

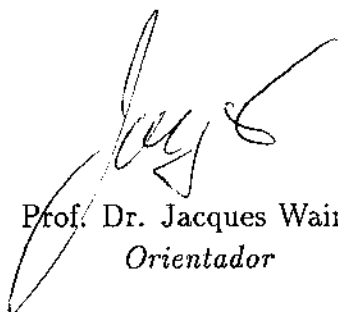
**Extensões Multiagentes
de duas Famílias de
Lógicas Não-monotônicas**

Ana Maria Monteiro

Extensões Multiagentes de duas Famílias de Lógicas Não-monotônicas

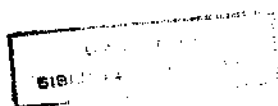
Este exemplar corresponde à redação final da tese devidamente corrigida e defendida pelo Sr. Ana Maria Monteiro e aprovada pela Comissão Julgadora.

Campinas, 27 de março de 1996.



Prof. Dr. Jacques Wainer
Orientador

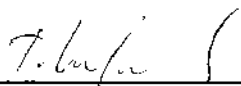
Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP, como requisito parcial para a obtenção do título de Mestre em Ciência da Computação.



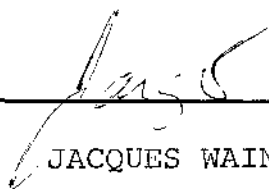
Tese de Mestrado defendida e aprovada em 27 de março de 1996
pela Banca Examinadora composta pelos Profs. Drs.



Prof (a). Dr (a). FLÁVIO SOARES CORRÊA DA SILVA



Prof (a). Dr (a). JOÃO MEIDANÊS



Prof (a). Dr (a). JACQUES WAINER

Extensões Multiagentes de duas Famílias de Lógicas Não-monotônicas¹

Ana Maria Monteiro²

Departamento de Ciência da Computação
IMECC – UNICAMP

Banca Examinadora:

- Dr. Jacques Wainer³ (Orientador)
- Dr. Flávio Soares Corrêa da Silva⁴
- Dr. João Meidanis⁵
- Dra. Ariadne M.B.R. Carvalho (Suplente)⁶

¹Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação da UNICAMP, como requisito parcial para a obtenção do título de Mestre em Ciência da Computação.

²A autora é graduada em Ciência da Computação pela Universidade Nacional de La Plata.

³Professor do Departamento de Ciência da Computação - IMECC - UNICAMP.

⁴Professor do Instituto de Matemática e Estatística - USP - SP.

⁵Professor do Departamento de Ciência da Computação - IMECC - UNICAMP.

⁶Professor do Departamento de Ciência da Computação - IMECC - UNICAMP.

À memória do meu avô Antonio.

Agradecimentos

A meu orientador Jacques pela confiança e paciência.

A meus amigos Clara e “el turco”, por seu carinho e apoio constantes.

A minha família. Em especial a minha avó Josefa e a minha tia Carolina.

Aos amigos Sueli, Ricardo, Mariano, Maria, Horacio, Sandra e Rosiane que fizeram com que não tivesse tantas saudades da minha terra.

A Jorge, Marcela e Zunilda, pelo carinho.

A meus amigos platenses Leda, Marta, Angela, Marisa, Adriana, Ruben, Claudia P., Claudia Q., Marcelo e Alina, por sua presença através do mail e seu incentivo constante.

Aos meus colegas das turmas 1993, 1994 e 1995.

A Diego, Raul, Lúcio, Maurício, Dinalva, Clevan, Mateus, Victor, Cris, Márcio Botelho, Márcio Stefani, Alexandre, Nuccio, Kareem, Verônica, George, Daniel, Cristina, Ricardo, Juliano, Luis Mariano, Cereja, Alex, Jerônimo, Walter, Maria Emilia, Anderson, Elaine, Helena, Bruno, “la Petty”, Márcia, Ximena, Pedro, Carlos e Pedrinho.

Aos professores do Departamento de Ciência da Computação, em especial à Ariadne por sua amizade.

Aos funcionários do departamento de computação, em especial a Roseli, a Solange e a Luiz.

A todos aqueles que me socorreram quando tive uma dúvida de português e a todos aqueles que corrigiram meus erros.

A Dona Rosa, Seu Agenor, Janete, Carla, Heloisa e Paula, por fazer meu dia-a-dia mais agradável.

Ao CNPq e à Faculdade de Ciências Exatas da Universidade Nacional de La Plata pelo apoio financeiro recebido.

Ao Brasil, por sua generosidade.

Resumo

O termo agente é usado com freqüência na área de representação de conhecimento e raciocínio. Diferentes sistemas foram propostos para modelar um agente raciocinando acerca de seu ambiente, um ambiente que, em muitas aplicações, pode incluir outros agentes. Assim, um agente, às vezes, deve interagir com outros agentes para levar a cabo uma tarefa ou atingir um objetivo. Neste caso, pode ser importante para o agente representar o conhecimento dos outros agentes com os quais ele interage e pode necessitar raciocinar acerca do conhecimento desses agentes para comunicar-se com eles.

Por outro lado, no contexto dos sistemas inteligentes, um agente pode necessitar realizar inferências na presença de informação incompleta ou obter conclusões em ausência de informação em contrário. Então o agente necessita dispor de mecanismos de raciocínio não-monotônicos.

O trabalho apresentado aqui relaciona-se com as duas questões anteriores: a proposta de lógicas não-monotônicas que não só modelam o agente raciocinando não-monotonicamente a respeito do mundo, mas também a respeito do conhecimento de outros agentes.

Este trabalho propõe a extensão de duas famílias de lógicas, para permitir o tratamento de fórmulas que contêm operadores de crenças. Mais especificamente:

- desenvolvemos duas lógicas condicionais que permitem raciocinar acerca das crenças de um outro agente (o qual raciocina também na lógica condicional). Discutimos também as propriedades de tais lógicas.
- propomos uma lógica preferencial multiagentes, baseada numa relação de preferência proposicional, de forma que tudo aquilo que é capturado por essa relação de preferência proposicional translada-se “corretamente” à lógica multiagentes.

Abstract

The term agent is frequently used in the area of knowledge representation and reasoning. Different systems have been proposed to model an agent reasoning about his environment, which in many applications includes other agents. Thus, an agent sometimes must interact with other agents to accomplish a task or a goal. If that is the case, it may be important for the agent to represent the knowledge of other agents with whom it interacts and may find it necessary to reason about their knowledge in order to communicate with them.

On the other hand, in the framework of an intelligent system, an agent must make certain inferences in the face of incomplete information or reach conclusions which rely, in part, on the “absence of evidence to the contrary”. Then the agent need to be invested with non-monotonic reasoning mechanisms.

The work presented here is concerned with the two issues above: the proposal of non-monotonic logics that not only model the agent reasoning non-monotonically about the world, but also about the knowledge of other agents.

This work proposes extensions of two families of logics in such a way that those extensions are able to deal with formulas that contain belief operators. More specifically:

- we develop two conditional logics that allow one to reason about the belief of another agent (which reasons in conditional logic himself). We also discuss the properties of such logics.
- we propose a multi-agent preferential logic, based on a propositional preference relation, in such a way that, whatever is captured by the propositional preference relation, it is carried over “correctly” to the multi-agent logic.

Conteúdo

1	Introdução	1
1.1	Lógica e Representação de Conhecimento	1
1.2	Sistemas de Raciocínio Não-Monotônico	3
1.2.1	Princípios Básicos	5
1.3	Descrição do Problema	6
1.4	Organização da Dissertação	7
2	Fundamentos Lógicos	9
2.1	Classificação das Lógicas Não-monotônicas	9
2.1.1	Interpretação dos Defaults	10
2.2	Lógica Default	11
2.2.1	Introdução	11
2.2.2	Teorias Default	11
2.2.3	Propriedades das Teorias Default	13
2.2.4	Teorias Default Normais	14
2.3	Lógica Modal	15
2.3.1	Semântica de Mundos Possíveis	16
2.3.2	Principais Sistemas Modais	18
2.3.3	Aplicações da Lógica Modal	22
2.4	Lógicas Condicionais	24
2.4.1	O Enfoque de Delgrande	25
2.4.2	Lógicas Condicionais de Normalidade	31
2.5	Lógicas Preferenciais	44
2.5.1	Lógica preferencial	44
2.5.2	Circunscrição	47
3	Lógicas Epistemicamente Estendidas	53
3.1	Introdução	53
3.2	Extensões Epistêmicas de uma Lógica	54
3.2.1	Requisitos para uma Lógica Epistemicamente Estendida	54

4	Extensões Epistêmicas das Lógicas Condicionais CT4O e CO	56
4.1	Introdução	56
4.2	A Extensão de CT4O	57
4.2.1	Resultados	60
4.3	A Extensão de CO	63
5	Uma Extensão Multiagentes de uma Lógica Preferencial	65
5.1	Introdução	65
5.2	A Lógica $\mathcal{L}^*$	65
5.3	A Lógica $\mathcal{L}_{\preceq}^*$	69
5.3.1	A Relação de Preferência $\preceq$	69
5.4	Interpretação Epistêmica da Lógica $\mathcal{L}_{\preceq}^*$	70
5.5	Propriedades da Lógica $\mathcal{L}_{\preceq}^*$	71
5.6	Exemplos	73
6	Conclusões e Extensões	76
A	Demonstrações dos Teoremas do Capítulo 4	78
5	Demonstrações dos Teoremas do Capítulo 5	85
	Bibliografia	97

Capítulo 1

Introdução

1.1 Lógica e Representação de Conhecimento

Um dos objetivos da Inteligência Artificial (IA) é compreender melhor o comportamento inteligente humano e utilizar essa compreensão na construção de sistemas computacionais artificialmente inteligentes.

Os pesquisadores desta área geralmente concordam com a importância do conhecimento e do raciocínio para a inteligência. Desta forma, o estudo de técnicas para representar conhecimento nos computadores converteu-se num dos temas centrais de IA.

Partindo da hipótese de que os sistemas inteligentes devem ter conhecimento, a forma mais conveniente para fornecer a eles esse conhecimento seria utilizando uma linguagem natural. Como isso ainda não é possível, é necessário utilizar uma linguagem artificial, uma linguagem formal para representar o conhecimento que poderá ser manipulado mais facilmente pelo computador. Essas linguagens formais para representar conhecimento deverão ser mais do que simples cadeias de caracteres sem significado. Elas devem ser capazes de mostrar como as expressões simbólicas estão relacionadas com a parte do mundo que está sendo modelada. Para tanto, é necessário definir uma semântica dessas linguagens; daí a importância da lógica formal na representação do conhecimento [Nil91], não obstante a falta de consenso quanto à relação entre o raciocínio humano e o raciocínio formal da lógica [Bir91].

Vários pesquisadores consideram que a lógica é muito importante para IA, já que é um método promissor para representação de conhecimento [Moo82]. Ela provê uma maneira clara e precisa de associar significado aos símbolos e um critério para distinguir entre as inferências justificadas e as não justificadas.

A IA compartilha, com outras áreas, a mais básica das motivações que levaram ao desenvolvimento da lógica clássica no fim do século passado: o propósito de formalizar um certo tipo de raciocínio. Mas o tipo de raciocínio que filósofos e matemáticos tentavam formalizar, e que deu origem à lógica desenvolvida para capturar o raciocínio matemático [Men79], difere em parte do tipo de raciocínio que as pessoas realizam diariamente. O conhecimento em que são baseadas as decisões das pessoas no dia-a-dia não é tão completo e preciso como o conhecimento de um matemático tentando estabelecer um resultado na sua área.

Uma das principais diferenças é que o raciocínio matemático tem a propriedade de ser *monotônico*, ou seja, se uma fórmula α se deduz de um conjunto de fórmulas Γ então α também é dedutível a partir de qualquer conjunto que inclua Γ . Esse tipo de raciocínio pode perfeitamente ser representado pela lógica clássica, que também possui essa propriedade (monotonicidade). Para formalizar o raciocínio de senso comum que os seres humanos realizam, geralmente não-monotônico, seria preciso uma lógica diferente. Muitas vezes, conclusões já feitas são descartadas na presença de nova informação ou conhecimento. Por exemplo, do fato que “os brasileiros, em geral, bebem cerveja” e que “João é brasileiro”, pode ser inferido, em ausência de outra informação (por “default”) que “João bebe cerveja”; mas se depois tem-se conhecimento de que “João não ingere bebidas alcoólicas”, então não se deveria deduzir que “João bebe cerveja”. Esta forma de raciocínio, que permite que informação adicional invalide conclusões anteriores é chamada não-monotônica.

Uma das perguntas que pode ser feita é se, na verdade, não é possível representar e raciocinar, como no exemplo de João, no contexto da lógica clássica. Por que não representar tal exemplo como:

$$\begin{aligned} &(\forall x)(\text{brasileiro}(x) \wedge \neg \text{brasileiro_excepcional}(x) \supset \text{bebe-cerveja}(x)) \\ &(\forall x)(\text{brasileiro_excepcional}(x) \leftrightarrow \text{crianca}(x) \vee \neg \text{bebe_álcool}(x) \vee \dots) \\ &\text{brasileiro}(\text{João})? \end{aligned}$$

Porque isto exige representar explicitamente todas as exceções, o que em muitos casos pode ser impossível. Mesmo que fosse possível representar explicitamente todas as exceções, para demonstrar que “João bebe cerveja” é preciso demonstrar que “João não é uma criança”, que “João não bebe álcool”, etc. O que se pretende é, por exemplo, demonstrar que “João bebe cerveja” quando não se tem informação que João é excepcional, sem ter que demonstrar que João não é criança, que bebe álcool, etc. E isto está fora dos limites da lógica clássica.

