Ingeniería y Ciencia, ISSN 1794–9165
Volumen 4, número 7, junio de 2008, páginas 23–45

Algunas lógicas modales asociadas al razonamiento de agentes inteligentes

Some modal logics associated to the reasoning of intelligent agents

Gloria Rúa M.[1] y Manuel Sierra A.[2]

Recepción: 15-may-2007/Modificación: 13-feb-2008/Aceptación: 18-feb-2008
Se aceptan comentarios y/o discusiones al artículo

Resumen

Se presentan las jerarquías de sistemas SCR–nT4, SCR–nT5 y SCR–nD45 con $n \geqslant 3$, en las cuales se formaliza la noción de creencia en el sentido de *creencia justificada*, de *conocimiento* y de *convicción* respectivamente, dando como resultado sistemas de lógicas doxásticas y epistémicas en los cuales el problema de la omnisciencia lógica puede ser parcialmente controlado. Los sistemas son caracterizados con semánticas al estilo Kripke, en las cuales, la longitud de las cadenas de mundos posibles se encuentra restringida en función del tipo de razonador. Así, la profundidad de un modelo corresponde a la longitud máxima de las cadenas de mundos posibles que figuren en el modelo, resultando que los modelos de profundidad n se encuentran asociados a los sistemas deductivos SCR–nT4, SCR–nT5 y SCR–nD45.

Palabras claves: agente, razonador, creencia, lógica modal, omnisciencia lógica, lógica doxástica, lógica epistémica, modelos de Kripke.

Abstract

The hierarchies of systems SCR–nT4, SCR–nT5 and SCR–nD45 with $n \geqslant 3$ are presented. In these systems the belief notion is formalized in the sense

[1] Ingeniera de Sistemas, gruamari@eafit.edu.co, egresada, Universidad EAFIT, Medellín–Colombia.
[2] Magíster en Matemáticas, msierra@eafit.edu.co, profesor integrante del grupo en Lógica y Computación, Universidad EAFIT, Medellín–Colombia.

of justified belief, of knowledge and of conviction respectively, giving systems of doxastic and epistemic logics as a result where the problem of the logical omniscience can be partially controlled. The systems are characterized with Kripke-style semantics. In these semantics, the length of the possible world chains is restricted in function of reasoner's type. Moreover, the depth of a model corresponds to the maxim length of the possible world chains that they figure in the model, being that the models of depth n are associated to the deductive systems SCR-nT4, SCR-nT5 and SCR-nD45.

Key words: agent, reasoner, belief, modal logic, logical omniscience, doxastic logic, epistemic logic, kripke models.

1 Presentación

En este trabajo, inicialmente se presenta un enfoque de los agentes inteligentes, en el cual, la noción de creencia juega un papel fundamental. Se describen algunas de las características que un agente podría tener, y se privilegia la descripción de los agentes como sistemas intencionales. Se indica cómo se formalizan los sistemas intencionales utilizando lógicas modales y semánticas de mundos posibles, y se presenta el llamado problema de la omnisciencia lógica, en donde se describen las características más problemáticas de las lógicas modales normales, cuando se utilizan como lógicas del conocimiento y la creencia.

En [1], como una aproximación a la solución del problema de la omnisciencia lógica, se presentan como extensiones del cálculo proposicional clásico, las jerarquías de sistemas deductivos SCR–$(n+1)$ y CP–n con $n \geqslant 0$. SCR–n es el *sistema de creencias para los razonadores de tipo n* y CP–n es el *cálculo proposicional asociado a los razonadores de tipo n*. Los teoremas de los sistemas SCR–n son interpretados como las creencias de un razonador de tipo–n, mientras en los sistemas CP–n se interioriza la noción de creencia mediante el operador $[R]$, en el siguiente sentido: X es una creencia de un razonador tipo–n (X es un teorema de SCR–n) si y solamente si $[R]X$ es un teorema de CP–n. La forma como se construyen los sistemas, permite que un razonador de tipo–$(n+1)$ sepa que es de tipo–n pero no siempre puede saber que es de tipo–$(n+1)$. Lo anterior significa que los razonadores de la jerarquía no son autoconscientes. La autoconciencia sólo se puede garantizar al extender los sistemas SCR–n al sistema modal K4. En los sistemas SCR–n el problema de la omnisciencia lógica se encuentra limitado, puesto que, estos sistemas carecen de la regla de inferencia 'de X se sigue $[R]X$', y en los teoremas del sistema el número de ocurrencias de operadores de creencia se encuentra limitado.

En [2], los sistemas de la jerarquía SCR–n con $n \geqslant 1$, son caracterizados con una semántica al estilo Kripke, en la cual, la longitud de las cadenas de mundos posibles se encuentra restringida en función del tipo de razonador. Así, la profundidad de un modelo corresponde a la longitud máxima de las cadenas de mundos posibles que figuren en el modelo, resultando que los modelos de profundidad n caracterizan el sistema deductivo SCR–$(n+1)$.

Finalmente, se presentan en este trabajo, como extensiones de los sistemas SCR–n, las jerarquías de sistemas SCR–nT4, SCR–nT5 y SCR–nD45, en los cuales se formaliza la noción de creencia en el sentido de *creencia justificada*, de *conocimiento* y de *convicción* respectivamente, dando como resultado sistemas de lógicas doxásticas y epistémicas en los cuales el problema de la omnisciencia lógica puede ser parcialmente controlado.

2 Marco teórico

2.1 Agentes inteligentes

La Inteligencia Artificial se ha ocupado de temas como el razonamiento, la búsqueda, la planificación, la gestión del conocimiento, el aprendizaje, los sistemas expertos, etcétera; siendo aplicable en diferentes disciplinas científicas. Las diferentes áreas que esta disciplina comprende se integran en la construcción de sistemas inteligentes como entidades que puedan actuar de forma autónoma y razonada, estos sistemas son llamados *Agentes Inteligentes.* Hacia este campo de estudio y desarrollo se ha enfocado la Inteligencia Artificial, de tal modo que se define como una disciplina orientada a la construcción de agentes inteligentes.

Con la tecnología de agentes, se pretende abordar de una manera más apropiada la construcción de sistemas complejos aplicados a diversos campos. En la mayoría de las ocasiones, los agentes no son desarrollados de forma independiente sino como un grupo de entidades que constituyen un sistema al cual se denomina multi–agente[3]. En este caso los agentes deben o pueden interactuar entre ellos. Las interacciones más habituales como informar o consultar a otros agentes permiten a los agentes "hablar" entre ellos, tener en cuenta lo que realiza cada uno de ellos y razonar acerca del papel jugado por los diferentes agentes que constituyen el sistema. La comunicación entre agentes se realiza por medio de un *lenguaje de comunicación de agentes.*

Desde las diferentes instancias interesadas en el desarrollo de esta tecnología, se avanza en la obtención de una metodología común para la construcción de agentes inteligentes, con tal metodología, la aplicabilidad de las técnicas de Inteligencia Artificial a cualquier tipo de problema podría abordarse de una manera más clara y unificada.

