de plano? En el próximo capítulo se abordará este asunto desde un punto de vista semántico y modelo-teorético.

CAPÍTULO 3 – PROPIEDADES CENTRALES

INTRODUCCIÓN

Habiendo alcanzado resultados que expresan metamatemáticamente la propiedad de no monotonía (**2.3.1-2**), y luego de indicar algunos contraejemplos y objeciones a las propiedades centrales de la LNM, en este capítulo se abordarán los siguientes ejes teóricos: en primer lugar, se ofrecerá un recorte temático y una síntesis de los fundamentos del enfoque semántico de la LNM, específicamente de la *teoría de modelos preferenciales* (sección **3.1**) y de la *lógica modal generalizada* (sección **3.2**). Dentro de este enfoque semántico, se expondrá de qué modo pueden superarse algunas de las dificultades que afrontan las propiedades centrales, mencionadas en el **CAPÍTULO 2**, junto a otros resultados metalógicos relacionados y/o de interés. Por otra parte, al presentar los fundamentos del enfoque semántico de LNM y mostrar de qué modo pueden restaurarse las propiedades centrales, se estarán sentando las bases para formular posteriormente una alternativa de la propiedad de no monotonía, basada en la *conjetura del razonador semimonótono*. Por último, a lo largo del capítulo se integrará un eje filosófico con digresiones y análisis de temas tales como la interpretación de la relación de consecuencia no monótona y la relación entre lenguaje objeto y metalenguaje. Este eje tendrá mayor importancia en sección **3.1**, por lo que recibirá un tratamiento relativamente más extenso y profundo, mientras que en la sección **3.2** se intentará minimizar las digresiones de este tipo.

Los orígenes del enfoque semántico de la LNM se remontan a la propuesta de Y. Shoham (1987) donde se introduce una teoría de modelos basada en *modelos preferenciales*. No obstante, el clásico artículo de S. Krauss, D. Lehmann y M. Magidor[79] (Kraus et al., 1990) resulta una referencia ineludible, en tanto ofrece una teoría semántica rigurosamente estructurada, junto a resultados de representación. El citado artículo contiene el primer abordaje sistemático de algunos de los principales problemas metalógicos de la LNM desde una perspectiva semántica. En el mismo, los autores presentan las construcciones semánticas y algebraicas básicas de su teoría de modelos preferenciales, que permite no sólo englobar muchos de los sistemas más importantes de LNM, sino también contrastar desde el punto de vista semántico las propiedades centrales, generar nuevas propiedades derivadas, y demostrar teoremas de corrección[80] y completitud.

Los cimientos básicos de la teoría de modelos preferenciales, primordial tanto en la sección **3.1** como en **3.2**, se construyen sobre *estructuras preferenciales* compuestas por conjuntos de mundos posibles y relaciones de preferencia. Para un razonador, un mundo m es *preferible* sobre un mundo n si m es *más normal* que n, y en el mundo m puede concluirse plausiblemente, en base a una proposición A, que otra proposición B es verdadera si todos los mundos que satisfacen A y son más normales entre los mundos que satisfacen A, también satisfacen B. En la propuesta de KLM presentada en (Kraus et al.,

[79] En adelante, *KLM.*

[80] En el original, *soundness*. Aquí se traduce este último término como *corrección*, tomando en cuenta que se busca generar un teorema análogo a lo que en lengua española se conoce como *teorema de corrección* para LE, por ejemplo, en (Quezada, 1995:95).

1990), se incorpora un lenguaje que contiene todas las conectivas proposicionales clásicas. En el artículo de referencia los autores proponen cinco sistemas, cada uno de cuales puede describirse mediante propiedades que definen una teoría de pruebas y una teoría de modelos, junto a teoremas de corrección, completitud y otros teoremas de representación. El sistema C cumulativo, que reconstruye la propuesta clásica de Gabbay, es considerado el sistema más débil que resulta de interés, y a su vez constituye una suerte de *piedra fundacional* para el diseño del resto de los sistemas. Debido a sus características fundacionales y su relevancia temática para los objetivos de esta tesis, en la sección **3.1** se analizarán las características del sistema C (omitiendo el resto de los sistemas de KLM, que son extensiones de este último). Como se mostrará posteriormente, en el marco de este sistema C (y sus extensiones), es posible restaurar las propiedades centrales de la LNM.

En la sección **3.2** se abordará la propuesta de K. Schlechta en su (2004). En el texto citado, el autor propone construir un fundamento semántico común para los sistemas de LNM basado en nociones como *preferencia*, *distancia* y *tamaño*, entre otras. A partir de estas nociones definidas en términos modelo-teoréticos, es posible demostrar *condiciones de coherencia*, propiedades que satisfacen los sistemas que pueden ser reconstruidos mediante aquel fundamento semántico. Entre estas propiedades, se destacan las propiedades centrales. La propuesta de Schlechta se ubica, en términos generales, dentro de la tradición de KLM, aunque con diferencias y matices varios, muchos de los cuales serán explayados posteriormente. En su perspectiva, en tanto las semánticas basadas en modelos preferenciales utilizan mundos posibles, el objetivo pasa por construir una *lógica modal generalizada* (LMG) que funcione como marco para el razonamiento de sentido común. Cabe destacar que esta teoría general no sólo permite restaurar las propiedades centrales, integrando los resultados de KLM, sino que además ofrece un marco conceptual para la formulación de la propiedad de *semimonotonía*, y su fundamento semántico, que serán presentados oportunamente en el **CAPÍTULO 4**. En resumen, las exposiciones subsiguientes intentan representar el núcleo de las propuestas de los autores mencionados, y si bien muchos de los formalismos son presentados con diferencias de notación, al igual que en los capítulos previos se respetará el espíritu del original. Asimismo, se ofrecerá una labor de interpretación y análisis de tales formalismos que no está presente en los textos de referencia. En ese contexto analítico, junto a los ejes fundamentales del capítulo mencionados previamente emerge un eje filosófico (de filosofía de la lógica) que puede resumirse en el debate sobre el uso e interpretación del TNM que realizan KLM y Schlechta. Se trata de una problemática que atraviesa el capítulo, y será desarrollada con cierta profundidad en tanto condiciona los propios *fundamentos filosóficos* de las propuestas de los autores, incluyendo la interpretación de las propiedades metalógicas y los resultados de representación.

3.1 LA PROPUESTA DE KRAUS, LEHMANN Y MAGIDOR

En primer lugar, los autores establecen un lenguaje $\mathcal{L}$ cerrado bajo conectivas proposicionales clásicas. Este lenguaje consiste en el conjunto de todas las fórmulas bien formadas (FBF) a partir de un conjunto de variables proposicionales y conectivas, con las reglas de formación habituales. Luego se establece una semántica sobre un conjunto $\mathcal{U}$ de mundos, y una relación de satisfacción $\vDash$ entre mundos y fórmulas. El conjunto $\mathcal{U}$ (universo de referencia) es “el conjunto de todos los mundos que consideraremos posibles” (Kraus et al., 1990:172), y se define del siguiente modo: entendiendo $\mathcal{L}$ como el conjunto de las FBF´s construidas a partir de un conjunto de variables proposicionales,

$\mathcal{U}$ será "un subconjunto del conjunto de todas las asignaciones de valores de verdad a variables proposicionales". Es decir, aquí se especifica *mundo posible* como una asignación clásica de valores de verdad a letras proposicionales (fórmulas atómicas). A continuación, los autores afirman que "nos reservamos el derecho de considerar universos de referencia que son subconjuntos propios del conjunto de todos los modelos de $\mathcal{L}$" (Kraus et al., 1990:172). Estos universos restringidos permiten modelar *restricciones estrictas* donde sólo se toman en cuenta aquellas asignaciones que satisfacen una FBF considerada. Intuitivamente, si se tiene una proposición que afirma que "Tweety es un ave", y una aserción como "las aves vuelan", se puede inferir plausiblemente que "Tweety vuela" considerando sólo los mundos que satisfacen "las aves vuelan".

En la propuesta de KLM resulta de extrema importancia el TNM ($|\sim$). Como se mencionó en el **CAPÍTULO 2**, en (Gabbay, 1985) se introduce por primera vez este símbolo metalingüístico para representar una relación de consecuencia (no monótona), sobre la que se definen y se demuestran propiedades en términos axiomáticos, análogamente al tratamiento habitual que recibe la relación de consecuencia clásica $\vdash$. Gabbay aclara que utiliza un nuevo tipo de relación de consecuencia "$|\sim$" para indicar que "el sistema es no monótono" (1985:440), y en ese sentido, una expresión de la forma $\alpha|\sim\beta$ debería leerse como "β es una consecuencia no monótona de α". Otras lecturas posibles se obtienen al reemplazar *consecuencia no monótona* por *consecuencia plausible*, *no deductiva*, *derrotable*, etc. Lo importante es notar que el sentido originario de este símbolo metalingüístico es el de una *relación de consecuencia* (sintáctica) en un sistema no monótono, es decir, que admite eliminación de consecuencias previamente inferidas. Gabbay no deja dudas al respecto, cuando en su (1985:440) se refiere a $|\sim$ como una *relación de consecuencia* y en repetidas ocasiones lo caracteriza como una *relación* entre FBF's. La adhesión parcial de KLM a este enfoque se hace patente en varios pasajes del texto de referencia, notablemente al momento de abordar los problemas de representación. Aun así, para los autores el TNM representa un concepto mucho más complejo, de hecho, lo caracterizan como una *metanoción*. Así, el símbolo considerado es interpretado y utilizado de manera mucho más *libre* por decirlo de algún modo, o en términos de los propios autores, de modo “contrario a la terminología habitual” (Kraus et al., 1990:173).

Como se analizará posteriormente, esto puede resultar en dificultades que se ubican a nivel de los fundamentos filosóficos de la propuesta de KLM, pues es difícil saber si el símbolo $|\sim$ representa un elemento del metalenguaje, del lenguaje objeto, o bien si siquiera existe una distinción de este tipo. Por supuesto, no debería construirse un dogma o un culto de los *usos y costumbres* tradicionales de la LE, pero la ausencia de un criterio claro en cuanto al significado de los símbolos utilizados en un texto sobre lógica no parece ser la mejor opción bajo ninguna circunstancia. En ocasiones, los autores utilizan el TNM al modo usual, como un símbolo metalingüístico (relación de consecuencia no monótona) en línea con la propuesta de Gabbay. Por ejemplo, en la definición **3.1** presentada en (Kraus et al., 1990:176). Sin embargo, también interpretan a $|\sim$ prácticamente como un símbolo del lenguaje objeto, cuyo significado es similar al atribuido en lógicas condicionales a la modalidad diádica “>”, que simboliza *implicaciones condicionales* (Kraus et al., 1990:170). Así, una expresión de la forma $\alpha|\sim\beta$, que denominan *aserción condicional*, es leída como “si α normalmente β”. En esta interpretación $|\sim$ funciona como una conectiva binaria (más precisamente, como un tipo de modalidad)[81], que podría interpretarse a grandes rasgos como una *implicación derrotable*. Ante este panorama, es factible calificar la situación como una *aporía del TNM* en la propuesta de KLM.

[81] Los autores rechazan el enfoque que simboliza “normalmente” como una modalidad unaria, y proponen usar $|\sim$ por su carácter binario.

Además de los usos mencionados para el símbolo $|\!\sim$, los autores parecen introducir un tercer significado en tanto definen "relación de consecuencia" (no monótona) como "conjuntos de aserciones condicionales" (Kraus et al., 1990:173), es decir, conjuntos de expresiones de la forma $\alpha|\!\sim\beta$ donde $|\!\sim$ es entendido como una modalidad binaria (normalmente). Sería aventurado conjeturar respecto a si este tercer significado ubica a $|\!\sim$ en el metalenguaje o en el lenguaje objeto, ya que no se ofrecen mayores precisiones al respecto. A lo largo del artículo, el símbolo $|\!\sim$ es utilizado en al menos otros tres sentidos diferentes: como sinónimo de "sistema", como un *operador no monótono* (tal como emerge de la definición de sistema C cumulativo, véase **3.1.1**); y también como una relación de consecuencia semántica (acompañada de un subíndice, como se especificará posteriormente). Pero entre las diferentes interpretaciones de $|\!\sim$ presentes en KLM, las más importantes parecen ser las primeras dos: *relación de consecuencia no monótona* (sentido originario de Gabbay, símbolo metalingüístico); y *modalidad binaria* ("normalmente", lenguaje objeto). El resto de las interpretaciones pueden reducirse, aproximadamente, a estas dos. En ese sentido, puede afirmarse que existe una *dualidad básica* en el significado que los autores atribuyen a $|\!\sim$ en su (Kraus et al., 1990).

Cabe resaltar que esta dualidad de significado no parece afectar la *teoría de modelos preferenciales*, de hecho, la teoría de KLM es ampliamente aceptada como un aporte fundamental para la construcción de una semántica para sistemas de LNM, y el artículo (Kraus et al., 1990) es considerado un clásico en la literatura del área. Resulta interesante que al principio del artículo los autores comienzan definiendo de manera *sui generis* el TNM, pero luego progresivamente a lo largo del texto se observa una *deriva* del significado del mismo que desemboca en una perspectiva más apegada a la tradición (Gabbay). Entre esos dos extremos, se abre la posibilidad de realizar varias lecturas o interpretaciones de algunos resultados, específicamente en lo que concierne a las propiedades de los sistemas y resultados de representación. Particularmente, la propiedad central de *reflexividad* de los sistemas no monótonos puede ser interpretada en (Kraus et al., 1990) como indicando tanto una *relación (binaria homogénea) reflexiva*, como también como una *relación arreflexiva*, según la interpretación que se atribuya al TNM. Posteriormente se regresará sobre este asunto. Finalmente, cabe resaltar que, en términos generales, la noción del TNM *contraria a la terminología habitual* no parece haber tenido el mismo éxito, en términos de recepción, que la propia teoría de modelos preferenciales en la que se inserta. En textos posteriores que abordan problemas de metalógica de LNM y que reconstruyen o al menos introducen referencias a la teoría de modelos preferenciales se observa, aún con matices, un cierto *retorno* hacia al sentido originario atribuido por Gabbay a $\mid\sim$, en tanto símbolo metalinguistico que representa una relación de consecuencia lógica no monótona.

En lo que sigue, se sintetizará la teoría sintáctica/axiomática y la teoría de modelos de KLM de un modo lo más apegado posible al original, aunque restringiendo el análisis al sistema C (cumulativo) el cual constituye el sistema básico sobre el cual se construyen el resto de los sistemas preferenciales. En la exposición subsiguiente se seguirá el orden propuesto por los autores, comenzando por la parte axiomática y continuando con la parte semántica. No podrá pasarse por alto la dualidad de significado de $\mid\sim$ en los casos en los que ésta afecta la interpretación de los axiomas y propiedades.

3.1.1 Una relación de consecuencia $\mid\sim$ es cumulativa syss contiene todas las instancias del axioma de reflexividad (**3.1.2**) y es cerrada bajo las reglas de inferencia de equivalencia lógica izquierda (**3.1.3**), debilitamiento derecho (**3.1.4**), corte (**3.1.5**) y monotonía cauta (**3.1.6**).

Notablemente, esta definición establece como condición para el sistema preferencial básico C, la satisfacción de las propiedades centrales tradicionales de la LNM: reflexividad, corte y monotonía cauta. Estas, junto a equivalencia lógica izquierda y debilitamiento derecho, definen una *relación de consecuencia cumulativa*. Así, las propiedades centrales resultan restauradas para sistemas basados en modelos preferenciales (extensiones del sistema C). Posteriormente se abordarán las demostraciones de las mismas. Por lo pronto, nótese que en la definición **3.1.1** todo indica que se utiliza el símbolo $|\!\sim$ en sentido tradicional, metalingüístico, en tanto se establecen las condiciones que debe cumplir una relación de consecuencia (no monótona) para ser considerada cumulativa. La idea de que la relación contiene *instancias de un axioma*, y la noción de *clausura bajo reglas de inferencia* remiten al sentido más estricto (menos libre) de $|\!\sim$.

En línea con este uso más estricto o apegado a la tradición, cabe señalar que el concepto de clausura usualmente refiere a un operador que permite obtener todas las consecuencias que es posible obtener bajo las reglas consideradas, a partir de un conjunto cualquiera de FBF's. En el caso de las clausuras no monótonas, el operador no monótono considerado debe ser entendido de manera global como una función (por ej. la función C de Makinson)[82]. A continuación, se presentarán las reglas mencionadas en la definición **3.1.1** en notación gentzeniana, al modo de reglas estructurales, tal como son formuladas en el original de KLM. Estas reglas constituyen la base axiomática del sistema C, donde $\vDash \alpha$ significa que α es válida (tautológica) y $\vDash \alpha \rightarrow \beta$ que β es una consecuencia deductiva de α.

[82] Debe entenderse por clausura el resultado final de los procesos de inferencia. Así, en el caso de las lógicas default basadas en Reiter, se trata del conjunto que representa el paso final de inferencia incluyendo los procesos de selección entre extensiones, en el caso de que exista más de una para un mismo conjunto de fórmulas.

3.1.2 $\alpha|\!\sim\alpha$

3.1.3 $\dfrac{\vDash\alpha\leftrightarrow\beta,\ \ \alpha|\!\sim\gamma}{\beta|\!\sim\gamma}$

3.1.4 $\dfrac{\vDash\alpha\rightarrow\beta,\ \ \gamma|\!\sim\alpha}{\gamma|\!\sim\beta}$

3.1.5 $\dfrac{\alpha\wedge\beta|\!\sim\gamma,\ \ \alpha|\!\sim\beta}{\alpha|\!\sim\gamma}$

3.1.6 $\dfrac{\alpha|\!\sim\beta,\ \ \alpha|\!\sim\gamma}{\alpha\wedge\beta|\!\sim\gamma}$

Cabe aclarar que para poder afirmar que las propiedades centrales han sido restauradas (para sistemas preferenciales), no resulta adecuado que los conceptos de *relación de consecuencia* y *normalmente* sean considerados idénticos. Tal lo indicado previamente, la definición **3.1.1** no parece admitir otra interpretación del TNM que la habitual (como relación de consecuencia). Lo mismo podría afirmarse para el resto de las propiedades **3.1.2-6**. De hecho, en los comentarios a las mismas, los autores utilizan terminología tradicional, y en ningún momento parecen retomar la idea de interpretar |~ como "normalmente". Por ejemplo, consideran que **3.1.2** (reflexividad) es el requisito mínimo para cualquier tipo de razonamiento basado en "alguna noción de consecuencia" (Kraus et al., 1990:177). En **3.1.3** (equivalencia lógica izquierda), se afirma que esta regla expresa la exigencia de que "las fórmulas lógicamente equivalentes tengan las mismas consecuencias" (plausibles) (Kraus et al., 1990:177). **3.1.4** (debilitamiento derecho) expresa que un razonador debe estar "dispuesto a aceptar como consecuencia plausible todo lo que es lógicamente implicado por una consecuencia plausible" (Kraus et al., 1990: 177). **3.1.5** (corte) expresa la posibilidad de inferir una consecuencia no monótona γ a partir de los hechos conocidos α aumentados con una hipótesis plausible β deducible no monotónicamente de los mismos hechos, y luego proceder a inferir la conclusión γ sin recurrir a la hipótesis adicional β. Los autores describen esta regla como la siguiente máxima: “una conclusión plausible es tan segura como los supuestos en los que se basa” (Kraus et al., 1990:177-178). No obstante, señalan que algunas interpretaciones probabilísticas pueden invalidar corte (según lo consignado en el **CAPÍTULO 2** sección **2**). Por último, **3.1.6** es una versión de monotonía cauta, que expresa que el aprendizaje de un nuevo hecho, cuya verdad puede ser concluida plausiblemente, no invalida conclusiones plausibles previas. No todos los sistemas de LNM satisfacen esta propiedad[83], aunque sí la satisfacen los sistemas preferenciales. Lo mismo ocurre con las propiedades derivadas del sistema C (véase **APÉNDICE 2**) las cuales, si bien son satisfacibles en sistemas preferenciales, para muchas existen contraejemplos en la bibliografía del área[84].

Como se indicó arriba hay razones fuertes para considerar que, tanto en la definición **3.1.1** como en las propiedades **3.1.2-6** corresponde interpretar $|\!\sim$ en un sentido similar a las propiedades presentadas en (Gabbay, 1985). A continuación, se mencionarán algunas de las dificultades que surgen al entender ese símbolo de un modo diferente (por ejemplo, *normalmente*), ejemplificadas en la propiedad **3.1.2** (reflexividad). En la propuesta de Gabbay (1985), la expresión **3.1.2** debería entenderse de un modo análogo a la propiedad de inclusión/reflexividad de LE (**1.3.1**, **1.2.3**), y es una propiedad que en cierto modo expresa la tradición de considerar las LNM’s como extensiones de la LE. En tanto el TNM expresa una *relación* (de consecuencia), en última instancia la propiedad metalógica de reflexividad tiene un anclaje teórico en la propiedad algebraica de

[83] Cfr. (Kraus et al., 1990:178).
[84] Cfr. (Kraus et al., 1990:180).

reflexividad (de una relación binaria homogénea). Este *significado estricto* puede ser captado por una función sobre conjuntos de FBF's, por ejemplo, el operador C de Makinson en la ecuación **2.2.5** (inclusión). En el marco de un sistema supraclásico como el analizado en (Gabbay, 1985), reflexividad es una propiedad demostrable de manera trivial, en tanto es satisfacible para sistemas de LE. En caso del sistema C de KLM no existe una regla estructural equivalente a supraclasicalidad, pero aun así reflexividad es demostrable recurriendo a construcciones semánticas (**3.1.16-17**).

En la tradición de Gabbay, la propiedad en cuestión puede entenderse simplemente como la aserción de que *toda FBF es una consecuencia plausible de sí misma*. Sin embargo, en el marco de la propuesta de KLM es posible aún otra lectura, dada la dualidad de significado del TNM. Si se entiende este símbolo como *normalmente* (modalidad del lenguaje objeto) el sentido de la propiedad de reflexividad **3.1.2** parece cambiar radicalmente, pues vendría a significar que toda FBF *normalmente* es una consecuencia plausible de sí misma, del mismo modo que las aves normalmente (pero no siempre) vuelan. Entendida de este modo, **3.1.2** sería una *versión débil* de la propiedad propuesta por Gabbay[85]. Lo mismo para el resto de las propiedades: MC se debería interpretar como "*normalmente*, la adición de consecuencias plausibles a las premisas no invalida conclusiones previamente obtenidas", etc. Aunque el debilitamiento generalizado de las propiedades de LNM podría sortear las muchas de las objeciones planteadas en la bibliografía de LNM de las cuales son reconstruidas en el **CAPÍTULO 2** de esta tesis (incluyendo el contraejemplo a inclusión/reflexividad **2.2.12**), el precio a pagar parece ser demasiado alto desde el punto de vista de una disciplina estricta como la metalógica, y en

[85] Aunque sólo se puede especular al respecto, tal debilitamiento de reflexividad estaría en línea con el hecho de que el sistema C de KLM no satisface supraclasicalidad, ya que para un sistema supraclásico se impondría la *versión fuerte* (Gabbay) de aquella propiedad.

particular teniendo en cuenta el objetivo de KLM de ofrecer una contrastación semántica de las propiedades centrales de Gabbay, pues ¿acaso no se estarían contrastando *otras propiedades*, diferentes a aquellas? Por otra parte, la noción de *reflexividad débil* parece entrar en conflicto con la propia definición **3.1.1**, ya que si sólo es posible afirmar que de α normalmente se sigue α, hay que suponer que *no siempre es así*, y por lo tanto |~ podría no contener todas las instancias de **3.1.2**. Además, si normalmente los elementos a la izquierda del TNM se relacionan consigo mismos, pero no se excluye que algunos elementos no se relacionen consigo mismos, lo más preciso en términos algebraicos sería denominar la propiedad de **3.1.2** como *arreflexividad* (de una relación binaria homogénea).