Também deve ser mencionado que há sistemas de IA tais como os baseados em Prolog, sistemas de planejamento ou algumas ferramentas para sistemas especialistas baseadas

em frames, regras ou redes semânticas [BW77, Bra77, RG77, SWC70], que são capazes de obter conclusões não-monotônicas. Mas, se tais sistemas existem, inclusive desde os primeiros dias da IA, por que procurar sistemas não-monotônicos baseados em lógica? Uma possível resposta a esta pergunta está relacionada ao fato de que as ferramentas existentes, ou são muito restritivas, pois foram desenhadas com um fim específico, ou quando são suficientemente gerais, pode acontecer que os resultados nem sempre sejam completamente entendidos ou que dependam da intuição da pessoa que desenvolve o sistema. Mas isto não é desejável, já que, em geral, a intuição não é suficiente. É importante, então, dar uma base formal à intuição, estabelecer definições precisas de inferência não-monotônica. Surge, então, a necessidade de uma semântica para essa inferência. A lógica formal é uma possível abordagem, e é essa a abordagem seguida neste trabalho.

1.2 Sistemas de Raciocínio Não-Monotônico

Na tentativa de formalizar o raciocínio de senso comum foram propostos diferentes sistemas com a capacidade de raciocinar não-monotonicamente: sistemas capazes de obter conclusões que logo possam ser descartadas em presença de nova informação ou sistemas com a capacidade de rever o conhecimento quando novo conhecimento é adquirido.

A necessidade de raciocinar de forma não-monotônica surge, por exemplo, quando o conhecimento disponível é incompleto e mesmo assim tem que ser tomada uma ação. Assim, de certa forma, o conhecimento é “completado” e conclusões não seguras são obtidas. Tais conclusões poderão ser revogadas quando nova informação estiver disponível.

Do ponto de vista das aplicações, este tipo de representação de conhecimento e raciocínio é utilizado para planejamento, diagnóstico, entendimento de linguagem natural, comunicações, visão, legislação, etc.. Em geral, a não-monotonicidade parece ser uma característica fundamental do raciocínio de senso comum em todas as áreas de aplicação.

A noção de raciocínio não-monotônico não é um fenômeno único e podem-se distinguir vários tipos: raciocínio default [Rei80], raciocínio autoepistêmico [Moo85b, Lev90], convenções de comunicação tais como a Hipótese de Mundo Fechado (“Closed-world Assumption”) [Rei84], sistemas de herança (redes semânticas e “frames”) [Sow84, Min81, Tou86], raciocínio probabilístico ou difuso [ZK92, KaL88].

Segundo a enumeração anterior, a noção de raciocínio não-monotônico é ampla e inclui o raciocínio probabilístico, onde evidência adicional pode diminuir a probabilidade condicional de um enunciado. Mas o raciocínio probabilístico, assim como o possibilístico ou o raciocínio difuso, é usualmente tratado numericamente: para representar o grau de plausi-

bilidade ou de certeza são utilizados números. Existem também algumas propostas onde a teoria das probabilidades infinitesimais é utilizada para prover de uma semântica aos formalismos não-monotônicos [Gef92]. Contudo, o uso do termo *raciocínio não-monotônico* está mais restrito aos enfoques baseados em lógica. Foram esses os enfoques considerados neste trabalho.

Os formalismos lógicos considerados procuram capturar distintos aspectos da não-monotonicidade do raciocínio humano. Entre esses aspectos podem ser mencionados:

- **A capacidade de raciocinar a partir de informação incompleta.** Tentando modelar estas capacidades surgiram os sistemas de *raciocínio default*, capazes de representar e raciocinar com enunciados da forma “Se α então normalmente β ” ou seja com defaults. Os *defaults* são afirmações ou regras gerais que admitem exceções ou propriedades prototípicas de classe ou indivíduos, do tipo “os brasileiros bebem cerveja” ou “os mamíferos geralmente vivem na terra”. Os defaults têm um papel importante no raciocínio de senso comum, já que permitem fazer predições úteis, na ausência de informação completa. Esse tipo de afirmações gerais que admitem exceções podem ser representadas e utilizadas para raciocinar nos sistemas defaults. A Lógica Default proposta por Reiter [Rei80] é um dos sistemas de raciocínio default mais conhecidos.
- **A capacidade de raciocinar não-monotonicamente sobre as próprias crenças.** Neste caso foi proposta a Lógica Autoepistêmica [Moo85b, Kon88] que permite representar e raciocinar com enunciados do tipo “se há um mamífero e não acredito (ou conheço) que não viva na terra, então esse mamífero vive na terra”. Embora este aspecto e o aspecto mencionado anteriormente sejam diferentes, muitos autores falam de representar defaults na Lógica Autoepistêmica.
- **A capacidade de estabelecer convenções na comunicação que permitam um intercâmbio de informação mais eficiente.** Esta é a capacidade do ser humano que faz com que uma pessoa saiba, por exemplo, que se num cartaz com os horários de partida de ônibus não aparece uma determinada hora, é porque nessa hora não parte nenhum ônibus. Uma das tentativas de formalizar este tipo de raciocínio é a Hipótese do Mundo Fechado [Rei84].
- **A capacidade de rever as próprias crenças quando é adquirido novo conhecimento.** Há varias propostas que tentam formalizar esta capacidade [AGM88, Gar88, Bou94a].
- **A capacidade de raciocinar com informação inconsistente.** Neste ponto pode-se mencionar a proposta de Poole [Poo88].

As capacidades enumeradas acima, em geral, não estão isoladas e aparecem frequentemente em combinação. O mesmo acontece com os distintos formalismos.

Neste trabalho serão considerados, fundamentalmente, os sistemas de raciocínio default.

1.2.1 Princípios Básicos

Embora existam diferentes propostas que tentam capturar certos aspectos relacionados com o raciocínio não-monotônico, há um conjunto de princípios básicos que seria desejável que estivessem presentes em qualquer abordagem formal da não-monotonicidade [Del94]:

- **Especificidade:** dado um conjunto de defaults que podem ser aplicados simultaneamente, e cuja aplicação individual produz resultados diferentes (*defaults conflitantes*), deveriam ser aplicados os mais específicos. Ou seja, deveria preferir-se os default mais específicos aos menos específicos.
- **Herança:** deveria ter-se herança de propriedades por default.
- **Irrelevância:** a informação irrelevante não deveria influir na dedução.

Para ilustrar estes princípios, suponha que se tenha as seguintes informações: “os mamíferos, geralmente, vivem na terra”, “os ursos são mamíferos” “as orcas são mamíferos”, “as orcas geralmente não vivem na terra” e “os mamíferos, em geral, têm pelo”.

- No exemplo, se há uma “orca” então, pelo princípio de especificidade, deveria concluir-se (por default) que “não vive na terra”, já que o default mais específico tem prioridade.
- Se há um “urso” deveria concluir-se (por default) que “tem pelos”, assumindo que não se tem informação em contrário.
- Do fato de ter uma “orca preta”, também deveria poder concluir-se que “não vive na terra”, já que não se tem informação em contrário, pois ser preto é irrelevante para o fato de que “as orcas vivem, geralmente, na terra”.

Ainda que desejável, a formalização de tal conjunto básico de princípios é difícil, sobretudo porque estão inter-relacionados e inclusive podem ser conflitantes. Em geral, as distintas lógicas não-monotônicas existentes só tratam adequadamente alguns destes aspectos, mas falham no conjunto total.

Levando em conta estes princípios, Delgrande [Del94] classifica as propostas para capturar aspectos relacionados com o raciocínio não-monotônico em duas categorias:

- **Sistemas fortes:** estes sistemas são permissivos, no sentido de que podem ser obtidas inferências não desejadas. Eles não tratam bem a especificidade dos defaults. Muitos dos trabalhos iniciais pertencem a esta categoria. A Lógica Autoepistêmica, Lógica Default e Circunscrição são exemplos destes sistemas.
- **Sistemas fracos:** em contraste com os anteriores, nestes sistemas algumas inferências default desejáveis não são obtidas. Embora tratem bem a especificidade dos defaults, são deficientes no tratamento da informação irrelevante. Pode-se dizer que estes sistemas têm um conjunto de propriedades que constitui um núcleo mínimo que deveria estar presente em todos os sistemas de inferência não-monotônica. Estes sistemas têm atraído, recentemente, a atenção dos pesquisadores da área. Entre outros pode ser mencionado o sistema N de Delgrande [Del88], ϵ -implicação [Pea88] e CT4 de Boutilier [Bou92].

1.3 Descrição do Problema

Os sistemas artificialmente inteligentes, tais como os baseados em conhecimento, podem ser vistos como *agentes* (artificiais) raciocinando e aplicando conhecimento para atingir objetivos. Esses agentes, em geral, têm que agir num ambiente onde nem sempre têm a sua disposição conhecimento completo sobre seu entorno, como acontece com as pessoas no mundo real.

No raciocínio que as pessoas realizam diariamente, elas em geral atribuem crenças às outras pessoas, e estas crenças ajudam a fazer predições sobre o que essas outras pessoas vão fazer.

Se um agente inteligente está interagindo com outros agentes, então necessitará raciocinar sobre esses agentes, e em particular raciocinar sobre as crenças, conhecimento e pontos de vista dos outros, para que a interação com eles seja efetiva. Assim, por exemplo, na comunicação com o usuário, um sistema de treinamento precisa conhecer não só sobre a área de treinamento como também a respeito do que conhece o sujeito que está sendo treinado (o outro agente no sistema).

A maioria das lógicas não-monotônicas propostas não consegue modelar um agente raciocinando acerca do conhecimento de outros agentes. Em geral, o estudo do raciocínio não-monotônico está limitado ao caso de um único agente, o que limita sua utilidade para modelar a comunicação entre eles. Por outro lado, as lógicas propostas para manipular o conhecimento ou as crenças de mais de um agente, em geral, têm a limitação de que

os agentes que elas modelam são monotônicos [FHV91, GD93, HM85, HM92] ou estão baseadas em lógicas autoepistêmicas [Mor90, Lak93, HM93, PJ95a]

Desta forma, para poder modelar o conhecimento não-monotônico de outros agentes, devem ser propostos novos formalismos ou adaptar os já existentes. Neste sentido, esta dissertação, levando em conta as limitações das lógicas não-monotônicas existentes, propõe a extensão de algumas delas, de maneira que possam representar fórmulas que façam referência ao conhecimento ou às crenças de outros agentes. Essas lógicas serão denominadas *lógicas epistemicamente estendidas* [Wai93]. Isto envolve estender a linguagem que essas lógicas utilizam e a semântica associada.

Nesta dissertação são propostas extensões de duas famílias de lógicas não-monotônicas. Estas extensões permitem representar crenças e raciocinar em um sistema multiagentes. Mais precisamente:

- Duas das lógicas condicionais propostas em [Bou94b] foram utilizadas como base para a obtenção de extensões que permitam representar as crenças e modelar o raciocínio de dois agentes.
- Foi definida uma lógica multiagentes não-monotônica, baseada numa lógica preferencial proposicional [Sho88].

No primeiro caso, foram escolhidas duas lógicas a família de de lógicas condicionais propostas por Butelier porque essas lógicas fornecem um conjunto mínimo de propriedades, desejáveis em toda lógica não-monotônica e permitem expressar as regras defaults na própria lógica. Com relação à lógica preferencial, ela foi escolhida como base para a definição da lógica multiagentes, pois é um formalismo suficientemente geral como para capturar os detalhes de muitos dos sistemas de raciocínio default.

1.4 Organização da Dissertação

Neste trabalho assume-se conhecidos os conceitos fundamentais da lógica proposicional clássica (LPC) e da lógica de primeira ordem (LPO). A dissertação está organizada da seguinte forma:

- No próximo capítulo é apresentada uma revisão bibliográfica das principais lógicas não-monotônicas relacionadas com o trabalho desenvolvido e os conceitos básicos da lógica modal.
- No capítulo 3 é introduzido o conceito de lógicas epistemicamente estendidas e são analisadas algumas das propriedades que seriam desejáveis nessas lógicas.

- No capítulo 4 são apresentadas as extensões epistêmicas de duas das lógicas propostas por [Bou94b].
- No capítulo 5 é apresentada uma lógica multiagentes não-monotônica, baseada na lógica proposta por [Sho88].
- Finalmente, no capítulo 6 são apresentadas as conclusões e possíveis extensões deste trabalho.

Capítulo 2

Fundamentos Lógicos

Neste capítulo apresentamos inicialmente uma classificação das lógicas não-monotônicas baseada no enfoque utilizado para desenvolvê-las. No contexto dessa classificação são apresentados os fundamentos lógicos do trabalho realizado. Também são apresentados os conceitos fundamentais da lógica modal.