Se pueden distinguir dos nociones extremas de agentes. Una noción define un agente como a una entidad capaz de *intercambiar mensajes utilizando un lenguaje* de comunicación de agentes. Esta es la definición más utilizada dentro de la ingeniería de software basada en agentes, cuyo fin es conseguir la interoperabilidad entre aplicaciones en el ámbito semántico. Denota un sistema de software o hardware que disfruta de propiedades como autonomía, sociabilidad, reactividad, iniciativa etcétera. Una segunda noción de agente más fuerte o restrictiva que la anterior es la que considera a un agente como un sistema que, además de las propiedades ya enumeradas, lo define e implementa usando *conceptos que normalmente se aplican a los seres humanos.* Por ejemplo, es muy común caracterizar a un agente utilizando conceptos mentales como conocimiento, creencia, intención y obligación. En [4] se define un agente como "una entidad cuyo estado se caracteriza por un conjunto de componentes mentales, tales como creencias, capacidades, intenciones y acuerdos".

2.2 Agentes como Sistemas Intencionales

En general, se caracteriza a los agentes como *tomadores de decisiones racionales.* Se establece la funcionalidad a través de construcciones "mentales", como creencias, deseos e intenciones. Es clara la dificultad de especificar un sistema utilizando este tipo de construcciones, sin embargo, puede ser fácil si se construyen las especificaciones como una serie de "reglas" (fórmulas lógicas) como, por ejemplo: "Si el agente–1 cree que el agente–2 cree que la opción–1 está vacía entonces el agente–1 debería creer que el agente–2 tiene información errónea, y el agente–1 debe intentar informar al agente–2 de su error".

Las *creencias* se usan generalmente para referirse a la información que el agente tiene sobre su entorno. Esta información puede ser incorrecta, de la misma manera que en un humano la información sobre el entorno (su creencia) puede ser incorrecta. El término *intención* se usa para referirse a una meta que el agente seguirá hasta que tenga éxito o falle completamente.

Al explicar una actividad humana, se suele utilizar oraciones como "no fue a la reunión porque creía que la habían cancelado" o "estudió mucho porque quería obtener la mejor nota". Estas oraciones usan la *psicología popular*, por la que se puede predecir y explicar el comportamiento humano atribuyendo actitudes, tales como creencia, deseo y esperanza. La mayoría de las personas que hayan leído las oraciones anteriores encontrarán su significado enteramente claro.

En [5] se argumenta que con la postura intencional probablemente se puede describir cualquier cosa. Por ejemplo, considerando un interruptor de la luz: "Es perfectamente coherente tratar a un interruptor de la luz como a un agente con la capacidad de transmitir corriente a voluntad, el cual transmite invariablemente la corriente cuando cree que nosotros queremos que la transmita y no de otra manera; mover el interruptor es simplemente la manera de comunicar nuestros deseos". La mayoría de las personas encontrarán tal descripción absurda e infantil. La razón puede ser que mientras la descripción intencional de la postura es perfectamente coherente con el comportamiento observado de un interruptor de la luz, y es internamente correspondiente, "...esto no aporta nada, puesto que, se comprende esencialmente su mecanismo de forma suficiente como para proporcionar una descripción mecánica más simple de su comportamiento".

En resumen, cuanto más se sabe sobre un sistema, menos se necesita abordar explicaciones intencionales de su comportamiento. En cambio, con sistemas muy complejos, aún cuando se disponga de una visión completa y exacta de la arquitectura del sistema y de su trabajo, puede no ser factible una explicación mecánica de su comportamiento. Considere un computador, aunque se tenga su descripción técnica completa disponible, no es muy factible acudir a tal descripción para explicar por qué aparece un menú cuando se hace clic con el ratón sobre un icono. En tales situaciones, puede ser más apropiado adoptar una postura de descripción intencional, si esa descripción es coherente, y más simple que las alternativas. Las nociones intencionales serán así herramientas de abstracción, que proveerán una manera conveniente y familiar para describir, explicar, y predecir el comportamiento de sistemas complejos.

La teoría de la intencionalidad presentada en [6] dice que para explicar y predecir el comportamiento de un sistema se pueden tomar tres posturas: "una *postura física*, con la que se deduce el comportamiento a partir de la estructura y las leyes de la Física, la Química, la Biología; una *postura de diseño*, con la que uno abstrae los detalles de la constitución física del sistema y, suponiendo que ha sido diseñado, puede predecir su comportamiento si se conocen las intenciones del diseñador, así, se puede predecir cuándo sonará el despertador, aunque no se sepa nada de su estructura interna; una *postura intencional*, con la que el comportamiento se deduce a partir de los deseos y creencias que se adscriben al sistema, se trata al sistema como un agente racional, y se imagina qué creencias y deseos podría tener el agente, dada su situación en el mundo, y se predice su comportamiento suponiendo que actuará para satisfacer esos deseos".

Resulta que "un sistema intencional es aquel *cuyo comportamiento puede predecirse mediante el método de atribuirle creencias, deseos y perspicacia racional,* y puede serlo en distintos grados: uno de primer orden, tiene simplemente estados intencionales (creencias, deseos, etcétera) propios, uno de segundo orden, tiene además, creencias, deseos, etcétera sobre los estados intencionales de otros".

"En la vida cotidiana se suele adoptar informalmente esta postura intencional respecto a artefactos: 'el carro *no quiere* arrancar', 'el corrector ortográfico *se empeña* en corregir esta palabra' ... Pero, obviamente, se hace en un sentido metafórico".

La postura intencional resulta justificada, cuando la *complejidad* del sistema impide que se pueda predecir su comportamiento mediante una postura física o de diseño; siendo los deseos y creencias, *herramientas de abstracción* útiles para predecir el comportamiento de un sistema.

Existe una gran variedad de actitudes intencionales que pueden adscribirse a los sistemas, actitudes *epistémicas* como el conocimiento y la sabiduría, actitudes *doxásticas* como la creencia y la duda, actitudes *teleológicas* como el deseo y la intención, actitudes *deónticas* como la obligación y el compromiso.

2.3 Formalización de las nociones intencionales

Los agentes inteligentes se consideran en la Inteligencia Artificial como sistemas cuya conducta se puede predecir atribuyendo creencias, deseos e intenciones. Para representar estos atributos, se emplean diversos formalismos lógicos como la teoría de la intención presentada en [7], la lógica multi–modal BDI (Creencia, Deseo e Intención; *Belief, Desire, Intention*) y la familia de lógicas, para especificar sistemas multiagente, propuesta en [8]. En [6] se acuñó el término *sistema intencional* para describir entidades cuyo comportamiento puede predecirse atribuyendo creencia, deseo y talento racional.