Ahora bien, si se realiza una lectura de las propiedades **3.1.2-6** más cercana a la tradición (Gabbay), esto a su vez repercute inevitablemente en las bases filosóficas del enfoque de KLM, pues implica que no debería utilizarse el símbolo |~ como una modalidad (normalmente) del lenguaje objeto, sino como una relación de consecuencia en el sentido habitual de esta noción metalingüística. Esta situación, al menos en el artículo de referencia, no es resuelta, y la mera afirmación de que |~ representa una *metanoción* no parece ser suficiente para ello. Pero en la parte semántica de la propuesta de los autores, en la práctica se impone una utilización de los símbolos más apegada a la tradición. Al parecer la dualidad de significado del TNM es abandonada progresivamente a medida que se despliega la teoría de modelos preferenciales, culminando en los resultados de representación. De hecho, como se mostrará posteriormente, al momento de fundamentar la propiedad **3.1.2** (reflexividad) mediante construcciones semánticas (**3.1.16-17**) ocurre algo interesante: el propio fundamento semántico de esa regla estructural parece impedir considerar el TNM de un modo *contrario a la terminología habitual*. Con lo dicho hasta aquí, es razonable defender la aserción de que las propiedades centrales pueden ser restauradas para sistemas preferenciales (en base a las construcciones semánticas presentadas a continuación), pero a condición de que se interprete el TNM en el sentido tradicional de Gabbay. Caso contrario, puede afirmarse que las propiedades del sistema C establecidas en la definición **3.1.1**, aun siendo *similares* nominal y formalmente a las propiedades centrales tradicionales de la LNM, poseen un significado diferente a estas últimas, por lo que literalmente serían *otras propiedades*, con lo cual no podría afirmarse sin lugar a dudas que han sido restauradas las propiedades centrales.

Hasta aquí, se han desarrollado algunos aspectos de la propuesta de KLM, junto a una digresión filosófica y metalógica relativamente extensa en relación con el uso del TNM y la posibilidad de restaurar las propiedades centrales para el sistema C (y sus extensiones). Por el momento se considera zanjado este asunto, y no se volverá sobre el mismo excepto por breves comentarios aclaratorios. Posteriormente, en el **CAPÍTULO 3** y en el **APENDICE 2** se retomará la temática de la aporía del TNM. A continuación, se intentarán sintetizar las bases de la semántica modal propuesta por los autores para el sistema C. El objetivo de la siguiente exposición es mostrar efectivamente de qué modo pueden ser restauradas las propiedades centrales mediante construcciones semánticas modales basadas en modelos preferenciales. Sólo se presentarán aquellas definiciones que se consideran indispensables para ese fin, tratando de respetar el sentido original de las mismas, aun con variantes de notación, y ofreciendo en todos los casos sólo comentarios, interpretaciones y análisis que contribuyan a la clarificación conceptual de las estructuras presentadas. En la concepción de KLM, un *modelo* consiste en un conjunto de *estados* y una relación binaria de preferencia entre estos estados. El razonador se describe mediante un modelo, que acepta $\alpha|\sim\beta$ syss todos aquellos estados que son *más preferidos* entre todos los estados que satisfacen α, satisfacen β.

3.1.7 Sea $\prec$ una relación binaria en un conjunto de mundos U. $\prec$ es asimétrica syss $\forall s, t \in U$ tal que $s \prec t, t \nprec s$.

3.1.8 Sea $V \subseteq U$, $t \in V$ es *minimal* en V syss $\forall s \in V$, $s \nprec t$. Además, $t \in V$ es un *mínimum* de V syss $\forall s \in V$ tal que $s \neq t, t \prec s$.

3.1.9 Sea $P \subseteq U$. P es *suave* (smooth) syss $\forall t \in P$, o bien $\exists s$ minimal en P tal que $s \prec t$, o bien t es minimal en P.

3.1.10 Un *modelo cumulativo* es una estructura $< S, l, \prec >$, donde S es un conjunto de elementos denominados *estados*, $l: S \mapsto 2^{\mathcal{U}}$ es una función que *etiqueta* a todo estado con un subconjunto (no vacío) de $\mathcal{U}$ y $\prec$ es una relación binaria en S, que satisface la condición de *suavidad* (smoothness) descripta en **3.1.12**.

3.1.11 Sea $< S, l, \prec >$, si α es una fórmula y $s \in S$, $s \Vdash \alpha$ syss para todo mundo $m \in l(s)$, $m \vDash \alpha$. El conjunto $\{s | s \in S, s \Vdash \alpha\}$ de todos los estados que satisfacen α se denota α^.

3.1.12 Una estructura $< S, l, \prec >$ satisface la *condición de suavidad* syss $\forall \alpha \in \mathcal{L}$, el conjunto α^ es suave.

3.1.13 Dado un modelo cumulativo $W =< S, l, \prec >$, una relación de consecuencia preferencial se denota $|\!\sim_W$ y se define del siguiente modo: $\alpha |\!\sim_W \beta$ syss para todo s minimal en α^, $s \Vdash \beta$.

3.1.14 Una tripleta $< S, l, \prec >$ es un *modelo cumulativo fuerte* syss la relación $\prec$ es asimétrica, y para cada fórmula α, el conjunto α^ tiene un mínimum.

La definición **3.1.7** introduce la relación de preferencia $\prec$ entre mundos posibles en un conjunto cualquiera de mundos U, y una definición de asimetría para esta relación. En

3.1.8 se afirma que una relación de preferencia hace minimal (en un conjunto de mundos posibles) a un mundo posible t si no existe otro mundo que sea preferible respecto a t. Asimismo, si un mundo es preferible sobre cualquier otro mundo distinto de sí mismo en el conjunto de mundos considerados, el mismo se denomina *mínimum* del conjunto. **3.1.9** afirma que un conjunto de mundos posibles es *suave* syss contiene al menos un mundo minimal (para todo mundo del conjunto existe otro mundo que es minimal y es preferible respecto al primero, o bien el primero es minimal). En **3.1.10** se define "modelo cumulativo" como una estructura compuesta por un conjunto S de estados, una función l que asigna a cada estado en S un subconjunto no vacío de $\mathcal{U}$ (es decir, del conjunto de modelos de $\mathcal{L}$). Asimismo, se establece el universo conjuntista de la relación binaria de preferencia $\prec \subseteq S$ X S (producto cartesiano).

De modo que la relación $\prec$ establece preferencias entre estados etiquetados con conjuntos de asignaciones de valores de verdad a letras proposicionales. Nótese que la def. **3.1.10** no incluye un conjunto de fórmulas ni una relación de satisfacción, a diferencia de los modelos clásicos para sistemas de LE y modales. Estos elementos son contemplados en la definición **3.1.11**, por lo que una definición *completa* de "modelo cumulativo" debería entenderse como **3.1.10-11**, y la definición **3.1.10** podría considerarse más precisamente como la definición de "marco cumulativo". En **3.1.11** se introduce una relación binaria $\Vdash$ de satisfacción preferencial. Allí se afirma que un estado de cosas s satisface preferencialmente a una fórmula α syss todos los mundos m que pertenecen al conjunto de mundos posibles asignados por la función l al estado s, satisfacen clásicamente a α. Asimismo, se define $\alpha^\wedge$ como el conjunto de todos los estados de cosas que satisface preferencialmente a una FBF α. En **3.1.12** se afirma que un modelo cumulativo es suave syss el conjunto $\alpha^\wedge$ es suave. **3.1.13** define una relación de consecuencia (semántica) preferencial $|\!\sim_W$: una FBF β es una consecuencia preferencial de otra FBF α syss todo estado s que satisface preferencialmente a α y es minimal (ningún estado entre en el conjunto de estados que satisfacen preferencialmente a α es preferido sobre s), también satisface a β. De las definiciones anteriores puede interpretarse un modo de inferencia no monótona: intuitivamente, $\prec$ representa la elección del agente de estados que son más normales que otros. Así, el agente puede inferir plausiblemente β a partir de α si todos los estados que son más normales para el agente y satisfacen α, también satisfacen β. Por último, **3.1.14** establece que un modelo cumulativo es *fuerte* si la relación de preferencia es asimétrica y para toda FBF α, el conjunto α^ de estados que satisface preferencialmente a α tiene un mínimum, es decir, contiene un estado que es preferible por sobre cualquier otro estado del conjunto.

Luego, los autores abordan el estudio de problemas de representación. Aquí se presentarán los resultados expuestos en (Kraus et al., 1990) que resultan temáticamente relevantes para los objetivos de esta tesis. Como se mencionó antes, KLM se desplazan hacia una lectura más apegada al sentido originario del símbolo $|\sim$. Esto no resulta sorprendente debido a la propia naturaleza de los problemas de representación, para cuyo abordaje prácticamente se ven forzados a considerar $|\sim$ en tanto relación de consecuencia sintáctica cumulativa que funciona en tándem con la relación de consecuencia semántica $|\sim_W$.

3.1.15 Para cualquier modelo cumulativo W, la relación de consecuencia $|\sim_W$ que él define es una "relación cumulativa", es decir, las relaciones de consecuencia definidas mediante modelos cumulativos satisfacen todas las reglas del sistema C.

Si se interpreta el símbolo $|\sim$ como una relación de consecuencia sintáctica y el signo $|\sim_W$ como una relación de consecuencia semántica, dejando a un lado cualquier dualidad o interpretación sui géneris de los símbolos, este resultado vendría a expresar en términos generales, que todo sistema de LNM con una semántica basada en modelos preferenciales satisface, entre otras, las propiedades axiomáticas centrales (reflexividad y cumulatividad). Los autores ofrecen pruebas de corte y monotonía cauta, mientras que la propiedad de reflexividad puede ser justificada en base a las construcciones **3.1.16-17**. Para corte, se supone que 1) todos los estados minimales que satisfacen preferencialmente una FBF α también satisfacen otra FBF β; y 2) todos los estados minimales que pertenecen a $(\alpha \wedge \beta)^\wedge$ satisfacen una FBF γ. Para un estado minimal cualquiera s que satisface α, s satisface β (según supuesto 1), y por ende satisface $\alpha \wedge \beta$. Dado que s es minimal en $\alpha^\wedge$ y $(\alpha \wedge \beta)^\wedge \subseteq \alpha^\wedge$ (por lema 3.15 en (Kraus et al., 1990:183), $(\alpha \wedge \beta)^\wedge = \alpha^\wedge \cap \beta^\wedge$ y por definición de intersección $\alpha^\wedge \cap \beta^\wedge \subseteq \alpha^\wedge$), s es minimal en $(\alpha \wedge \beta)^\wedge$ y por lo tanto satisface γ (según supuesto 2). En consecuencia, todos los estados minimales que satisfacen α satisfacen también γ.

Para monotonía cauta, se supone 1) $\alpha|\!\sim_W \beta$; y 2) $\alpha|\!\sim_W \gamma$. Sea un s minimal en $(\alpha \wedge \beta)^\wedge$ tal que s satisface γ. Además, $s \in \alpha^\wedge$. Si no existiera tal s, por condición de suavidad debería haber un s' minimal en $\alpha^\wedge$ tal que $s' \prec s$. Pero en ese caso, dado que s' satisface β (por supuesto 1) y en consecuencia $s' \in \alpha^\wedge \cap \beta^\wedge$, puede concluirse que $s' \in (\alpha \wedge \beta)^\wedge$ lo cual significa que $s' \in \alpha^\wedge$ lo cual contradice la minimalidad de s en $\alpha^\wedge$. En consecuencia, s es minimal en $\alpha^\wedge$, en $(\alpha \wedge \beta)^\wedge$ y dado el supuesto 2), s satisface γ. Posteriormente, se establece que, en el marco de una semántica basada en modelos preferenciales, para una relación cumulativa $|\!\sim$ puede construirse un modelo cumulativo W “que define una relación $|\!\sim_W$ que es exactamente $|\!\sim$” (Kraus et al., 1990:183). Es decir, se establece una propiedad de completitud entre las relaciones (sintácticas) cumulativas y las relaciones (semánticas) preferenciales.

3.1.16 Un mundo $m \in \mathcal{U}$ es *normal* para una FBF α syss $\forall \beta \in \mathcal{L}$ tal que $\alpha|\!\sim\beta$, $m \vDash \beta$.

3.1.17 Para una relación de consecuencia cumulativa, todos los mundos normales que satisfacen α satisfacen β syss $\alpha|\sim\beta$.

3.1.18 $\alpha\sim\beta$ (α es *equivalente* a β) syss $\alpha|\sim\beta$ y $\beta|\sim\alpha$.

3.1.19 $\alpha^{\sim} \leq \beta^{\sim}$ syss $\exists\alpha' \in \alpha^{\sim}$ tal que $\beta|\sim\alpha'$

En **3.1.16** se afirma que los modelos normales de una fórmula son aquellos que satisfacen clásicamente a todas sus consecuencias plausibles, y en **3.1.17** se establecen las bases del teorema de corrección/completitud **3.1.22.** Según lo explayado previamente, las definiciones **3.1.16-17** permiten justificar semánticamente la propiedad **3.1.2** (reflexividad). Pero a continuación los autores afirman que esta debe entenderse del siguiente modo: "un mundo normal para α satisface α" (Kraus et al., 1990:184). Esta afirmación reabre interrogantes respecto a la interpretación de la propiedad en cuestión, ¿debería entenderse **3.1.2** como reflexividad (Gabbay)? ¿se trata de una propiedad diferente, un tipo de *reflexividad débil* o arreflexividad? ¿significa que una formula se sigue de sí misma del mismo modo que las aves normalmente vuelan? Para zanjar esta cuestión, es suficiente tomar en cuenta el propio fundamento semántico de **3.1.2**: puesto que el conjunto de todos los mundos que satisfacen clásicamente a un α es idéntico al conjunto de los mundos normales para α, α es tanto una consecuencia plausible como clásica de α, con lo cual puede afirmarse que la versión de reflexividad de KLM no se diferencia de la versión de Gabbay o "versión fuerte" (en la que $|\sim$ funciona como una relación de consecuencia (no monótona) en el metalenguaje). Esto excluye cualquier interpretación *débil* de **3.1.2**, y restaura el sentido tradicional (Gabbay) de la misma.

Por otra parte, la aserción de que *un mundo normal para* α *satisface* α no debería entenderse como la afirmación de que *una fórmula normalmente se sigue de sí misma* del mismo modo que las aves normalmente vuelan (interpretación que se desprende de la noción de $|\!\sim$ como significando *normalmente*), sino como la afirmación de que en un sistema basado en modelos preferenciales, está garantizada la satisfacción de reflexividad, pues (trivialmente) los mundos normales para una fórmula cualquiera satisfacen esa fórmula. Aquí se hace patente lo mencionado antes respecto a la *deriva del TNM* a lo largo del texto citado, pues para probar reflexividad en base a las construcciones semánticas **3.1.16-17** se impone una concepción del TNM *en línea* con la terminología habitual. A esta altura, no parece conveniente ni necesario seguir interpretando el símbolo $|\!\sim$ de un modo diferente al sentido originario de Gabbay.

Con los resultados expuestos hasta aquí se completa el objetivo de mostrar de qué modo es posible restaurar las propiedades centrales de la LNM desde un punto de vista semántico para sistemas basados en modelos preferenciales. A fin de ofrecer un cuadro más completo del sistema C de KLM, particularmente en lo que concierne a los resultados de representación, se continuará con el comentario de las definiciones subsiguientes **3.1.18-19** y se presentarán y analizarán otras que resultan relevantes para los fines mencionados. En **3.1.17** la parte después del "si…" se sigue de la definición **3.1.16**, mientras que para la parte después del "solo sí…" se ofrece una prueba en (Kraus et al., 1990:184). En **3.1.18** se establece una relación de equivalencia (cumulativa), mientras que **3.1.19** introduce una relación $\leq$ entre clases de equivalencias (bajo $\sim$) que es reflexiva, antisimétrica (si $\alpha^{\sim} \leq \beta^{\sim}$ y $\beta^{\sim} \leq \alpha^{\sim}$, entonces $\alpha^{\sim} = \beta^{\sim}$) y no transitiva.

3.1.20 $W =_{def} < S, l \prec >$ donde $S =_{def} \mathcal{L}/\!\sim$, $l(\alpha^{\sim}) =_{def} \{m \mid m$ es un mundo normal para $\alpha\}$; y $\alpha^{\sim} \prec \beta^{\sim}$ syss $\alpha^{\sim} \leq \beta^{\sim}$ y $\alpha^{\sim} \neq \beta^{\sim}$, con $\prec$ asimétrica.

3.1.21 Para cualquier $\alpha \in L$ el estado $\alpha^{\sim}$ es un mínimum de $\alpha^{\wedge}$.

3.1.22 $\alpha|\!\sim\beta$ syss $\alpha|\!\sim_W\beta$

En **3.1.20** se amplía la definición de marco cumulativo, especificando a S como el conjunto de todas las clases de equivalencias de FBF´s bajo $\sim$. La función l etiqueta clases de equivalencia de FBF´s cualesquiera α con mundos normales de α, es decir, con mundos $m \in \mathcal{U}$ que satisfacen todas las consecuencias plausibles de α. Así, los mundos normales respecto a α son mundos normales respecto a cada fórmula que es una equivalencia (cumulativa) de $\alpha,$ es decir, para cada β tal que $\alpha|\sim\beta$ y $\beta|\sim\alpha$. Por último, se define $\prec$ como una relación entre clases de equivalencias de FBF´s. En **3.1.21** se afirma que la clase de equivalencias de una fórmula es un mínimum de la misma, y por lo tanto el modelo W definido en **3.1.20** es un modelo cumulativo fuerte. Finalmente, se establece el teorema fundamental de corrección/completitud **3.1.22**, cuya prueba se ofrece en (Kraus et al., 1990:185).

A modo de breve recapitulación sobre el sistema C y algunos aspectos más generales de la propuesta de KLM, cabe resaltar en primer lugar que la dualidad de significado de $|\sim$ parece condicionar (desde un punto de vista filosófico) la construcción axiomática del sistema C, pero mayormente no afecta su teoría de modelos. Particularmente, la interpretación de las propiedades centrales de la LNM como *versiones débiles* de las propiedades de Gabbay resulta problemática. Sin embargo, nada impide realizar una lectura de las propiedades del sistema C más apegada al sentido originario de $|\sim$, como *relación de consecuencia no monótona*. Dando ese paso, se eliminan los problemas internos de la teoría de KLM, y se conservan sus resultados axiomáticos y semánticos (sintetizados en **3.1.22** para el sistema C). Asimismo, con ello es posible afirmar que las propiedades centrales de Gabbay, para las cuales existen contraejemplos en la bibliografía de LNM que afectan la originaria *pretensión universalista* de las mismas, son demostrables en un marco restringido: para la familia de sistemas de LNM basados en (o que pueden ser reconstruidos mediante) modelos preferenciales. Siempre y cuando se interprete el TNM del modo habitual, esto constituye una *restauración parcial de las propiedades centrales*. Más aun, como se explicitará en la próxima sección, es posible restaurar supraclasicalidad para sistemas basados en LMG, como así también la propiedad de conservación de la consistencia; ambas propiedades no satisfacibles por el sistema C de KLM. Finalmente, estos resultados constituyen un *suelo firme* para las propiedades centrales, ganado a partir de la propuesta de KLM (liberada de sus problemas filosóficos), y al parecer indican una cierta primacía de la semántica sobre la axiomática en sistemas de LNM, tal como sugiere explícitamente K. Schlechta cuando afirma que "la estrategia correcta parece comenzar por la semántica" (Schlechta, 2004:10). A continuación, se abordará la propuesta de este último autor.

3.2 LOS *SISTEMAS DE COHERENCIA* DE K. SCHLECHTA

En (Schlechta, 2004) el autor defiende la idea de que un gran número de LNM's y otras lógicas orientadas al razonamiento de sentido común pueden abarcarse y reconstruirse mediante unas pocas nociones semánticas: *preferencia*, *tamaño,* y *distancia* (entre las más importantes). Estas nociones, que en algunos casos son interdefinibles (particularmente, preferencia y tamaño), expresan de un modo cualitativo el grado de certeza (o incertidumbre) de la información disponible en las premisas de una inferencia. Según el autor, la mencionada reconstrucción de sistemas mediante aquellas nociones semánticas apunta hacia una *lógica modal generalizada* (LMG). Esta lógica modal generalizada, a su vez, funcionaría como "marco general para el razonamiento de sentido común" (Schlechta, 2004:22). En la obra citada se ofrece una suerte de "caja de herramientas" de enorme complejidad, que combina construcciones lógicas sintácticas y semánticas, junto a estructuras matemáticas de teoría de modelos y algebraicas, orientadas a la reconstrucción de sistemas de LNM tanto en su aspecto sintáctico como en el semántico, pero con clara primacía de la semántica. Asimismo, se propone establecer un conjunto de propiedades genéricas o *condiciones de coherencia* para los sistemas, entre las que se encuentran las propiedades centrales tradicionales de la LNM.

Se trata de una obra que en parte retoma el estilo de exposición de la *vieja escuela* metamatemática de Tarski, donde los límites entre teoría de modelos, álgebra y lógica resultan difusos, pero con menos elegancia y ordenamiento en relación con (Tarski, 1969a) o (Tarski, 1969b), como así también con menor solidez filosófica que estos últimos. Aun así, esto no impide considerar a (Schlechta, 2004) como una obra de gran complejidad filosófica y metalógica. Particularmente el autor aborda, aunque de manera indirecta o subsidiaria, la aporía del TNM originada por la identificación entre este símbolo con la noción de *normalidad*. Esta aporía surge, tal como se explayó en la sección anterior, de la propuesta presente en (Kraus et al., 1990). En la concepción de Schlechta sobre este tópico se observa cierta oscilación entre tres posturas (que aparecen diseminadas a lo largo del texto en la forma de breves digresiones intercaladas entre los desarrollos modelo-teoréticos): 1) restaurar la tradición clásica de considerar el TNM como una relación de consecuencia no monótona, tal como en la propuesta original de (Gabbay, 1985); 2) conceder cierta vigencia en este aspecto al enfoque de KLM; y 3) una propuesta ecléctica entre las anteriores que implica una virtual disolución entre metalenguaje y lenguaje objeto, guiada por la estrategia de “poner tantas cosas como sea posible en el lenguaje objeto” (Schlechta, 2004:4), incluyendo la noción de normalidad simbolizada indistintamente por el TNM, por funciones de elección, por el operador modal unario N, por un cuantificador semigeneralizador ∇, etc. Estas oscilaciones teóricas pueden resultar un tanto extrañas a simple vista, pero debe considerarse el contexto más general de la propuesta del autor para comprender caritativamente su postura respecto a este asunto.

En primer lugar, claramente el mayor interés de Schlechta no pasa por resolver la aporía del TNM (ni cualquier otro problema de filosofía de la lógica), sino, en última instancia, por ofrecer diferentes interpretaciones y reconstrucciones de la noción de *normalidad*, desde diversos enfoques y con numerosas y complejas herramientas matemáticas, a fin de reconstruir una diversidad de sistemas de LNM. El concepto de normalidad es sin duda el hilo conductor que atraviesa toda la obra, que de conjunto puede considerarse como un intento de ofrecer un fundamento modelo-teorético general para el mismo. Paralelamente, el autor parece buscar la mayor amplitud posible en el aspecto sistemático, y al intentar englobar muchos sistemas dentro de un mismo enfoque general, es esperable que emerjan algunas aporías o dificultades entre la diversidad de enfoques particulares. Desde este punto de vista, no resulta sorprendente que puedan encontrarse en (Schlechta, 2004) razones para atribuirle al autor cualquiera de las 3 posturas mencionadas respecto al TNM, según el sistema que se esté reconstruyendo, según la noción de normalidad considerada, etc. Es un hecho que el mayor interés del autor no es filosófico, pero no puede dejar de observarse que interpretar el TNM de un modo diferente a la tradición de Gabbay genera dificultades metateóricas insoslayables (tal como se consignó en la sección **3.1**).

En síntesis, la propuesta de Schlechta es de una complejidad inmensa, en todos los aspectos, tal es así que D. Makinson advierte en el *Prefacio* del libro citado que, si bien se trata de una obra motivada por consideraciones lógicas, debería ser entendida como una exposición que transcurre a un nivel *prelingüístico*, de carácter mayormente algebraico y modelo-teorético. Este libro requiere no sólo de "madurez matemática" y “perseverancia” para ser comprendido, tal como afirma certeramente Makinson[86], sino también un dominio de temas de filosofía de la lógica y metalógica igualmente robusto para dimensionar su alcance e implicancias en tanto *texto sobre lógica*. En definitiva, aun cuando la obra citada deja abiertas algunas cuestiones filosóficas, aun cuando puedan cuestionarse los intentos de resolución de la aporía del uso del TNM (heredada de la tradición de KLM), se trata sin dudas de un trabajo monumental, abstruso y multifacético, llamado a convertirse en una referencia ineludible para temas de LNM y de lógica en general.