2.1 Classificação das Lógicas Não-monotônicas

Na tentativa de formalizar o raciocínio de senso comum foram propostos diferentes sistemas com a capacidade de raciocinar não-monotonicamente. As lógicas não-monotônicas mais importantes podem ser agrupadas segundo diferentes critérios. Um deles é o tipo de enfoque utilizado para desenvolvê-las:

- **Enfoque baseado em regras de inferência não-monotônicas:** estas regras de inferência especificam diferentes maneiras de estender o conjunto de teoremas clássicos de um conjunto de premissas, de forma a incluir algumas crenças adicionais que eventualmente podem ser descartadas na presença de informação adicional. O exemplo mais importante deste tipo de enfoque é a Lógica Default proposta por Reiter [Rei80].
- **Enfoque modal:** uma linguagem lógica clássica é estendida agregando um operador modal, que é usado, por exemplo, para representar que uma determinada fórmula é consistente. Exemplos baseados neste enfoque são a Lógica Não-Monotônica proposta por McDermott e Doyle [McD80] e a Lógica Autoepistêmica proposta por Moore [Moo85b].
- **Enfoque baseado em condicionais:** neste enfoque a linguagem é aumentada com um conectivo binário. Esse conectivo binário é introduzido para representar

os defaults. Entre as diferentes propostas podem ser mencionadas a de Delgrande [Del87, Del88], Boutilier [Bou94a] e Lamarre [Lam91].

- **Enfoque preferencial:** neste enfoque, em vez de se considerar todos os modelos de um determinado conjunto de premissas, só é considerado um subconjunto desses modelos, aqueles “preferidos” segundo determinado critério. Pode ser mencionada aqui a proposta de Shoham [Sho88] e um caso particular deste, Circunscrição [McC80].

2.1.1 Interpretação dos Defaults

Os defaults têm um papel importante no raciocínio de senso comum, já que muitas afirmações acerca do mundo são expressas como defaults do tipo “os mamíferos, em geral, vivem na terra”.

Basicamente existem duas interpretações dos defaults conceitualmente diferentes [Gef92]:

- **Interpretação extensional:** nesta interpretação os defaults são considerados como prescrições que estendem um conjunto de crenças, na ausência de evidência “conflitante”, que pode originar-se em fatos conhecidos ou na interrelação dos mesmos defaults. Os diferentes defaults podem dar origem a distintas extensões que nem sempre são igualmente plausíveis. Por este motivo, a codificação de defaults nas lógicas que realizam esta interpretação deve ser usualmente estendida com axiomas de cancelamento apropriados, prioridades, ou mecanismos similares para que as extensões mais plausíveis sejam as preferidas.
- **Interpretação condicional:** nesta interpretação, introduzida mais recentemente, os defaults são vistos como crenças cuja validade está ligada a um determinado contexto. As lógicas que adotam esta interpretação dos defaults tentam solucionar alguns dos problemas presentes nas lógicas que adotam a outra interpretação.

A seguir é apresentada uma revisão bibliográfica das principais lógicas não-monotônicas relacionadas com o trabalho apresentado. Essas lógicas pertencem às famílias de lógicas desenvolvidas utilizando um enfoque condicional e preferencial. Também são apresentados os conceitos fundamentais da Lógica Default [Rei80], que pertence à família de lógicas desenvolvidas utilizando um enfoque baseado em regras de inferência não-monotônicas. A Lógica Default, embora não esteja relacionada diretamente com o trabalho desenvolvido é utilizada como ponto de referência e comparação com as outras lógicas.

2.2 Lógica Default

2.2.1 Introdução

A Lógica Default (LD), proposta por Reiter [Rei80] provê de uma formalização na qual um default do tipo *A's são tipicamente B's* é interpretado como: se x é um A e é consistente assumir que x é um B , então x é um B . Nesta proposta os defaults não são representados na linguagem, ao contrário, eles são considerados regras de inferência adicionais (não-monotônicas). Estas regras são usadas como base para a obtenção de um conjunto de *extensões* plausíveis à base de conhecimento, portanto a interpretação dos defaults realizada nesta lógica é uma interpretação extensional.

2.2.2 Teorias Default

A linguagem da LD é qualquer linguagem $\mathcal{L}$ de primeira ordem, clássica. As *regras default* ou simplesmente os *defaults* são definidos como segue:

Definição 2.1 Um default δ é uma configuração da forma:

$$\alpha(x) : \beta_1(x) \dots \beta_n(x) / \gamma(x)$$

onde:

- $\alpha(x)$, $\beta(x)$ e $\gamma(x)$ são fórmulas de primeira ordem cujas variáveis livres estão contidas em $x = \{x_1 \dots x_m\}$.
- $\alpha(x)$ é chamada de pré-requisito do default δ .
- $\beta_1(x) \dots \beta_n(x)$ é chamada de justificativa do default δ .
- $\gamma(x)$ é chamada de conseqüente do default δ .

O significado da regra default é o seguinte: se para um x específico $\alpha(x)$ pode ser demonstrado, e $\neg\beta_1(x), \dots, \neg\beta_n(x)$ não podem ser demonstrados, então conclua $\gamma(x)$.

Assim, por exemplo, o default “os mamíferos geralmente vivem na terra” pode ser representado como segue:

$$\text{mamífero}(x) : \text{vive_terra}(x) / \text{vive_terra}(x)$$

e é lido “se x é um mamífero e não é conhecido que seja falso que x viva na terra, então pode-se inferir que x vive na terra”.

Definidos os defaults, serão introduzidas as definições de *teoria default* e *teoria default fechada*:

Definição 2.2 *Uma teoria default é um par (D, W) , onde:*

- *D é um conjunto de defaults.*
- *W é um conjunto de fórmulas de primeira ordem.*

W contém os fatos e regras que são verdadeiros no mundo, e D contém os fatos e regras nos quais é razoável acreditar, sempre que forem consistentes com o que é conhecido. W é a base de conhecimento lógica que vai ser estendida (“completada”) utilizando-se as regras presentes em D .

Definição 2.3 *Uma teoria default fechada é uma teoria default onde os pré-requisitos, justificativas e conseqüentes dos defaults não contêm variáveis livres.*

Para completar W , Reiter introduz um operador de ponto fixo Γ . A idéia intuitiva é aplicar a W tantas regras defaults quanto seja possível, sem violar a consistência.

Definição 2.4 *Seja S um conjunto de fórmulas fechadas, $T = (D, W)$ uma teoria default fechada. Então $\Gamma(S)$ é o menor conjunto tal que:*

- D1.** $W \subseteq \Gamma(S)$.
- D2.** $Th^1(\Gamma(S)) = \Gamma(S)$.
- D3.** Se $\alpha : \beta_1, \dots, \beta_n / \gamma \in D$, $\alpha \in \Gamma(S)$ e $\neg\beta_i \notin S$, para $i \in \{1, \dots, n\}$ então $\gamma \in \Gamma(S)$.

Definição 2.5 *Uma extensão de uma teoria $T = (D, W)$ é qualquer subconjunto E de fórmulas ($E \subseteq \mathcal{L}$) tal que $\Gamma(E) = E$.*

A condição D1 garante que os fatos e regras conhecidos sobre o mundo estejam contidos em cada extensão, D2 garante que as extensões sejam dedutivamente fechadas e D3 garante que são aplicados tantos defaults quantos forem possíveis. A minimalidade, por sua vez, garante que a extensão não tem fórmulas que não estejam baseadas em W ou em D ; apenas expressões geradas pela aplicação das regras de inferência à base de conhecimento são permitidas numa extensão; adições gratuitas não são permitidas.

É claro que esta abordagem é não-monotônica. Por exemplo, a teoria $T = (D, W)$ onde D e W estão definidos como

$$\begin{aligned} D &= \{mamífero(x) : vive_terra(x)/vive_terra(x)\} \\ W &= \{mamífero(Moby)\} \end{aligned} \tag{2.1}$$

tem uma única extensão que contém a fórmula $vive_terra(Moby)$. No entanto, se $\neg vive_terra(Moby)$ é agregado a W , então a extensão desta nova teoria já não tem a

¹Onde $Th(A)$ está definida da maneira usual, ou seja, $Th(A) = \{\alpha : A \vdash \alpha\}$.

fórmula *vive_terra(Moby)*.

As definições apresentadas só estão definidas para teorias default fechadas. Os defaults que contêm variáveis livres são interpretados como esquemas que representam todas as instâncias sem variáveis livres dos defaults.

No exemplo (2.1) a teoria $T = (D, W)$ tem uma única extensão. Mas se $T = (D, W)$ onde

$$\begin{aligned} D &= \{ \text{mamífero}(x) : \text{vive_terra}(x) / \text{vive_terra}(x), \\ &\quad \neg \text{pelo}(x) : \neg \text{vive_terra}(x) / \neg \text{vive_terra}(x) \} \\ W &= \{ \text{mamífero}(\text{Moby}), \neg \text{pelo}(\text{Moby}) \} \end{aligned}$$

então T tem duas extensões, uma que contém $\neg \text{vive_terra}(\text{Moby})$ e outra que contém $\text{vive_terra}(\text{Moby})$. Conclui-se assim que as teorias default podem ter mais de uma extensão.

As extensões podem ser vistas como as diferentes maneiras plausíveis que tem um agente para estender seu conjunto inicial de crenças W com relação às regras default. A idéia original de Reiter foi que um agente poderia considerar qualquer uma das extensões de uma teoria como um modelo de trabalho do mundo e raciocinar a partir dessa extensão, embora geralmente se considere que as *conseqüências de uma teoria default* sejam as fórmulas que estão na intersecção de todas as extensões.

2.2.3 Propriedades das Teorias Default

As teorias default têm as seguintes propriedades, que não demonstraremos aqui²:

- Se $T = (D, W)$ é uma teoria tal que $D = \emptyset$, então T tem uma única extensão que é $Th(W)$.
- Se uma teoria tem uma extensão inconsistente, então essa é sua única extensão.
- Se $T = (D, W)$ e W é inconsistente, então a única extensão de T é o conjunto de todas as fórmulas da linguagem $\mathcal{L}$.
- As extensões são mínimas com relação à inclusão de conjuntos, ou seja, se E_1 e E_2 são duas extensões tal que $E_1 \subseteq E_2$, então $E_1 = E_2$.
- Há teorias que não têm extensões. Por exemplo, $T = (\emptyset, \{ : p / \neg p \})$.

²As demonstrações destas propriedades acham-se em [Bes89]

2.2.4 Teorias Default Normais

Na proposta original, Reiter estava convencido de que os únicos defaults interessantes eram os *defaults normais*.

Definição 2.6 Um default é normal se ele é um default da forma

$$\alpha(x) : \beta(x) / \beta(x),$$

Ou seja, um default é normal se a justificativa é equivalente ao conseqüente.

As teorias que só têm esse tipo defaults são chamadas de *teorias default normais*. Estas teorias têm as propriedades seguintes:

- As teorias default normais sempre têm uma extensão.
- Se uma teoria default normal tem uma extensão E , e um outro default normal é agregado a essa teoria, então a nova teoria tem uma extensão E' tal que E' contém E (semi-monotonicidade): ou seja, a agregação de defaults não destrói as extensões anteriores.

Em alguns casos, a agregação de defaults normais produz extensões não desejadas. Assim, por exemplo, se a teoria T é tal que:

$$W = \{estudante_pos(Paulo), adulto(Paulo)\}$$

$$D = \{estudante_pos(x) : \neg trabalha(x) / \neg trabalha(x), adulto(x) : trabalha(x) / trabalha(x), \}$$

então os dois defaults da teoria são conflitantes e T tem duas extensões, uma que contém $trabalha(Paulo)$ e uma outra que contém $\neg trabalha(Paulo)$. Mas neste caso, provavelmente, deseja-se dar prioridade ao primeiro default porque é mais específico (princípio de especificidade). É preciso então, em certas circunstâncias, bloquear a aplicação do default menos específico, mas isto pode tornar a formulação longa e confusa.

Para tentar solucionar os problemas do tipo apresentado anteriormente surgiram, por exemplo, as teorias default semi-normais [RC81] e a Lógica Default com prioridades [Bre94].

Em geral, a LD apresenta algumas dificuldades na tentativa de formalizar raciocínio default. Na LD os defaults são regras de inferência e não expressões da linguagem. Assim, elas não podem ser manipuladas por outras regras de inferência, nem podem ser deduzidas novas regras default.

A LD não tem problemas com a informação irrelevante e permite representar e raciocinar sobre herança de propriedades, na medida que tal representação não tenha defaults conflitantes, já que, como foi visto, o princípio de especificidade não é válido nesta lógica. Em geral, os problemas da LD estão associados com as prioridades dos defaults. Recentemente Brewka [Bre94] formulou uma proposta referente ao raciocínio sobre prioridades na LD.

2.3 Lógica Modal

A teoria da lógica modal (LM) foi desenvolvida originalmente para capturar a diferença entre os fatos e as regras que podem ser verdadeiros (verdades contingentes) e os fatos e as regras que não podem ser falsos (verdades necessárias). Assim, por exemplo, a maioria das pessoas concorda que “ $1+1 = 2$ ” é necessariamente verdadeiro, já que, para que seja falso deveria ter que mudar o significado dos símbolos 1, 2, = e +. No entanto, o fato “Ana tem cabelo comprido” é algo que tranqüilamente pode não ser verdadeiro, ou seja, é uma verdade contingente.