Las actitudes más apropiadas para representar agentes se pueden clasificar en dos categorías: las *actitudes de información* tales como creencia y conocimiento y las *pro–actitudes* tales como deseo, intención, obligación, propósito, preferencia. Entonces, las actitudes de información serán las relativas a la información que un agente tiene sobre el mundo que él ocupa, y las pro–actitudes serán esas que de alguna manera guían las acciones de los agentes. Se tiene como tema de importancia, determinar qué *combinación* de actitudes sea la más apropiada para caracterizar a un agente.

A las nociones intencionales, tales como creencia y deseo, no se pueden aplicar las reglas estándar de sustitución de la *lógica clásica de primer orden.* En lógica clásica, el significado o valor semántico de una fórmula depende únicamente de los significados de sus sub–fórmulas. En contraste, las nociones intencionales tales como creencia no son funciones de verdad, las sustituciones por términos equivalentes no preservan el significado. Esto es lo que se ha denominado *opacidad referencial.* Por esta razón, la lógica clásica no es apropiada para razonar sobre nociones intencionales, y se requieren formalismos alternativos.

Hay dos aspectos a tener en cuenta cuando se construye un formalismo lógico para nociones intencionales: la aproximación sintáctica y la aproximación semántica. Cualquier formalismo debe caracterizarse en términos de estas dos aproximaciones diferentes: su *sistema deductivo,* y sus *modelos semánticos*[9].

Respecto a la aproximación sintáctica, existen dos enfoques fundamentales. El primero es el uso de un lenguaje *modal*, que contiene *operadores modales* que no son funciones de verdad. Un enfoque alternativo implicaría el uso de un *meta–lenguaje*, el cual es un lenguaje de primer orden que contiene términos que denotan fórmulas de algún otro *lenguaje objeto*, de tal forma que las nociones intencionales pueden representarse utilizando un predicado del meta–lenguaje. Ambos enfoques tienen sus ventajas y desventajas.

Respecto a la aproximación semántica, se tiene un enfoque básico, y consiste en adoptar una semántica de *mundos posibles*, donde las creencias, conocimientos, metas, y demás características del agente, se caracterizan con base en un conjunto de *mundos posibles* y una relación de *accesibilidad* entre ellos. La semántica de mundos posibles tiene asociada una *teoría de correspondencia*[10], la cual permite caracterizar axiomáticamente algunas propiedades que podrían ser, bajo ciertas circunstancias, semánticamente interesantes.

La semántica de los mundos posibles para lógicas de conocimiento y creencia fue propuesta originalmente por Hintikka en 1962, [11]. El aporte de Hintikka fue ver que las creencias de los agentes se podían caracterizar como un conjunto de *mundos posibles*. Hintikka acuñó el término *alternativas epistémicas* para describir los mundos posibles dadas unas creencias determinadas. En la actualidad, es más común utilizar la semántica de los mundos posibles, en la forma de una lógica modal normal que utiliza las técnicas desarrolladas por Kripke en 1963, [12].

Las *lógicas epistémicas* se definen comúnmente como ciertas *lógicas modales normales* que utilizan la semántica desarrollada por Kripke y que satisfacen algunos axiomas específicos relacionados con el conocimiento. Una lógica modal normal simple es esencialmente una lógica proposicional clásica, extendida mediante la adición del operador de necesariedad '$\Box$', donde la fórmula $\Box X$ puede leerse "necesariamente X", y la fórmula $\Box X$ será cierta en el mundo posible específico si X es cierta en cada mundo posible accesible desde el mundo posible específico.

Para utilizar la lógica descrita anteriormente como lógica epistémica, la fórmula $\Box X$ se lee como: "se sabe que X". Los mundos en el modelo se interpretan como alternativas epistémicas, la relación de accesibilidad define qué alternativas están disponibles desde cualquier mundo determinado, y además se deben satisfacer ciertos axiomas.

Los axiomas D, T, 4, y 5 son claves para lógicas de conocimiento y la creencia. El axioma D dice que las creencias de un agente son no–contradictorias; puede ser escrito como $\Box X \rightarrow \sim \Box \sim X$, que se lee "si el agente cree algo, entonces el agente no cree lo contrario". Este axioma parece una propiedad razonable de conocimiento y la creencia. El axioma T se conoce frecuentemente como el *axioma de conocimiento*, puede ser escrito como $\Box X \rightarrow X$, y dice que "lo que el agente conoce (o sabe) es cierto", se considera como el axioma que distingue el conocimiento de la creencia, parece razonable creer algo que es falso, pero habría duda en decir que se sabe algo que es falso. Así, el conocimiento frecuentemente se define como una creencia cierta, "el agente sabe algo si el agente lo cree y además es cierto". El axioma 4, $\Box X \rightarrow \Box\Box X$, "si el agente cree algo entonces el agente cree que lo cree", dice que un agente es consciente de lo que él sabe, se conoce como el axioma de *introspección positiva.* La introspección es el proceso por el que un observador se provee de creencias, y se discute de forma detallada en [9]. De forma similar, el axioma 5 es el *axioma de introspección negativa,* $\sim\Box X \rightarrow \Box\sim\Box X$, y dice que un agente es consciente de lo que él no sabe. Juntas, la introspección positiva y negativa, implican que un agente tiene conocimiento completo sobre que sabe y lo que no. Se acepta generalmente que la introspección positiva es una propiedad menos exigente que la introspección negativa, al ser una propiedad más razonable como recurso para razonar de forma definida. Por otro lado, el axioma K, $\Box(X \rightarrow Y) \rightarrow (\Box X \rightarrow \Box Y)$, dice que si el agente cree un condicional y cree su antecedente entonces el agente también cree su consecuente, es decir, el agente cree las consecuencias de su conocimiento. Los axiomas K, T, D, 4 y 5 se eligen frecuentemente como una lógica idealizada del *conocimiento,* y K, D, 4 y 5 como una lógica idealizada de la *creencia.*

Las lógicas mencionadas corresponden al conocimiento o creencia de un único agente. Para tratar con el conocimiento de múltiples agentes, se adiciona a la estructura del modelo un conjunto indexado de relaciones de accesibilidad, una para cada agente. El lenguaje se extiende entonces reemplazando el operador modal único, $\Box$, por un juego indexado de operadores modales unarios $\{\Box_i\}$, donde $i \in \{1, \ldots, n\}$, y cada operador $\Box_i$ tiene las mismas propiedades que '$\Box$', y la fórmula $\Box_i X$ se lee como "el agente i cree X".