Dicho esto, y dada la amplitud y complejidad de la obra citada, aquí sólo podrá realizarse una exposición sintetizada basada en un recorte temático de algunos de los fundamentos de la propuesta de Schlechta: a saber, aquellos que resultan de interés para los fines de esta tesis (mencionados en **INTRODUCCIÓN Y OBJETIVOS**). Pero tal recorte temático supone ciertas opciones metodológicas, que se harán explícitas a continuación. Como se mencionó previamente, desde un punto de vista filosófico el abordaje del problema del TNM y la relación entre lenguaje objeto y metalenguaje en (Schlechta, 2004) *admite*, de manera más o menos forzada según el contexto, 3 lecturas posibles. Además, existe en el texto citado una tendencia bastante marcada hacia una suerte de disolución de cualquier distinción entre lenguaje objeto y metalenguaje, que se expresa en una notación *prelingüística* (en términos de Makinson). Aquí se optará por defender la idea de que la entera propuesta de Schlechta en su (2004) no sólo es perfectamente compatible con la opción de restaurar la tradición clásica de considerar el TNM como una relación de consecuencia no monótona, sino que de ese modo se gana en solidez de

[86] Cfr. *Prefacio* de D. Makinson en (Schlechta, 2004:vi).

fundamentos filosóficos, en tanto se evitan los problemas mencionados en la sección **3.1** de esta tesis.

Asimismo, siempre que sea posible se optará por considerar admisible una interpretación de la notación utilizada por el autor que distinga con mayor claridad entre metalenguaje y lenguaje objeto. Esta no sólo es una opción conservadora, en línea con la tradición tarskiana que dio origen a tal distinción (que originalmente buscaba evitar paradojas dentro del lenguaje objeto), sino que, en un sentido amplio, prácticamente es una condición de posibilidad para desarrollar reflexiones metalógicas. Habitualmente, las reflexiones metadisciplinarias se desenvuelven en un nivel lingüístico diferente a la disciplina a la que refieren, pero debido a las características propias de la propuesta de Schlechta, no siempre será posible mantener una distinción de ese tipo. Establecidas estas opciones metodológicas, en lo que sigue se reducirán al mínimo el desarrollo de digresiones filosóficas al respecto. Particularmente, el tópico de la interpretación del TNM resultaba de la mayor importancia en la sección **3.1** en tanto afectaba la interpretación misma de las propiedades centrales, pero en gran parte fue saldado en esa misma sección, por lo que no se requiere continuar profundizando aquí. No obstante, en el **APÉNDICE 2** se recolectarán algunos ejemplos de la postura de Schlechta sobre estos asuntos y se desarrollarán algunas reflexiones filosóficas.

El eje temático de esta sección **3.2** se encolumna con el eje más general de la tesis en dos aspectos esenciales: en la presentación y discusión sobre las propiedades de coherencia (que en parte constituye una relectura, desde el enfoque de Schlechta, de algunos de los desarrollos presentados en la sección anterior), y en la exposición y comentario de algunas de las construcciones modelo-teoréticas vinculadas a la noción de *tamaño*. Ambos aspectos contribuirán posteriormente en el **CAPÍTULO 4** para los fines de formular y fundamentar semánticamente la propiedad de semimonotonía.

Por otra parte, se presentarán algunos fundamentos generales de la LMG, como así también de la versión de Schlechta de los sistemas preferenciales, que se diferencia en importantes aspectos de la versión KLM (tal como se hará patente posteriormente). Sobre estos últimos tópicos sólo se abordarán los rudimentos mínimos, aquellos necesarios para encuadrar metodológicamente las exposiciones de las temáticas mencionadas en el párrafo anterior (particularmente de las condiciones de coherencia), que resultan de mayor interés para los fines de esta tesis. Por último, las construcciones específicas para la noción de *distancia* no serán tenidas en cuenta. En general, se intentará respetar al máximo el espíritu del original, pero no se seguirá necesariamente el orden de exposición del autor, ni el énfasis que este pone en cada una de las construcciones. Por ejemplo, la primer definición general consignada y comentada aquí (**3.2.1**) a simple vista resulta de enorme trascendencia, pero el autor la presenta casi *al pasar*, sin numeración (a diferencia de la mayoría de las definiciones que ofrece), dentro del cuerpo de un párrafo introductorio en (Schlechta, 2004:3).

Es importante notar, asimismo, que a lo largo del texto citado, capítulo a capítulo el autor alterna entre abordajes en ocasiones más cercanos a la teoría de modelos puramente matemática, con definiciones y recursos algebraicos, para luego realizar un viraje puramente lógico-sintáctico, y renglón seguido volver a combinar teoría de modelos y semántica clásica. La propia definición **3.2.1** es una muestra de ello, como así también la mayoría de las definiciones presentadas en esta sección. Además, Schlechta intercala desarrollos de temáticas (tradicionalmente) metalógicas como el estudio de propiedades de sistemas, problemas de representación, etc., combinando disquisiciones filosóficas breves aquí y allá. Este estilo tan particular de exposición en ocasiones dificulta la lectura y reconstrucción de su propuesta. En síntesis, lo que sigue es un recorte temático y una exposición un tanto libre de algunos de los fundamentos de la LMG, según los fines de esta tesis, junto a comentarios y especificaciones no presentes en el original. Entrando de lleno en tal exposición, se ofrece la siguiente definición:

3.2.1 $T|\!\sim\alpha$ syss $f(M(T)) \vDash \alpha$

Aquí se hace patente de inmediato el anclaje modelo-teorético del pensamiento de Schlechta. En su enfoque, las semánticas se fundan en *funciones de elección* f sobre conjuntos, para las cuales se da por supuesto un uso sin reservas del axioma de elección de E. Zermelo (Schlechta, 2004:26). Una función de elección f define una lógica que cumple la condición **3.2.1** (siendo T una teoría, conjunto de FBF's), y constituye la base para fundamentar, junto a otras estructuras algebraicas, cualquier noción de normalidad en términos modelo-teoréticos. Por ejemplo, una lectura admisible de esta definición, interpretando el TNM en la tradición de Gabbay y estableciendo el fundamento semántico en la tradición de Shoham, es decir, sobre mundos posibles entendidos como modelos clásicos, podría ser la siguiente: una FBF α es una consecuencia no monótona de una teoría T syss la imagen de la función de elección f sobre los modelos de T satisface clásicamente α. Las funciones de elección permiten reconstruir a nivel modelo-teorético las tres nociones semánticas principales: *preferencia*, *tamaño* y *distancia*[87].

A su vez, estas nociones son tres modos diferentes de entender el concepto de normalidad desde el punto de vista semántico. Intuitivamente, $f(M(T))$ fija los mundos posibles considerados por el razonador, ya sea por preferencia (como en la propuesta de KLM), por tamaño (*la mayoría* o una gran parte de los mundos posibles) o bien por distancia (por *cercanía* entre los mundos posibles). En añadidura, según la interpretación realizada aquí del TNM, la definición **3.2.1** establece una equivalencia entre los resultados de las

[87] El autor también aborda otros enfoques semánticos que no serán tomados en cuenta aquí, como preservación de definibilidad, algoritmos de Farkas, sistemas de Markov, y otros. Debido a su importancia para una LMG, en el **Apéndice 2** se presentará una definición básica de preservación de definibilidad.

inferencias sintácticas no monótonas, y los resultados modelo-teoréticos basados en elecciones sobre mundos posibles. Aunque no parece ser la intención del autor establecer un resultado de representación en esos términos, bajo la interpretación tradicional del TNM, la expresión **3.2.1** puede ser considerada como un teorema metamatemático *análogo* al teorema fundamental del sistema C de KLM **3.1.22** (corrección + completitud).

En la perspectiva de Schlechta, desde el punto de vista sintáctico resulta importante diseñar simbolismos para revertir la extrapolación de la noción de normalidad al metalenguaje (efectuada por KLM). Para sistemas preferenciales (es decir, sistemas basados en el sistema C de KLM), el autor propone simbolizar las aserciones condicionales, por ej. "normalmente las aves vuelan", como $\vdash N(A) \rightarrow V$, que se lee aproximadamente del siguiente modo "los casos normales A implican clásicamente V" (Schlechta, 2004:4). En primer lugar, se introduce un operador modal unario N (normalidad). Además, se complementa esta noción de normalidad en términos sintácticos utilizando el torniquete clásico y la implicación material. De ese modo, se evita escribir las aserciones condicionales como expresiones metalingüísticas $\alpha|\!\sim\beta$ (en la tradición de KLM). Posteriormente, el autor introduce, para lógicas default, un nuevo tipo de cuantificador denominado *cuantificador generalizado* el cual funciona de un modo similar al operador N. Este cuantificador, que aquí se nombrará como *semigeneralizador*, se representa mediante el símbolo ∇ (Schlechta, 2004:368), significa "casi todos" o "en la mayoría de los casos" (Schlechta, 2004:9).

El semigeneralizador constituye un modo de interpretar sintácticamente la noción de normalidad desde el punto de vista del *tamaño* de los subconjuntos de objetos considerados, que permite simbolizar aserciones condicionales tales como "normalmente las aves vuelan” del siguiente modo: $\triangledown\, x(Ax \to Vx)$. También, definiendo previamente el universo de discurso al conjunto de las aves, se simboliza $\triangledown\, xVx$ (como ocurre en (Schlechta, 1995))[88]. Nótese que, en ambos casos, tanto cuando se utiliza el operador N como cuando se utiliza el semigeneralizador, se recurre de manera auxiliar a formalismos de LE para complementar la simbolización de aserciones condicionales. Aunque el autor mayormente opta por utilizar el símbolo $\triangledown$ para la reconstrucción de reglas default (según los criterios de una LMG, por ej. en el Capítulo 7 de su (2004)), también lo intercambia libremente, dependiendo del contexto, con el operador N (por ej. en el Capítulo 8 de su (Schlechta, 2004)). Esto expresa, a nivel sintáctico, la interdefinibilidad entre las nociones semánticas de tamaño y preferencia.

Tanto la propuesta de simbolizar la noción de normalidad (en el lenguaje objeto) mediante la modalidad unaria N, como la de utilizar un semigeneralizador $\triangledown$, redundan en innegables ventajas respecto de la utilización del TNM para el mismo fin (tradición KLM). Schlechta menciona las siguientes: es posible negar la modalidad N, lo cual permite simbolizar los *casos anormales* de α como $\alpha \wedge \neg N(\alpha)$; y permite establecer operaciones de contraposición: de $\neg\beta \to \neg N(\alpha)$ y $\neg\beta$ se sigue $\neg N(\alpha)$, esquema que resulta ilustrativo de los casos de retracción de conclusiones. Además, el enriquecimiento del lenguaje (objeto) con la incorporación de la modalidad N junto a otras nociones semánticas y algebraicas, permiten de conjunto un formalismo más expresivo, donde es posible definir una noción de consistencia, negaciones y contraposiciones, modalidades anidadas, combinaciones booleanas y relativizaciones; permite definir con precisión la interpretación de los símbolos y operadores, y clarificar la interacción entre los mismos.

[88] Cabe aclarar que estos ejemplos son ilustrativos, y no aparecen en el original (Schlechta, 2004), donde se da preferencia a las construcciones abstractas axiomáticas y semánticas.

Similares ventajas pueden obtenerse con el semigeneralizador ∇: este último puede ser negado, anidado, y mezclado con cuantificadores tradicionales (Schlechta, 2004:372).

De modo que los sistemas basados en LMG optan por reconstruir la noción de normalidad mediante simbolismos diseñados específicamente en el *lenguaje objeto*, principalmente el operador N y el semigeneralizador. Ahora bien, cabe destacar para Schlechta no deberían construirse sintaxis lógicas directamente, pues desde su punto de vista es muy difícil hallar *buenas lógicas* sin una semántica (al menos implícita). Se necesita tanto la lógica (sintáctica) como la semántica, "pero la estrategia correcta parece ser comenzar con la semántica" (Schlechta, 2004:10). Las *axiomáticas* o sistemas sintácticos serían el segundo paso, en tanto trabajan sobre objetos de las semánticas formales (funciones de elección, filtros, ideales, etc.). Posteriormente, debería abordarse el problema matemático de la adecuación entre semántica y axiomática (corrección, completitud). Así, para el autor, la vara para evaluar una lógica es justamente su adecuación con cierta semántica. Además, Schlechta critica la concepción común en LNM según la cual las extensiones o puntos fijos de un sistema constituyen su *semántica*, pues si esto es correcto, entonces podría afirmarse, análogamente, que la semántica de LE es el conjunto de sus teoremas. Esta crítica ubica al autor en la perspectiva de Shoham y KLM de construir semánticas especiales para LNM's, y a su vez lo acerca a la tradición modelo-teorética de la LE.

Una de las teorías semánticas/modelo-teoréticas de mayor importancia en la propuesta de Schlechta se basa en *teoría de filtros*, cuyas estructuras serán presentadas y comentadas en las siguientes definiciones. Tanto la noción de *filtro* como la de *ideal* pueden distinguirse en estructuras *estrictas*, por un lado, y estructuras *débiles*, por otro. Las primeras corresponden a las definiciones clásicas de teoría de filtros, por ejemplo en (Manzano, 1999), mientras que las segundas tienen un antecedente en la noción de *filtro propio* de Bell – Slomson (1969):

3.2.2 Sea un conjunto base X, un filtro sobre X es un conjunto $\mathcal{F} \subseteq \mathcal{P}(X)$ que satisface las siguientes condiciones: 1) $X \in \mathcal{F}$; 2) si $A \subseteq B \subseteq X$, $A \in \mathcal{F}$ implica $B \in \mathcal{F}$; 3) $A, B \in \mathcal{F}$ implica $A \cap B \in \mathcal{F}$. El filtro se denomina "débil" si en lugar de satisfacer 3) satisface 3') $A, B \in \mathcal{F}$ implica $A \cap B \neq \emptyset$.

Un filtro es un conjunto de subconjuntos de X, que por condición 1) incluye a X, por condición 2) para cualquier $X' \subseteq X$ que pertenece al filtro todos los superconjuntos de X' que son subconjuntos de X pertenecen al filtro. Dados dos elementos (subconjuntos de X) cualquiera del filtro, este contiene a su vez el conjunto intersección entre ambos. Si esto no ocurre, pero la intersección entre ambos no es vacía, el filtro es débil. En otras palabras, si el filtro satisface 3), incluye al conjunto vacío mientras que si satisface 3') no incluye al conjunto vacío, ya que, siempre y cuando la estructura de base sea un álgebra booleana[89], se cumple una condición especial de absorción de la intersección conjuntista tal que para cualesquier A, $A \cap \emptyset = \emptyset$ (por absorción de la intersección), y esto es justamente lo que impide la cláusula 3') para conjuntos que pertenecen al filtro. De manera que los filtros débiles excluyen $\mathcal{F} = \mathcal{P}(X)$, el *filtro degenerado*, y en todos los casos $\mathcal{F} \subset \mathcal{P}(X)$. Es decir, la cláusula 3´) contempla el caso límite cuando sólo *atraviesa el filtro* el conjunto vacío.

En el enfoque de Schlechta, los filtros débiles resultan el correlato modelo-teorético del semigeneralizador ∇. Como el filtro es no degenerado, el semigeneralizador viene a significar "en la mayoría de los casos, *pero no en todos los casos*". Por supuesto, si la (semi)generalización incluye todos los casos, no tendría sentido construir lógicas diferentes a la LE. En este punto es interesante citar la aserción del autor respecto a que un filtro débil es "menos idealista" que un filtro en sentido estricto (Schlechta, 2004:27).

---

[89] Véase **CAPÍTULO 4** sección **2**.

Si bien esta afirmación no es explayada más allá de su mera enunciación, argüiblemente, el asunto pasa por la consideración de que, en conjuntos de objetos del mundo *a la mano*, no parece tener sentido (más precisamente, es trivial) construir semigeneralizaciones que atribuyen determinadas características comunes a una mayoría de los subconjuntos de un conjunto de objetos del mundo, incluyendo al vacío. Como es sabido, las predicaciones sobre elementos del conjunto vacío se cumplen trivialmente, pero en la práctica es como afirmar que la mayoría de las aves vuelan, incluso las no existentes. Así, excluir $\emptyset$ hace a la noción de filtro débil *menos idealista*, pero no *lo suficientemente realista*.

Nótese que, dado que $\mathcal{F} \subseteq \mathcal{P}(X)$, el filtro estricto o clásico admite ser interpretado como "todos los casos" cuando $\mathcal{F} = \mathcal{P}(X)$. En el caso de los filtros débiles, excluyendo el conjunto vacío técnicamente no se consideran todos los elementos de $\mathcal{P}(X)$. Así, el filtro débil capta mejor la intuición de una *mayoría de los casos* excluyendo *todos los casos* (semigeneralización) puesto que, si se cumple la cláusula 3') se garantiza $\mathcal{F} \subset \mathcal{P}(X)$. En este punto, se visualiza una posible aporía filosófica. La cláusula 3') excluye $\emptyset$ del filtro, y esto es metamatemáticamente importante para un sistema de LNM, pues, nuevamente, si se consideran todos los casos en una generalización, es fútil utilizar herramientas sintácticas y semánticas diferentes a las de LE. Pero en términos de realismo filosófico, es un mero tecnicismo, ya que podría argumentarse que cuando se trata de *objetos del mundo a la mano*, si la semigeneralización excluye del filtro sólo al subconjunto vacío, aun cuando garantiza $\mathcal{F} \subset \mathcal{P}(X)$ en la práctica se estarían considerando todos los casos (porque el único subconjunto que se filtra es vacío).

Esto puede tener repercusiones en la propuesta del autor de reconstruir lógicas default mediante sistemas de LMG con semigeneralizador como herramienta para simbolizar reglas default. Se trata de un problema filosófico (de filosofía de la lógica) de difícil resolución, que no podrá ser profundizado aquí. Ciertamente la aporía se basa un caso extremo, dentro de un contexto realista y tomando en cuenta consideraciones extralógicas, pero no puede dejar de mencionarse, quedando abierta a futuras indagaciones. Por otra parte, esta situación no resulta problemática en términos puramente metamatemáticos, cuando se trabaja con objetos ideales. En lo inmediato, la exclusión del conjunto vacío tampoco resulta problemática desde un punto de vista modelo-teorético para la finalidad de especificar que un filtro es débil syss es un subconjunto propio del conjunto potencia del conjunto base $\mathcal{F} \subset \mathcal{P}(X)$, lo cual permite capturar sin mayores dificultades la idea de una mayoría de los casos (pero no todos los casos) para *objetos matemáticos*. Continuando con los fundamentos de la teoría de filtros versión Schlechta, se presentarán las siguientes definiciones, que serán oportunamente comentadas:

3.2.3 Un filtro es principal syss existe un $X' \subseteq X$ tal que $\mathcal{F} = \{A: X' \subseteq A \subseteq X\}$.

Intuitivamente, un filtro principal fija un *subconjunto mínimo* X' donde el *máximo* es el propio conjunto base del filtro, y el filtrado de subconjuntos de X excluye los elementos que están fuera de esos *márgenes*. Esta definición admite, en el caso límite, que el mínimo y el máximo sean conjuntos idénticos, con lo cual $\mathcal{F}(X) = \mathcal{P}(X)$.

3.2.4 Un ideal sobre X es un conjunto $\mathcal{I} \subseteq \mathcal{P}(X)$ que satisface las siguientes condiciones: 1) $\emptyset \in \mathcal{I}$; 2) si $A \subseteq B \subseteq X$, $B \in \mathcal{I}$ implica $A \in \mathcal{I}$; 3) $A, B \in \mathcal{I}$ implica $A \cup B \in \mathcal{I}$. El ideal se denomina “débil” si en lugar de satisfacer 3) satisface 3’) $A, B \in \mathcal{I}$ implica $A \cup B \neq X$.

Un ideal es un conjunto de subconjuntos de X, por condición 1) incluye al conjunto vacío, por condición 2) para cualquier $X' \subseteq X$ que pertenece al ideal, todos los subconjuntos de X' pertenecen al ideal. Dados dos elementos del ideal, este contiene a su vez el conjunto unión entre ambos. Si esto no ocurre, pero la unión entre ambos no es igual a X, el ideal es débil. Un ideal es la conversa de un filtro, si $\mathcal{F}$ es un filtro (débil) sobre X, entonces se define un ideal (débil) sobre X como $\mathcal{I} \coloneqq \{X - A: A \in \mathcal{F}\}$; y si $\mathcal{I}$ es un ideal (débil) sobre X, entonces $\mathcal{F} \coloneqq \{X - A: A \in \mathcal{I}\}$ es un filtro (débil) sobre X[90]. Los filtros e ideales débiles pueden entenderse, en la perspectiva de Schlechta, como una noción abstracta de *tamaño*: un filtro débil sobre X representa *subconjuntos grandes* de X, mientras que un ideal débil sobre X representa *subconjuntos pequeños* de X, y el resto de los subconjuntos de X representan elementos de *tamaño medio* (los cuales pueden ser expresados como filtros principales).

Sin embargo, esta intuición no aplica para filtros e ideales clásicos, que admiten (en el caso extremo) ser iguales al conjunto base y por lo tanto representan sólo dos modos o métodos diferentes de construir subconjuntos de cualquier tamaño, que pueden tener el mismo resultado por diferentes vías (particularmente, representar subconjuntos grandes del conjunto base, incluyendo al propio conjunto base). Hecha esta aclaración, que tomará relevancia posteriormente en el **CAPÍTULO 4** sección **2**, cabe destacar que los filtros débiles pueden utilizarse además para reconstruir de manera bastante cercana a la intuición, el concepto de "modelos preferenciales" (interdefinibilidad de las nociones de tamaño y preferencia). Por un lado, un filtro débil sobre un conjunto de modelos de una teoría T expresa la noción de *normalidad*, los modelos que satisfacen T. Por otra parte, un ideal débil expresa la noción de *anormalidad*, los modelos que no satisfacen T o *excepciones*.

A continuación, se presentarán algunos rudimentos de la teoría de modelos preferenciales (versión Schlechta). Más precisamente, se trata de una tricotomización de las estructuras preferenciales clásicas, que se discriminan en tres tipificaciones modelo-teoréticas: 1)

[90] En el original, debido a un error de tipeo el segundo condicional de esta definición aparece como "si $\mathcal{I}$ es un ideal (débil) sobre X, entonces $\mathcal{F} \coloneqq \{X - A: A \in \mathcal{F}\}$ es un filtro (débil) sobre X". Cfr. (Schlechta, 2004:28).

versión minimal (similar a la propuesta de KLM, pero con diferencias tales como el uso de funciones de elección); 2) versión con copias (de elementos de conjuntos arbitrarios); y 3) versión límite (con recurso a segmentos iniciales minimizadores). Resulta importante sintetizar la teoría de modelos preferenciales (versión Schlechta) por varios motivos, en primer lugar, permite ilustrar de qué modo interactúan las estructuras preferenciales con las funciones de elección. Asimismo, tomadas de conjunto, la teoría de filtros débiles y la teoría de modelos preferenciales (versión Schlechta) constituyen una base adecuada a los fines de presentar un recorte, incluso una *versión minimalista* de la LMG, en tanto ambas teorías ofrecen las construcciones modelo-teoréticas sin las cuales sería muy difícil pretender *decir algo* respecto de tal propuesta. Por último, sin estos rudimentos mínimos, la posterior presentación de las condiciones de coherencia (entre ellas las propiedades centrales de la LNM) resultaría mera *letra muerta*.

En una inicial aproximación a la teoría de modelos preferenciales (versión Schlechta) puede considerarse un punto de partida tradicional (en la tradición de Shoham y KLM), tomando una definición de "estructura preferencial" $\mathcal{M} =< M_{\mathcal{L}}, \prec>$, como "un conjunto de modelos clásicos, con una relación binaria" de satisfacción, que se distingue de las estructuras Kripke por el tipo de relación utilizada (Schlechta, 2004:71). Tal lo consignado en la sección **3.1**, las estructuras preferenciales trabajan con un lenguaje proposicional $\mathcal{L}$, donde el conjunto de todos sus modelos clásicos se denota $M_{\mathcal{L}}$. Sin embargo, no puede pasarse por alto que el tipo de construcciones modelo-teoréticas utilizado por Schlechta en su teoría de modelos preferenciales admite en ocasiones, como es el caso de la definición **3.2.5** o en general, en todas las estructuras que trabajan con copias, una interpretación *libre de restricciones* respecto a los objetos considerados[91]

[91] En el **Capítulo 4** se notará la importancia de este enfoque libre de restricciones para los fines de la metalógica y la metamatemática, para la reformulación y fundamentación semántica de la propiedad de no monotonía **2.3.1**.

(conjuntos, modelos, mundos posibles, relaciones, etc.). En este punto existe cierta ruptura con toda la tradición previa dentro del *enfoque preferencial* de LNM.