A diferença entre verdades contingentes e verdades necessárias, assim como determinadas intuições acerca das relações entre ambas, do tipo “se α é necessariamente verdadeiro então α é contingentemente verdadeiro”, é o que tenta formalizar a lógica modal.

Mas, ao analisar uma sentença, o que faz com que ela seja considerada necessária ou contingente? A resposta a esta pergunta tem a ver com o fato de se considerar possível a existência de mundos onde a sentença possa ser falsa. Assim, por exemplo, dadas as sentenças:

1. Só os objetos físicos têm cores
2. João está escrevendo a sua tese
3. Nada pode viajar mais rápido do que a luz

pode ser imaginado um mundo onde 1 seja falsa? Fortes mudanças deveriam ser feitas com relação ao conceito de objeto físico e ao que significa ter cor para mostrar que 1 pode ser falsa. No caso 2 é perfeitamente possível pensar num mundo onde João não esteja escrevendo a sua tese. Com relação a 3, para a maioria dos físicos contemporâneos, parece ser uma verdade necessária em mundos suficientemente parecidos com o atual, onde palavras como luz e viajar conservam o mesmo significado. No entanto, escritores de ficção científica têm considerado 3 como uma verdade contingente que pode ser descartada sem que os leitores considerem que suas histórias não têm sentido. Então 3 parece ser uma

verdade necessária em alguns contextos e contingente em outros.

Assim, na lógica modal, decidir quando alguma sentença é verdadeira ou não, pode ser feito inspecionando-se o mundo tal qual ele é. Por outro lado, decidir quando uma sentença é necessariamente verdadeira só pode ser feito levando-se em conta todos os mundos que podem ser considerados possíveis. O conceito de necessidade está relacionado então com o conjunto de mundos que podem ser considerados possíveis. Restrições referentes a que mundos podem ser vistos como possíveis, a partir de um determinado mundo, induzem a uma *relação de acessibilidade* entre os mundos. Esta relação de acessibilidade é a característica mais importante desses mundos. Diferentes propriedades dessa relação estão relacionadas com as diferentes lógicas modais que se podem definir. Assim, a introdução de um axioma particular numa lógica modal está associada com determinadas propriedades da relação de acessibilidade.

A seguir serão apresentados formalmente os conceitos básicos associados com a família de lógicas modais.

2.3.1 Semântica de Mundos Possíveis

Para poder realizar um tratamento rigoroso do conceito de necessidade é preciso ter uma linguagem na qual se possa fazer referência às verdades necessárias ou contingentes e prover um conjunto de axiomas e regras de inferência junto com uma semântica adequada para esses axiomas e regras [HC68, Che80].

Uma linguagem modal $\mathcal{L}$ é uma linguagem proposicional clássica (LPC)³ formada a partir de um conjunto de variáveis proposicionais $P = \{p_0, p_1, \dots\}$ e dos conectivos $\neg$ (negação) e $\supset$ (implicação material), aumentado com o conectivo ou *operador modal* $\Box$ (necessidade). Formalmente:

Definição 2.7 *As fórmulas bem formadas (fbf) de $\mathcal{L}$ são o menor conjunto tal que:*

1. $p_i \in P$ é uma fbf (fórmula bem formada atômica).
2. Se α e β são fbf então $\neg(\alpha)$, $(\alpha \supset \beta)$ e $\Box(\alpha)$ são fbf.

Conjunção ($\wedge$), disjunção ($\vee$) e bicondicional ($\leftrightarrow$) são introduzidos da maneira usual, por definição:

³Aqui só vai ser apresentada a versão proposicional da lógica modal, uma versão de primeira ordem pode-se ver em [HC68].

Definição 2.8

1. $\alpha \wedge \beta \equiv_{def} \neg(\alpha \supset \neg\beta)$.
2. $\alpha \vee \beta \equiv_{def} \neg\alpha \supset \beta$.
3. $\alpha \leftrightarrow \beta \equiv_{def} (\alpha \supset \beta) \wedge (\beta \supset \alpha)$.

Um novo operador modal é introduzido por definição:

Definição 2.9 $\Diamond\alpha \equiv_{def} \neg(\Box\neg\alpha)$.

Este conectivo é introduzido para expressar possibilidade. A fórmula $\Diamond\alpha$ é lida como “ α é possível”.

Uma vez definida a linguagem, vai ser apresentada a semântica da lógica modal em termos das estruturas de mundos possíveis introduzidas por Kripke [Kri63]. Estas estruturas⁴ são formadas a partir de um conjunto W e uma relação binária R definida em W . Os elementos de W são denominados *mundos possíveis*. Estes mundos possíveis correspondem às diferentes situações ou estados das coisas que um agente pode conceber ou imaginar como possíveis. R é a *relação de acessibilidade* entre esses mundos. Assim, se wRv , então se diz que v é acessível a w ou que w “enxerga” a v . Se wRv , então v é um possível estado alternativo das coisas, se w fosse a descrição da realidade.

A relação R é, inicialmente, qualquer relação binária definida em W . Especificando diferentes restrições para essa relação podem ser obtidas diferentes *lógicas* ou *sistemas modais*. Formalmente as *estruturas de mundos possíveis* ou *modelos modais*⁵ podem ser definidas como:

Definição 2.10 Um modelo modal é um par $M = (W, R)$ onde W é um subconjunto de 2^P e R é uma relação binária em W .

Um mundo possível é qualquer subconjunto de P .

Dado um modelo modal M , a validade de uma fórmula num mundo w de um modelo M (denotado por $M, w \models \alpha$) é definida como segue:

Definição 2.11 Seja $M = \langle W, R \rangle$ um modelo modal. A validade de uma fórmula modal num mundo w do modelo M é definida como:

⁴Estas estruturas são, na lógica modal, o conceito equivalente às atribuições de valores de verdade às variáveis proposicionais na LPC, ou às interpretações na LPO.

⁵Também denominadas estruturas de Kripke ou modelos de Kripke.

1. $M, w \models \alpha$ se e somente se $\alpha \in w$, para α fórmula atômica.
2. $M, w \models \neg\alpha$ se e somente se $M, w \not\models \alpha$.
3. $M, w \models \alpha \supset \beta$ se e somente se $M, w \not\models \alpha$ ou $M, w \models \beta$.
4. $M, w \models \Box\alpha$ se e somente se $M, v \models \alpha$, para todo v tal que wRv .

Definição 2.12 Se $M, w \models \alpha$ então se diz que M satisfaz α . α é válida em M (denotado por $M \models \alpha$) se $M, w \models \alpha$ para todo $w \in W$.

Na lógica modal o fato de um modelo satisfazer uma fórmula está associado não só com esse modelo, mas também com um mundo particular dessa estrutura, diferentemente da lógica clássica no qual só tem que ser considerado o modelo.

O símbolo $\|\alpha\|^M$ será utilizado para representar o conjunto $\{w \in W : M, w \models \alpha\}$ e de maneira abreviada $\|\alpha\|$ quando esteja claro, pelo contexto, o modelo M no qual é considerada a validade de α .

Se α é uma fórmula, então um α - mundo é um mundo possível na qual α é verdadeira.

2.3.2 Principais Sistemas Modais

A seguir serão introduzidos os axiomas e as regras de inferência de diferentes sistemas modais. À medida que esses axiomas são introduzidos são investigadas também as propriedades da relação de acessibilidade que fazem com que eles sejam válidos. Os axiomas têm nomes que são utilizados, em geral, para denominar os sistemas que surgem agregando esses axiomas ao cálculo proposicional. Assim, por exemplo, a lógica **K** é a que surge agregando o axioma **K** ao cálculo proposicional, a lógica **T** é a obtida agregando o axioma **T** a **K**, e assim por diante.

Desta forma, um sistema ou lógica modal pode ser definido como:

Definição 2.13 Um sistema ou lógica modal **S** é qualquer subconjunto de sentenças de $\mathcal{L}$ que contém as tautologias proposicionais e é fechado segundo o esquema de axioma **K** e as regras de inferência **MP** (Modus Ponens) e **Nec** (Necessitação):

K : $\Box(\alpha \supset \beta) \supset (\Box\alpha \supset \Box\beta)$.

MP : a partir de α e $\alpha \supset \beta$ infira β .

Nec : a partir de α infira $\Box\alpha$.

Definição 2.14 *O sistema modal K é o menor dos sistemas modais.*

Para que o axioma K seja válido não é preciso colocar nenhuma restrição na relação de acessibilidade. K é válido em qualquer modelo modal, da mesma forma que são válidos todos os axiomas da LPC.

Antes de definir os outros sistemas modais, será introduzido um conjunto de conceitos que se aplicam a qualquer um desses sistemas.

Primeiro é definido o conceito de *dedução* num sistema modal:

Definição 2.15 *Dado um sistema modal S e uma fórmula $\alpha \in \mathcal{L}$ se dirá que:*

- α pode ser deduzida em S (denotado por $\vdash_S \alpha$) se e somente se $\alpha \in S$.
- α é dedutível a partir de um conjunto $\Gamma \in \mathcal{L}$ (denotado por $\Gamma \vdash_S \alpha$) se existe algum subconjunto de Γ finito $\{\alpha_1, \dots, \alpha_n\}$ tal que $\vdash_S (\alpha_1 \wedge \dots \wedge \alpha_n) \supset \alpha$.

Os sistemas modais, em geral, estão associados com classes de modelos por meio de resultados de caracterização.

Definição 2.16 *Se $\mathcal{C}$ é uma classe de modelos, então uma fórmula α é válida com relação a $\mathcal{C}$ se $M \models \alpha$ para todo $M \in \mathcal{C}$ (denotado por $\models_{\mathcal{C}} \alpha$).*

Definição 2.17 *Se S é um sistema modal então:*

- S é correta com relação a $\mathcal{C}$ se e somente se $\vdash_S \alpha$ implica $\models_{\mathcal{C}} \alpha$
- S é completo com relação a $\mathcal{C}$ se e somente se $\models_{\mathcal{C}} \alpha$ implica $\vdash_S \alpha$

Definição 2.18 *$\mathcal{C}$ está caracterizado por S se e somente se S é correto e completo com relação a $\mathcal{C}$.*

Assim, por exemplo, a classe de todos os modelos está caracterizada pela lógica K , já que o axioma K é verdadeiro num sistema modal com qualquer relação de acessibilidade.

O Sistema T

O próximo esquema de axioma considerado é o esquema de *axioma T*, que diz que qualquer fórmula que é necessariamente verdadeira é, na verdade, verdadeira:

$$\text{T: } \Box(\alpha) \rightarrow \alpha.$$

T é válido na classe de modelos onde a relação de acessibilidade é reflexiva, ou seja, onde qualquer mundo é acessível a partir de si próprio.

Definição 2.19 O sistema **T** é o sistema obtido a partir de *K*, agregando o esquema de axioma **T**.

O Sistema S4

A seguir será considerado o esquema de axioma **4**:

$$\text{4: } \Box\alpha \supset \Box\Box\alpha.$$

Definição 2.20 O sistema obtido agregando ao sistema **T** o esquema de axioma **4**, é o sistema **S4**.

O sistema **S4** corresponde à classe de modelos onde a relação de acessibilidade é reflexiva e transitiva. Este sistema é importante quando a relação de acessibilidade é utilizada para representar uma *ordem* entre os mundos possíveis.

Dado um modelo modal no qual a relação de acessibilidade *R* é transitiva, um *grupo* ("cluster") é qualquer subconjunto *U* de *W* maximal⁶, tal que todos os elementos de *U* são mutuamente acessíveis⁷, ou seja, *R* restringida a *U* é uma relação de equivalência maximal. Assim, qualquer relação de acessibilidade reflexiva e transitiva, num modelo modal pode ser vista como uma ordem parcial no conjunto de grupos desse modelo. Com essa interpretação de *R*, os grupos são os conjuntos maximais de mundos cujos membros têm todos a mesma *classificação*, segundo a relação de ordem parcial entre grupos.

O Sistema S5

Um outro sistema é o sistema **S5**.

⁶Com relação à inclusão de conjuntos.

⁷Dois mundos *u*, *v* ∈ *U* são mutuamente acessíveis se *uRv* e *vRu*.

Definição 2.21 O sistema **S5** é obtido a partir de **S4** agregando o esquema de axioma:

$$5 : \neg \Box \alpha \supset \Box \neg \Box \alpha.$$

Este sistema é caracterizado pelas classes de modelos onde R é uma relação de equivalência (reflexiva, simétrica e transitiva).

Os Sistemas **S4.5** e **KD4.5**

O próximo esquema de axioma considerado é:

$$D : \Box(\Box \alpha \supset \beta) \vee \Box(\Box \beta \supset \alpha)$$

Definição 2.22 O sistema obtido agregando ao sistema **S4** o axioma **D** é o sistema **S4.5**.

No sistema **S4.5** é exigido que a relação R seja, embora reflexiva e transitiva, totalmente conexa⁸. Os modelos então, nesta classe, consistem de conjuntos de grupos de mundos totalmente ordenados.

Definição 2.23 O sistema obtido agregando ao sistema **K** os esquemas de axioma 4, 5 e **D** é o sistema **KD45**.

O sistema **KD45** também chamado de **S5 fraco**.