2.4 El problema de la omnisciencia lógica

En los sistemas de lógicas modales normales se tienen dos propiedades básicas: la primera es la validez del esquema axiomático K: $\Box(X \rightarrow Y) \rightarrow (\Box X \rightarrow \Box Y)$, la segunda propiedad, conocida como *regla de necesariedad,* dice que: si X es válida entonces $\Box X$ también es válida. Dado que K es válido, será un teorema de cualquier axiomatización completa de una lógica modal normal, y la segunda propiedad aparecerá en general como una regla de inferencia de cualquier axiomatización de una lógica modal normal. Estas dos propiedades resultan ser las *características más problemáticas de las lógicas modales normales* cuando se utilizan como lógicas de conocimiento y la creencia.

La regla de necesariedad dice que un agente sabe todas las fórmulas válidas. Entre otras cosas, esto significa que un agente sabrá todas las tautologías proposicionales, y dado que hay un número infinito de éstas, un agente tendrá un número infinito de elementos de conocimiento. Ahora considerando el axioma K, el cual dice que el conocimiento de un agente se cierra bajo la implicación, junto con la regla de necesariedad este axioma implicará que el conocimiento de un agente se cierra bajo la consecuencia lógica, es decir, un agente creerá todas las consecuencias lógicas de sus creencias.

Estos dos aspectos: el saber todas las fórmulas válidas, y el del conocimiento o creencia cerrado bajo la consecuencia lógica; juntos constituyen el famoso problema de la *omnisciencia lógica.* Se ha argumentado ampliamente que este problema provoca que el modelo de los mundos posibles sea inapropiado para representar creencias en sistemas reales, dado que estos cuentan con recursos finitos.

Los formalismos para razonar sobre agentes han seguido un largo camino desde el primer trabajo sobre lógicas de conocimiento y creencia debido a Hintikka en 1962 [11]. Dentro de la Inteligencia Artificial, quizás el énfasis principal de los trabajos subsiguientes ha sido intentar desarrollar formalismos que capten la relación entre los diversos elementos que comprenden a un estado cognitivo de un agente. A pesar de que se han realizado progresos, todavía siguen pendientes muchas cuestiones y problemas fundamentales.

A nivel técnico, el problema de la omnisciencia lógica no puede considerarse resuelto. La semántica de los mundos posibles sigue siendo la elección para muchos investigadores, pero ésta no representa, en general, un modelo realista para agentes con recursos limitados.

En las secciones siguientes, se presenta una aproximación, con la cual se plantea una solución parcial al problema de la omnisciencia lógica. Dicha aproximación se inicia construyendo una jerarquía básica de formalismos, cada uno construido con base en el anterior, de tal forma que en cada nuevo sistema de la jerarquía el agente incremente su capacidad de razonamiento. En esta jerarquía la regla de necesariedad no tiene validez general, y su aplicación depende de la capacidad de razonamiento del agente. La caracterización de cada uno de los sistemas de la jerarquía, se logra con una semántica de mundos posibles, restringida según la capacidad de razonamiento del agente. La jerarquía básica consta de los sistemas SCR–n (sistema de creencias para un razonador de tipo–n, donde n es un entero positivo que indica la capacidad de razonamiento del agente), y a partir de ella se construyen y caracterizan semánticamente, las jerarquías de sistemas del conocimiento y la creencia SCR–nT4, SCR–nT5 y SCR–nD45 donde $n \geqslant 3$.

3 Sistemas básicos

3.1 Sistemas deductivos SCR–n ($n \geqslant 1$)

El lenguaje de todos los sistemas de la jerarquía SCR–n ($n \geqslant 1$) consta de los conectivos binarios $\rightarrow$, $\vee$, $\wedge$, $\leftrightarrow$; y los conectivos unarios $\sim$, $[R]$. El conjunto de formulas del cálculo proposicional clásico CP es generado recursivamente a partir de un conjunto de formulas atómicas utilizando los conectivos de la siguiente forma:

1. Si A es una fórmula atómica entonces A es una fórmula.

2. Si A es una fórmula entonces $\sim(A)$ es una fórmula.

3. Si A y B son fórmulas entonces $(A) \rightarrow (B)$, $(A) \leftrightarrow (B)$, $(A) \wedge (B)$ y $(A) \vee (B)$ son fórmulas.

El *sistema deductivo para el cálculo proposicional clásico*(CP), consta de los siguientes axiomas:

Ax 0.1	$A \to (B \to A)$
Ax 0.2	$(A \to (B \to C)) \to ((A \to B) \to (A \to C))$
Ax 0.3	$A \to (A \vee B)$
Ax 0.4	$B \to (A \vee B)$
Ax 0.5	$(A \to C) \to ((B \to C) \to ((A \vee B) \to C))$
Ax 0.6	$(A \wedge B) \to A$
Ax 0.7	$(A \wedge B) \to B$
Ax 0.8	$(A \to B) \to ((A \to C) \to (A \to (B \wedge C)))$
Ax 0.9	$A \to (\sim A \to B)$
Ax 0.10	$A \vee \sim A$
Ax 0.11	$(A \leftrightarrow B) \to (A \to B)$
Ax 0.12	$(A \leftrightarrow B) \to (B \to A)$
Ax 0.13	$(A \to B) \to [(B \to A) \to (A \leftrightarrow B)]$.

Como única *regla de inferencia* se tiene el *Modus Ponens* (MP): de A y $A \to B$ se infiere B.

Los sistemas SCR–n *sistema de creencias para los razonadores de tipo–n* se construyen de la siguiente manera:

SCR–1 es un *cálculo proposicional clásico* (CP).

Para $n \geqslant 1$, SCR–$(n+1)$ es el mismo sistema SCR–n adicionando $[R]A$ como axioma a cada axioma A, y representando internamente las reglas de inferencia primitivas, es decir:

A es una fórmula de SCR–$n \Rightarrow A$ y $[R]A$ son fórmulas de SCR–$(n+1)$

A, B son fórmulas de SCR–$(n+1) \Rightarrow A \wedge B$, $A \vee B$, $A \to B$, $A \leftrightarrow B$ y $\sim A$ son fórmulas de SCR–$(n+1)$

A es un axioma de SCR–$n \Rightarrow A$ y $[R]A$ son axiomas de SCR–$(n+1)$

$[R](A \to B) \to ([R]A \to [R]B)$ es un axioma de SCR–$(n+1)$.

Los sistemas tienen como única regla de inferencia el *modus ponens* (MP).

Se dice que una fórmula A es un *teorema de SCR–n*, o que *el razonador cree o acepta A (A es cierto sobre el sistema SCR–n)*, denotado $\Vdash_n A$, si y solamente si A es la última fórmula de una sucesión finita de fórmulas, tales que cada una de ellas es un axioma de SCR–n o se infiere de dos fórmulas anteriores utilizando la regla de inferencia MP. Cuando A es un teorema de SCR–1, es decir de CP, se utiliza la notación $\vdash A$.