La versión minimal 1) es la que tiene mayores similitudes con la teoría de modelos preferenciales tradicional (en la vertiente de Shoham y KLM). Dado un conjunto de modelos $X \subseteq M_{\mathcal{L}}$, un modelo $x \in X$ se denomina $\prec$-minimal en X syss no existe un $x' \in X$ tal que $x' \prec x$. El conjunto de todos los elementos minimales de X se denota $\mu_{\prec}(X)$ (donde μ es una función de elección de modelos). Se define $T \vDash_{\mathcal{M}} \alpha$ syss $\mu(M(T)) \vDash \alpha$ (clásicamente), es decir syss todos los modelos minimales de T satisfacen α, en otras palabras, syss $\mu(M(T)) \subseteq M(\alpha)$. En esta versión, se puede reemplazar $T \vDash_{\mathcal{M}} \alpha$ por $T|\!\sim\alpha$ siempre que el contexto lo permita. Por otra parte, $\overline{\overline{T}}$ (clausura no monótona de T, véase definición **3.2.12)** simboliza el conjunto de todas las $\vDash_{\mathcal{M}}$-consecuencias. El principal problema de esta versión es que incluso si T es consistente, $\overline{\overline{T}}$ será inconsistente si $\mu\big(M(T)\big) = \emptyset$ (si no hay modelos mínimos de T). Es decir, el conjunto de elementos preferenciales de un conjunto no vacío puede ser vacío, esto ocurre cuando se presentan cadenas de preferencias descendientes infinitas (no hay elementos minimales, o más precisamente, debajo de cada elemento de la cadena hay un elemento minimal). Así, se generan clausuras inconsistentes a partir de teorías consistentes.

Tales situaciones, donde los sistemas preferenciales tradicionales pueden "introducir inconsistencias allí donde no había ninguna" (Schlechta, 2004:75) ya había sido señalada por Makinson en su (1994), tal como se mencionó en el **CAPÍTULO 2**. Para lidiar con este tipo de situaciones, Schlechta introduce *versiones límite* (con segmentos iniciales minimizadores) de las principales estructuras modelo-teoréticas, como se verá en las siguientes definiciones. Por otra parte, es importante aclarar, como lo hace el autor, que en general "el razonamiento preferencial no funcionará en una base de datos clásicamente inconsistente, por la razón trivial de que cualquier subconjunto del conjunto vacío es vacío" (Schlechta, 2004:74). Esta aseveración, que en los hechos impone una *cláusula de consistencia de las bases de conocimiento* para sistemas basados en LMG, resulta importante a la luz de la denominada paradoja de Israel (**2.2.11**), que establece la inconsistencia generalizada de las bases de conocimiento de los agentes racionales. Posteriormente en la sección **3.3** y en el **CAPÍTULO 4** se retornará sobre este asunto.

En la versión con copias 2), un modelo m puede ocurrir muchas veces en la estructura, con diferentes posiciones respecto a $\prec$. Se define $\mu(X)$ como el conjunto de $x \in X$ para los cuales existe al menos una copia de x que es minimal en X. Para ilustrar las estructuras con copias, se ofrece el siguiente ejemplo en (Schlechta, 2004:73): sea el modelo m y sus copias equivalentes m_1 y m_2. Sea $m' \prec m_1$ y $m'' \prec m_2$. De ese modo, $m \in \mu(\{m, m'\})$ pues m_2 es una copia de m y es minimal en $\{m, m'\}$; y $m \in \mu(\{m, m''\})$ pues m_1 es una copia de m y es minimal en $\{m, m''\}$. No obstante, $m \notin \mu(\{m, m', m''\})$ ya que ninguna de sus copias es minimal en $\{m, m', m''\}$. En relación con la función μ, la parte lógica se vuelve más o menos trivial. Ante tal nivel de abstracción, según el autor el siguiente paso es dejar a un lado la temática de modelos clásicos y lógica y "sólo hablar acerca de conjuntos arbitrarios y relaciones, con y sin copias de los elementos" (Schlechta, 2004:74). Es decir, de estructuras libres de restricciones.

La versión limite 3) trabaja con conjuntos de modelos $Y \subseteq X$ tales que para todo $x \in X$ existe un $y \in Y$ tal que $y \preccurlyeq x$ ($y \prec x$ o $y = x$), y si $y \in Y$, $x \in X$, $y \prec x$, entonces $y \in Y$ (es decir, para un modelo cualquiera en X y es preferible a cualquier otro modelo en X, este pertenece a un subconjunto Y de modelos preferenciales de X). Así, Y es "cerrado hacia abajo" en X, y "minimiza" todos los elementos de X. Los conjuntos Y se denominan "segmentos iniciales minimizadores" (SIM). Se define $T \vDash_{\mathcal{M}} \alpha$ syss existe un SIM Y de

$M(T)$ tal que $Y \vDash \alpha$ clásicamente, y el conjunto $\overline{\overline{T}} \coloneqq \{\alpha : T \vDash_{\mathcal{M}} \alpha\}$. Según esta última definición, si T es consistente $\overline{\overline{T}}$ no contendrá $\perp$ (será consistente). De este modo, con la versión limite se supera el problema de las posibles inconsistencias generadas en sistemas preferenciales tradicionales. Esta solución requiere, a su vez, de una especificación de *subconjunto definible* y de *funciones preservadoras de definibilidad* (véase **APÉNDICE 2**), como así también de una definición de *modelos preservadores de definibilidad – variante límite* (en definición **3.2.6**).

Luego de esta breve introducción a la teoría de modelos preferenciales (versión Schlechta), se presentará la siguiente definición **3.2.5**, donde se introducen las estructuras y construcciones básicas. Estas estructuras son presentadas como libres de restricciones, trabajando en un universo arbitrario U con una relación binaria $\prec$ en U o un conjunto de copias de elementos de U. En ocasiones es útil considerar una función de elección $\mu(X)$ incluso cuando $X \nsubseteq U$, para indicar que la estructura tiene "agujeros". En todas las definiciones la normalidad es relativa, no absoluta. No interesan los elementos más normales del universo, sino de un conjunto dado:

3.2.5 Dado $U \neq \emptyset$ y un X arbitrario, donde X no necesariamente tiene algo que ver con U o con $\mathcal{U}$ (definido abajo). Las funciones de elección $\mu_{\mathcal{M}}$ son funciones desde V a V (donde V es el universo conjuntista considerado)

1) Modelos (o estructuras) preferenciales

Versión sin copias

Un par $\mathcal{M} \coloneqq< U, \prec>$ con U siendo un conjunto arbitrario, y $\prec$ una relación binaria arbitraria se denomina modelo o estructura preferencial.

Versión con copias

Un par $\mathcal{M} \coloneqq < \mathcal{U}, \prec >$ con $\mathcal{U}$ siendo un conjunto arbitrario de pares, y $\prec$ una relación binaria arbitraria se denomina modelo o estructura preferencial. $< x, i > \in \mathcal{U}$ donde x es el elemento de U e i es el índice de la copia.

2) Elementos minimales, las funciones $\mu_{\mathcal{M}}$ y segmentos iniciales minimizadores (SIM)

Versión minimal sin copias

Sea $\mathcal{M} \coloneqq < U, \prec >$, se define $\mu_{\mathcal{M}}(X) \coloneqq \{x \in X : x \in U \wedge \neg\exists x' \in X \cap U . x' \prec x\}$. $\mu_{\mathcal{M}}(X)$ es el conjunto de elementos minimales de X (en $\mathcal{M}$).

Versión minimal con copias

Sea $\mathcal{M} \coloneqq < \mathcal{U}, \prec >$, se define $\mu_{\mathcal{M}}(X) \coloneqq \{x \in X : \exists < x, i > \in \mathcal{U} . \neg\exists < x', i' > \in \mathcal{U}(x' \in X \wedge < x', i' > \prec < x, i >)\}$.

Cuando $< x, i > \prec < y, j >$ se afirma que $< x, i >$ "mata" o minimiza $< y, j >$. Asimismo, cuando para todo $< y, j > \in \mathcal{U}, y \in Y$ existe un $< x, i > \in \mathcal{U}, x \in X$ tal que $< x, i > \prec < y, j >$ se afirma que el conjunto X mata o minimiza al conjunto Y.

Versión límite sin copias

Sea $\mathcal{M} \coloneqq < \mathcal{U}, \prec >$, se define $Y \subseteq X \subseteq U$ como "segmento inicial minimizador" (SIM) de X syss $(\forall x \in X \exists y \in Y . y \preccurlyeq x) \wedge \forall y \in Y, \forall x \in X(x \prec y \rightarrow x \in Y))$

Versión límite con copias

Sea $\mathcal{M} \coloneqq < \mathcal{U}, \prec >$, se define $Y \subseteq X \subseteq \mathcal{U}$ como SIM syss $(\forall < x, i > \in X \exists < y, j > \in Y . < y, j > \preccurlyeq < x, i >) \wedge \big(\forall < y, j > \in Y, \forall < x, i > \in X(< x, i > \prec < y, j > \rightarrow < x, i > \in Y)\big)$

3) $\mathcal{M}$ es inyectivo (o 1-copia) syss existe al menos una copia $< x, i >$ para cada x. $\mathcal{M}$ es reflexivo, transitivo, etc. syss $\prec$ también lo es. Un conjunto $\mathcal{X}$ de SIM es cofinal en otro conjunto de SIM $\mathcal{X}'$ (para el mismo conjunto base X) syss para todo $Y' \in \mathcal{X}'$, existe un $Y \in \mathcal{X}, Y \subseteq Y'$.

Esta definición básica **3.2.5** constituye el núcleo modelo-teorético y abstracto, libre de restricciones, de la teoría de modelos preferenciales (versión Schlechta). En la misma se especifican las estructuras básicas como las funciones de elección y los modelos preferenciales, bajo un universo conjuntista arbitrario. Nótese que, en la versión con copias, las relaciones de preferencia relacionan pares de elementos a su izquierda y a su derecha, en vez de elementos individuales a ambos lados (como era el caso en la propuesta de KLM). El primer elemento del par es el elemento *copiado*, y el *índice* es otro elemento del conjunto considerado. Luego se define el conjunto de elementos minimales de un conjunto dado, con recurso a funciones de elección y relaciones de preferencia. Además, los modelos minimales también son *minimums* del conjunto X. Asimismo, se establece el mecanismo mediante el cual conjuntos completos de elementos minimizan a otros conjuntos, y sobre esta base se definen los conjuntos SIM (con y sin copias). Por último, se instituyen (entre otras condiciones) los requisitos para considerar a un modelo preferencial como inyectivo, es decir cuando cada elemento del universo tiene al menos una copia en el modelo. La siguiente definición disminuye el nivel de abstracción de las estructuras respecto a la definición previa (puramente modelo-teorética), y se sitúa en un nivel estrictamente lógico. En la misma se reintroducen las restricciones, y se define una relación de consecuencia lógica semántica para sistemas preferenciales:

3.2.6 Se define una relación de consecuencia para una estructura preferencial con un lenguaje proposicional $\mathcal{L}$.

1) Si m es un modelo clásico de $\mathcal{L}$, $< m, i > \vDash \alpha$ syss $m \vDash \alpha$, y si X es un conjunto de tales pares, $X \vDash \alpha$ syss para todo $< m, i > \in X$, $m \vDash \alpha$

Si $\mathcal{M}$ es una estructura preferencial, y X es un conjunto de $\mathcal{L}$-modelos (o un conjunto de pares $< m, i >$), $\mathcal{M}$ es una estructura preferencial clásica.

2) Validez en una estructura preferencial (relación de consecuencia semántica)

Sea una estructura preferencial clásica $\mathcal{M}$,

Variante minimal

Se define $T \vDash_{\mathcal{M}} \alpha$ syss $\mu_{\mathcal{M}}(M(T)) \vDash \alpha$, es decir $\mu_{\mathcal{M}}(M(T)) \subseteq M(\alpha)$. Asimismo, se determina $T^{\mathcal{M}} \coloneqq \{\alpha: T \vDash_{\mathcal{M}} \alpha\}$.

Un modelo $\mathcal{M}$ es *preservador de definibilidad* syss para todo $X \in D_{\mathcal{L}}$ (donde $D_{\mathcal{L}} \subseteq \mathcal{P}(M_{\mathcal{L}})$), $\mu_{\mathcal{M}}(X) \in D_{\mathcal{L}}$.

Variante límite

Se define $T \vDash_{\mathcal{M}} \alpha$ syss existe un SIM $Y \subseteq \mathcal{U} \upharpoonright M(T)$ tal que $Y \vDash \alpha$

(donde $\mathcal{U} \upharpoonright M(T) \coloneqq \{< x, i > \in \mathcal{U}: x \in M(T)\}$

Una SIM X es definible syss $\{x: \exists < x, i > \in X\} \in D_{\mathcal{L}}$

En **3.2.6** se define una relación de consecuencia (semántica) preferencial dentro del enfoque o teoría de modelos preferenciales (versión Schlechta). La cláusula 1) establece que una copia satisface una FBF syss el *original* satisface esa FBF, y un conjunto de copias satisface una FBF syss el original de tal conjunto de copias satisface la FBF. La cláusula 2) (variante minimal) introduce una relación de consecuencia semántica preferencial $\vDash_{\mathcal{M}}$ que se define mediante la noción de satisfacción clásica de una FBF y una función de elección sobre el conjunto de modelos clásicos de tal FBF. Asimismo, se especifica la clausura preferencial $T^{\mathcal{M}}$ como el conjunto de las $\vDash_{\mathcal{M}}$-consecuencias de una teoría T. Por otra parte, se determina que un modelo es preservador de definibilidad syss los modelos minimales de un conjunto de modelos definibles X, también son definibles (véase **APÉNDICE 2**). La variante límite de esta cláusula es adecuada para situaciones donde existen cadenas descendentes infinitas sin elementos minimales, permitiendo superar el problema de la generación de inconsistencias a partir de teorías consistentes, pero según reconoce el autor, es una variante excesivamente complicada (Schlechta, 2004:78). La siguiente definición ofrece una versión modelo-teorética de *suavidad*:

3.2.7 Dado $\mathcal{Z} \subseteq \mathcal{P}(U)$, donde posiblemente $\mathcal{Z} = D_{\mathcal{L}}$, una estructura preferencial $\mathcal{M}$ es $\mathcal{Z}$-suave syss en todo $X \in \mathcal{Z}$ todo elemento $x \in X$ o bien es minimal en X o bien *está por encima* de un elemento minimal en X. Es decir, (versión sin copias) si $x \in X \ \in \mathcal{Z}$, entonces o bien $x \in \mu_{\mathcal{M}}(X)$ o bien existe un $x' \in \mu_{\mathcal{M}}(X)$ tal que $x' \prec x$ (nótese que suavidad supone transitividad de $\prec$). Considerando copias, si $x \in X \ \in \mathcal{Z}$, y $< x, i > \in \mathcal{U}$, entonces o bien no existe $< x', i' > \in \mathcal{U}$ con $x' \in X$ tal que $< x', i' > \prec < x, i >$ o bien existe un $< x', i' > \in \mathcal{U}$ tal que $< x', i' > \prec < x, i >$ con $x' \in X$, es decir no existe $< x'', i'' > \in \mathcal{U}$ tal que $x'' \in X$ y $< x'', i'' > \prec < x', i' >$.

Según el autor, "en el caso finito sin copias, suavidad es una consecuencia trivial de transitividad y falta de ciclos" (Schlechta, 2004:79). La noción de suavidad resulta de gran importancia para Schlechta en tanto expresa semánticamente la propiedad de *cumulatividad*, presentada posteriormente en la definición **3.2.20**. Una vez establecidas las estructuras básicas de la teoría de modelos preferenciales (versión Schlechta), se procederá a presentar algunas de las propiedades de las funciones de elección que fundamentan tales estructuras. Estas propiedades **3.2.8-11** constituyen a su vez una base modelo-teorética para muchas de las propiedades metalógicas presentadas posteriormente en **3.2.13-21** (condiciones de coherencia). La investigación sobre las propiedades de las funciones de elección comienza por la búsqueda de una respuesta a la siguiente pregunta planteada en términos modelo-teoréticos: dado $f(M(T))$, ¿puede elegirse libremente $f(M(T)')$, o debe existir algún tipo de coherencia entre $f(X)$ y $f(Y)$? En otras palabras, ¿debe existir alguna relación entre X e Y, los conjuntos de modelos de T?. Por ejemplo, en estructuras preferenciales clásicas $f(Y) \cap X \subseteq f(X)$ siempre que $X \subseteq Y$. Para el autor, las propiedades de las funciones de elección son las propiedades fundamentales de la LMG, que resultan comunes a las distintas lógicas reconstruidas dentro del enfoque. En primer lugar, se establece que las estructuras preferenciales generales se caracterizan por una función de elección de modelos minimales $\mu : \mathcal{Y} \rightarrow \mathcal{Y}$ tal que para todo $X, Y \in \mathcal{Y}$ se cumplen las siguientes condiciones (libres de restricciones):

3.2.8 $\mu(X) \subseteq X$

3.2.9 $X \subseteq Y \rightarrow \mu(Y) \cap X \subseteq \mu(X)$

3.2.10 $\mu(X) \subseteq Y \subseteq X \rightarrow \mu(X) = \mu(Y)$

3.2.11 $Y \neq \emptyset \rightarrow \mu(Y) \neq \emptyset$

La condición **3.2.8** es *análoga* a (o más precisamente, una versión modelo-teorética de) la propiedad tradicional de *supraclasicalidad* (véase **3.2.17** versión C)), donde una función de elección asigna a un conjunto X de modelos clásicos, un subconjunto de modelos minimales de X. La condición **3.2.9** expresa que si X es un subconjunto de Y, entonces la intersección entre los modelos preferenciales de Y y el conjunto X es un subconjunto de los modelos preferenciales de X. Asimismo, implica que “todo elemento que no es minimal en algún conjunto X, no será minimal en algún superconjunto Y de X”

(Schlechta, 2004:106). **3.2.10** es una versión modelo-teorética de *cumulatividad* (véase **3.2.20** versión C)), y **3.2.11** implica que los conjuntos de modelos de teorías consistentes tienen modelos minimales. Esta última propiedad puede interpretarse como una *condición de consistencia* para funciones de elección sobre conjuntos de modelos minimales, y constituye un fundamento para la propiedad **3.2.18**.

Ahora bien, dada una función de elección que satisface tales propiedades, se plantea el problema del pasaje desde el nivel modelo-teorético al nivel lógico-semántico, problema que se reduce a hallar una estructura semántica preferencial para representar tal función de elección. La clave del asunto es el establecimiento de condiciones para minimizar x, si $x \in X - \mu(X)$. O bien un elemento $x \in X - \mu(X)$ es minimizado por algún elemento en X (quizá por x mismo en una cadena/ciclo descendiente infinita/o), o bien se requiere un conjunto de elementos en X para minimizar x. El primer caso trabaja directamente con elementos x, el segundo caso supone copias: deben *destruirse* todas las copias y por ello se necesita un conjunto completo de elementos. El autor presenta primero las construcciones algebraicas para el caso general sin copias en sus variantes transitiva y no transitiva. La variante transitiva no impone nuevas condiciones, si $a \succ b \succ c$, entonces si la relación es transitiva puede reemplazarse b por c para minimizar a. Es decir, si a es minimizado por un conjunto que contiene b, es posible sustituir b por cualquier conjunto que minimiza b. Sin embargo, es posible que b esté presente en cualquier conjunto que minimiza b (él puede minimizarse a si mismo), y no se sabe en general qué c minimiza b. Así, se reemplaza b por un conjunto que contiene b, y no es sorprendente que el nuevo conjunto minimiza a.

Pero si se sabe que c es un elemento o singletón que minimiza b, la situación es diferente, y transitividad puede imponer condiciones suplementarias. El problema principal tanto en el caso transitivo como en el no transitivo es elegir los elementos que minimizan un determinado elemento. Cuando no hay información suficiente para decidir, se utilizan conjuntos *completos* de elementos para minimizar un determinado elemento. Esta falta de información se codifica en copias (del elemento a minimizar), que son minimizadas por tal conjunto de elementos. En términos generales, se trata de hallar elementos adecuados $x' \in X, x' \prec x$ si $x \in X - \mu(X)$. Los x' son elementos cualesquiera de X, y se debe encontrar entre ellos, los que son minimales. Ese es el núcleo del problema, que explica el recurso a copias en la teoría de modelos preferenciales de Schlechta: dado un x no minimal, ¿Cuáles x' minimales minimizan a x? (teniendo en cuenta que la definición **3.2.5** de modelos minimales en su versión con copias establece que un modelo minimal es uno que tiene al menos una copia tal que ningún otro modelo es preferencial respecto a ella). La solución es construir una copia $< x, x' >$ para cada $x' \in X$, que es minimizada por $x' : < x, x' > \succ x'$. "Esto garantiza que el conjunto completo X minimizará todas las copias de x" (Schlechta, 2004:132).

En otras palabras, dado un x no minimal, se mapean todos los elementos de X para construir pares (copias) donde el primer elemento es x y el *índice* es cada uno de los elementos de x, alternativamente. Intuitivamente, los elementos que no son minimales serían los mundos menos normales, las excepciones. Estos (se supone que) son minimizados por mundos más normales. Si no se conoce de antemano cuáles son los mundos/modelos que minimizan las excepciones, se genera una copia que consiste en un par cuyo primer elemento es una copia del modelo *menos normal* indexada con cualesquiera otros modelos de X (cada uno de ellos), y se establece que esta copia es minimizada por el modelo cualesquiera de X (el segundo elemento del par). Se trata de una opción que puede parecer un tanto arbitraria, pero que es totalmente legítima en términos estrictamente modelo-teoréticos. Esta opción asegura que cada mundo *menos* *normal* sea minimizado por una copia de cada elemento de X, entonces se justifica que las excepciones no son modelos minimales porque ninguna de sus copias son minimales. En ese sentido, la minimización de cada elemento no minimal es una *gran disyunción* de minimizaciones. Nótese que las copias son elementos de $\{x : x \in X - \mu(X)\}$ X X (producto cartesiano). Para un ejemplo de minimización con copias véase **APÉNDICE 2.**

Las propiedades de las funciones de elección **3.2.8-11** resultan de enorme importancia para los fines de la fundamentación metalógica de los sistemas de LNM, y, por lo tanto, para los fines de esta tesis. Específicamente, estas constituyen la base modelo-teorética que permite demostrar las *condiciones de coherencia* **3.2.13-21** de los sistemas reconstruidos dentro del enfoque de LMG. Antes de la exposición de estas condiciones, se requiere una definición precisa del concepto de *clausura no monótona* desde el punto de vista sintáctico. Tomando como referencia la noción de clausura deductiva $\overline{T} \subseteq \mathcal{L}$, $\overline{T} \coloneqq \{\alpha : T \vdash \alpha\}$, se define $\overline{\overline{T}}$ como el conjunto de consecuencias cerrado bajo una relación de consecuencia no monótona, representada por $|\!\sim$ (Schlechta, 2004:29).

3.2.12 $\overline{\overline{T}} \coloneqq \{\alpha : T |\!\sim \alpha\}$.