Teorema 2.1 (Completeness e Correção dos Sistemas Modais)

Seja $A = \{Ax_1, Ax_2, \dots\}$ o conjunto de axiomas de um sistema modal **S** e C a classe de todos os modelos modais de A , então **S** é correta e completa com relação a C .

Além dos axiomas considerados, existem outros possíveis axiomas para definir o comportamento do operador modal $\Box$, cada um dos quais impõe suas próprias restrições na relação de acessibilidade. Lewis [Lew73] mostrou que existem 42 configurações diferentes de tais restrições, por exemplo, reflexiva e simétrica mas não transitiva, reflexiva e transitiva mas não simétrica, e assim por diante.

⁸Uma relação binária R definida num conjunto W é *totalmente conexa* quando para todo $u, v \in W$ uRv ou vRu .

2.3.3 Aplicações da Lógica Modal

Embora os conectivos $\Box$ e $\Diamond$ tenham sido introduzidos para formalizar as idéias de necessidade e possibilidade, a importância da lógica modal para a IA está relacionada ao fato de que podem ser consideradas interpretações alternativas desses operadores. Pode-se atribuir qualquer significado intuitivo aos mundos e aos operadores, que seja consistente com a definição formal dada.

Um uso comum da lógica modal é como lógica do tempo. Nesta interpretação os mundos possíveis são interpretados como pontos no tempo, e a relação de acessibilidade como a precedência temporal. Assim, t_2 é acessível a partir de t_1 se $t_1 \leq t_2$.

No contexto da IA, a principal aplicação da lógica modal é como lógica do conhecimento. Esta aplicação tem sua origem no trabalho de Hintikka [Hin62] e foi usado por Moore para raciocinar sobre agentes inteligentes [Moo80] e mais recentemente por Halpern, Moses e outros, para raciocinar sobre computação distribuída [Hal87]. No contexto particular da computação distribuída, os mundos podem representar a história global da computação ou a sequência de estados de cada processador.

Como já foi dito, um agente num sistema inteligente pode necessitar raciocinar sobre mais do que o estado atual do mundo. Pode precisar raciocinar sobre outros possíveis estados, sobre o próprio conhecimento ou crenças e o conhecimento e crenças de outros agentes.

Então como é necessária a habilidade de raciocinar sobre o próprio conhecimento, parece ser adequado analisar se algum dos sistemas lógicos conhecidos provê um bom modelo para esse tipo de raciocínio. A lógica modal é um modelo apto, já que os mundos possíveis podem ser interpretados como os mundos alternativos compatíveis com o que o agente conhece ou acredita. Assim, um agente conhece ou acredita numa determinada proposição α , num determinado mundo, se α é verdadeira em todos os mundos que são acessíveis a partir desse mundo, ou seja, quando nesse mundo se satisfaz $\Box\alpha$. Nesta interpretação da lógica modal o símbolo $\Box$ é substituído, em geral, pelo símbolo K (em inglês "Knowledge"). Neste contexto a fórmula $K\alpha$ é lida como "o agente conhece α ".

Se o operador K vai capturar as idéias intuitivas sobre conhecimento deve satisfazer várias propriedades "ideais" geralmente atribuídas ao conhecimento:

1. Se o agente conhece α e conhece $\alpha \supset \beta$, então conhece β . Em símbolos $K(\alpha \supset \beta) \supset (K\alpha \supset K\beta)$ (Axioma **K**).

2. Não podem ser conhecidas coisas que não sejam verdadeiras. Ou seja $K\alpha \supset \alpha$ (Axioma **T**).
3. Se o agente conhece α então conhece que conhece α . Em símbolos $K\alpha \supset K K \alpha$ (Axioma **4**, denominado neste contexto axioma de *introspecção positiva*).
4. Também pode-se pedir que se o agente não conhece algo, então o agente conhece que não conhece. Em símbolos $\neg K\alpha \supset K \neg K\alpha$ (Axioma de *introspecção negativa*).

Várias combinações dos esquemas de axiomas 1-4 junto com os esquemas de axiomas da LPC e as regras **MP** e **Nec** parecem ser adequadas para representar conhecimento, já que permitem capturar as propriedades mencionadas acima. Estas combinações constituem as chamadas de *lógicas modais de conhecimento* ou *lógicas modais epistêmicas*. Mas, geralmente, a lógica associada ao conhecimento é a lógica **S5**.

No caso de ter que representar as crenças de um agente, o fato de que um agente acredite em α , em geral, é representado como $B\alpha$ (em inglês “Belief”), onde B é o operador que substitui o operador modal $\Box$. Alguns pesquisadores consideram que a diferença essencial entre conhecimento e crenças é que o conhecimento são crenças justificadamente verdadeiras, ou seja, que um agente pode ter crenças que são falsas. Então uma lógica para modelar crenças vai se diferenciar de uma lógica que modela conhecimento pelo fato de que não será válido o axioma **T**. Nas lógicas de crenças, em geral, é agregado o axioma **D** que garante que as crenças de um agente são consistentes. Neste caso, a lógica associada, geralmente, as crenças é a lógica **KD45**.

Tanto os sistemas que permitem representar conhecimento, quanto os sistemas que permitem representar crenças, impõem omnisciência lógica dos agentes, ou seja, se um agente conhece (resp. acredita) num conjunto de fatos, então ele conhece (resp. acredita) suas conseqüências lógicas. Algumas tentativas para evitar esta propriedade podem ser vistos em [Kon86, FH88].

Uma outra aplicação possível da lógica modal é para raciocinar sobre conhecimento e ação [Moo80].

As lógicas modais também podem ser generalizadas com a inclusão de mais de um operador modal primitivo. Lógicas multimodais com um número finito de $\Box$ (ou K ou B) são usadas para representar o conhecimento ou crenças de múltiplos agentes [HM85, GD93].

2.4 Lógicas Condicionais

A *lógica condicional*, que surgiu inicialmente para representar expressões condicionais da linguagem natural que não podiam ser representadas aplicando o condicional material⁹, também pode ser de utilidade para aplicações onde é preciso certa capacidade de raciocínio default. As propriedades desta lógica fazem com que ela seja apropriada não só para expressar sentenças do tipo “Se.... então normalmente....”, ou seja, para expressar sentenças gerais que admitem exceções, mas também para resolver certos problemas, presentes em outros formalismos tais como a Lógica Default, causados, principalmente, pela interação entre defaults. Surgiram, assim, as interpretações condicionais das expressões default.

As diferentes propostas da lógica condicional estão baseadas, na sua maioria, numa linguagem proposicional clássica ou numa linguagem de primeira ordem aumentada com um *conectivo condicional*, denotado geralmente por “ $\Rightarrow$ ”. A linguagem provê uma forma natural e econômica para expressar o fato de que um determinado β segue normalmente de um determinado α , simplesmente basta representar $\alpha \Rightarrow \beta$. Neste formalismo as sentenças condicionais podem ser interpretadas como regras default. As regras default são afirmações acerca do mundo que se está modelando, representadas na própria linguagem, permitindo então a interação com outras regras default ou outros fatos. É por isso que a lógica condicional oferece algumas vantagens com relação à formalização de alguns aspectos do raciocínio default.

Apesar dessas vantagens mencionadas, as interpretações condicionais não têm algumas das propriedades desejáveis do raciocínio default que estão presentes nas interpretações extensionais. A ausência de determinadas propriedades surge fundamentalmente pela forma como é tratada a informação irrelevante nas interpretações condicionais. Por exemplo, dado o default $\alpha \Rightarrow \beta$, β será deduzido por default na presença de α ; mas no enfoque condicional a partir do default $\alpha \Rightarrow \beta$, β não pode ser deduzido na presença de α e γ , onde γ é informação adicional que é irrelevante. Isto acontece porque as interpretações condicionais são “conservadoras” demais, elas consideram toda evidência como relevante, a menos que seja demonstrado o contrário.

Mas apesar destes problemas, em geral, como é sugerido em [Pea89], as interpretações condicionais têm um conjunto de propriedades que constitui um núcleo mínimo que deveria estar presente em qualquer sistema de raciocínio default.

Esse conjunto mínimo de propriedades está relacionado com a capacidade dos sistemas condicionais de:

⁹Por exemplo, expressões do tipo “se você tivesse assistido ao filme, você teria gostado”.

- representar naturalmente regras default,
- não ter o problema de múltiplas extensões,
- tratar satisfatoriamente com a especificidade dos defaults.

De entre as propostas que surgiram, baseadas na lógica condicional, podem ser mencionadas as de Delgrande [Del87, Del88], Boutilier [Bou92, Bou94b] e Lamarre [Lam91], entre outras.

A seguir serão apresentadas as proposta de Delgrande e a família de lógicas condicionais de normalidade introduzida por Boutilier¹⁰ que será utilizada como base para as lógicas desenvolvidas no capítulo 4.

2.4.1 O Enfoque de Delgrande

Um dos sistemas propostos para formalizar raciocínio default baseado num condicional para representar as regras default é o de Delgrande[Del88], que desenvolve uma semântica de mundos possíveis e os axiomas apropriados para dotar ao operador $\Rightarrow$ das propriedades necessárias para realizar raciocínio default.

Na maioria dos enfoques apresentados até agora é impossível raciocinar adequadamente sobre defaults. Por exemplo, a partir de “todos os A são B” deveria-se deduzir que “os A não são tipicamente B” é falso. No entanto, em todas as lógicas vistas até aqui, ambas afirmações podem ser feitas simultaneamente de maneira consistente.

Para evitar este problema, Delgrande diferencia raciocinar sobre defaults de raciocinar com defaults. O primeiro tem a ver com a dedução (monotônica) de defaults a partir de outros defaults, e o segundo com a dedução (não-monotônica) de conclusões plausíveis a partir de defaults. Ele introduz a lógica condicional N, que permite expressar e raciocinar acerca de defaults. Logo, os defaults deduzidos na lógica são usados para realizar inferências não-monotônicas.

A lógica N

A linguagem $\mathcal{L}$ da lógica N é a da lógica de primeira ordem, aumentada com um conectivo binário “ $\Rightarrow$ ”. O significado intuitivo de uma fórmula do tipo $\alpha \Rightarrow \beta$ é a usual das

¹⁰As lógicas propostas por Delgrande e Boutilier, embora possuam uma linguagem modal, foram desenvolvidas utilizando um enfoque condicional.

interpretações condicionais para raciocínio default. Com esta linguagem, pode-se então representar sentenças do tipo “os pássaros normalmente voam” ou “se chove, em geral, o jogo de futebol é suspenso”.

O conjunto de fórmulas bem formadas (fbf) é especificado da maneira usual, só que a linguagem da lógica N exclui fórmulas com ocorrências aninhadas do *conectivo condicional* ($\Rightarrow$).

A semântica desta lógica é uma semântica de mundos possíveis onde “ $\alpha \Rightarrow \beta$ ” é verdadeira quando β é verdadeira nos mundos menos excepcionais onde α é verdadeira. A determinação de quais mundos são mais ou menos excepcionais é feita em função da relação de acessibilidade E entre mundos. Assim, dados dois mundos possíveis w_1 e w_2 , a relação $w_1 E w_2$ se verifica se o mundo w_1 é no mínimo tão excepcional quanto w_2 ou w_2 é tão uniforme quanto w_1 .

As fórmulas da linguagem $\mathcal{L}$ são interpretadas em termos de um modelo $M = \langle W, E, DI, V \rangle$, onde:

- W é um conjunto de mundos, o conjunto de mundos possíveis.
- E , a relação de acessibilidade entre mundos possíveis, é uma relação definida no conjunto W reflexiva, transitiva e tal que se $w_1 E w_2$ e $w_1 E w_3$, então $w_2 E w_3$ ou $w_3 E w_2$.
- DI é o domínio dos indivíduos.
- V é uma função definida no conjunto de termos e símbolos de predicado tal que:
 1. para todo termo t , $V(t) \in DI$.
 2. para um predicado n -ário P , $V(P)$ é um conjunto de $(n+1)$ -tuplas $\langle t_1, \dots, t_n, w \rangle$, onde cada $t_i \in DI$ e $w \in W$.

V mapeia um símbolo de predicado em relações em mundos.

No contexto da lógica N, a fórmula α é identificada com o conjunto $\|\alpha\|^M$ de mundos onde ela é verdadeira.

Com a finalidade de definir a interpretação de uma fórmula é introduzida uma *função de seleção de mundos* f que, dado um mundo w e o conjunto de mundos identificados com a fórmula α , retorna o conjunto de mundos menos excepcionais onde α é verdadeira. Formalmente:

Definição 2.24 $f(w, \|\alpha\|^M) = \{w_1 : wEw_1 \text{ e } M, w_1 \models \alpha \text{ e para todo } w_2 \text{ tal que } w_1Ew_2 \text{ e } M, w_2 \models \alpha, \text{ tem-se } w_2Ew_1\}.$

Dado um modelo $M = \langle W, E, DI, V \rangle$, a verdade de uma fórmula num mundo do modelo é definida como:

Definição 2.25

- (i) Para um símbolo de predicado n -ário P , termos $t_1, \dots, t_n$ e $w \in W$, $M, w \models P(t_1, \dots, t_n)$ se e somente se $\langle V(t_1), \dots, V(t_n), w \rangle \in V(P)$.
- (ii) $M, w \models \neg\alpha$ se e somente se $M \not\models \alpha$.
- (iii) $M, w \models \alpha \supset \beta$ se e somente se $M, w \not\models \alpha$ ou $M, w \models \beta$.
- (iv) $M, w \models \alpha \Rightarrow \beta$ se e somente se $f(w, \|\alpha\|^M) \subseteq \|\beta\|^M$.
- (v) $M, w \models (\forall x)\alpha$ se e somente se para todo V' que coincide com V em todos os valores atribuídos as variáveis, exceto possivelmente no valor atribuído a x , tem-se que $M', w \models \alpha$ onde $M' = \langle W, E, DI, V' \rangle$.