Un resultado de interés, el cual será utilizado más adelante, y cuya prueba se encuentra presentada en [2], es el siguiente.

Proposición 3.1 (Creencia de la conjunción)**.** *Para cada $n \geqslant 1$ se tiene que, un razonador de tipo–n cree una conjunción si y solamente si el razonador cree cada uno de los coyuntos*

$$\Vdash_n [R](A \wedge B) \leftrightarrow ([R]A \wedge [R]B), \textit{para } n \geqslant 2\,.$$

3.2 n–modelos

Los marcos y modelos para los sistemas de la jerarquía SCR–n ($n \geqslant 1$) tienen similitudes con los marcos y modelos de las lógicas normales, pero también tienen diferencias, además en los nuevos marcos y modelos se cambia la presentación al hacer explícito en el mundo actual.

Definición 3.1. $M = (S, M_1, R)$ es un *marco* si y solamente si S es un conjunto, M_1 es un elemento de S y R es una relación binaria sobre S. Los elementos de S son llamados *mundos posibles*, el mundo posible M_1 es llamado el *mundo actual*, y la relación R es llamada *relación de accesibilidad* (la relación asociada al razonador R).

Si $M_1, M_2, \ldots, M_{n+1} (n \geqslant 0)$ son mundos posibles diferentes entre sí y $C = M_1 M_2 \ldots M_n M_{n+1}$ entonces, C es una *cadena* de M si y solamente si $(\forall k, 1 \leqslant k \leqslant n)(M_k R M_{k+1})$. En el caso anterior se dice que la *profundidad de* C es n ($prof(C) = n$), por lo que, la cadena formada por un único mundo, el mundo actual M_1, tiene profundidad 0. Observar que una cadena tiene profundidad n significa que la cadena esta formada por $n+1$ mundos posibles.

Si M_x es un mundo posible del marco M, se dice que: la *profundidad de* M_x es $k (prof(M_x) = k)$ si y solamente si $k = max\{p : prof(C) = p$, C es una cadena de M y M_x es el extremo final de $C\}$. Se dice que la *profundidad del marco* $M = (S, M_1, R)$ es $k (prof(M) = k)$ si y solamente si $k = max\{p : prof(M_x) = p$ y $M_x \in S\}$.

Definición 3.2. Sea $M = (S, M_1, R)$ un marco y F el conjunto de todas las fórmulas, $M = (S, M_1, R, V)$ es un n*–modelo* (*modelo de profundidad* n) si y solamente si $prof(M) \leqslant n$ y además V es una función (valuación) de $S \times F$ en $\{0, 1\}$ la cual satisface, para cada mundo posible M_x, las siguientes reglas o condiciones:

$Vat.$	$V(M_x, p) = 1$ ó $V(M_x, p) = 0$ cuando p es una fórmula atómica
$V\sim.$	$V(M_x, {\sim}A) = 1 \Leftrightarrow V(M_x, A) = 0$
$V\wedge.$	$V(M_x, A \wedge B) = 1 \Leftrightarrow V(M_x, A) = 1 = V(M_x, B)$
$V\vee.$	$V(M_x, A \vee B) = 0 \Leftrightarrow V(M_x, A) = 0 = V(M_x, B)$
$V \rightarrow.$	$V(M_x, A \rightarrow B) = 0 \Leftrightarrow V(M_x, A) = 1$ y $V(M_x, B) = 0$
$V \leftrightarrow.$	$V(M_x, A \leftrightarrow B) = 1 \Leftrightarrow V(M_x, A) = V(M_x, B)$
$V[R].$	$V(M_x, [R]A) = 1 \Leftrightarrow (\forall M_y \in S)(M_x R M_y \Rightarrow V(M_y, A) = 1)$, cuando $prof(M_x) < n$.

Observar que, respecto a los conectivos $\rightarrow, \leftrightarrow, \vee, \wedge$ y $\sim$, las valuaciones en cada mundo posible son valuaciones tradicionales en el sentido de la lógica clásica.

Definición 3.3. Sea A una fórmula,

A es *verdadera en el* n*–modelo* $M = (S, M_1, R, V)$ (denotado $M \vDash_n A$) $\Leftrightarrow$ $V(M_1, A) = 1$.

A es n*–válida* (denotado $\vDash_n A$) $\Leftrightarrow$ $(\forall M, M$ un n–modelo$)(M \vDash_n A)$.

Observar que cuando se habla de 0–validez, se hace referencia a los 0–modelos, pero en estos modelos no aplica la regla $V[R]$, es decir, los 0–modelos son los modelos del cálculo proposicional clásico CP, por lo que 0–validez coincide con validez en CP.

Resulta entonces que una fórmula A *no* es n–válida si y solamente si $(\exists M$, M un n–modelo$)(M$ no $\vDash_n A)$, es decir, $(\exists M = (S,\ M_1,\ R,\ V)$, M un n–modelo$)(V(M_1,\ A) = 0)$. Por lo que, si la fórmula A no es válida, utilizando las reglas $V\sim$, $V\wedge$, $V\vee$, $V \rightarrow$, $V \leftrightarrow$ y $V[R]$ a partir de $V(M_1, A) = 0$, se construye un n–modelo $M = (S, M_1, R, V)$ que refute la validez de la fórmula A, este modelo es llamado n*–modelo refutador*. Pero si la fórmula A es válida, entonces la construcción del n*–modelo refutador* fracasará puesto que, en alguno de los mundos posibles (bien sea M_1 o un mundo generado por la aplicación de la regla $V[R]$) del modelo en construcción se presentará una inconsistencia.

En resumen, para probar la n–validez de una fórmula A, se supone que la fórmula A no es n–válida, es decir, es falsa en el mundo actual M_1, y a partir de esta información se construye el n–modelo refutador. Si tal n–modelo no existe entonces se concluye que la fórmula A es válida.

Para finalizar esta sección, se presenta una proposición, cuya prueba se encuentra en [2], y en la cual se establece la conexión entre los sistemas deductivos y los n–modelos.

Proposición 3.2 (Caracterización semántica de SCR–$(n+1)$)**.** *Para cada $n \geqslant 0$, y para cada fórmula X, los teoremas del sistema SCR–$(n+1)$ son fórmulas n–válidas y sólo ellas.*

$$\Vdash_{n+1} X \Leftrightarrow \vDash_n X\,.$$

4 Sistemas doxásticos y epistémicos

4.1 Sistemas deductivos SCR–nD45, SCR–nT4, SCR–nT5 $(n \geqslant 1)$

Definición 4.1.

D es el esquema $[R]X \rightarrow \sim[R] \sim X$ (consistencia)

T es el esquema $[R]X \rightarrow X$ (conocimiento)

4 es el esquema $[R]X \rightarrow [R][R]X$ (introspección positiva)

5 es el esquema $\sim[R]X \rightarrow [R] \sim[R]X$ (introspección negativa).