Las condiciones de coherencia, o *propiedades metalógicas de los sistemas de LNM reconstruidos bajo el enfoque de LMG*, son presentadas por el autor bajo tres versiones diferentes, siempre que esto resulta posible: A) versión metalógica con una FBF a la izquierda de $|\!\sim$; B) versión metalógica con una teoría (conjunto de FBF's) a la izquierda de $|\!\sim$ y con el uso de clausuras; C) versión modelo-teorética con funciones de elección. En el siguiente listado, cuando la versión A) coincida con una propiedad del sistema C de KLM **3.1.2-6**, se omitirá la misma indicando su numeración (la única diferencia vendrá dada por el uso de $\vdash$ en vez de $\vDash$ cuando corresponda). Para Schlechta, un sistema P

(preferencial) se define por las siguientes propiedades: conjunción derecha (**3.2.13**), disyunción izquierda (**3.2.14**), equivalencia lógica izquierda (**3.2.15**), debilitamiento derecho (**3.2.16**), supraclasicalidad (**3.2.17**), preservación de la consistencia (**3.2.18**) monotonía cauta (**3.2.19**), y cumulatividad (**3.2.20**). Asimismo, la adición de monotonía racional (**3.2.21**) genera el sistema *R* (racional) (Schlechta, 2004:32). Por último, es importante tomar en cuenta el significado de la noción de *coherencia*, que se ancla en una distinción entre propiedades que dejan el lado izquierdo de la relación de consecuencia sin cambios (por ej., conjunción derecha **3.2.7** versiones A) y B)) y propiedades que modifican el lado izquierdo (por ej., monotonía cauta en versión KLM **3.1.6**). Las propiedades que modifican el lado izquierdo son "propiedades de coherencia" (Schlechta, 2004:92). Su rol de transferencia es muy importante para las LNM's, donde no se cumple la propiedad de *coherencia fuerte* de monotonía, que permite *transferencias poderosas*. En general, cuanto mayor cantidad de propiedades de coherencia tenga un sistema, permitirá al agente tener mayores regularidades para razonar con eficiencia, es decir, para ser más *estable*:

3.2.13 A) $\frac{\alpha|\sim\beta,\ \ \alpha|\sim\gamma}{\alpha|\sim\beta\wedge\gamma}$

B) $\frac{T|\sim\beta\ \ \ T|\sim\gamma}{T|\sim\beta\wedge\gamma}$

C) No posee

3.2.14 A) $\frac{\alpha|\sim\gamma\ \ \ \beta|\sim\gamma}{\alpha\vee\beta|\sim\gamma}$

B) $\frac{T|\sim\gamma\ \ \ T'|\sim\gamma}{T\vee T'|\sim\gamma}$

C) $f(A \cup B) \subseteq f(A) \cup f(B)$

3.2.15 A) KLM **3.1.3**

B) $\dfrac{\overline{T}=\overline{T'}}{\overline{\overline{T}}=\overline{\overline{T'}}}$

C) no posee

3.2.16 A) KLM **3.1.4**

B) $\dfrac{\vdash\alpha\rightarrow\beta,\ \ T|\!\sim\alpha}{T|\!\sim\beta}$

C) no posee

3.2.17 A) $\dfrac{\alpha\vdash\gamma}{\alpha|\!\sim\gamma}$

B) $\overline{T} \subseteq \overline{\overline{T}}$

C) $f(X) \subseteq X$

3.2.18 A) $\dfrac{\alpha|\!\sim\bot}{\alpha\vdash\bot}$

B) $\dfrac{T|\!\sim\bot}{T\vdash\bot}$

C) $\dfrac{f(X)=\emptyset}{X=\emptyset}$

3.2.19 A) KLM **3.1.6**

B) $\dfrac{T\subseteq\overline{T'}\subseteq\overline{\overline{T}}}{\overline{\overline{T}}=\overline{\overline{T'}}}$

C) $\frac{f(X)\subseteq Y\subseteq X}{f(Y)\subseteq f(X)}$

3.2.20 A) $\dfrac{\alpha|\!\sim\beta}{\frac{\alpha|\sim\gamma}{\alpha\wedge\beta|\sim\gamma}\quad\frac{\alpha\wedge\beta|\sim\gamma}{\alpha|\sim\gamma}}$

B) ídem **3.2.19** B)

C) $\frac{f(X)\subseteq Y\subseteq X}{f(Y)=f(X)}$

3.2.21 A) $\frac{\alpha|\!\sim\beta \quad \alpha|\!\not\sim\neg\gamma}{\alpha\wedge\gamma|\!\sim\beta}$

B) $\frac{T|\!\sim\beta \quad T|\!\not\sim\neg\gamma}{T\cup\{\gamma\}|\!\sim\beta}$

C) $\frac{X\subseteq Y \quad Y\cap f(X)\neq\emptyset}{f(X)=f(Y)\cap X}$

Como puede observarse, algunas de estas propiedades coinciden con propiedades del sistema C de KLM, entre ellas *monotonía cauta*, una de las propiedades centrales de la LNM. No obstante, el sistema P de Schlechta satisface además las propiedades *supraclasicalidad* (**3.2.17**) y *conservación de consistencia* (**3.2.18**), que no están presentes en el sistema C. Ambas propiedades están contempladas en la propuesta de (Makinson, 1994), y a su vez remiten a propiedades expresadas por primera vez en (Gabbay, 1985). Supraclasicalidad, expresada previamente en **2.2.9**, se basa en la propiedad de *compatibilidad de* $|\!\sim$ *con* $\vdash$ de Gabbay; mientras que preservación de la consistencia (**2.2.10**), es una versión de la condición de *consistencia de la información nueva con la información vieja* de este último autor. Tal como se aseveró previamente, la presencia de **3.2.18** entre las condiciones de coherencia se fundamenta en **3.2.11**, y en última instancia, en las versiones límite de las construcciones modelo-teoréticas de la LMG. Estas propiedades **3.2.18** (axiomática) y **3.2.11** (modelo-teorética) junto a la propiedad de preservación de definibilidad (véase **APENDICE 2**) permiten sortear los obstáculos relativos a la generación de inconsistencias en sistemas preferenciales, y por ese motivo garantizan que, en sistemas de LMG, si la base de conocimiento es consistente también lo será su clausura no monótona.

Ahora bien, tomando en cuenta que el sistema C no satisface **3.2.17** ni **3.2.18**, puede afirmarse que es más débil que el sistema $P,$ aún cuando el primero satisface las propiedades centrales de reflexividad y cumulatividad. En primer lugar, supraclasicalidad hace trivial a reflexividad, pero es una propiedad más fuerte o de mayor alcance que esta última en tanto garantiza no sólo que las propias FBF´s de una teoría T son consecuencias no monótonas de T, sino también que la clausura deductiva de T es un subconjunto de la clausura no monótona de T. Por otra parte, si el sistema satisface conservación de consistencia la clausura no monótona de T será inconsistente solo si la clausura deductiva lo es. Así, el sistema no generará inconsistencias, lo cual es conveniente debido a que las inconsistencias pueden forzar retracción de consecuencias, y por ello un sistema que satisface conservación de consistencia es más fuerte (más estable) que un sistema que no satisface esa propiedad. Con criterios similares, podría afirmarse que el sistema R sería más fuerte que P y C, pues además de las propiedades centrales, supraclasicalidad y conservación de consistencia, satisface monotonía racional, que es una propiedad más audaz que monotonía cauta, dado que esta última garantiza conservación de consecuencias de T sólo cuando se amplía T agregando FBF's que pertenecen a la clausura no monótona de T, mientras que monotonía racional garantiza conservación de consecuencias de T cuando se amplía T agregando FBF's cuya negación no pertenece a la clausura no monótona de T.

Ahora bien, los criterios para definir la fortaleza o debilidad de un sistema lógico pueden considerarse parcialmente (o completamente) arbitrarios, pues como todo criterio de clasificación depende de compromisos teóricos previos (en este caso, de filosofía de la lógica) posiblemente no explicitados e incluso injustificados. No se puede construir una prueba para una clasificación de sistemas lógicos, como si se estuviera demostrando (por ejemplo) la propiedad de cumulatividad de un sistema basado en modelos preferenciales. De hecho, en (Schlechta, 2004:40), el autor afirma que los sistemas (de LNM) basados en modelos preferenciales son *más fuertes* incluso que los sistemas de LE, debido a que satisfacen supraclasicalidad, pues permiten "conjeturar más allá de cierto conocimiento (formalizado mediante lógica clásica)". Se trata de un debate que no podrá ser profundizado aquí, pero sólo se mencionará que si la vara para medir la fortaleza de un sistema es la cantidad de consecuencias que pueden obtenerse a partir de una base, el sistema universal $\mathcal{L}$ tarskiano (con las dificultades mencionadas en el **CAPÍTULO 2** relativas a su inconsistencia) sería el más fuerte de todos los sistemas, lo cual resulta problemático.

Por otra parte, Schlechta destaca la relación conceptual estrecha entre la definición modelo-teorética de suavidad y la propiedad axiomática de cumulatividad. Esta relación ya era patente en la demostración de esta última propiedad efectuada por KLM en el marco de la demostración de corrección del sistema C (reconstruída en **3.1.15**). De hecho, para Schlechta, cumulatividad y suavidad son dos herramientas imprescindibles, que permiten establecer muchas de las condiciones de coherencia consignadas previamente. Además, existe un vínculo estrecho entre estas nociones y propia noción de coherencia, que se define, en términos generales como la transferencia de conclusiones desde una situación T a una situación T'. En LE $T \vdash \alpha$ implica $T' \vdash \alpha$ si $T \subseteq T'$ (monotonía). Las operaciones de transferencia permiten pensar de manera *relevante*, al transferir conclusiones desde una situación a otra sin necesidad de comenzar de nuevo cada vez. En LNM, tal transferencia es también posible, "cuando la lógica es cumulativa" (Schlechta, 2004:38). Las condiciones de coherencia son *imprescindibles* en el sentido de que constituyen "más de la mitad de una lógica, en particular, las propiedades de coherencia son a menudo la llave para los resultados de representación" (Schlechta, 2004:39), siendo resultados de representación aquellos relativos a demostraciones de corrección, completitud y problemas afines como numerabilidad, etc.[92]. En este punto, teniendo en cuenta que el interés por encontrar un conjunto de condiciones o propiedades que caracterizan una LNM (la "definición 18" de Gabbay, las propiedades *interesantes* de Makinson, las reglas KLM…) se ha desplazado en Schlechta a las mencionadas condiciones de coherencia, cabe preguntarse respecto a la cuestión de la "pretensión de universalidad" o el alcance de tales condiciones, dentro del enfoque del autor.

En el texto citado, se ofrecen algunas pistas al respecto. En primer lugar, se afirma que las LNM's fueron diseñadas para formalizar aspectos del razonamiento acerca de los *casos normales* (interesantes, mayoritarios, etc.). Según el autor, hay muchas LNM's porque hay muchas intuiciones respecto a cómo razonar no deductivamente, y cada intuición se codifica en un formalismo. De modo que "ser no monótono es una propiedad, no el nombre de una lógica única" (Schlechta, 2004:40), tal cual reza el epígrafe de la **INTRODUCCIÓN** de esta tesis. Y justamente, en el contexto de esta tesis, no puede dejar de destacarse que el autor no ofrece una definición estrictamente formal de esta propiedad fundamental (tampoco se observa una formulación de ese tipo en el más reciente (D. Gabbay & Schlechta, 2016)). Pero a simple vista, el enfoque de Schlechta no debería excluir, en principio, la posibilidad de la existencia de LNM's que no cumplan

[92] En esta sección se omiten los temas de representación, considerándose suficiente para los fines de esta tesis con lo expuesto en la sección **3.1** para el sistema de KLM.

específicamente las condiciones de coherencia listadas previamente, en tanto el autor no identifica la propiedad de *no monotonía* con la noción de *coherencia*, y reconoce la existencia de múltiples intuiciones (y formalismos) relativos al razonamiento práctico o de sentido común (no monótono). Por otra parte, es interesante notar que (en el marco de una digresión filosófica) el autor postula un criterio general para las LNM´s cuando establece una condición modelo-teorética *sine qua non* para que un sistema sea considerado no monótono. Esta condición es la propiedad modelo-teorética **3.2.8**, o, en términos metalógicos, supraclasicalidad (**3.2.17** C)): "todas las LNM's pueden ser representadas por una función de elección de modelos f tal que $f(X) \subseteq X$" (Schlechta, 2004:40). Este criterio puede resultar un tanto sorprendente, en tanto argüiblemente excluye al sistema C y al resto de los sistemas de KLM, ya que estos no satisfacen ninguna forma de supraclasicalidad.

No podrá profundizarse aquí sobre esta nueva aporía filosófica. Sólo se mencionará que posiblemente la ausencia en la bibliografía del área de una expresión formal de no monotonía (tal como la propuesta en **2.3.1**) es lo que genera diferencias de criterios y dificultades en la clasificación de sistemas como *no monótonos*. A favor Schlechta, hay que mencionar que el estudio de temas metadisciplinarios como la clasificación de sistemas y la investigación sobre propiedades metalógicas (así como otros temas de filosofía de la lógica como el uso del TNM y la relación entre metalenguaje y lenguaje objeto, entre otros que se han abordado aquí) constituyen ejes subsidiarios en el marco de una obra cuyos principales objetivos pasan mayormente, como se mencionó antes, por el eje estrictamente sistemático de ofrecer una de *caja de herramientas* de teoría de modelos para los fines de la construcción de sistemas de LNM. Aun así, queda claro que, por definición, cualquier sistema reconstruido o construido mediante el enfoque LMG de Schlechta cumple con aquella condición mínima **3.2.8** / **3.2.17** C).

3.3 PROPIEDADES CENTRALES RESTAURADAS

A modo de conclusión de este capítulo, se ofrecerán breves reflexiones algunas de las cuales serán retomadas en el **CAPÍTULO 4**. En primer lugar, en las secciones previas, se ha intentado ofrecer una exposición sintética basada en un recorte temático que apunta a reconstruir las bases de las propuestas de KLM y Schlechta, a fin de mostrar la estructura que *debería* tener un sistema de LNM para superar las objeciones y contraejemplos a las propiedades centrales de Gabbay. En ese sentido, se hace patente el doble carácter descriptivo y prescriptivo inherente a las reflexiones metalógicas y metamatemáticas, mencionado en el **CAPÍTULO 1** sección **2**. Con la convicción de que no es suficiente con afirmar que *las propiedades centrales pueden ser restauradas bajo el enfoque semántico de la LNM*, sino que es necesario explicitar (al menos de modo abreviado) las bases y fundamentos modelo-teoréticos que permiten tal restauración, se ha intentado sintetizar y comentar algunos de los aspectos centrales de la teoría de modelos preferenciales de KLM y de la (mucho más compleja) LMG de K. Schlechta. En ese camino, se ha realizado un esfuerzo por evitar detenerse excesivamente en tecnicismos, y mantener un enfoque equilibrado entre lo metalógico, lo filosófico y lo estrictamente sistemático de la propuesta de cada autor. Pero no se ha podido evitar una discusión medianamente extensa respecto al uso de TNM en la sección **3.1**, debido a la importancia mayúscula de este tema para la interpretación las propiedades centrales en la concepción de KLM.

Con esa perspectiva, ha quedado expuesto a lo largo del capítulo que las propiedades centrales de la LNM (básicamente reflexividad + cumulatividad[93]) son demostrables para

[93] Cfr. (Strasser & Antonelli, 2019).

sistemas preferenciales basados en el sistema C de KLM (véase las demostraciones en la sección **3.1**); y para sistemas basados en LMG (Schlechta), en este último caso con un fundamento modelo-teorético anclado en las propiedades de las funciones de elección (véase sección **3.2**). Además, en el marco de LMG, se han restaurado dos propiedades adicionales, las cuales resultan de enorme importancia en la propuesta originaria de Gabbay (1985): *supraclasicalidad* y *conservación de la consistencia*. La satisfacción de la primera de ellas, según la propuesta de Schlechta, constituye el criterio para clasificar un sistema como no monótono. Por otra parte, conservación de la consistencia refleja el hecho de que los sistemas reconstruidos en base al enfoque LMG superan el problema (que afecta a los sistemas preferenciales tradicionales) de la generación de inconsistencias a partir de teorías consistentes. De manera que, en términos generales, y en tanto la LMG de Schlechta *absorbe*, o permite reconstruir los sistemas de KLM, puede afirmarse que las propiedades centrales no sólo son restauradas bajo el enfoque semántico de la LNM, sino que además son *ampliadas*. Las mencionadas en (Strasser & Antonelli, 2019) pueden ser aumentadas y reconfiguradas en una tripleta de propiedades, a saber: supraclasicalidad, cumulatividad y conservación de la consistencia. Como ha sido mencionado previamente, supraclasicalidad supone reflexividad, y cumulatividad puede desglosarse en monotonía cauta y corte.

Aquí, finalmente, se halla un suelo firme (dentro de ciertos límites) para la familia de teorías de LNM representadas bajo el enfoque semántico: se trata de sistemas que reflejan el razonamiento de sentido común, el fenómeno de eliminación de conclusiones, y que satisfacen un conjunto ampliado de propiedades centrales que hacen a la *estabilidad* del razonamiento de los agentes. Así, el resultado principal de este capítulo es la explanación de la *restauración parcial* y *reconfiguración* de las propiedades centrales tradicionales de la LNM, con lo cual se consideran salvadas (para la familia de teorías de LMG, en resumidas cuentas, para el enfoque semántico de la LNM) las objeciones y contraejemplos planteados en el **CAPÍTULO 2**.

Sin embargo, en el transcurso de tal explanación han surgido una serie de tópicos y problemas de orden filosófico y metalógico, muchos de los cuales han quedado abiertos por el momento, y otros sólo serán planteados en esta sección. Junto a la aporía del TNM, no puede dejar de mencionarse que ni KLM ni Schlechta generan formulaciones estrictamente simbólicas de la propiedad de no monotonía. Aun así, en (Schlechta, 2004) se ofrecen herramientas modelo-teoréticas y sintácticas que pueden resultar de interés para ese fin, específicamente, la noción de semigeneralización y la teoría de filtros, que serán utilizadas para la formulación y cimentación semántica de una variante de **2.3.1** en el **CAPÍTULO 4**. Además, es importante mencionar las dificultades relativas a la consistencia de las bases de conocimiento de los agentes racionales. Particularmente, permanece abierto para la LMG el problema generado por la paradoja de Israel. Como se indicó en el **CAPÍTULO 2**, esta paradoja pone a la propiedad de conservación de la consistencia en *tensión filosófica*.

En el caso de los sistemas basados en LMG, simplemente los razonadores no podrán obtener conclusiones plausibles cuando la teoría es inconsistente (véase sección **3.2**). Esto parece indicar una *exigencia implícita* de consistencia de las bases de conocimiento para trabajar en sistemas de LMG. Esta exigencia por un lado resulta muy conveniente para la propuesta global de Schlechta, pues los sistemas de LMG satisfacen supraclasicalidad como condición de coherencia fundamental. Pero a su vez, marca un límite para este tipo de sistemas, en tanto un sistema supraclásico debería lidiar, de un modo u otro, con la ley de explosión lógica. Lisa y llanamente, ante una contradicción, ¿el sistema será capaz de evitar la explosión de conclusiones puramente deductivas? En el texto de referencia (Schlechta, 2004) no se hallan indicios respecto al modo de abordar este problema. En este punto, se plantean un par de alternativas igualmente *incómodas*: si el sistema no es explosivo, se genera una tensión con la propiedad de supraclasicalidad, y si es explosivo, se aleja del comportamiento de los agentes de sentido común (tanto como los sistemas de LE), en tanto los razonadores comunes simplemente no se comportan de manera explosiva. Adicionalmente, en general en LNM cualquier exigencia de consistencia de las bases de conocimiento resulta *ad hoc*, teniendo en cuenta los objetivos de la IA de modelar agentes racionales tal como se comportan en la vida real, y la *paradoja de Israel* **2.2.11** que afecta a tales agentes.

En el próximo **CAPÍTULO 4** se indicará una posible vía para la superación (al menos parcial) de la paradoja de Israel mediante la formulación de una propiedad que captura la noción de semimonotonía (**4.2.1**), aunque en principio tal solución no podría aplicarse para los sistemas de LMG *tal como son* en la propuesta de Schlechta, debido a la exigencia implícita de consistencia de las bases de conocimiento. No obstante, la propiedad de semimonotonía y las condiciones de coherencia (particularmente las propiedades centrales) parecen tener una raíz filosófica común en la búsqueda de describir/prescribir condiciones que deberían satisfacer los sistemas que se proponen modelar agentes racionales de sentido común, a saber, en la noción de *estabilidad* de tales razonadores. Posteriormente se retornará sobre este asunto.

Otra dificultad filosófica notable que se ha presentado en este capítulo es el problema de la distinción entre lenguaje objeto y metalenguaje, directamente relacionado con la aporía del TNM. Al respecto, se ha intentado balancear la tradición con la letra original de los autores, y procurando interpretaciones caritativas desde el punto de vista filosófico. Pero estos tópicos no han sido clausurados completamente, y aún puede problematizarse y profundizarse sobre los mismos. En el **APÉNDICE 2** se ofrecerá una breve digresión filosófica que, según espera el autor, pueda contribuir al debate. Finalmente, aun cuando puedan señalarse algunos problemas filosóficos y metalógicos, sin lugar a duda las propuestas de KLM y Schlechta resultan de enorme e indiscutible importancia desde el punto de vista estrictamente sistemático y disciplinario, en tanto constituyen una referencia ineludible para cualquier intento de fundamentación semántica de sistemas de LNM: de hecho, la LMG puede considerarse sin más como *el enfoque semántico de LNM*. Asimismo, ambas propuestas juegan un rol similar en asuntos metalógicos de LNM, pues además de restaurar (dentro de ciertos límites) e incluso ampliar las propiedades centrales, sus construcciones modelo-teoréticas han sido utilizadas exitosamente para lograr sólidos resultados de representación, como teoremas de corrección y completitud (véase sección **3.1**).

CAPÍTULO 4 – SEMIMONOTONÍA.

INTRODUCCIÓN

En este último capítulo, se ofrecerá en la sección **4.1** una breve recapitulación de algunos de los tópicos más importantes abordados a lo largo de la tesis, especificando los resultados alcanzados en cada caso. Asimismo, en relación con estos tópicos, se presentará una conjetura relativa al comportamiento racional de ciertos agentes de sentido común, denominada *conjetura del razonador semimonótono*. La misma servirá de anclaje filosófico para la sección **4.2**, en la cual se formulará una nueva propiedad denominada *semimonotonía* (**4.2.1**), que es una variante de la propiedad de no monotonía (**2.3.1**). El estudio y formulación de **4.2.1** persigue tres objetivos fundamentales: a) reflejar idealmente el comportamiento de un tipo particular de razonadores de sentido común, los *razonadores semimonótonos*, dotando a los sistemas que los modelan de un fundamento metamatemático rudimentario, con un componente axiomático y un componente modelo-teorético asociados[94]; b) rescatar e integrar la intuición intelectual detrás de las nociones de *estabilidad* y *coherencia*, que se encuentran en la raíz filosófica de las propiedades centrales tradicionales de la LNM; c) ofrecer una respuesta al desafío planteado por la

[94] Y que, claro está, pueden ser ampliados según las necesidades específicas de cada sistema.

paradoja de Israel (problema de la inconsistencia generalizada de las bases de conocimiento). Por último, de manera subsidiaria a los objetivos del capítulo, se ofrecerá una conclusión provisoria respecto a la posibilidad de reemplazar la LNM por la lógica paraconsistente, tal como se propone en (Estrada & Garcia, 2009).

4.1 RECAPITULACIÓN Y CONJETURA DEL *RAZONADOR SEMIMONÓTONO*

En los capítulos **1 - 3**, se ha seguido un camino que intenta integrar un enfoque metalógico/metamatemático y una mirada filosófica (de filosofía de la lógica), respecto a temáticas y problemas que emergen del ámbito estrictamente disciplinario y sistemático de la LNM. Se ha tomado como punto de partida un análisis en profundidad del concepto de monotonía, en un **CAPÍTULO 1** de carácter axiomático que puede caracterizarse como netamente *clásico* y mayormente metalógico/metamatemático. El mismo comienza con una simple pregunta: ¿qué significa *monotonía*? (en LE), y a partir de esa pregunta se abre un ámbito de indagación que lleva a una profundización y un análisis de las propuestas clásicas que definen tal concepto como propiedad, regla estructural, como concepto e intuición intelectual. En el camino, se abordaron problemas subsidiarios como el significado estricto de varias de las principales propiedades y axiomas de la lógica clásica, la propia noción de sistema, el tópico de la relación entre información e inferencias y el problema de la inconsistencia del sistema universal $\mathcal{L}$ tarskiano, entre otros. Este primer capítulo, con un desarrollo puramente *analítico*, sentó las bases metamatemáticas para abordar una segunda pregunta que guía el **CAPÍTULO 2**: ¿qué significa *no monotonía*? En este caso, se trató de un abordaje también axiomático y metalógico, pero con mayor anclaje en la filosofía de la lógica (si se lo compara con el capítulo previo). El punto de partida es la problematización de un concepto, el de no monotonía, que se ha incorporado al léxico de la lógica de un modo tan familiar que no parece requerirse ya ninguna aclaración. No obstante, el mismo aún puede ser puesto en cuestión y enfocado de un modo estrictamente filosófico y metalógico. No se trata de una temática clausurada, como ha quedado claro a lo largo del capítulo citado.

De hecho, de tal concepto sólo se tienen, en la bibliografía de referencia, algunas intuiciones, no siempre claras, algunas definiciones semi-formales, y una aproximación indirecta mediante la postulación de ciertas propiedades (denominadas *centrales*) que originariamente se pretendía que funcionaran (de conjunto) como una definición de no monotonía, pero que luego fueron fuertemente objetadas. Junto a ello, la ausencia de una definición estrictamente formal de no monotonía constituye una importante dificultad filosófica y metalógica del campo considerado. Reduciendo y simplificando el asunto *in extremis*, justamente esa problemática es la que ha sido atacada a lo largo de este trabajo. Como resultado, el autor de esta tesis propone una definición estrictamente formal, que integra simbolismos de LE, lógica modal tradicional, notación conjuntista, el operador C no monótono (Makinson), y que toma como *premisa* la definición **2.2.4** de Gabbay, globalmente aceptada en la bibliografía del área. La definición formal **2.3.1** (junto a su equivalente **2.3.2**) se propone, de este modo, unificar y expresar metamatemáticamente las intuiciones primordiales sobre *la propiedad fundamental* de la LNM, muy probablemente la única propiedad común a todos los sistemas del área.