Do ponto de vista sintático, a lógica condicional **N** é o menor conjunto de fórmulas de $\mathcal{L}$ que contém a lógica de primeira ordem clássica e é fechado segundo os esquemas de axiomas **N1** – **N6** e a regra de inferência **R_N**:

- N1.** $\alpha \Rightarrow \alpha$
- N2.** $((\alpha \Rightarrow \beta) \wedge (\alpha \Rightarrow \gamma)) \supset (\alpha \Rightarrow (\beta \wedge \gamma))$
- N3.** $(\alpha \Rightarrow \beta) \supset (((\alpha \wedge \beta) \Rightarrow \gamma) \supset (\alpha \Rightarrow \gamma))$
- N4.** $\neg(\alpha \Rightarrow \beta) \supset ((\alpha \Rightarrow \gamma) \supset ((\alpha \wedge \neg\beta) \Rightarrow \gamma))$
- N5.** $((\alpha \Rightarrow \gamma) \wedge (\beta \Rightarrow \gamma)) \supset ((\alpha \vee \beta) \Rightarrow \gamma)$
- N6.** $(\forall x)(\alpha \Rightarrow \beta) \supset (\alpha \Rightarrow (\forall x)\beta)$ se x não está livre em α .
- R_N.** A partir de $\beta \supset \gamma$ inferir $(\alpha \Rightarrow \beta) \supset (\alpha \Rightarrow \gamma)$.

As noções de teorema, dedução e consistência são definidas da maneira usual. O fato de que uma fórmula α é dedutível a partir de um conjunto de fórmulas Γ na lógica **N** é simbolizado por $\Gamma \vdash_N \alpha$.

Na lógica assim definida o seguinte resultado é válido:

Teorema 2.2 $\models \alpha$ se e somente se $\vdash_N \alpha$.

Inferência Não-monotônica na Lógica **N**

O objetivo de Delgrande no seu trabalho [Del88] é definir um operador default $\rightsquigarrow$ onde, a partir de uma teoria T , $T \rightsquigarrow \alpha$ é interpretado como α pode ser concluído por default

a partir de T , utilizando a lógica N para representar os defaults e considerações extra-lógicas para fazer inferências default. Embora a lógica N seja monotônica, Delgrande apresenta dois enfoques equivalentes baseados nela que permitem realizar inferências não-monotônicas.

O ponto de partida para ambos enfoques é uma teoria $T = (D, C)$, onde D é um conjunto de fórmulas da lógica N e C é um conjunto consistente de fórmulas de uma linguagem de primeira ordem. D representa relações gerais entre objetos, leis e expressões do tipo “os mamíferos têm pulmões” ou “os pássaros tipicamente voam”, ou seja, um conjunto de fórmulas necessárias, defaults ou negações de defaults. C contém os fatos contingentes que restringem o mundo que está sendo modelado. Nos dois enfoques propostos, ou C é aumentado segundo regras específicas que dependem de D ou D é aumentado segundo regras que dependem de C . As hipóteses feitas para aumentar alternativamente C ou D são as seguintes:

Hipótese de Normalidade: o mundo que se está modelando é um dos menos excepcionais, segundo D , onde as fórmulas de C são verdadeiras.

Hipótese de Relevância: só se supõe que têm influência na validade de um condicional aquelas expressões das quais essa influência é conhecida.

Baseados nestas hipóteses, são propostos dois enfoques onde a idéia subjacente é a de só considerar um subconjunto dos modelos de uma teoria T para realizar inferências default.

Primeiro Enfoque

O primeiro enfoque consiste em aumentar D não-monotonicamente. Assim, por exemplo, se temos

$$\text{mamífero} \Rightarrow \text{vive_terra},$$

segundo a hipótese de relevância, se nenhuma coisa indica que a propriedade de ser *marrom* é relevante para a propriedade dos mamíferos de viver na terra, então

$$\text{mamífero} \wedge \text{marrom} \Rightarrow \text{vive_terra}$$

pode ser agregado a D .

Os condicionais agregados a D contêm pré-condições adicionais que, segundo a hipótese mencionada, não são relevantes.

Formalmente, isto é feito segundo a seguinte definição:

Definição 2.26 *Seja Γ um conjunto de fórmulas de N . $\alpha \Rightarrow \gamma$ tem suporte em Γ se existe uma fórmula β tal que:*

- (1) $\vdash_{LPO} \alpha \supset \beta$,
- (2) $\Gamma \vdash_N \beta \Rightarrow \gamma$,
- (3) *se existe uma fórmula β' tal que $\vdash_{LPO} \alpha \supset \beta'$ e $\Gamma \vdash_N \neg(\beta' \Rightarrow \gamma)$ então $\vdash_{LPO} \beta \supset \beta'$.*

A idéia intuitiva desta definição é a de que, $\alpha \Rightarrow \gamma$ está suportado por Γ , se existe um condicional $\beta \Rightarrow \gamma$ onde β se deduz de α e não há nenhum outro condicional em Γ com um antecedente mais forte que negue γ .

Em função desta idéia pode ser definido um procedimento para obter uma extensão de um conjunto de defaults que leve em consideração todas as propriedades irrelevantes para esses defaults:

Definição 2.27 *Seja (D, C) uma teoria default e $\beta_1, \beta_2, \dots$ uma ordenação das fórmulas da linguagem de primeira ordem a partir da qual é formada a linguagem $\mathcal{L}$. Uma extensão default maximal $E(D_C)$ de D é definida por:*

$$\begin{aligned}
 E_0 &= D, \\
 E_{i+1} &= E_i \cup \\
 &\quad \{(\alpha \wedge \beta_i) \Rightarrow \gamma : D \vdash_N (\alpha \Rightarrow \gamma) \text{ e } (\alpha \wedge \beta_i) \Rightarrow \gamma \text{ é suportado por } C \cup D\} \cup \\
 &\quad \{(\alpha \wedge \neg \beta_i) \Rightarrow \gamma : D \vdash_N (\alpha \Rightarrow \gamma) \text{ e } (\alpha \wedge \beta_i) \Rightarrow \gamma \text{ não é suportado por } C \cup D\}, \\
 E(D_C) &= \bigcup_{i=0}^{\infty} E_i.
 \end{aligned}$$

Independentemente da ordem das fórmulas β_i , em cada caso, só uma extensão é gerada e $E(D_C)$ é consistente se D for consistente.

Os teoremas de uma teoria default são definidos como:

Definição 2.28 *Seja $T = (D, C)$ uma teoria default. Uma fórmula α é dedutível a partir de T se $E(D_C) \vdash_N C \Rightarrow \alpha$ (denotado por $T \leadsto \alpha$).*

Segundo Enfoque

No outro enfoque, em lugar de agregar suposições a D para limitar os modelos considerados ao realizar inferências default, supõe-se que o mundo que está sendo modelado é um dos mais simples e consistente com os fatos e os defaults. O que se deseja é agregar

informação consistente a C . Assim, a fórmula $\alpha \supset \gamma$ será agregada a C se ela é *contingentemente suportada*. Várias extensões podem resultar do fato de agregar informação adicional a C . Neste caso, uma fórmula vai ser deduzida por default de T se ela é dedutível (da maneira usual na lógica de primeira ordem) de todas as extensões de T . Isto pode ser formalizado como segue:

Definição 2.29 $\alpha \supset \gamma$ está suportado contingentemente numa teoria default $T = (D, C)$ se e somente se

1. $D \vdash_N \alpha \Rightarrow \gamma$,
2. $C \cup D \cup \{\alpha \supset \gamma\}$ é consistente,
3. se existe α' tal que $\vdash_{LPO} C \supset \alpha'$ e $C \cup D \vdash_N \neg(\alpha' \Rightarrow \gamma)$, então $\vdash_{LPO} \alpha \supset \alpha'$.

Isto significa que $\alpha \supset \gamma$ é agregado a C se $C \cup D \cup \{\alpha \supset \gamma\}$ é consistente e, para qualquer fórmula α' que é implicada por C e que tem conflito com γ então, α' é implicada por α . Assim, como no primeiro enfoque, só os condicionais mais fortes “razoáveis” são contingentemente suportados.

Baseada nesta última definição é introduzida a noção de extensão maximal consistente:

Definição 2.30 Seja $T = (D, C)$ uma teoria default. Então, E é uma extensão maximal contingente de T se e somente se existe uma seqüência $C_0, C_1, \dots$ tal que:

1. $C_0 = C$,
2. $C_{i+1} = C_i \cup \{\alpha \supset \gamma\}$, onde $\alpha \supset \gamma$ é contingentemente suportado em (D, C_i) ,
3. $E = \bigcup_{i=0}^{\infty} C_i$.

Definição 2.31 Seja $T = (D, C)$ uma teoria default. Seja EXT o conjunto de extensões maximais contingentes de T . Então α é dedutível a partir de T (denotado por $T \rightsquigarrow' \alpha$) se e somente se $\alpha \in Th(E)$ para todas as extensões $E \in EXT$.

Os dois enfoques propostos são equivalentes:

Teorema 2.3 $T \rightsquigarrow \alpha$ se e somente se $T \rightsquigarrow' \alpha$.

Exemplos

No artigo [Del88], Delgrande mostra que muitos dos exemplos típicos de raciocínio default são tratados adequadamente na lógica por ele proposta.

Defaults conflitantes e transitividade são tratados de maneira razoável.

Diferentemente do enfoque de Reiter e McCarthy, não é necessário expressar explicitamente anormalidades para predicados ou agregar outros defaults para restringir as possíveis interações entre os defaults.

Assim, nesta proposta defaults mais específicos prevalecem sobre os mais gerais. Por exemplo, numa teoria $T = (D, C)$, onde D e C estão definidos como:

$$\begin{aligned} D &= \{mamífero(x) \Rightarrow vive_terra(x), orca(x) \Rightarrow \neg vive_terra(x), orca(x) \Rightarrow mamífero(x)\} \\ C &= \{orca(a)\}, \end{aligned}$$

então pode-se deduzir por default a partir de T que $\neg vive_terra(a)$.

Com relação à transitividade, se na teoria $T = (D, E)$, D e C estão definidos como:

$$\begin{aligned} D &= \{orca(x) \Rightarrow mamífero(x), mamífero(x) \Rightarrow tem_pulmões(x)\} \text{ e} \\ C &= \{orca(a)\}, \end{aligned}$$

então pode-se concluir por default $tem_pulmões(a)$.

2.4.2 Lógicas Condicionais de Normalidade

Em [Bou92, Bou94b], Boutilier apresenta uma família de lógicas condicionais que permitem expressar diferentes formas de raciocínio tentativo, tais como raciocínio default, revisão de crenças, raciocínio autoepistêmico e subjuntivo entre outros.

Nesta família de lógicas, as regras default tomam a forma de um condicional $\alpha \Rightarrow \beta$, e são interpretadas da seguinte maneira: nas situações mais normais nas quais α se verifica, β também se verifica ou α normalmente implica β . Assim, as expressões default podem ser interpretadas como expressões de normalidade ou prototípicas, ou como generalizações que admitem exceções como em “os brasileiros bebem cerveja”.

Nas seções seguintes vão ser apresentadas as diferentes componentes da família de lógicas condicionais apresentadas por Boutilier, começando pela mais simples de todas que é a lógica CT40.

A Lógica CT4O

A linguagem desta lógica é uma linguagem bimodal¹¹ $\mathcal{L}_C$ e a semântica está baseada nas estruturas de mundos possíveis, onde a relação de acessibilidade $\geq$ captura a idéia de normalidade, assumindo que o conjunto de mundos possíveis W está ordenado, segundo a relação $\geq$, de acordo com uma certa medida de normalidade.

Mas, que significa dizer que um mundo seja mais normal que outro? Por que, por exemplo, um mundo onde todos os brasileiros bebem cerveja é mais normal que um onde alguns brasileiros não bebem? Na realidade não existe nenhuma propriedade intrínseca dos mundos que faz com que uns sejam mais normais que outros. A ordenação entre os mundos não depende de relações existentes a priori entre eles, mas sim das crenças de cada agente: é uma ordenação puramente subjetiva que o agente faz considerando certos aspectos que ele considera mais relevantes.

Os mundos mais normais serão os menores de acordo com a relação de acessibilidade. Assim, para um agente particular os mundos mais normais são aqueles que não violam nenhuma de suas expectativas. Se um agente não tem expectativa nenhuma, todos os mundos são igualmente normais. Se ele tem como única expectativa que “os brasileiros normalmente bebem cerveja”, então qualquer mundo onde alguns brasileiros não bebem cerveja será menos normal que um mundo onde todos bebem. Mais expectativas impõem mais estrutura na ordenação que um agente considera possível.