Observar que el esquema D, por implicación material y negación de la conjunción es equivalente a $\sim([R]X \wedge [R] \sim X)$, lo cual por la proposición 3.1, creencia de la conjunción, significa $\sim[R](X\wedge \sim X)$; lo anterior significa que el esquema D garantiza la *consistencia de las creencias que el razonador acepta.*

En términos del operador de creencia $[R]$, se define el *operador de posibilidad* (R) de la forma

$$(R)X =_{def} \sim[R] \sim X \,,$$

donde $(R)X$ se lee *el razonador R considera posible X.*

Observar que por la definición de posibilidad se tiene $(R) \sim X \leftrightarrow \sim[R] \sim\sim X$, lo cual por *doble negación*[1] es $(R) \sim X \leftrightarrow \sim [R]X$, y por *transposición en el bicondicional* resulta $\sim(R) \sim X \leftrightarrow [R]X$. Se tiene de esta forma que el operador de creencia puede ser definido en términos del operador de posibilidad. Esta última equivalencia será también referenciada como *definición de la posibilidad.*

Definición 4.2. Los sistemas SCR–nT4, SCR–nT5 y SCR–nD45 donde $n \geqslant 3$ se obtienen a partir del sistema SCR–n de la siguiente forma:

Los sistemas SCR–2T y SCR–2D se obtienen a partir del sistema SCR–2 agregando como axiomas los esquemas T y D respectivamente. Los sistemas SCR–3T y SCR–3D se construyen de la siguiente forma:

- X es una fórmula de SCR–2T $\Rightarrow$ X y $[R]X$ son fórmulas de SCR–3T
- X es una fórmula de SCR–2D $\Rightarrow$ X y $[R]X$ son fórmulas de SCR–3D
- X, Y son fórmulas de SCR–3T $\Rightarrow$ $X \wedge Y$, $X \vee Y$, $X \rightarrow Y$, $X \leftrightarrow Y$ y $\sim X$ son fórmulas de SCR–3T
- X, Y son fórmulas de SCR–3D $\Rightarrow$ $X \wedge Y$, $X \vee Y$, $X \rightarrow Y$, $X \leftrightarrow Y$ y $\sim X$ son fórmulas de SCR–3D
- X es un axioma de SCR–2T $\Rightarrow$ X y $[R]X$ son axiomas de SCR–3T
- X es un axioma de SCR–2D $\Rightarrow$ X y $[R]X$ son axiomas de SCR–3D.

Los sistemas SCR–3T4 y SCR–3T5 se obtienen a partir del sistema SCR–3T agregando como axiomas los esquemas 4 y 5 respectivamente. Los sistemas

[1] Los siguientes son resultados muy conocidos de CP, los cuales son utilizados en este trabajo. Doble negación: $A \leftrightarrow \sim\sim A$. Transposición: $(A \rightarrow B) \leftrightarrow (\sim B \rightarrow \sim A)$. Transposición en el bicondicional: $(A \leftrightarrow B) \leftrightarrow (\sim B \leftrightarrow \sim A)$. Equivalencia material: $(A \leftrightarrow B) \leftrightarrow ((A \rightarrow B) \wedge (B \rightarrow A))$. Silogismo hipotético: $((A \rightarrow B) \wedge (B \rightarrow C)) \rightarrow (A \rightarrow C)$.

SCR–$(n+1)$T4 y SCR–$(n+1)$T5 donde $n \geqslant 3$ se construyen de la siguiente forma:

- X es una fórmula de SCR–nT4 $\Rightarrow$ X y $[R]X$ son fórmulas de SCR–$(n+1)$T4
- X es una fórmula de SCR–nT5 $\Rightarrow$ X y $[R]X$ son fórmulas de SCR–$(n+1)$T5
- X, Y son fórmulas de SCR–$(n+1)$T4 $\Rightarrow X \wedge Y$, $X \vee Y$, $X \rightarrow Y$, $X \leftrightarrow Y$ y $\sim X$ son fórmulas de SCR–$(n+1)$T4
- X, Y son fórmulas de SCR–$(n+1)$T5 $\Rightarrow X \wedge Y$, $X \vee Y$, $X \rightarrow Y$, $X \leftrightarrow Y$ y $\sim X$ son fórmulas de SCR–$(n+1)$T5
- X es un axioma de SCR–nT4 $\Rightarrow$ X y $[R]X$ son axiomas de SCR–$(n+1)$T4
- X es un axioma de SCR–nT5 $\Rightarrow$ X y $[R]X$ son axiomas de SCR–$(n+1)$T5.

El sistema SCR–3D45 se obtiene a partir del sistema SCR–3D agregando como axiomas los esquemas 4 y 5. Los sistemas SCR–$(n+1)$D45 donde $n \geqslant 3$ se construyen de la siguiente forma:

- X es una fórmula de SCR–nD45 $\Rightarrow$ X y $[R]X$ son fórmulas de SCR–$(n+1)$D45
- X, Y son fórmulas de SCR–$(n+1)$D45 $\Rightarrow X \wedge Y$, $X \vee Y$, $X \rightarrow Y$, $X \leftrightarrow Y$ y $\sim X$ son fórmulas de SCR–$(n+1)$D45
- X es un axioma de SCR–nD45 $\Rightarrow$ X y $[R]X$ son axiomas de SCR–$(n+1)$D45.

Para indicar que una fórmula X es un teorema de alguno de estos sistemas se utiliza la notación $\Vdash_{nT4} X$, $\Vdash_{nT5} X$ y $\Vdash_{nD45} X$ respectivamente.

Mientras no se diga lo contrario, en lo que sigue siempre que se haga referencia a los nuevos sistemas, se entiende que $n \geqslant 3$.

Proposición 4.1 (Esquema T como posibilidad)**.** *En los sistemas SCR–nT4 y SCR–nT5 se infiere para cada fórmula X*

$$X \to (R)X.$$

Prueba. Por el esquema T se tiene $[R] \sim X \to \sim X$, lo cual por *transposición* equivale a $X \to \sim [R] \sim X$, y esto por la definición de posibilidad significa $X \to (R)X$. □

Proposición 4.2 (Esquema 5 como posibilidad)**.** *En los sistemas SCR–nT5 y SCR–nD45 se infieren para cada fórmula X:*

$$(R)X \to [R](R)X$$
$$(R)[R]X \to [R]X\,.$$

Prueba. Del esquema 5 se tiene $\sim [R] \sim X \to [R] \sim [R] \sim X$, lo cual, por la definición de posibilidad, significa $(R)X \to [R](R)X$. Para probar la segunda parte de la proposición, notar que de la primera parte se obtiene $(R) \sim X \to [R](R) \sim X$, por *transposición* se infiere $\sim [R](R) \sim X \to \sim (R) \sim X$, por *doble negación* resulta $\sim [R] \sim\sim (R) \sim X \to \sim (R) \sim X$, lo cual por la definición de posibilidad significa $(R)[R]X \to [R]X$. □