El autor de esta tesis postula tales expresiones metamatemáticas con la esperanza de que funcionen como fórmulas que definan la propiedad universal de no monotonía para todos los sistemas independientemente de sus características particulares, de otras propiedades y condiciones que estos puedan satisfacer, de sus construcciones, herramientas y aplicaciones. Como se mencionó en **INTRODUCCIÓN Y OBJETIVOS**, la finalidad de explorar las vías para la construcción de una expresión puramente formal de no monotonía es *contribuir a la fundamentación metalógica de la LNM*.

Pero inmediatamente se abre otra cuestión, pues a fin de cuentas ¿qué otras propiedades satisfacen las LNM's? Particularmente, ¿qué ocurre con las propiedades centrales? Estas se encontraban en el origen de la reflexión metalógica sobre LNM, y aún resultan de una enorme importancia en la bibliografía del área. Independientemente del debate sobre su status metadisciplinario y las objeciones planteadas, las mismas expresan una indagación filosóficamente relevante, que pregunta por las condiciones bajo las cuales puede considerarse a un agente racional como *estable*, constructor y conservador de conocimiento (dentro de ciertos límites marcados por la propiedad de no monotonía). De allí la importancia de indagar con mayor profundidad sobre las propiedades centrales. En el **CAPÍTULO 3** se han abordado estos asuntos desde una perspectiva más semántica que axiomático/sintáctica, y con un estilo de exposición mayormente *sistemático*, con un menor grado de abstracción metadisciplinaria en relación con los capítulos anteriores. No obstante, la exposición se encuentra atravesada por inevitables asuntos de filosofía de la lógica relativos al uso del símbolo del TNM, la distinción entre lenguaje objeto y metalenguaje, problemas de consistencia, entre otros.

La conclusión del capítulo es que resulta posible restaurar las cuestionadas propiedades centrales, siempre y cuando se interprete el TNM de manera tradicional (es decir, sin debilitar las propiedades), pero pagando un precio relativamente alto: el de la *universalidad* atribuida por Gabbay a estas propiedades en relación con los sistemas de LNM, pretensión que no puede seguir sosteniéndose. Al parecer, ni siquiera puede sostenerse la *deseabilidad* (cierta *generalidad restringida)* de Makinson, simplemente, las propiedades centrales lo son para sistemas basados en (construidos o reconstruidos mediante) el enfoque semántico de LNM, particularmente la teoría de modelos preferenciales tradicional (KLM) y la LMG de Schlechta. En este último enfoque, además, es posible restaurar otras propiedades importantes, que reconfiguran la clase de las propiedades centrales en una tripleta compuesta por supraclasicalidad, cumulatividad y conservación de la consistencia (véase sección **3** del **CAPÍTULO 3**).

Pero fuera de esta familia de teorías no puede garantizarse ni parece ser necesaria la satisfacción de tales propiedades. De hecho, resulta perfectamente aceptable que un sistema probabilístico sea no monótono y no cumulativo, o que un sistema que restaura consistencia se no monótono y no supraclásico. A su vez, estos sistemas pueden satisfacer otras propiedades, que pueden no ser condiciones aceptables para sistemas basados en LMG o en cualquier otro enfoque. Así, una de las principales conclusiones del capítulo es que las denominadas *propiedades centrales* pueden ser restauradas (y ampliadas) en el marco de sistemas basados en LMG, pero se trata de una *restauración parcial*, es decir, que estas no pueden seguir siendo consideradas propiedades universales, genéricas o únicas, ni siquiera parecen ser privilegiadas (incluso podría cuestionarse el mote de *centrales*). En los hechos se trata de propiedades importantes sólo para una familia de teorías de LNM, y esto no hace más que reafirmar que en este campo teórico todo apunta a la *diversidad* y cualquier pretensión universalista por fuera de la propiedad fundamental (no monotonía, expresada en **2.3.1-2**) debería ser abandonada.

De cualquier modo, ninguno de los autores analizados en el **CAPÍTULO 3** renuncia explícitamente a cierta pretensión de universalidad o, al menos, de generalidad de las propiedades centrales. En KLM, las propiedades de reflexividad, corte y monotonía cauta constituyen las condiciones mínimas sin las cuales "un sistema no debe considerarse un sistema lógico" (Kraus et al., 1990:176), y Schlechta postula supraclasicalidad como condición sin la cual un sistema no puede considerarse no monótono. En ambos casos se trata de afirmaciones que, voluntariamente o no, establecen un criterio de clasificación de sistemas lógicos. Lamentablemente, estos criterios no son desarrollados en profundidad por los autores, por lo que deberían ser tomados más bien como *declaraciones de principios filosóficos*, *opciones metodológicas* o incluso *propuestas no justificadas* de criterios de clasificación. A fin de cuentas, los autores deberían justificar por qué motivos un sistema debería satisfacer obligatoriamente *esas* propiedades y no otras, para considerarse no monótono.

Además, tales criterios resultan excesivamente restrictivos, e incluso generan dificultades de clasificación, por ejemplo, la propuesta de Schlechta excluye los sistemas preferenciales basados en el sistema C de KLM, ya que estos no satisfacen ninguna forma de supraclasicalidad. La postura más razonable al respecto parece ser aceptar que las semánticas construidas por KLM y Schlechta delimitan una familia de teorías de LNM que satisfacen ciertas propiedades, pero esto no excluye de ningún modo la posibilidad de lógicas no monótonas que no satisfagan algunas o ninguna de ellas, e incluso puedan resultar estables (en mayor o menor medida). Como alternativa a los mencionados criterios no justificados de estos autores, quien suscribe esta tesis propone las propiedades **2.3.1-2** como lineamientos de clasificación de sistemas no monótonos, opción fundada en la definición semi-formal de no monotonía de Gabbay (**2.2.4**). Así, de manera accesoria, **2.3.1-2** pueden funcionar también como un *criterio de clasificación formal* de sistemas lógicos.

Por otra parte, la restauración de la propiedad de conservación de la consistencia en LMG constituye un claro avance teórico para el enfoque semántico de la LNM, ya que las construcciones basadas en versiones límite con segmentos iniciales minimizadores, junto a las de preservación de definibilidad, permiten superar uno de los problemas de los sistemas preferenciales tradicionales, a saber: la generación de inconsistencias a partir de teorías consistentes (cuando se producen cadenas de preferencias descendentes infinitas). Aun así, tal como se indicó en la sección **3.3**, esta propiedad se encuentra en *tensión filosófica*. Si bien es muy importante para los sistemas lógicos basados en LMG superar el problema de la generación de inconsistencias, estos sistemas aún son susceptibles de fallar ante teorías inconsistentes (por ejemplo, en un sistema preferencial una teoría inconsistente no tendrá modelos, y por lo tanto no tendrá modelos minimales). Como se especificó previamente, esta situación en la práctica impone una cláusula ad hoc de consistencia de las bases de conocimiento, pero al precio de alejarse del comportamiento de los razonadores reales, los cuales normalmente utilizan bases de conocimiento inconsistentes. Este problema no podrá ser profundizado aquí para sistemas LMG. Sobre la situación puntual de inconsistencia generalizada planteada en **2.2.11** (paradoja de Israel), en el **CAPÍTULO 2** se consignó que existen para las LNM´s al menos dos opciones razonables: implementar mecanismos que permitan *tolerar* esas inconsistencias (lo cual las acercaría a alguna forma de paraconsistencia), o bien implementar mecanismos de restauración de la consistencia (lo cual puede tener repercusiones a nivel de las propiedades centrales de reflexividad y supraclasicalidad, como se ejemplificó en **2.2.12**).

Pero podría agregarse una opción más compleja, que depende de algunas consideraciones filosóficas. Esta tercera opción vendría a combinar restauración de la consistencia con tolerancia de la inconsistencia, en un marco de *estabilidad*, en el sentido que da Schlechta a este término, es decir, *regularidad para razonar con eficiencia* (véase **CAPÍTULO 3** sección **2**), que se asemeja al comportamiento habitual de un tipo particular de *agentes de sentido común*. Se trata de aquellos razonadores que no transitan su vida cotidiana con posturas dogmáticas extremas (tolerando todas sus inconsistencias), ni con posturas escépticas igualmente extremas (eliminando todas sus inconsistencias), ni siendo explosivos (obteniendo todas las conclusiones a partir de una contradicción). Se trata de describir (o incluso construir, recuérdese el doble status descriptivo – prescriptivo de la metalógica) un razonador *pragmático*, que suele tolerar algunas inconsistencias y eliminar otras, en virtud de consideraciones mayormente prácticas (y, por ende, extralógicas). Si la paradoja de Israel es correcta y refleja el hecho fáctico de la inconsistencia generalizada de las bases de conocimiento, aún cabe preguntarse por el comportamiento de los razonadores ante tal situación: ¿la eliminación de conclusiones es la excepción o la norma? ¿Qué ocurre a su vez con la explosión lógica? Según el punto de vista del autor de esta tesis, el comportamiento humano puede variar enormemente, pero el primer punto que debería quedar zanjado casi automáticamente es la explosión lógica: como se mencionó en la sección **3.3**, los razonadores de sentido común no se comportan de ese modo.

Por otra parte, existe un tipo de razonador (no explosivo) que podría denominarse *semimonótono* en el sentido de que parece tener un comportamiento normal que tiende más bien a la búsqueda de incrementar o mantener su conocimiento, y de obtener conclusiones tanto deductivas como plausibles, sosteniéndolas hasta tener evidencia práctica de la necesidad de su eliminación. Así, estos razonadores semimonótonos si bien eliminan conclusiones, sólo lo hacen cuando se ven forzados a ello *por la realidad*, por la práctica o por exigencias extralógicas, independientemente de la inconsistencia generalizada de las bases de conocimiento. Un ejemplo de este tipo de razonadores puede ser representado mediante la idea filosófica del *hombre de término medio*, moderadamente dogmático y poco propenso a revisar sus creencias. Salvando las distancias entre disciplinas, incluso algunas de las teorías epistemológicas más importantes describen un comportamiento de los científicos que podría asemejarse, quizá de manera un poco más atenuada respecto al *hombre de término medio*, a este tipo de *racionalidad semimonótona*. Autores como T. Kuhn en su (1962) o I. Lakatos en su (1978), en debate con el falsacionismo popperiano, sostienen que los científicos intentan *salvar* sus teorías de las objeciones y contraejemplos, antes que revisarlas o retractarlas, incluso cuando estas chocan con los datos empíricos generándose inconsistencias evidentes en el corpus teórico considerado.

Ahora bien, de la mera descripción de un tipo de razonadores semimonótonos o *estables*, no explosivos, que eliminan conclusiones sólo en casos más o menos excepcionales, se desprende trivialmente que su comportamiento racional es, en términos generales, *no monótono* en el sentido de las fórmulas **2.3.1-2**. Del mismo modo, en la medida en que un agente tenga un comportamiento mayormente *crítico* o *autocrítico*, en la medida en que sea *propenso habitualmente* a revisar sus conclusiones previas, se alejará progresivamente de esta idea de semimonotonía, aunque podrá seguir siendo considerado no monótono en sentido estricto. En cuanto a los razonadores semimonótonos, hay que agregar que, al no ser explosivos, no resulta conveniente que un sistema que modele su comportamiento sea supraclásico, pues en ningún caso podrá tolerar inconsistencias sin generar una explosión lógica. Pero quizá pueda precisarse aún más la descripción de esta conjetura del razonador semimonótono. Para finalizar esta tesis, en la próxima sección se explorará esta posibilidad.

4.2 PROPIEDAD DE SEMIMONOTONÍA

La expresión **2.3.1** (propiedad de no monotonía) implica que en un sistema que la satisface para un agente será accesible al menos un mundo donde no se cumple monotonía. En tanto en esa expresión se toman como referencia (o se extrapolan) algunas características del sistema T, los mundos accesibles al agente pueden o no coincidir con el conjunto de todos los mundos posibles, y no se establecen especificaciones adicionales respecto a sus características. En otros términos, no se sabe *cuándo* emergerá ese mundo que hará fallar monotonía, sólo se sabe que será accesible *en algún momento* para el agente. Una caracterización análoga puede construirse desde la tradición de KLM: la propiedad de no monotonía garantiza que existe un mundo $m' \prec m$ que hace fallar monotonía, pero a priori no se puede definir cuál es ese mundo minimal (y el recurso a estructuras con copias sólo permite construir una *gran disyunción de posibilidades* entre todos los mundos posibles, lo cual no resulta muy informativo. Véase **CAPÍTULO 3** sección **2**).

Esta característica hace que la definición **2.3.1** sea lo suficientemente amplia como para abarcar todo tipo de sistemas que admiten eliminación de consecuencias, tanto si el sistema admite retracción de conclusiones en *la mayoría menos uno* de mundos posibles, en *sólo en uno* de ellos, o incluso en todos. Así, esa formulación de no monotonía engloba una amplia gama de *tipos de razonadores*, desde casos extremos, como por ejemplo el de un razonador pirroniano comprometido con la destrucción de todas sus certezas partiendo de la inconsistencia generalizada de su propio conocimiento, tanto como razonadores *casi dogmáticos* que procuran mantener al mínimo la retracción de conclusiones, y *en el medio*, un amplio rango de razonadores más o menos *críticos*.

En textos metalógicos de LNM como los citados a lo largo de esta tesis, desde el fundacional (Gabbay, 1985), pasando por (Makinson, 1994), (Kraus et al., 1990) y (Schlechta, 2004) entre otros, en general se buscan propiedades que permitan establecer una cierta *estabilidad* de los sistemas ante la falla de la propiedad de monotonía, búsqueda que puede vincularse, con reservas, a la conjetura del razonador semimonótono[95]. Es que, para estos autores, al no cumplirse monotonía, al menos deberían cumplirse otras condiciones (por ejemplo, las propiedades centrales o las condiciones de coherencia) que garanticen una persistencia parcial del conocimiento, donde el razonador no sea constantemente forzado a destruir conclusiones. Esta idea es la que Schlechta intenta rescatar mediante la noción de *coherencia*, que fija ciertos requisitos para lo que puede ocurrir a la derecha del TNM (en las conclusiones), cuando se realizan cambios a la izquierda (en las premisas). Tal búsqueda de *propiedades de estabilidad* por parte de los autores mencionados parece reflejar un intento por reconstruir el comportamiento de razonadores de sentido común que normalmente conservan sus conclusiones previas, o al menos, en la mayoría o en subconjuntos grandes de los casos, mientras que en casos más o menos excepcionales retractan algunas de ellas. Al parecer, la intuición detrás de esta noción de estabilidad es análoga a la conjetura de los razonadores semimonótonos[96].

El comportamiento semimonótono de los agentes puede reflejarse formalmente, utilizando las herramientas adecuadas, como una propiedad de sistemas no monótonos (que satisfacen **2.3.1-2**). En la construcción se utilizará como operador modal un tipo de semigeneralizador, que refleja una intuición similar al símbolo $\triangledown$ utilizado por Schlechta en su (2004). Sin embargo, en tanto el semigeneralizador $\triangledown$ se vincula originalmente con filtros débiles en el contexto de la reconstrucción de las lógicas default, aquí se optará por introducir el símbolo "◬" como operador modal para expresar una intuición similar (aunque no idéntica) al semigeneralizador de Schlechta. El operador modal ◬ posee un

[95] Particularmente, la conjetura del razonador semimonótono excluye la propiedad de supraclasicalidad, que es una de las condiciones de coherencia en (Schlechta, 2004).

[96] De cualquier modo, los siguientes desarrollos no dependen en absoluto de la aceptación de estas últimas reflexiones respecto a las propiedades centrales y las propiedades de coherencia.

anclaje semántico en la estructura de ideales (clásicos), definida posteriormente en **4.2.4**, y trabaja sobre conjuntos arbitrarios de sentencias de un sistema:

4.2.1 ⩓$\Delta\ \forall\Gamma((\Delta \subseteq \Gamma) \longrightarrow (C(\Delta) \subseteq C(\Gamma)))$

Esta definición es claramente menos amplia que **2.3.1**. La interpretación intuitiva es la siguiente: para la mayoría de las bases de conocimiento y para toda ampliación, se conservan consecuencias. En otras palabras, se captura la idea de una cierta *estabilidad* (semimonotonía) de las bases de conocimiento de los agentes no monótonos, aun cuando puedan fallar otras propiedades de coherencia, incluso algunas de las propiedades centrales restauradas para sistemas de LMG como supraclasicalidad (en el caso de sistemas que restauran consistencia) o cumulatividad (en sistemas probabilísticos). El semigeneralizador ⩓ también puede entenderse como indicando que para subconjuntos grandes de las bases de conocimiento se garantiza una condición análoga a monotonía, y ante ampliaciones arbitrarias sólo casos más o menos excepcionales entran en conflicto con alguna sentencia de la ampliación considerada, activando mecanismos de eliminación de sentencias que hacen fallar la condición análoga a monotonía.

En el caso de bases de conocimiento inconsistentes, si un sistema satisface **4.2.1** y no es explosivo, esto implicaría que no es supraclásico, y que se aproxima a alguna forma de *paraconsistencia*, en el sentido de que normalmente o en la mayoría de los casos, el razonador toleraría inconsistencias. Es decir, el agente semimonótono admitiría dentro de ciertos límites (en la mayoría de los casos, o al menos en un subconjunto grande de casos) las inconsistencias. De ese modo podría resolverse parcialmente la paradoja de Israel para razonadores semimonótonos, pues si bien existe una inconsistencia generalizada de las bases de conocimiento, esto no afecta a este tipo de razonadores (al menos en la mayoría de los casos). Por ejemplo, las inconsistencias relacionadas con la mencionada paradoja no activarían eliminación de sentencias salvo en casos excepcionales, y los agentes semimonótonos convivirían más o menos cómodamente con ellas. Por supuesto, para otro tipo de razonadores filosóficos, escépticos o extremadamente críticos, menos propensos a tolerar inconsistencias, la paradoja no pierde vigencia. Cabe aclarar, aunque no se podrá profundizar aquí, que la concepción de semimonotonía **4.2.1** no implica que los agentes sean *paraconsistentes* en el sentido tradicional del término, sino algo bastante más complejo. Se trata de reflejar un comportamiento tal que los agentes, si bien mayormente toleran las inconsistencias (y no son explosivos), en ocasiones fuerzan retracción de conclusiones, fenómeno que no es tomado en cuenta en la lógica paraconsistente tradicional[97]. Así, en principio las LNM no podrían reducirse a las lógicas paraconsistentes, como se postula en (Estrada & Garcia, 2009).

En síntesis, **2.3.1-2** parecen ser mejores candidatas a expresiones universales y estrictas de no monotonía (en comparación con **4.2.1**), pues establecen las condiciones mínimas para que un sistema sea considerado no monótono, y en ese sentido constituyen una verdadera *herramienta para la clasificación de sistemas lógicos*. Por su parte, **4.2.1** expresa un modo diferente de enfocar la noción de no monotonía, en un sentido más específico, a saber: para sistemas que intentan reconstruir el comportamiento de *razonadores estables*, que mayormente conservan consecuencias y no siempre retractan sus conclusiones ante inconsistencias o por otros motivos. Para finalizar, se ofrecerá una construcción modelo-teorética que sirve como fundamento semántico para la expresión **4.2.1**. Lo primero a recalcar es que si bien tal propiedad incluye un operador modal similar a ▽, aquí se opta por utilizar un símbolo diferente para expresar una intuición similar la que expresa el semigeneralizador de Schlechta, el símbolo ⟁. Se trata de una diferencia de notación que tiene un fundamento en la interpretación semántica de cada uno de los

[97] Para una introducción al tema, Cfr. (Priest et al., 2018).

símbolos, ya que $\triangledown$ remite a una estructura de filtro débil, mientras que ⟁ se fundamenta en la estructura de ideal (en sentido clásico, como es presentado por ejemplo en (Manzano, 1999) y (Bell & Slomson, 1969)). El motivo de esta divergencia se halla en las dificultades para extrapolar $\triangledown$ y filtros débiles tal como son utilizados por Schlechta, es decir, para reconstruir lógicas default y operando sobre conjuntos de elementos de un dominio, al terreno metalógico, donde **4.2.1** opera sobre conjuntos de sentencias. En el primer caso, se trata de objetos que cumplen propiedades, mientras que en el segundo se trata de fórmulas de las cuales se obtienen consecuencias. Otra forma de entender el problema es afirmando que $\triangledown$ es un operador que cumple una función semejante a un cuantificador de primer orden, mientras que ⟁ se asemeja más a un cuantificador de segundo orden. Por ello, para los fines de fundamentar semánticamente **4.2.1**, conviene adoptar la estructura de ideal clásico (definición **4.2.2**), que refleja mejor a la intuición detrás de la propiedad de semimonotonía y del operador ⟁.

Cabe recordar que en el **CAPÍTULO 3** sección **2** se mencionó que la definición de "ideal clásico" admite la construcción de conjuntos que son subconjuntos de cualquier tamaño respecto al conjunto base, y por lo tanto no se requiere necesariamente vincular la noción de ideal con la idea de subconjuntos pequeños. Esta flexibilidad teórica no está reñida con la tradición de Schlechta, quien por ejemplo justifica la introducción de la estructura de filtro débil del siguiente modo: "la noción de filtro fue creada para otros propósitos diferentes a los nuestros, por lo que a priori no hay necesidad de considerarla sin cambios" (Schlechta, 2004:84). Lo importante para el autor es rescatar las *intuiciones esenciales* detrás de las estructuras, más que su expresión formal estricta, aspecto en el cual se toma la libertad de introducir cambios allí donde lo considera necesario. En esa tesitura, aquí se opta por adaptar la semántica basada en teoría de filtros a otros fines, utilizando la estructura de ideal clásico cuya definición se presenta a continuación:

4.2.2 Un ideal sobre el conjunto de todas las sentencias S es un conjunto $\mathcal{I}$ tal que satisface: 1) $\emptyset \in \mathcal{I}(S)$; 2) si $A \in \mathcal{I}(S)$ y $B \subseteq A$ entonces $B \in \mathcal{I}(S)$; 3) Si $A, B \in \mathcal{I}(S)$ entonces $A \cup B \in \mathcal{I}(S)$.

Respecto a la cláusula 1), del conjunto vacío solo se concluyen tautologías, y éstas se preservan como consecuencias de cualquier superconjunto, por lo tanto, vale monotonía para $\emptyset$. La cláusula 2) establece que si para A se cumple monotonía también se cumple para cualquier subconjunto de A. Por último, la condición 3) indica que si para A y B se cumple monotonía entonces también se cumple para el supremo entre A y B. Intuitivamente, los ideales son subconjuntos grandes de la base de conocimiento de un razonador semimonótono, para los que se cumple monotonía ante cualquier ampliación. Si $\mathcal{I}(S) = \Delta$, todos los subconjuntos de Δ satisfacen monotonía ante la ampliación considerada, y también la cota superior entre dos subconjuntos cualquiera de ellos. Si para alguna ampliación ocurre $\mathcal{I}(S) \subset \Delta$, esto implica que algún subconjunto de Δ entra en conflicto con algún subconjunto de la ampliación, y se produce una falla de monotonía. En este caso, el ideal representa el máximo subconjunto propio de Δ para el cual se cumple monotonía. A su vez, se garantiza que todos los subconjuntos de este *subconjunto límite* cumplen monotonía, como así también el supremo entre dos subconjuntos cualquiera de ellos.

Ahora bien, resulta interesante analizar cómo se comportan las definiciones **4.2.1** y **4.2.2** ante la ampliación nula $\Gamma = \Delta$. Desde el punto de vista del proceso del razonamiento, corresponde al *momento* en el que el razonador es dotado con un conjunto no vacío de sentencias, previo a cualquier inferencia y a cualquier ampliación de dicho conjunto. En este caso inicial se cumple $C(\Delta) = C(\Delta)$ y se satisface trivialmente la condición análoga a monotonía. Aquí no hay repercusiones respecto a **4.2.1**, ya que por definición **4.2.2** se garantiza monotonía para un ideal sobre Δ, y puesto que en este caso la condición se cumple trivialmente para Δ, también se cumplirá para cualquier subconjunto de Δ. Asimismo, para este resultado es irrelevante que Δ sea consistente o inconsistente. En el segundo caso, aun cuando el sistema pueda hacer fallar la propiedad de inclusión, como se ejemplificó en **2.2.12**, no debería fallar forzosamente la condición de monotonía ya que independientemente de que $\Delta \supseteq \Gamma \supset C(\Delta)$, se verifica $C(\Delta) \subseteq C(\Delta)$ (nuevamente, de manera trivial, pues para cualquier conjunto de sentencias X, $C(X) = C(\mathrm{X})$).