A interpretação da relação $\geq$ é a seguinte: $w \geq v$ se v é no mínimo tão normal quanto w . Se $w \geq v$ e $v \not\geq w$, então v é menos excepcional ou mais normal que w (denotado por $w > v$).

Reflexividade (todo mundo é tão normal quanto si mesmo) e transitividade serão os requisitos mínimos da relação $\geq$ da lógica CT4O para capturar coerentemente a noção de normalidade. Mais restrições sobre esta relação vão definir outras lógicas da família.

Será definida a seguir a lógica condicional CT4O, sua linguagem, a semântica associada e seus axiomas e regras de inferência.

A linguagem de CT4O, $\mathcal{L}_C$, é a linguagem da lógica proposicional clássica (LPC) formada a partir de um conjunto de variáveis proposicionais $P = \{p_0, p_1, \dots\}$ e dos conectivos $\neg$ (negação), $\supset$ (implicação material) ampliado com os conectivos unários $\Box$ (necessidade) e $\Box$ (inacessibilidade).

¹¹Neste caso a linguagem tem dois operadores modais primitivos.

Definição 2.32 As fórmulas bem formadas (fbf) de $\mathcal{L}_c$ são o menor conjunto tal que:

1. Se $\alpha \in \mathbf{P}$, então α é uma fbf (fórmula bem formada atômica).
2. Se α e β são fbf então $\neg(\alpha)$, $(\alpha \supset \beta)$, $\Box(\alpha)$ e $\Box(\alpha)$ são fbf.

Conjunção($\wedge$), disjunção($\vee$), bicondicional($\leftrightarrow$) e possibilidade ($\Diamond$) estão definidos da maneira usual [HC68].

A semântica da lógica bimodal está baseada nas mesmas estruturas usadas para as lógicas monomodais.

Definição 2.33 Um CT4O-modelo é um par $M = \langle W, \geq \rangle$, onde W é um conjunto de mundos possíveis (subconjuntos de $\mathbf{P}$), $\geq$ (a relação de normalidade) é uma relação binária reflexiva e transitiva definida em W .

Definição 2.34 Seja $M = \langle W, \geq \rangle$ um CT4O-modelo. A validade de uma fórmula α no mundo w do modelo M (denotado por $M, w \models \alpha$) é definida como:

1. $M, w \models \alpha$ se e somente se $\alpha \in w$, para α fórmula atômica.
2. $M, w \models \neg\alpha$ se e somente se $M, w \not\models \alpha$.
3. $M, w \models \alpha \supset \beta$ se e somente se $M, w \not\models \alpha$ ou $M, w \models \beta$.
4. $M, w \models \Box\alpha$ se e somente se $M, v \models \alpha$, para todo v tal que $w \geq v$.
5. $M, w \models \Box(\alpha)$ se e somente se $M, v \models \alpha$, para todo v tal que $w \not\geq v$.

Assim, a fórmula $\Box\alpha$ será interpretada como “ α se verifica em todos os mundos tão ou mais normais” e $\Box(\alpha)$ será interpretada como “ α se verifica em todos os mundos menos normais ou inacessíveis”.

Definição 2.35 Para qualquer CT4O-modelo $M = \langle W, \geq \rangle$, α é válida em M (denotado por $M \models \alpha$) se e somente se $M, w \models \alpha$, para cada $w \in W$. Uma fórmula α é CT4O-válida (denotado por $\models_{CT4O} \alpha$) quando $M \models \alpha$ para todo CT4O-modelo M . Uma fórmula se satisfaz se $\neg\alpha$ não é válida.

A seguir serão definidos novos conectivos, próprios de $\mathcal{L}_c$

Definição 2.36

1. $\Diamond \alpha \equiv_{def} \neg \Box \neg \alpha.$
2. $\Box \alpha \equiv_{def} \Box \alpha \wedge \Box \alpha.$
3. $\Diamond \alpha \equiv_{def} \Diamond \alpha \vee \Diamond \alpha.$

No contexto da lógica condicional de normalidade, a fórmula $\Diamond \alpha$ é interpretada como “ α é verdadeira em algum mundo menos normal ou incomparável”, $\Box \alpha$ como “ α é verdadeira em todos os mundos mais normais ou não” e $\Diamond \alpha$ se interpreta como “ α é verdadeira em algum mundo mais ou menos normal”.

Em função da definição anterior, tem-se o seguinte resultado [?]:

Teorema 2.4

1. $M, w \models \Diamond \alpha$ se para algum $v \in W$ tal que $w \not\geq v$, $M, v \models \alpha$.
2. $M, w \models \Box \alpha$ se para todo $v \in W$, $M, v \models \alpha$.
3. $M, w \models \Diamond \alpha$ se para algum $v \in W$, $M, v \models \alpha$.

Uma vez definida a semântica para a lógica CT4O, e com o objetivo de fornecer uma semântica para defaults condicionais baseada numa ordenação de mundos possíveis segundo o grau de normalidade, resta agora definir o condicional normativo $\alpha \Rightarrow \beta$.

Intuitivamente, pode ser dito que o condicional $\alpha \Rightarrow \beta$ é verdadeiro se nos mundos “mais normais” (os mundos minimais segundo a relação $\geq$) onde α é verdadeira β também é verdadeira. Mas nada nas estruturas escolhidas garante a existência de tais mundos minimais onde α seja verdadeira, já que poderia existir uma seqüência infinita de mundos mais e mais normais onde α seja verdadeira. Algumas lógicas proposta, como a de Kraus, Lehmann e Magidor [KLM90] ou a de Shoham [Sho88] consideram só aqueles modelos onde sempre existem os mundos mais normais para qualquer α .

Mas outras condições de verdade podem ser atribuídas a $\alpha \Rightarrow \beta$, seguindo a idéia intuitiva, de maneira a não ter que supor a existência de tais mundos minimais:

- Se existem mundos minimais onde α for verdadeira, então $\alpha \Rightarrow \beta$ será verdadeira quando β é verdadeira nesses mundos.

- Se existe uma cadeia infinita de mundos mais e mais normais onde α é verdadeira, e se algum mundo dessa cadeia é tal que α e β são verdadeiras nele, e em todos os mundos mais normais que ele β é verdadeira sempre que α é verdadeira, então $\alpha \Rightarrow \beta$ pode ser considerada verdadeira. Neste caso, embora não existam mundos minimais onde α seja verdadeira, β é verdadeira na “limite” dessa cadeia de mundos, ou seja, se $\alpha \Rightarrow \beta$ é verdadeira, então se w é um mundo onde α é verdadeira, então existe um mundo v mais normal do que w ($w \geq v$) tal que $v \models \alpha$, e $u \models \alpha \supset \beta$, para todo mundo u tal que $v \geq u$. Expressadas essas condições de verdade na linguagem bimodal, um primeiro intento de definição de $\alpha \Rightarrow \beta$ poderia ser $\Box \Diamond (\alpha \wedge \Box (\alpha \supset \beta))$.

Mas esta tentativa de definição não considera os modelos onde, num mundo dado, é verdadeira $\Box \neg \alpha$. Neste caso, $\alpha \Rightarrow \beta$ vai ser considerada verdadeira, independentemente do que acontece com β .

Então a definição formal do *condicional normativo* é a seguinte:

Definição 2.37 $(\alpha \Rightarrow \beta) \equiv_{def} \Box (\Box \neg \alpha \vee \Diamond (\alpha \wedge \Box (\alpha \supset \beta)))$.

Deve ser assinalado que a validade do condicional não depende dos mundos em que é avaliado, mas da completa ordenação dos mundos. Esta visão é consistente com a interpretação da ordem de normalidade, já que tal ordenação é uma classificação arbitrária de situações, feitas por um agente segundo sua visão. Tal ordenação deve ser considerada como parte do estado de crenças do agente. Um mundo particular só tem influência indireta na avaliação de um condicional pelo fato de contribuir ao estado global de crenças do agente.

Também é introduzida a seguinte abreviatura:

Definição 2.38 $\alpha \not\Rightarrow \beta \equiv_{def} \neg (\alpha \Rightarrow \beta)$

A seguir serão introduzidos os axiomas e regras de inferência que caracterizam a lógica CT4O:

Definição 2.39 A lógica condicional CT4O é o menor conjunto $S \subseteq \mathcal{L}_C$ tal que S contém a lógica proposicional clássica e os esquemas de axiomas **K**, **K'**, **T**, **4**, e **H**, e é fechado segundo as regras de inferência **MP** e **Nec.**:

K: $\Box (\alpha \supset \beta) \supset (\Box \alpha \supset \Box \beta)$.

K': $\Box (\alpha \supset \beta) \supset (\Box \alpha \supset \Box \beta)$.

T: $\Box \alpha \supset \alpha$.

4: $\Box \alpha \supset \Box \Box \alpha$.

H: $\Diamond (\Box \alpha \wedge \Box \beta) \supset (\Box (\alpha \vee \beta))$.

MP: A partir de $\alpha \supset \beta$ e α inferir β .

Nec.: A partir de α inferir $\Box \alpha$.

Definição 2.40 A fórmula α pode ser demonstrada em $CT4O$ (denotado por $\vdash_{CT4O} \alpha$) se e somente se $\alpha \in CT4O$. α é dedutível a partir de um conjunto $\Gamma \subseteq \mathcal{L}_c$ (denotado por $\Gamma \vdash_{CT4O} \alpha$) se existe um subconjunto finito $\{\alpha_1, \dots, \alpha_n\}$ de Γ , tal que $\vdash_{CT4O} (\alpha_1 \wedge \dots \wedge \alpha_n) \supset \alpha$.

Teorema 2.5 O sistema $CT4O$ é caracterizado pela classe de $CT4O$ -modelos; ou seja, $\vdash_{CT4O} \alpha$ se e somente se $\models_{CT4O} \alpha$.

Propriedades da Lógica $CT4O$. Raciocínio default em $CT4O$

Pode-se considerar a lógica $CT4O$ um sistema “minimal” no qual realiza-se raciocínio default. Aqui o qualificativo “minimal” está relacionado, como já foi mencionado, com o fato de que ela possui um conjunto de propriedades que pode ser considerado mínimo em relação às propriedades que deveriam estar presente em qualquer sistema de inferência não-monotônica.

Esta lógica é apropriada para representar enunciados sobre normalidade e raciocinar sobre generalizações que admitem exceções, pois permite tratar satisfatoriamente a especificidade dos defaults.

Assim, nesta lógica, os seguintes resultados são válidos [?]:

Teorema 2.6 Em $CT4O$ são válidas as seguintes fórmulas:

ID: $\alpha \Rightarrow \alpha$.

LLE: $\Box (\alpha \equiv \beta) \supset ((\alpha \Rightarrow \gamma) \equiv (\beta \Rightarrow \gamma))$.

And: $((\alpha \Rightarrow \beta) \wedge (\alpha \Rightarrow \gamma)) \supset (\alpha \Rightarrow \beta \wedge \gamma)$.

RT: $(\alpha \Rightarrow \beta) \supset ((\alpha \wedge \beta \Rightarrow \gamma) \supset (\alpha \Rightarrow \gamma))$.

Or: $((\alpha \Rightarrow \gamma) \wedge (\beta \Rightarrow \gamma)) \supset (\alpha \vee \beta \Rightarrow \gamma)$.

RCM: $\Box (\beta \supset \gamma) \supset ((\alpha \Rightarrow \beta) \supset (\alpha \Rightarrow \gamma))$.

CM: $((\alpha \Rightarrow \beta) \wedge (\alpha \Rightarrow \gamma)) \supset (\alpha \wedge \beta \Rightarrow \gamma)$.

And e **RCM** garantem que o conjunto de fatos implicados normalmente por uma fórmula α é fechado por consequência lógica.

Levando em consideração as regras anteriores, pode-se ver que $CT4O$ permite raciocinar em casos simples de herança com exceções. Assim, considere, por exemplo, que as crenças de um agente sejam “os mamíferos vivem geralmente na terra”, “as baleias geralmente não vivem na terra”, “as baleias são mamíferos” e “as baleias têm pulmões” e que estejam representadas na base de conhecimento BC como segue:

$$BC = \{\text{mamífero} \Rightarrow \text{vive_terra}, \text{baleia} \Rightarrow \neg \text{vive_terra}, \text{baleia} \Rightarrow \text{mamífero},$$

$$baleia \Rightarrow tem_pulmões\}.$$

Se o agente tem o conhecimento de que algo é uma *baleia* e um *mamífero* então a herança de propriedades é feita a partir da classe *baleia* que é a mais específica. Então, neste caso, em virtude de CM pode-se deduzir a partir de BC:

$$(baleia \wedge mamífero) \Rightarrow \neg vive_terra \text{ e } (baleia \wedge mamífero) \Rightarrow tem_pulmões.$$

Ou seja,

$$BC \vdash (baleia \wedge mamífero) \Rightarrow \neg vive_terra \text{ e }$$

$$BC \vdash (baleia \wedge mamífero) \Rightarrow tem_pulmões$$

O mesmo pode ser deduzido se na BC $baleia \Rightarrow mamífero$ é substituído por $\exists (baleia \supset mamífero)$. Mas se em BC é agregado $\exists (baleia_azul \supset baleia)$, então, neste caso, não se pode deduzir $baleia_azul \Rightarrow \neg vive_terra$ ou $baleia_azul \Rightarrow tem_pulmões$.