Proposición 4.3 (Esquema 4 en SCR–nT5)**.** *En el sistema SCR–nT5 se infiere el esquema 4.*

$$\Vdash_{nT5} 4\,.$$

Prueba. Por el esquema T se tiene $[R](R)[R]Y \to (R)[R]Y$, y como de la proposición 4.2 se tiene $(R)[R]Y \to [R](R)[R]Y$, entonces por *equivalencia material* resulta $(R)[R]Y \leftrightarrow [R](R)[R]Y$, y como de la proposición 4.1 se tiene $[R]Y \to (R)[R]Y$, se infiere $[R]Y \to [R](R)[R]Y$. De la proposición 4.2 se tiene $(R)[R]Y \to [R]Y$, y como se tiene $[R]Y \to (R)[R]Y$, por *equivalencia material* resulta $[R]Y \leftrightarrow (R)[R]Y$, y como también se tiene $[R]Y \to [R](R)[R]Y$, entonces resulta $[R]Y \to [R][R]Y$, es decir, el esquema 4. □

Definición 4.3. Se dice que un razonador R de tipo–n *sabe que acepta (o sabe que cree)* X, denotado $[[R]][R]X$ si y solamente si el razonador cree X y además cree que acepta X, es decir,

$$R \textit{ sabe que cree } X \Leftrightarrow [[R]][R]X \Leftrightarrow ([R]X \wedge [R][R]X)\,.$$

Se dice que un razonador R de tipo–n *sabe que no acepta (o sabe que no cree)* X, denotado $[[R]] \sim [R]X$ si y solamente si el razonador no cree X y además acepta que no cree X, es decir,

$$R \textit{ sabe que no cree } X \Leftrightarrow [[R]] \sim [R]X \Leftrightarrow (\sim [R]X \wedge [R] \sim [R]X)\,.$$

Proposición 4.4 (Creer que no se acepta es saber que no se acepta)**.** *Un razonador de tipo–nT5 cree que no acepta algo si y solamente si sabe que no lo acepta.*

$$\Vdash_{nT5}\sim[R]X \Leftrightarrow \Vdash_{nT5} [[R]] \sim[R]X$$
$$\Vdash_{nT5}\sim[R]X \leftrightarrow [R] \sim[R]X\,.$$

Prueba. Supóngase $\Vdash_{nT5}\sim[R]X$, por el esquema 5 resulta $\Vdash_{nT5} [R] \sim[R]X$, al tener $\Vdash_{nT5}\sim[R]X$ y $\Vdash_{nT5} [R] \sim[R]X$, por la definición 4.3 se obtiene $\Vdash_{nT5} [[R]] \sim[R]X$. Para la recíproca, basta notar que por el esquema T, de $\Vdash_{nT} [R] \sim[R]X$ se sigue $\Vdash_{nT5}\sim[R]X$. La segunda parte resulta por *equivalencia material* en los esquemas 5 y T. □

Proposición 4.5 (Creer que se acepta es saber que se acepta)**.** *Un razonador de tipo–nT4 o tipo–nT5 cree que acepta algo si y solamente si sabe que lo acepta.*

$$\Vdash_{nT4} [R]X \Leftrightarrow \Vdash_{nT4} [[R]][R]X$$
$$\Vdash_{nT5} [R]X \Leftrightarrow \Vdash_{nT5} [[R]][R]X$$
$$\Vdash_{nT4} [R]X \leftrightarrow [R][R]X$$
$$\Vdash_{nT5} [R]X \leftrightarrow [R][R]X\,.$$

Prueba. Supóngase $\Vdash_{nT4} [R]X$, por el esquema 4 resulta $\Vdash_{nT4} [R][R]X$, al tener $\Vdash_{nT4} [R]X$ y $\Vdash_{nT4} [R][R]X$, por la definición 4.3 se obtiene $\Vdash_{nT4} [[R]][R]X$. Para la recíproca, basta notar que por el axioma T, de $\Vdash_{nT4} [R][R]X$ se sigue $\Vdash_{nT4} [R]X$. La segunda parte, gracias a la proposición 4.3 se prueba como la primera. Las dos últimas partes resultan por *equivalencia material* en los esquemas 4 y T.

De las proposiciones 4.3 y 4.5 se tiene que un razonador de tipo–nT5, cuando cree que acepta algo realmente sabe que lo acepta (introspección positiva), y cuando cree que rechaza algo realmente sabe que lo rechaza (introspección negativa), por lo que el operador de creencia de los razonadores de tipo–nT5 puede ser interpretado como un operador de *conocimiento*, es decir, los sistemas de la jerarquía SCR–nT5 donde $n \geqslant 3$ son *lógicas epistémicas*. Los razonadores de tipo–nT4 sólo poseen introspección positiva, por lo que el operador de creencia de los razonadores de tipo–nT4 no puede ser interpretado como en los razonadores de tipo–nT5, en [13] es interpretado como *creencia justificada*. Los razonadores de tipo–nD45 no poseen ningún tipo de introspección, pero al tener el esquema D, son consistentes respecto a las creencias que aceptan, por lo que el operador de creencia es interpretado como creencia en el sentido de *convicción*. Por lo que, los sistemas de las jerarquías SCR–nT4 y SCR–nD45 donde $n \geqslant 3$ son *lógicas doxásticas*. □

Proposición 4.6 (Consistencia de las creencias de los razonadores)**.** *Los razonadores de tipo–nT4 y tipo–nT5 son consistentes respecto a las creencias que aceptan.*

$$\Vdash_{nT4} D$$
$$\Vdash_{nT5} D\,.$$

Prueba. Por el esquema T se tiene $[R]X \to X$, y por la proposición 4.1 se tiene $X \to (R)X$, por *silogismo hipotético* se infiere $[R]X \to (R)X$, es decir, el esquema D. □

4.2 Semántica de los sistemas SCR–nD45, SCR–nT4, SCR–nT5 ($n \geqslant 1$)

Definición 4.4. Sea $\supset$ una relación de accesibilidad,

$\supset$ es *reflexiva* si y sólo si todo mundo accede a sí mismo, es decir, $\forall s(s \supset s)$

$\supset$ es *serial* si y sólo si todo mundo accede a algún mundo, es decir $\forall s \exists t(s \supset t)$

$\supset$ es *transitiva* si y sólo si un primer mundo accede a un segundo y el segundo accede a un tercero entonces el primer mundo también accede al tercero, es decir $\forall s \forall t \forall u((s \supset t \wedge t \supset u) \to s \supset u)$

$\supset$ es *euclidiana* si y sólo si cuando un mundo accede a otros dos entonces estos mundos acceden entre sí, es decir, $\forall s \forall t \forall u((s \supset t \wedge s \supset u) \rightarrow t \supset u)$

Definición 4.5.