BREVES CONCLUSIONES

Con **2.3.1-2** y **4.2.1-2** el autor espera haber cumplido el objetivo planteado en **INTRODUCCIÓN Y OBJETIVOS**, de *ofrecer una alternativa clara para contribuir a la resolución del problema de la ausencia de una expresión estrictamente formal de la propiedad de no monotonía*, como así también haber contribuido a fortalecer las bases filosóficas y metalógicas de las LNM's, introduciendo conjeturas y propiedades novedosas vinculadas a la noción de *semimonotonía*. La propiedad de no monotonía **2.3.1-2** puede ser interpretada semánticamente en dos registros diferentes: según una construcción modelo-teorética que extrapola a la metalógica estructuras similares a las del sistema T de la lógica modal tradicional; o bien según construcciones que extrapolan estructuras análogas a las utilizadas en sistemas preferenciales en la tradición de KLM y Schlechta. Como se mencionó previamente, esta propiedad tiene un carácter universal, y funciona como criterio de clasificación estricto entre sistemas lógicos, que permite distinguir entre sistemas monótonos (clásicos o no clásicos) y sistemas no monótonos.

Por su parte, la propiedad de semimonotonía **4.2.1** se ancla en **2.3.1-2** y en la conjetura del razonador semimonótono, y puede ser interpretada semánticamente con recurso a la teoría de modelos basada en ideales clásicos. Esta última propiedad pretende describir las condiciones que debería cumplir un sistema que se proponga modelar el comportamiento de un tipo bastante común de razonadores, que normalmente construyen y conservan conocimiento, aun tolerando inconsistencias u otro tipo de dificultades, y en menor medida eliminando conclusiones previas: *los razonadores semimonótonos*. Se trata de una propiedad novedosa que parte de una conjetura que intenta rescatar algo del pensamiento de M. Minsky en su (1974), cuando proponía como uno de los objetivos fundacionales de la IA y de la LNM la construcción de un *agente racional artificial* que emule el razonamiento de los agentes comunes. De modo que, en gran parte, tal labor pasa por elucidar el comportamiento racional de los agentes, lo cual es materia de debate científico y filosófico. Desde la perspectiva de la lógica, el autor propone la propiedad de semimonotonía como un aporte a la mencionada elucidación, como una descripción que pretende emular a un tipo particular de *razonadores estables*.

Tanto las definiciones metamatemáticas de no monotonía como las de semimonotonía constituyen los resultados fundamentales de esta tesis, que responden a problemas de fundamento de la LNM. Asimismo, algunas otras certezas importantes han sido sacadas a superficie y otros tantos problemas han quedado abiertos (los cuales han sido mencionando oportunamente). Entre las primeras, cabe destacar la especificación de una *restauración parcial* de las propiedades centrales de la LNM, fuertemente objetadas, para una familia de teorías basadas en el enfoque de LMG (Schlechta); y la certeza de que la LNM no puede ser reducida a lógica paraconsistente.

APÉNDICE 1 – CONTRAEJEMPLOS AL POSTULADO 1.4.4

Teniendo en cuenta que tanto el metateorema de deducción, como el teorema **1.4.8** de la TIS no han sido refutados, en este apéndice se abordará el problema relativo a si existen efectivamente, en el marco de lógica estándar, ejemplos de razonamientos deductivos que sean informativos. En lo que sigue, se da por supuesto que cuando se habla de la no informatividad de razonamientos deductivos, se trata de razonamientos cuyo conjunto de premisas es consistente. En primer lugar, en el apartado **A)** se tomará en cuenta la propuesta popperiana al respecto, mientras que en **B)** se abordarán brevemente algunos desarrollos de M. D'agostino, en línea con el pensamiento de J. Hintikka.

A) Entre las críticas actuales a la TCI se encuentran los cuestionamientos de matriz popperiana, resumidos en un par de comunicaciones breves firmadas por D. Frederick[98] donde se recuperan algunos de los desarrollos teóricos realizados por K. Popper en su libro "Unended Quest"[99]. Partiendo de una versión no cuantitativa del PRI, Popper distingue entre el "contenido informativo" (CI) y el "contenido lógico" (CL) de una teoría T. El CL coincide con lo que Tarski llama "clase de las consecuencias" de T, mientras

[98] Cfr. (Frederick, 2011) y (Frederick, 2014).
[99] Cfr. (Popper, 1976).

que el CI es la información semántica acarreada por T y se define como "el conjunto de sentencias que son incompatibles con la teoría" (Popper, 1976:24). En términos generales, dada una teoría cualquiera T, si CLT consiste en $\{a, b, c\}$, entonces CIT será $\{\neg a, \neg b, \neg c\}$, y a la inversa. Del mismo modo, si CIT es $\{a, b, c\}$, CLT será $\{\neg a, \neg b, \neg c\}$ y a la inversa[100]. Ahora bien, dado que el CL de cualquier teoría es, por definición, infinito[101]; también lo será su CI. Esta última afirmación provee la clave para la presentación de un contraejemplo a la TCI (postulado **1.4.4**). El contraejemplo parte de la afirmación de que las teorías de la gravitación de Newton (N) y Einstein (E) son incompatibles entre sí, por lo que cada una pertenece al CI de la otra. Particularmente, dado que E pertenece a CIN, $\neg E$ pertenece a CLN y por lo tanto $\neg E$ es lógicamente implicado por N. Como $\neg E$ dice "algo nuevo" desde el punto de vista de los contemporáneos de Newton[102], quedaría derrotado el postulado **1.4.4**.

Resulta notable que este contraejemplo tiene un problema bastante serio, el cual, si bien es parcialmente reconocido por los autores, no se le da demasiada importancia en los textos de referencia. El problema puede dividirse en dos partes interrelacionadas. La primera parte es relativa a la *infinitud* del CL de una teoría, mientras que la segunda tiene que ver con la *asimetría temporal* entre teorías rivales. La primera parte del asunto pasa por la ausencia de una especificación adecuada de lo que se entiende por "infinito", pues, a fin de cuentas, ¿se trata acaso de un *infinito potencial* o, por el contrario, de un *infinito actual*? Esta pregunta remite a un viejo problema filosófico y matemático. En pocas palabras, el infinito potencial es el infinito como *posibilidad* mientras que el infinito

[100] Refiriéndose al contenido informativo, Popper afirma: "…the elements of this set and the elements of the logical content stand in a one-one correspondence: to every element which is in one of the sets, there is in the other set a corresponding element, namely its negation" (Popper, 1976:24).
[101] Por ejemplo, tomando una sentencia cualquiera A, se puede construir una secuencia infinita de sentencias de la forma $A \vee B$, $A \vee C$, $A \vee D$…, etc.
[102] Cfr. (Frederick, 2011).

actual es el infinito como *realidad*[103]. La crítica popperiana deja este problema abierto, aunque parece inclinarse hacia una concepción del CL como *actualmente infinito*. De hecho, Popper afirma que dada una teoría T y dada la infinitud de CLT, puede afirmarse que CIT es infinito de un "modo no trivial", y por lo tanto cualquier teoría T' incompatible con T pertenece a CIT, incluyendo a cualquier teoría futura que pueda reemplazar a T algún día[104]. Esta última afirmación resulta, al menos, polémica, ya que una "teoría futura" es un conjunto de sentencias no existentes, una especie de *teoría fantasma* desde el punto de vista de la teoría actual considerada.

La objeción más obvia a una concepción "actualista" (por decirlo de algún modo) de la infinitud del CL y del CI es que, dado un razonador deductivo concreto (humano o máquina) que razona a partir de una teoría T resulta claro que, en su práctica inferencial, en cada momento del proceso el razonador dispondrá de un conjunto finito de consecuencias efectivamente inferidas a partir de T. Esto excluye la posibilidad de atribuir infinitud actual a CLT, ya que si tuviera ese atributo el contenido lógico estaría disponible *en su totalidad*, y en todo momento. Pero esto simplemente no ocurre. Por otra parte, el *factum* de la finitud del contenido lógico en cada momento del proceso inferencial no es incompatible con la atribución de infinitud potencial a CLT, ya que en la práctica el razonador estaría actualizando sucesivamente (quizá indefinidamente) un subconjunto finito de la clase potencialmente infinita de consecuencias, a saber, el conjunto de sentencias efectivamente inferidas.

La circunscripción del CLT a un subconjunto finito de la clase potencialmente infinita de consecuencias de la teoría considerada, también limita la extensión del CIT, entendiendo este conjunto según la definición de Popper. En cada momento del proceso inferencial, el

---

[103] Para un breve recuento teórico sobre este tópico, (Zhu et al., 2008).
[104] Cfr. (Popper, 1976: 26).

razonador puede acceder sólo a un subconjunto finito del conjunto potencialmente infinito de las sentencias que excluye el CLT. Por lo tanto, el CIT disponible efectivamente también será un subconjunto finito del conjunto potencialmente infinito CIT. De este modo, en tanto la *existencia* (a falta de un término mejor) de teorías fantasma dentro de CIT se fundamenta en última instancia en la supuesta infinitud actual de este conjunto, la descripción de CIT como actualmente finito – potencialmente infinito pone en cuestión la existencia de tales teorías, y con ello, la posibilidad de inferir deductivamente conclusiones novedosas a partir de T[105].

En el caso del ejemplo de las teorías N y E, se excluye la posibilidad de que E pertenezca a CIN antes de la creación de E, y por lo tanto se excluye cualquier inferencia que vaya de N a $\neg E$ en la época previa a la creación de E, aportando (supuesta) información novedosa. Esta situación constituye la segunda parte del problema: en breves palabras la concepción popperiana no lleva hasta las últimas consecuencias el elemental hecho de que existe una asimetría temporal entre las teorías N (publicada en 1687) y E (1915). Si se toma en cuenta esa asimetría, es claro que en un momento histórico ulterior a 1915 un razonamiento como $N, N \rightarrow \neg E \vdash \neg E$ podría ser efectivamente formulado, dentro de los marcos de la LE, pero por supuesto no resultaría adecuado como contraejemplo al postulado **1.4.4**[106], debido a que la conclusión se encuentra ya indicada entre las premisas. Por otra parte, si fuera posible viajar en el tiempo al año 1800, resultaría impracticable una deducción de $\neg E$ a partir de N. De modo que el supuesto contraejemplo al postulado

[105] Esto es reconocido, con otras palabras, por el propio Frederick cuando afirma que las conclusiones deductivamente válidas efectivamente obtenidas de un conjunto de premisas deben estar contenidas en las premisas. Sin embargo, a continuación, insiste en que “existen” otras conclusiones que actualmente son imposibles de inferir a partir de las premisas disponibles (y que, en algunos casos, permanecen en esa situación indefinidamente), y son estas las que aportan información novedosa. Cfr. (Frederick, 2011).
[106] Esta observación coincide con lo planteado en (McBride, 2014).

1.4.4 basado en las teorías N y E es un *oxímoron*: cuando era relevante como contraejemplo no existía, y cuando existía, no era relevante como contraejemplo[107].

B) Una segunda crítica actual al postulado **1.4.4** es realizado por M. D'agostino en base al pensamiento de J. Hintikka, y a herramientas de la teoría de la computabilidad. En los escritos de éste último de la década de 1970's, (Hintikka, 1970a), (Hintikka, 1970b) y (Hintikka, 1973), el autor formula su versión del PRI en términos cuantitativos (que no difiere demasiado de los postulados de (Carnap & Bar-Hillel, 1952)), junto a una teoría de la información que distingue entre "información de profundidad" e "información de superficie". Sintéticamente, información de profundidad es la totalidad de la información que puede extraerse de un conjunto de sentencias aplicando todas las herramientas de la LE, y coincide con el concepto de información semántica de la propuesta de Carnap – Bar – Hillel. Por definición, la información de profundidad de una teoría T no puede incrementarse mediante operaciones deductivas. Por otra parte, la información de superficie es sólo la parte de aquella totalidad que está disponible inmediata y explícitamente en las sentencias consideradas. Así, la información de superficie puede incrementarse mediante operaciones deductivas[108]. Para Hintikka, las inferencias que incrementan la información de superficie son sólo aquellas que involucran pasos de deducción suposicionales donde se introducen nuevas variables de individuos[109]. A simple vista, estas consideraciones conllevan una crítica al postulado **1.4.4**, pues en tanto algunas inferencias deductivas incrementan la información de superficie, aportan algún grado de novedad respecto a la información disponible inmediatamente[110]. Las ideas de Hintikka son recuperadas, aún con reservas, por D'agostino en sus artículos (D'agostino

---

[107] Por lo tanto, puede afirmarse que el postulado **1.4.4** ha resistido con éxito un intento de falsación por parte del mismísimo creador del criterio de falsación.
[108] Cfr. (Hintikka, 1973: 22-23).
[109] Cfr. (D'Agostino, 2014: 278).
[110] Posteriormente se volverá sobre esta distinción entre información de superficie/profundidad.

& Floridi, 2009) y (D'Agostino, 2014), a fin de criticar la TCI. La crítica de esas ideas comienza por la explicitación del problema conocido como "paradoja Cohen – Nagel" o "paradoja de la inferencia"

Paradoja de Cohen - Nagel: *Si en una inferencia la conclusión no está contenida en las premisas, la inferencia no puede ser válida, y si la conclusión no es diferente de las premisas, es inútil; pero la conclusión no puede estar contenida en las premisas y a su vez ser novedosa, en consecuencia, las inferencias no pueden ser válidas y útiles al mismo tiempo*[111].

El camino que traza D'agostino para superar esta paradoja comienza por rechazar las ideas del postulado **1.4.4** de un modo similar a la propuesta de Hintikka, afirmando que existen inferencias deductivas que aportan algún grado de novedad. Sin embargo, el autor va un paso más allá, mientras que Hintikka defendía la no informatividad de la lógica proposicional (pero señalaba que el razonamiento suposicional que involucra reglas de cuantificación ofrece información novedosa), D'agostino asevera que todo razonamiento deductivo suposicional, ya sea en el marco de la lógica proposicional como en lógica de predicados, involucra información (en algún grado) novedosa. Para fundamentar su postura, afirma que se debe distinguir entre información *efectiva* y *potencial.* En sentido lato, la información efectiva es información que real y actualmente está disponible al razonador, mientras que la potencial es información que el razonador virtualmente podría poseer. Así, un razonamiento deductivo no suposicional involucraría sólo información efectiva, mientras que un razonamiento deductivo suposicional involucraría información efectiva y también información potencial. Esta distinción lleva a su vez a la distinción entre inferencias deductivas puramente analíticas (no novedosas), aquellas donde no

[111] Cfr. (D'Agostino, 2014: 408).

interviene la información potencial, e inferencias deductivas con diferentes grados de sinteticidad, las cuales involucran información potencial y aportan diferentes grados de novedad. Así, en tanto una inferencia deductiva suposicional aportaría información novedosa, resultaría "útil" (a diferencia de las puramente analíticas).

Si bien no se podrá profundizar aquí en la teoría de D'agostino, es posible aislar inmediatamente el supuesto clave de su propuesta: la idea de que el razonamiento deductivo suposicional (en general, no sólo en lógica de predicados, como en el caso de Hintikka) aportaría información novedosa. De modo que cualquier razonamiento en el marco de LE, con al menos un paso de derivación suposicional, sería un ejemplo efectivo de un razonamiento deductivo informativo. La primera objeción a esta idea se dirige a la ausencia de una definición de "razonamiento suposicional" por parte de D'agostino. Esta omisión[112] resulta problemática, ya que la propuesta global del autor depende en gran parte de un concepto cuyo significado se da por supuesto. Una definición simple de "razonamiento suposicional" o "hipotético" puede encontrarse en (Marraud, 2007: 45), la cual parte de distinguir entre *inferencias*, entendidas como argumentos simples compuestos de premisas y conclusión; y *razonamientos suposicionales*, entendidos como argumentos complejos compuestos por subargumentos que funcionan como premisas y ofrecen fundamento a la conclusión. Ahora bien, la información contenida en la conclusión de un razonamiento suposicional no emerge "de la nada", sino que invariablemente puede rastreársela en el subargumento que funciona como razón para tal conclusión. Esta condición de la lógica clásica se hace patente mediante el Hauptsatz gentzeniano y la propiedad de subfórmula. De modo que, sin necesidad de problematizar respecto a la *potencialidad* o *efectividad* de la información[113], puede afirmarse que no

[112] En (D'agostino & Floridi, 2009) y (D'agostino, 2014).
[113] Claramente, la justificación que ofrece D'agostino para esa distinción en los textos de referencia resulta insuficiente.

resulta claro que la información disponible en la conclusión de un razonamiento suposicional sea novedosa en relación con su premisa (subargumento).

No obstante, aún si se concede -provisoriamente- certeza al supuesto que atribuye novedad al razonamiento suposicional, hay que mencionar que el tratamiento que realiza el autor sobre la paradoja de la inferencia presenta otro problema. En (D'Agostino, 2014) se cita a Cohen y Nagel como creadores de tal paradoja (aunque estos últimos no reivindican esa creación, sino que mencionan la paradoja como una crítica habitual a la lógica deductiva), pero se omite mencionar que inmediatamente después de la presentación de la misma en el capítulo IX de (Cohen & Nagel, 1934), los autores proceden a denunciar su carácter de *pseudoparadoja* o paradoja aparente, atacando así las críticas basadas en ella. Entre otras razones, Cohen y Nagel critican la (pseudo)paradoja de la inferencia por confundir la *novedad psicológica* y la *novedad lógica* de una conclusión relativamente a las premisas que la fundamentan[114]. Una conclusión de un argumento deductivo puede poseer novedad psicológica pero no novedad lógica, por ejemplo, las consecuencias de las demostraciones euclideanas pueden resultar sorprendentes aun teniendo en vista los supuestos de los que parten[115]. Lo mismo ocurre para demostraciones que involucran largas cadenas de razonamiento deductivo. Para Cohen y Nagel resulta claro que, si un razonamiento es deductivo y por lo tanto válido, la conclusión no puede (en tanto carece de independencia respecto a las premisas) poseer novedad lógica relativamente a estas[116]. Pero sí puede ocurrir que un argumento deductivo posea novedad psicológica, por ejemplo, una inferencia que concluya que "todos los

[114] Esta distinción fue presentada en la sección **1.4**.
[115] Cfr. (Cohen & Nagel, 1934: 174).
[116] Cfr. (Cohen & Nagel, 1934: 175).

inmortales son no humanos” a partir de la premisa de que “todos los humanos son mortales”.

A lo afirmado por Cohen – Nagel puede agregarse lo siguiente: que la paradoja de la inferencia supone un compromiso previo no explicitado respecto de lo que se considera *útil*. En primer lugar, habría que justificar que sólo lo *novedoso* resulta útil. Aun así, si se lograra justificar ese supuesto, todavía podría contraargumentarse que el ejemplo del párrafo anterior, que concluye que los inmortales son no humanos, constituye un caso donde la conclusión es una consecuencia lógica de su premisa y es novedosa (psicológicamente) respecto a esta última, por lo que es una inferencia válida y útil (según aquel criterio no justificado) al mismo tiempo. Desde el punto de vista de esta distinción entre novedad lógica y novedad psicológica, volviendo al caso de los razonamientos suposicionales resulta claro que la conclusión de cualquier argumento de este tipo carece de independencia lógica respecto del subargumento que funciona como premisa, por lo que carecen también de novedad lógica. Por último, en tanto el metateorema de la deducción vale para LE[117], siempre podrá aplicarse el mismo a cualquier razonamiento suposicional obteniéndose una tautología, a la cual a su vez puede aplicarse el teorema **1.4.8** para corroborar que la medida de su informatividad es igual a 0 bits[118]. Desde este punto de vista, puede afirmarse sin lugar a duda que los razonamientos suposicionales no constituyen contraejemplos al postulado **1.4.4**.

C) A modo de conclusión, interesa aquí mencionar que si bien las críticas al postulado **1.4.4** citados arriba no parecen haber llegado a buen puerto en su afán de presentar ejemplos concretos de razonamientos deductivos informativos (en el marco de LE), no

---

[117] Cfr. Teorema 43.1 en (Hunter, 1973: 170).

[118] Esto ocurre independientemente de cualquier consideración filosófica respecto a la efectividad o potencialidad de la información contenida en el subargumento del razonamiento suposicional considerado.

pueden pasarse por alto las críticas dirigidas contra la TCI. Pero aun reconociendo que esta última teoría se encuentra bajo presión crítica constante y rechazo generalizado, esta situación por sí sola no constituye una buena razón para rechazarla. Actualmente, se trata de un problema abierto. Al respecto, es interesante mencionar el caso de Hintikka, particularmente las reflexiones que realiza en uno de sus últimos artículos publicados en vida. En (Hintikka & Sandu, 2007) la distinción entre información de superficie e información de profundidad aparece considerada desde una perspectiva diferente. En primer lugar, los autores afirman que no existe una *noción viable* de información, es decir, una *teoría de la información* que permita dar cuenta adecuadamente de ese fenómeno. Esta afirmación, realizada casi *al pasar*, resulta de gran importancia ya que conlleva un cuestionamiento de la TIS y el reconocimiento de que la relación entre inferencias e información constituye un problema abierto. Posteriormente, en lo que parece ser una revisión de sus posturas previas, Hintikka reconoce que de existir una teoría de ese tipo podrían identificarse las inferencias deductivas con aquellas que no aportan información novedosa, y las inferencias ampliativas como aquellas que sí aportan información novedosa[119]. Sin embargo, como no existe una teoría de la información bien establecida, y dado que la idea de que las inferencias deductivas son no informativas ha sido puesta criticada, el autor recupera su distinción entre información de profundidad e información de superficie, pero argumentando que las inferencias deductivas son no informativas en cierto sentido, a saber, en tanto no aportan información de profundidad (aunque sí pueden aportar información de superficie).

Tomando en cuenta tanto los escritos de la década de 1970 como su (Hintikka & Sandu, 2007), es evidente que la propuesta de este autor permanece, a lo largo de su trayectoria intelectual, en cierta tensión y diálogo constante con la TCI y con la TIS. Particularmente,

[119] Cfr. (Hintikka & Sandu, 2007: 26).

en cuanto a la idea de que las inferencias deductivas son no informativas, finalmente no parece rechazarla por completo. Es que, a fin de cuentas, los conceptos de *información de profundidad* e *información de superficie* remiten, en el plano metafórico, a un caudal informativo que de algún modo se encuentra *ya presente* en un conjunto de enunciados, pero el cual se compone de una parte disponible inmediatamente, y otra parte "sumergida", que debe ser sacada a luz mediante operaciones lógicas (deductivas). Según esa metáfora, la parte sumergida si bien no sería accesible inmediatamente, sí estaría disponible en el caudal informativo inicial (en la "profundidad" del mismo). Por otra parte, el autor afirma en el mismo texto que si una información es totalmente novedosa, su verdad no puede quedar garantizada por la verdad precedente[120]. Así, deja claro que la noción misma de validez no puede separarse de la idea de que las inferencias deductivas no aportan información novedosa. Si una inferencia aporta información novedosa, no puede ser válida, y por lo tanto no es deductiva. De este modo, lo que resulta claro es que para Hintikka la crítica de la TCI nunca fue un tema clausurado (del mismo modo que no lo era para Popper, cuyas consideraciones al respecto fueron publicadas en un texto titulado "Unended Quest").

[120] Cfr. (Hintikka & Sandu, 2007: 16).

APÉNDICE 2 – PROPIEDADES, DEFINICIONES Y DIGRESIONES VARIAS

En este apéndice se abordarán algunas de las propiedades del sistema C de KLM, como así también definiciones y estructuras de LMG, junto una digresión filosófica relativa al uso del TNM en (Schlechta, 2004). En primer lugar, se presentarán las propiedades derivadas del sistema C de equivalencia (A2.1), conjunción derecha (**A2.2**), modus ponens derecho (A2.3), disyunción (**A2.4**), monotonicidad (**A2.5**), teorema de deducción débil (**A2.5**), transitividad (**A2.6**) y contraposición (**A2.6**). Tales propiedades son demostrables dentro del mencionado sistema de KLM, y se derivan directamente de las propiedades básicas consignadas en la sección **3.1**.