Un *nT4–modelo* para $n \geqslant 3$ es un n–modelo (modelo de profundidad n) en el cual la relación de accesibilidad es reflexiva y transitiva.

Un *nT5–modelo* para $n \geqslant 3$ es un n–modelo en el cual la relación de accesibilidad es reflexiva y euclidiana.

Un *nD45–modelo* para $n \geqslant 3$ es un n–modelo en el cual la relación de accesibilidad es serial, transitiva y euclidiana.

Una fórmula X es *nT4–válida* ($\models_{nT4}$ X) si y solamente si X es verdadera en todos los nT4–modelos.

Una fórmula X es *nT5–válida* ($\models_{nT5} X$) si y solamente si X es verdadera en todos los nT5–modelos.

Una fórmula X es *nD45–válida* ($\models_{nD45} X$) si y solamente si X es verdadera en todos los nD45–modelos.

Proposición 4.7 (Caracterización semántica de SCR–nT4, SCR–nT5 y SCR–nD45)**.** *Para cada fórmula X se tiene:*

$$\models_{nT4} X \Leftrightarrow \Vdash_{(n+1)T4} X$$
$$\models_{nT5} X \Leftrightarrow \Vdash_{(n+1)T5} X$$
$$\models_{nD45} X \Leftrightarrow \Vdash_{(n+1)D45} X\,.$$

Prueba. Se tiene como consecuencia inmediata de la proposición 3.2 y de los siguientes resultados de la teoría de la correspondencia presentada en [10]: el esquema T se encuentra semánticamente caracterizado por los marcos reflexivos, el esquema 4 por los marcos transitivos, el esquema 5 por los marcos euclidianos, y el esquema D por los marcos seriales. $\square$

5 Anotaciones finales

Para finalizar, observando los resultados obtenidos, se hacen las siguientes anotaciones:

Primera: cuando se tiene un razonador de tipo–X (donde X es nT4, nT5 ó nD45 con $n \geqslant 3$), la noción de profundidad del sistema SCR–X, en cierto sentido, mide la capacidad del razonador para hacer inferencias en lo que respecta al operador de creencia, es decir, razonadores de distinto tipo tienen diferente poder de razonamiento y este poder de razonamiento puede ser medido. Lo anterior puede ser útil en el modelamiento de agentes, puesto que a partir de las limitaciones reales del agente, podría determinarse su poder de razonamiento, es decir, su tipo, y por lo tanto decidir con cual sistema de las jerarquías SCR–nT4, SCR–nT5 y SCR–nD45 con $n \geqslant 3$ se debe iniciar su modelamiento.

Segunda: para construir los sistemas SCR–nT4, SCR–nT5 y SCR–nD45 con $n \geqslant 3$, a partir de las lógicas de la jerarquía SCR–n con $n \geqslant 1$, fueron utilizadas las metodologías de la teoría de la correspondencia, dando como resultado sistemas de lógicas doxásticas y epistémicas en los cuales el problema de la omnisciencia lógica es parcialmente controlado.

Tercera: en la práctica, cuando se quiere modelar el conocimiento o la creencia de varios agentes, que de alguna manera están interactuando, se utilizan las lógicas multi–modales, en las cuales cada agente tiene asociado un operador de creencia (para más información ver [14]).

Teniendo en cuenta lo anterior, y aceptando la presentación de los sistemas de las jerarquías SCR–nT4, SCR–nT5 y SCR–nD45 (donde $n \geqslant 3$), parece natural continuar con la construcción de *jerarquías de sistemas multi–agente*, donde los razonadores o agentes pueden ser de diferente tipo.

Referencias

[1] Manuel Sierra. *Tipos de razonadores*, Revista Universidad EAFIT, ISSN 0120–341X, **43**(146), 2007. Referenciado en 24

[2] Manuel Sierra. *Caracterización semántica de la jerarquía SCR–n*. Revista Boletín de Matemáticas, ISSN 0120–0380, **XIV**(2), 110–128 (2007). Referenciado en 25, 35, 37

[3] M. N. Huhns, N. Jacobs, T. Ksiezyk, W. M. Shen, M. R. Singh and P. E. Cannata. *Integrating enterprise information models in Camot.* In Proceedings of the International Conference on Intelligent and Cooperative Information Systems. ISBN 0–8186–3135–X, 32–43 (1992). Rotterdam, The Netherlands. Referenciado en 26

[4] Yoav Shoham. *Agent Oriented Programming* in Knowledge Representation and Reasoning Under Uncertainty: Logic at Work (Lecture Notes in Computer Science), ISBN 3540580956, Editors: Michael Masuch and Laszlo Polos, 123–130 (1993). Referenciado en 26

[5] Y. Shoham. *Agent–oriented programming*, Technical Report STAN–CS–1335–90, Computer Science Department. Stanford University, Stanford, CA 94305, 1990. Referenciado en 27

[6] D. C. Dennett. *The Intentional Stance.* ISBN 0–262–04093–X. The MIT Press. 1987. Cambridge, MA. Referenciado en 28, 29

[7] P. R. Cohen and H. J. Levesque. *Rational interaction as the basis for communication.* In Intentions in Communication, Edited by Philip R. Cohen, Jerry Morgan and Martha E. Pollack, ISBN 978–0–262–03150–9. The MIT Press, Cambridge, MA, 1990. Referenciado en 29

[8] Michael John Wooldridge. *The Logical Modelling of Computational Multi–Agent Systems*. PhD thesis, Department of Computation, UMIST, Manchester, UK. (Also available as Technical Report MMU–DOC–94–01). Department of Computing, Manchester Metropolitan University., Chester St., Manchester, UK, 1992. Referenciado en 29

[9] K. A. Konolige. *A deduction model of belief.* Pitman Publishing: London and Morgan Kaufmann. San Mateo, CA, 1986. Referenciado en 29, 31

[10] Brian F. Chellas. *Modal logic: an introduction*, ISBN 978–0521295154. Cambridge University Press, Cambridge, 1980. Referenciado en 30, 43

[11] Jaakko Hintikka. *Knowledge and Belief.* Cornell University Press, lthaca, NY, 1962. Referenciado en 30, 32

[12] Saul Kripke. *Semantical analysis of modal logic.* Zeitschrift für Mathematische Logik und Grundlagen der Mathematik, **9**, 1963. Referenciado en 30

[13] Wolfgang Lenzen. *Recent work in epistemic logic.* Acta Philosophica Fennica, **30**(1), 1978. Referenciado en 42

[14] Max Freund. *Lógica epistémica*. Enciclopedia iberoamericana de filosofía, ISBN 84–8164–045–X, **7**, Editorial Trotta S. A., Madrid, 1995. Referenciado en 44