A2.1 $\dfrac{\alpha|\!\sim\beta,\ \ \beta|\!\sim\alpha,\ \ \alpha|\!\sim\gamma}{\beta|\!\sim\gamma}$

A2.2 $\dfrac{\alpha|\!\sim\beta,\ \ \alpha|\!\sim\gamma}{\alpha|\!\sim\beta\wedge\gamma}$

A2.3 $\dfrac{\alpha|\!\sim\beta\rightarrow\gamma,\ \ \alpha|\!\sim\beta}{\alpha|\!\sim\gamma}$

A2.4 $\dfrac{\alpha\vee\beta|\!\sim\alpha,\ \ \alpha|\!\sim\gamma}{\alpha\vee\beta|\!\sim\gamma}$

A2.5 $\dfrac{\vDash\alpha\to\beta,\ \ \beta|\!\sim\gamma}{\alpha|\!\sim\gamma}$

A2.6 $\dfrac{\alpha|\!\sim\beta\to\gamma}{\alpha\wedge\beta|\!\sim\gamma}$

A2.7 $\dfrac{\alpha|\!\sim\beta,\ \ \beta|\!\sim\gamma}{\alpha|\!\sim\gamma}$

A2.8 $\dfrac{\alpha|\!\sim\beta}{\neg\beta|\!\sim\neg\alpha}$

Las siguientes definiciones abordan el concepto de "preservación de definibilidad" en el enfoque de LMG de Schlechta, que resultan de gran importancia para la solución del problema de generación de inconsistencias a partir de teorías consistentes en sistemas preferenciales tradicionales. Las definiciones de "subconjunto definible" (**A2.9**) y "función preservadora de definibilidad" (**A2.10**) se presentan a continuación:

A2.9 $\mathcal{D}_{\mathcal{L}} \subseteq \mathcal{P}(M_{\mathcal{L}})$ es el conjunto de subconjuntos definibles de $M_{\mathcal{L}}$, es decir $A \in \mathcal{D}_{\mathcal{L}}$ syss existe algún $T \subseteq \mathcal{L}$ tal que $A = M_T$.

A2.10 Para un $X \subseteq \mathcal{P}(M_{\mathcal{L}})$, una función $\mu: X \to \mathcal{P}(M_{\mathcal{L}})$ se denomina preservadora de definibilidad syss $\mu(Y) \in \mathcal{D}_{\mathcal{L}}$ para todo $Y \in \mathcal{D}_{\mathcal{L}} \cap X$.

A2.9 expresa que un conjunto de modelos es definible syss satisface una teoría, y **A2.10** establece que una función de elección sobre conjuntos de modelos preserva definibilidad syss asigna un conjunto de modelos definible a todo conjunto de modelos definible.

A continuación, se presentará un ejemplo de minimización con copias, que permite ilustrar el funcionamiento de las estructuras modelo-teoréticas correspondientes a la versión con copias de la teoría de modelos preferenciales de Schlechta:

A2.11 Sea $x \in X - \mu(X)$. Si μ se representa mediante una estructura preferencial z, x (más precisamente, todas las copias de x) no puede ser minimal en la estructura. Así, para cada copia $< x, i >$ de x, debe existir algún $< x', i' >$ tal que $x' \in X$ y $< x', i' >< < x, i >$, es decir que minimiza $< x, i >$. Un elemento $x \in X$ puede ser no minimal en X y en X'. Al mismo tiempo, X podría necesitar dos elementos, y e y', para minimizar (todas las copias de) x, y X'podría necesitar dos elementos, z y z', para minimizar x. Pero no se sabe qué copia es minimizada por qué elemento. Si estas son todas las posibilidades de minimizar las copias de x, cualquier Y que minimiza x debe contener $\{y, y'\}$ o $\{z, z'\}$. Pero esto es equivalente al hecho de que el rango de alguna función de elección en el producto cartesiano $\{y, y'\}$ x $\{z, z'\}$ tiene una intersección no vacía con Y. [Fin de ejemplo]

Por último, se consignará una breve digresión filosófica respecto al uso del TNM en (Schlechta, 2004). En primer lugar, es destacable que con la introducción en el lenguaje objeto del operador modal N y del semigeneralizador $\triangledown$ como simbolismos para representación de la noción de "normalidad", podría pensarse que la dualidad en el uso del TNM ha sido resuelta de una vez por todas. Pero (lamentablemente) el autor deja una puerta abierta a la ambigüedad en tanto no reemplaza (en el texto citado) todas las apariciones de expresiones de la forma $\alpha|\!\sim\beta$, cuando estas parecen utilizarse para simbolizar "normalidad", por expresiones como $\vdash N(\alpha) \rightarrow \beta$ o $\triangledown x(Ax \rightarrow Vx)$, por poner dos ejemplos meramente esquemáticos. Schlechta considera que esta transformación es "demasiado obvia, por lo que sólo se la realizará implícitamente en la mayoría de los casos" (2004:4). Esta última aseveración puede resultar confusa y complicar la lectura de las condiciones de coherencia. Por otra parte, en el texto citado existen numerosos fragmentos donde el autor utiliza el TNM de un modo más cercano a la tradición (Gabbay), por ejemplo en (Schlechta, 2004:11) se refiere a este símbolo como una "relación de inferencia", lo mismo ocurre cuando define clausura no monótona (Schlechta, 2004:29) y en la tabla de propiedades o condiciones de coherencia (Schlechta, 2004:32). Incluso, en (Schlechta, 2004:40) el autor describe directamente al TNM como una "relación de consecuencia no monótona", y en los resultados de completitud y corrección para sistemas preferenciales (Schlechta, 2004:187-188) esta última parece ser la única interpretación posible.

Este (tedioso) asunto de la dualidad de significado del TNM, heredado de la propuesta de KLM, involucra más que una mera discusión terminológica: se trata una verdadera aporía filosófica (de filosofía de la lógica) y sólo por eso justifica su abordaje en esta tesis, que tal como lo indica su subtítulo, intenta ofrecer un abordaje filosófico de algunos de los tópicos siempre que sea necesario para los fines más generales de la misma (fundamentación y formalización de la *propiedad de las propiedades* de los sistemas de LNM, no monotonía, y de las propiedades centrales). En general, la estrategia de "poner tantas cosas como sea posible en el lenguaje objeto" (Schlechta, 2004:4) puede entenderse de (al menos) dos modos: a) el autor intenta borrar los límites entre metalenguaje y lenguaje objeto, más precisamente, intenta disolver el primero en el segundo, y con ello, la metalógica en la lógica. En este sentido, se trataría de una radicalización de la postura de KLM; b) el autor intenta enriquecer el lenguaje objeto con conceptos y nociones que, según su interpretación, en la tradición de KLM se encontraban en el metalenguaje (primordialmente, la noción de "normalidad" simbolizada mediante el TNM), y reservar al metalenguaje la investigación de las propiedades de la relación de consecuencia y los problemas de representación. A lo largo del texto citado, parece haber razones para fundamentar ambas interpretaciones, si bien, aún con reservas, se considera que Schlechta no parece estar tan interesado en introducir el concepto mismo de *relación de consecuencia* en el lenguaje objeto, sino más bien la noción de *normalidad*, y en ese sentido su crítica parece estar dirigida primordialmente a la propuesta de KLM de *utilizar el TNM como representación de la noción de normalidad*. Para los fines de esta tesis, esta última consideración resulta suficiente y justifica dejar abierto el problema más general de la relación entre lenguaje objeto y metalenguaje en el pensamiento de K. Schlechta. A fin de cuentas, en tanto su enfoque es prelingüístico (según Makinson), ni siquiera parecen haber razones lo suficientemente fuertes como para inducir a pensar que *este problema debe ser resuelto*, al menos en el contexto de la obra citada.

La siguiente cita refuerza esta última idea, pero a la vez introduce nuevas complejidades y abre nuevos interrogantes: "si ponemos no sólo nociones como "normalidad" en el lenguaje objeto, sino también los conceptos semánticos generados, tales como preferencia, distancia, tamaño, etc., podemos unir los dos conceptos en un lenguaje objeto de manera análoga a un teorema de corrección y completitud" (Schlechta, 2004:5). Aparentemente, el autor se aproxima la idea de fusionar metalenguaje y lenguaje objeto, pero aun así no parece dar ese paso hasta el final pues no afirma que el lenguaje objeto ampliado contendrá efectivamente los resultados de representación, sino sólo que será posible unir conceptos axiomáticos y semánticos y generar resultados *análogos* a los de representación. Ahora bien, en virtud de la anterior cita, resulta interesante mencionar que, a lo largo del texto de referencia, los resultados de representación son expuestos *en cuanto tales*, y no parece haber indicios de aquellas "expresiones análogas" a los mismos. A esta altura, queda claro que la obra de Schlechta no sólo es de una enorme complejidad matemática, lógica y metalógica. Además de ofrecer una plétora de herramientas modelo-teoréticas, fragmentos filosóficos esclarecedores y análisis plenos de matices e intuiciones geniales, resulta completamente abstrusa en algunos pasajes, y contiene fragmentos de difícil comprensión e incluso bastante oscuros, por ejemplo el siguiente donde el autor se refiere a su sistema como "un modo de razonamiento y descripción multinivel, donde las complejidades innecesarias pueden ser escondidas desde los niveles superiores (aquí, desde normalidad), y son visibles sólo en los niveles inferiores (por ej., preferencia de una situación sobre otra) siempre que sea necesario"(Schlechta, 2004:5-6).

Pero, a fin de cuentas, el asunto se reduce simplemente de poner las cosas en su lugar, como indica la metáfora, si se permite una licencia lingüística, de poner *el inodoro en el baño y la heladera en la cocina*. Esto tiene aún mayores implicancias que las de simplemente ordenar y delimitar las esferas lingüísticas en la exposición de un sistema. Lo que Schlechta parece pasar por alto es que existe un precio a pagar por cualquier intento de disolución del metalenguaje o fusión de este en el lenguaje objeto, y ese precio es demasiado oneroso: lisa y llanamente esto equivale a aniquilar la metalógica, es decir, renunciar a cualesquier reflexión y problematización en términos metadisciplinarios. A fin de cuentas, la mera posibilidad de una *metalógica* supone un enfoque metalingüístico. Bloquear la metalógica aniquilando el metalenguaje no parece tener ninguna ventaja, por el contrario, entraña una serie de dificultades de difícil o imposible solución, cuyo mero planteo (en términos generales) excede los límites de esta tesis, más allá de consideraciones obvias relativas a la viabilidad de investigaciones sobre propiedades de sistemas (¡por ejemplo, las propiedades centrales de la LNM!). En el caso puntual de la propuesta de Schlechta, cabe preguntarse, por ejemplo: si *todo es lenguaje objeto*, ¿cuál sería el status de las propiedades o condiciones de coherencia de un sistema? ¿tiene sentido molestarse en construir demostraciones o justificaciones para las mismas? Pues,

¿acaso se requiere justificar las reglas de un sistema dentro del propio sistema? Preguntas similares pueden plantearse para los resultados de representación. En suma, el problema del uso del TNM puede ser resuelto sin recurrir a soluciones extremas, de manera simple y efectiva, tal como lo hace el autor en algunos pasajes del texto: separando la noción de normalidad del símbolo TNM (y enviándola al lenguaje objeto, con otras simbolizaciones tales como N o $\triangledown$), e interprendo el TNM según la tradición (en este caso, del creador del símbolo, D. Gabbay).

Pero en otros pasajes, por ejemplo en (Schlechta, 2004:40), cuando se retoma el asunto de las interpretaciones del TNM y el asunto de la simbolización de la noción de normalidad, vuelven a plantearse dificultades filosóficas. En el fragmento mencionado, el autor recapitula algunas interpretaciones "razonables" que ha recibido el TNM dentro de diferentes enfoques de LNM. Sintéticamente, se detectan dos interpretaciones principales para una expresión de la forma $\alpha|\sim\beta$: a) cuando se la interpreta como una aserción condicional de la forma "α normalmente β" (por ej. en KLM), específicamente se la está interpretando como un predicado sobre elementos individuales o pares de elementos; b) cuando se la ha entendido como significando "la mayoría de los α son β", realmente se está predicando sobre conjuntos de elementos. Tanto para a) como para b) el autor ofrece simbolizaciones alternativas (en el lenguaje objeto) que evitan la extrapolación del TNM. Para la opción a), Schlechta reconstruye a grandes rasgos la propuesta de KLM, pero con algunas diferencias: en primer lugar, la semántica es *modal generalizada*, es decir, puede aplicarse a mundos posibles entendidos como modelos clásicos, pero su anclaje modelo-teorético no restringe la noción de mundo posible y admite trabajar libremente de manera conjuntista sin pronunciarse sobre las características de los mundos posibles. Los sistemas que parten de la opción a) son denominados *sistemas basados en preferencia*, y a su vez se engloban dentro de la propuesta de una LMG de Schlechta. En la opción b), como se mencionó antes, se opta por mantener para el TNM el sentido clásico de relación de consecuencia no monótona, y se envía la noción de normalidad al lenguaje objeto simbolizándola mediante un operador modal unario N o bien mediante el semigeneralizador. El autor indica que la noción de "la mayoría de los casos" es reflejada de manera precisa por el semigeneralizador ∇, que permite construir predicaciones sobre conjuntos de individuos. Ahora bien, es notable que el autor se refiera a estas interpretaciones del TNM, incluyendo la de KLM (y otras que no han sido consignadas aquí), como "razonables". En esto, se observa por parte del autor un excesivo "respeto" por la tradición, un afán de no debatir frontalmente que dificulta la comprensión de su propia propuesta. Más aun teniendo en cuenta que posteriormente Schlechta en la práctica no toma en cuenta la interpretación a), y su propuesta se centra en ofrecer formalismos alternativos que evitan el uso del TNM para simbolizar predicados sobre una mayoría de los casos o que incluyen la noción de normalidad. [Fin de digresión]

BIBLIOGRAFÍA

Alchourrón, C. (1995). Concepciones de la lógica. In *Enciclopedia Iberoamericana de Filosofía* (pp. 11–48). Editorial Trotta.

Antonelli, A. (2005). *Grounded Consequence for Defeasible Logic*. Cambridge University Press.

Befenhart, S., Duvois, D., & Prade, H. (1997). Some Syntactic Approaches to the Handling of Inconsistent Knowledge Bases: A Comparative Study. *Studia Logica*, *58*, 17–45.

Bell, J., & Slomson, A. (1969). *Models and Ultraproducts: AN INTRODUCTION*. Elsevier.

Bimbó, K. (2015). *Proof theory, sequent calculi and related formalism*. CRC Press.

Brewka, G. (1991). *Nonmonotonic Reasoning: Logical Foundations of Commonsense*. Cambridge University Press.

Carnap, R., & Bar-Hillel, Y. (1952). An outline of a theory of semantic information. In *Technical Report No. 247*. Research Laboratory of Electronics - MIT.

Cohen, M., & Nagel, E. (1934). *An Introduction to Logic and Scientific Method*. Routledge.

Copi, I. (2014). *Introduction to logic* (Pearson).

D'Agostino, M. (2014). Analitical inference and the informational meaning of the logical operators. *Logique & Analyse*, *227*, 407–437.

D'agostino, M., & Floridi, L. (2009). The enduring scandal of deduction. *Synthese*, *167*, 271–315.

Estrada, L., & Garcia, C. (2009). Can Paraconsistency Replace Non-Monotonicity? *Proceedings of the Fifth Latin American Workshop on Non-Monotonic Reasoning*, 217–224.

Floridi, L. (2009). Philosophical Conceptions of Information. In *Formal Theories of Information* (pp. 13–53). Springer-Verlag.

Fréchet, M. (1924). Des familles et fonctions additives d'ensembles abstraits. *Fundamenta Mathematicae*, *5*, 206–251.

Frederick, D. (2011). Deduction and Novelty. *The Reasoner*, *5*(4), 56–57. https://www.kent.ac.uk/secl/researchcentres/reasoning/TheReasoner/vol5/TheReasoner-5(5).pdf

Frederick, D. (2014). Deduction and Novelty Again. *The Reasoner*, *8*(5), 51–52.

Frege, G. (1967). Begriffsschrift, a formula language, modeled upon that of arithmetic, for pure thought. In J. van Heijenoort (Ed.), *From Frege to Gödel - A source book in mathematical logic, 1879-1931* (pp. 1–82). Harvard University Press.

Gabbay, D., & Schlechta, K. (2016). *A New Perspective on Non Monotonic Logics*. Springer.

Gabbay, Dov. (1985). Theoretical foundations of non-monontonic reasoning in expert systems. In K. Apt (Ed.), *Logics and models of concurrent systems* (pp. 439–457). Springer.

Gabbay, Dov. (1994). Preface. In Dov Gabbay (Ed.), *Handbook of Logic in Artificial Intelligence and Logic Programming - Vol. 3, Nonmonotonic Reasoning and Uncertain Reasoning* (pp. v–xii). Oxford University Press.

GAMUT, L. T. F. (2009). *Lógica, lenguaje y significado (vol. 2)* (Eudeba (ed.)).

Gentzen, G. (1969). Investigations into logical deduction. In M. Szabo (Ed.), *The collected papers of Gerhard Gentzen* (pp. 68–131). North-Holland Publishing Company.

Ginsberg, M. (1994). AI and Nonmonotonic Reasoning. In Dov Gabbay (Ed.), *Handbook of Logic in Artificial Intelligence and Logic Programming - Vol. 3, Nonmonotonic Reasoning and Uncertain Reasoning* (pp. 1–33). Oxford University Press.

Hintikka, J. (1970a). On Semantic Information. In *Information and Inference* (pp. 3–27). Reidel Publishing Company.

Hintikka, J. (1970b). Surface Information and Depth Information. In *Information and Inference* (pp. 263–297). Reidel Publishing Company.

Hintikka, J. (1973). *Logic, Language-Games and Information.* Clarendon Press.

Hintikka, J., & Sandu, G. (2007). What is logic? In D. Jacquette (Ed.), *Philosophy of Logic* (pp. 13–39). Elsevier.

Hughes, G., & Cresswel, M. (1996). *A new introduction to modal logic*. Routledge.

Hunter, G. (1973). *METALOGIC - An Introduction to the Metatheory of Standard First Order Logic*. University of California Press.

Israel, D. (1980). What's wrong with non-monotonic logic? *AAAI-80 Proceedings*, 99–101.

K. Simmons. (2009). Tarski's Logic. In *Handbook of the history of logic - Vol. 5 Logic form Russell to Church* (pp. 511–616). North-Holland Publishing Company.

Kraus, S., Lehmann, D., & Magidor, M. (1990). Nonmonotonic Reasoning, Preferential Models and Cumulative Logics. *Artificial Intelligence*, *44*, 167–207.

Kuhn, T. (1962). *LA ESTRUCTURA DE LAS REVOLUCIONES CIENTÍFICAS*. Fondo de Cultura Económica.

Lakatos, I. (1978). *La metodología de los programas de investigación científica*. Alianza Editorial.

Lazzer, S. (2006). El signo de aserción fregeano. *Epistemología e Historia de La Ciencia*, 328–334.

Legris, J. (1999). Reglas estructurales y análisis de la consecuencia lógica. *Epistemología e Historia de La Ciencia*, *5*(5), 234–241.

Lehmann, D., & Magidor, M. (1992). What does a conditional knowledge base entail? *Artificial Intelligence*, *55*, 1–60.

Lukaszewicz, W. (1990). *Non-monotonic reasoning- formalization of commonsense reasoning*. Ellis Horwood Limited.

Makinson, D. (1994). General patterns in nonmonotonic reasoning. In D. Gabbay (Ed.), *Handbook of Logic in Artificial Intelligence and Logic Programming - Vol. 3, Nonmonotonic Reasoning and Uncertain Reasoning* (pp. 35–110). Oxford University Press.

Makinson, D. (1989). General theory of cumulative inference. In M. Reinfrank (Ed.), *General theory of cumulative inference* (pp. 1–18). Springer.

Manzano, M. (1999). *Model Theory*. Oxford University Press.

Marraud, H. (2007). *Methodus Argumentandi*. UAM Ediciones.

Martin-Löf, P. (1996). On the meanings of the logical constants and the justifications of the logical laws. *Nordic Journal of Philosophical Logic*, *1*(1), 11–60.

Mc Dermott, D., & Doyle, J. (1980). Non-monotonic logic I. *Artificial Intelligence*, *13*(1–2), 41–72.

McCarthy, J. (1959). Programs with commonsense. *Proceedings of the Tedding- Ton Conference on the Mechanization of Thought Processes*, 75–91.

Minsky, M. (1974). *Artifical Intelligence Memo No. 306 – A framework for representing knowledge*.

Peckhaus, V. (2004). Calculus ratiocinator versus characteristica universalis? The two traditions in logic, revisited. *History and Philosophy of Logic*, *25*(1), 3–14.

Peirce, C. (1994). Collected Papers of Charles Sanders Peirce. In C. Hartshorne, P.

Weiss, & W. Burks (Eds.), *Collected Papers of Charles Sanders Peirce . Vol 1 - 8*. Thoemmes Continuum.

Perlis, D. (1987). On the consistency of commonsense reasoning. In M. Ginsberg (Ed.), *Readings in Nonmonotonic Reasoning* (pp. 56–66). Morgan Kaufmann.

Pollock, J. (1987). Defeasible Reasoning. *Cognitive Science*, *11*(4), 481–518.

Poole, D. (1998). *Computational intelligence - A logical approach*. Oxford University Press.

Popper, K. (1976). *Unended Quest*. Fontana.

Porte, J. (1981). Fifty years of deduction theorems. *PROCEEDINGS OF THE HERBRAND SYMPOSIUM LOGIC COLLOQUIUM '81*, 243–250.

Prakken, H., & Vreeswijk, G. (2001). Logics for defeasible argumentation. In Dov Gabbay (Ed.), *Handbook of Philosophical Logic - Vol. 4* (pp. 219–318). Kluwer.

Priest, G., Tanaka, K., & Weber, Z. (2018). Paraconsistent Logic. In *Stanford Encyclopedia of Philosophy* (pp. 1–48). The Metaphysics Research Lab Center for the Study of Language and Information.

Quezada, D. (1995). Lógica clásica de primer orden. In *Enciclopedia Iberoamericana de Filosofía* (pp. 71–104). Editorial Trotta.

Rescher, N., & Manor, R. (1970). On inference from inconsistent premisses. *Theory and Decision*, *1*, 179–217.

Roetti, J. (2016). *Reglas y diálogos. Una discusión lógica*. Academia Nacional de Ciencias de Buenos Aires.

Rosen, K. (2011). *Discrete mathematics and its applications* (7°). McGraw-Hill.

Russell, S., & Norvig, P. (1994). *Artificial Intelligence: A Modern Approach*. Prentice Hall.

Schlechta, K. (1995). “Defaults as generalized quantifiers.” *Journal of Logic and Computation*, *5*(4), 473–494.

Schlechta, K. (2004). *Coherent Systems*. Elsevier.

Shannon, C. (1948). A mathematical theory of communications. *The Bell System Technical Journal*, *27*, 379–432.

Shoham, Y. (1987). A semantical approach to nonmonotonic logics. *Proceedings Logics in Computer Science*, 275–279.

Strasser, C., & Antonelli, A. (2019). Non-Monotonic Logic. In *The Stanford Encyclopedia of Philosophy* (pp. 1–66). The Metaphysics Research Lab Center for the Study of Language and Information.

Tarski, A. (1969a). Fundamental concepts of the methodology of the deductive sciences. In *Logic, semantics, metamathematics - Papers from 1923 to 1938* (pp. 60–109). Oxford University Press.

Tarski, A. (1969b). On some fundamental concepts of metamathematics. In *Logic, semantics, metamathematics - Papers from 1923 to 1938* (pp. 30–37). Oxford University Press.

Thomason, R. (2011). Non-monotonicity in linguistics. In J. Van Benthem (Ed.), *Handbook of logic and language* (pp. 781–853). Elsevier.

van Heijenoort, J. (1967). Logic as Calculus and Logic as Language. In Cohen (Ed.), *Proceedings of the Boston Colloquium for the Philosophy of Science 1964/1966* (pp. 440–446). Reidel Publishing Company.

von Plato, J. (2014). From Axiomatic Logic to Natural Deduction. *Studia Logica*, *102*(Special Issue: Gentzen's and Ja´skowski's Heritage 80 Years of Natural Deduction and Sequent Calculi), 1167–1184.

Zhu, W., Lin, Y., Gong, N., & Du, G. (2008). Descriptive definitions of potential and actual infinities. *Kybernetes*, *37*(3–4), 424–432.

UNIVERSIDAD NACIONAL DEL SUR

TESIS DE DOCTORADO EN FILOSOFÍA

MONOTONÍA, NO MONOTONÍA Y SEMIMONOTONÍA. *Una perspectiva metalógica y filosófica*

Damián Olivarez Stagnaro

BAHÍA BLANCA, ARGENTINA

2021