Assim, nesta lógica, o princípio de especificidade relacionado com o fato de que os defaults mais específicos são preferidos aos menos específicos é válido, mas não é válida a herança de propriedades por default, como foi mostrado no exemplo anterior das baleias azuis. Isto está relacionado ao fato de que nesta lógica não são válidos, em general, raciocínios do tipo:

$$(\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \gamma) \supset (\alpha \Rightarrow \gamma) \text{ (Transitividade),}$$

já que, se isto fosse válido, seria válido um raciocínio do tipo:

$$((estudante_pos \Rightarrow adulto) \wedge (adulto \Rightarrow trabalha)) \supset (estudante_pos \Rightarrow trabalha).$$

Esta é uma lógica conservadora onde inferências logicamente não desejadas, como a anterior, não podem ser obtidas.

Para tratar com herança, surgiram diferentes propostas baseadas em normalidade, aumentadas com princípios extra lógicos. Tem-se, por exemplo, a proposta de Delgrande [Del90] e a de Boutilier [Bou93], nos quais é agregada uma relação de preferência no conjunto de interpretações, relação que tenta formalizar intuições acerca da herança de propriedades numa hierarquia.

Nesta lógica, como em outras lógicas condicionais, a seguinte fórmula também não é válida:

$$(\alpha \Rightarrow \beta) \supset (\alpha \wedge \gamma \Rightarrow \beta) \text{ (Reforço do antecedente).}$$

Pedir que isto não seja válido é razoável, já que, em geral, não é razoável, por exemplo, que a afirmação “os brasileiros bebem cerveja” implique que “os brasileiros que não ingerem bebidas alcoólicas bebem cerveja”, ou seja, a seguinte fórmula não deveria ser válida:

$$(\text{brasileiro} \Rightarrow \text{cerveja}) \supset ((\text{brasileiro} \wedge \neg \text{álcool}) \Rightarrow \text{cerveja}).$$

Assim, como o reforço do antecedente não é válido, então é consistente afirmar que $\{\text{brasileiro} \Rightarrow \text{cerveja}, (\text{brasileiro} \wedge \neg \text{álcool}) \Rightarrow \neg \text{cerveja}, (\text{brasileiro} \wedge \neg \text{álcool} \wedge \text{cerveja_sem_álcool}) \Rightarrow \text{cerveja}, \dots\}$.

Também não é válida a regra Modus Ponens aplicada ao conectivo $\Rightarrow$, isto é, $(\alpha \wedge (\alpha \Rightarrow \beta)) \supset \beta$ não é uma fórmula válida em CT4O. Entretanto, uma versão mais fraca de Modus Ponens é válida em CT4O.

Teorema 2.7 $\models (\alpha \wedge (\alpha \Rightarrow \beta)) \Rightarrow \beta$.

Assim, por exemplo, do fato que “os mamíferos normalmente vivem na terra” e que existe “um mamífero”, pode-se dizer então que esse mamífero normalmente “vive na terra”.

Raciocínio Default

O teorema anterior sugere que o raciocínio default pode ser modelado como um processo no qual se interroga o que é que normalmente se deduz a partir de uma base de conhecimento BC , que inclui expressões condicionais, utilizando a versão fraca de Modus Ponens. Ou seja, pode ser inferido por default a partir da base de conhecimento BC aqueles α tais que $BC \Rightarrow \alpha^{12}$ é válida.

Se uma base de conhecimento BC tem, por exemplo, as seguintes sentenças:

$$BC = \{\text{mamífero} \Rightarrow \text{vive_terra}, \Box(\text{baleia} \supset \text{mamífero}), \text{baleia} \Rightarrow \neg \text{vive_terra}, \text{baleia}\} \quad (2.2)$$

¹²Aqui BC representa a conjunção de todas as fórmulas da base de conhecimento.

então resulta válido:

$$\models BC \Rightarrow \neg \text{vive_terra}.$$

Pode ser observado, neste exemplo, que esta lógica não tem problemas com a interrelação dos defaults, já que, por exemplo, se se tem que algo é uma “baleia”, então pode-se deduzir que normalmente não vive na terra, o que não acontece com, por exemplo, a Lógica Default apresentada anteriormente.

Da base de conhecimento anterior também pode ser deduzido que “os mamíferos são tipicamente não-baleias” ($\models BC \supset (\text{mamífero} \Rightarrow \neg \text{baleia})$). Esta última afirmação é diferente da afirmação “os mamíferos não são tipicamente baleias” ($\text{mamífero} \not\Rightarrow \text{baleia}$). Em CT4O ambas afirmações podem ser representadas, da mesma forma como pode ser representado que as baleias são mamíferos excepcionais porque elas não vivem na terra ($\text{baleia} \not\Rightarrow \text{vive_terra}$), sem a necessidade de afirmar que as baleias são não-terrestres ($\text{baleia} \Rightarrow \neg \text{vive_terra}$). A primeira sentença é uma *negação fraca* e a segunda uma *negação forte* do condicional¹³.

A Lógica CO

Os CT4O-modelos refletem as condições mínimas com relação à ordem de normalidade. Pode querer-se impor mais restrições na relação $\geq$. Pedir que $\geq$ seja uma relação totalmente conexa é uma possível restrição. Isto está relacionado com o fato de que todos os mundos são comparáveis com relação à normalidade. Os modelos consistem de conjuntos totalmente ordenados de grupos (“clusters”) de mundos igualmente normais.

A lógica CT4O pode ser estendida para se obter a lógica CO, simplesmente mudando a definição de modelo:

Definição 2.41 Um CO-modelo é um par $M = \langle W, \geq \rangle$, onde W é um conjunto de mundos possíveis, $\geq$ é uma relação binária reflexiva, transitiva e totalmente conexa definida em W .

Do ponto de vista dos axiomas, só tem que ser agregado o axioma S à lógica CT4O para obter-se a lógica CO.

Definição 2.42 A lógica condicional CO é o menor conjunto $S \subseteq \mathcal{L}_c$ tal que S contém a LPC e os esquemas de axiomas K, K', T, 4, S e H, e é fechado segundo as regras de inferência MP e Nec:

¹³Se o antecedente do condicional é possível, a negação forte implica a fraca.

K: $\Box(\alpha \supset \beta) \supset (\Box\alpha \supset \Box\beta)$.

K': $\Box(\alpha \supset \beta) \supset (\Box\alpha \supset \Box\beta)$.

T: $\Box\alpha \supset \alpha$.

4: $\Box\alpha \supset \Box\Box\alpha$.

S: $\alpha \supset \Box\alpha$.

H: $\Diamond(\Box\alpha \wedge \Box\beta) \supset (\Box(\alpha \vee \beta))$.

MP: A partir de $\alpha \supset \beta$ e α inferir β .

Nec: A partir de α inferir $\Box\alpha$.

Os conceitos referentes a quando uma fórmula pode ser deduzida em *CO* são definidos da maneira usual.

Em *CO*, como a relação de acessibilidade é totalmente conexa, $\Diamond\alpha$ pode ser lido como “ α é verdade em algum mundo menos normal”, já que em *CO* não há mundos inacessíveis.

Diferentemente dos *CT4O*-modelos, nos *CO*-modelos só existe um caminho de normalidade. Isto significa que não têm que ser consideradas diferentes alternativas e isto influi na definição de $\Rightarrow$ que pode ser reformulada como:

Definição 2.43 $(\alpha \Rightarrow \beta) \equiv_{def} \Box \neg\alpha \vee \Diamond(\alpha \wedge \beta \wedge \Box(\alpha \supset \beta))$.

Pode ser demonstrado [?] que as duas definições coincidem em *CO*:

Teorema 2.8 $\models_{CO} \Box(\Box\neg\alpha \vee \Diamond(\alpha \wedge \beta \wedge \Box(\alpha \supset \beta))) \leftrightarrow \Box \neg\alpha \vee \Diamond(\alpha \wedge \beta \wedge \Box(\alpha \supset \beta))$

Também pode ser demonstrado que:

Teorema 2.9 O sistema *CO* está caracterizado pela classe de *CO*-modelos; ou seja, $\vdash_{CO} \alpha$ se e somente se $\models_{CO} \alpha$.

Propriedades de CO

A lógica *CO*, tal como a lógica *CT4O*, é uma lógica onde é possível representar e raciocinar com defaults. As fórmulas que são válidas em *CT4O* também são válidas em *CO*. Nesta lógica também não são válidos o Modus Ponens condicional, o reforço do antecedente e a transitividade. Mas *CO* possui propriedades que não são válidas em *CT4O*.

Assim, como foi visto no último exemplo da seção 2.4.2 da base de conhecimento

$BC = \{mamífero \Rightarrow vive_terra, \Box(baleia \supset mamífero), baleia \Rightarrow \neg vive_terra, baleia\}$,

a fórmula $\text{mamífero} \Rightarrow \neg \text{baleia}$ era dedutível em $CT4O$. Mas se em BC , a fórmula $\text{baleia} \Rightarrow \neg \text{vive_terra}$ é substituída pela negação fraca $\text{baleia} \nRightarrow \text{vive_terra}$, então também deveria ser possível deduzir que “as baleias são tipicamente não-mamíferos” ($\text{baleia} \Rightarrow \neg \text{mamífero}$), já que as baleias continuam sendo mamíferos excepcionais, embora BC só tenha a negação fraca. No entanto, esta dedução não é válida em $CT4O$.

Em CO também são válidas [?] as seguintes fórmulas gerais:

Teorema 2.10 *As seguintes são fbf válidas em CO*

RM : $((\alpha \Rightarrow \gamma) \wedge (\alpha \wedge \beta \nRightarrow \gamma)) \supset (\alpha \Rightarrow \neg \beta)$

CV : $(\alpha \nRightarrow \beta) \supset ((\alpha \Rightarrow \gamma) \supset (\alpha \wedge \neg \beta \Rightarrow \gamma)).$

As Lógicas CO^* e $CT4O^*$

Tentando solucionar o problema da informação irrelevante, Boutilier introduz a lógica CO^* , uma extensão da lógica CO onde é considerada a classe de CO -modelos onde, o conjunto de mundos possíveis inclui todos os mundos possíveis que podem ser considerados a partir do conjunto P de variáveis proposicionais, ou seja, o conjunto 2^P .

Assim, para obter a lógica CO^* só é necessário modificar a definição de modelo:

Definição 2.44 *Um CO^* -modelo é um par $M = \langle W, \geq \rangle$, onde M é um CO -modelo tal que $W = 2^P$.*

Axiomaticamente:

Definição 2.45 *CO^* é a menor extensão de CO fechada segundo as regras de CO e que contém o seguinte esquema de axioma:*

LP. $\Diamond \alpha$, para toda fórmula proposicional não contraditória α .

Teorema 2.11 *O sistema CO^* está caracterizado pela classe de CO^* -modelos; ou seja, $\vdash_{CO^*} \alpha$ se e somente se $\models_{CO^*} \alpha$.*

A lógica CO^* é uma extensão própria da lógica CO , já que, todas as fórmulas que são válidas em CO continuam sendo válidas em CO^* , mas há fórmulas que são válidas somente em CO^* , por exemplo, se α é a fórmula p_0 então $\Diamond \alpha$ é válida só em CO^* .

A lógica $CT4O$ pode ser estendida da mesma forma que a lógica CO considerando a seguinte classe de modelos:

Definição 2.46 Um $CT4O^*$ -modelo é um par $M = \langle W, \geq \rangle$, onde M é um $CT4O$ -modelo e $W = 2^P$.

Axiomaticamente:

Definição 2.47 $CT4O^*$ é a menor extensão de $CT4O$ fechada segundo as regras de $CT4O$ e que contém o seguinte esquema de axioma:

LP. $\Diamond \alpha$ para toda fórmula proposicional não contraditória α .

Teorema 2.12 O sistema $CT4O^*$ está caracterizado pela classe de $CT4O^*$ -modelos; ou seja, $\vdash_{CT4O^*} \alpha$ se e somente se $\models_{CT4O^*} \alpha$.

Como no caso anterior, a extensão $CT4O^*$ obtida considerando estes modelos, é uma extensão própria de $CT4O$.

Relevância

Foi visto na seção 2.4.2 que a versão fraca de Modus Ponens poderia ser utilizada para determinar o que normalmente é deduzido de uma base de conhecimento. Mas existem determinados exemplos onde isto não acontece. Assim, do fato de se ter “os mamíferos vivem geralmente na terra” e de se ter “um mamífero cujo pelo é marrom”, nas lógicas da família apresentadas anteriormente não pode ser deduzido que “vive na terra”; ou seja, a fórmula

$$((\text{mamífero} \Rightarrow \text{vive.terra}) \wedge \text{mamífero} \wedge \text{marrom}) \Rightarrow \text{vive.terra}$$

não é logicamente válida nem em $CT4O$ nem em CO (resp. $CT4O^*$ e CO^*).

Isto acontece porque a lógica condicional, como já foi dito, é uma lógica muito “conservadora”. Então, como é possível que ter pelo marrom seja relevante para o fato de viver ou não na terra, nesta lógica não pode ser deduzido que “um mamífero com pelo marrom normalmente vive na terra”, a partir de que “os mamíferos geralmente vivem na terra”. E isto deriva do fato de que não é válido, como também foi visto, o reforço do antecedente. Mas, neste caso, parece razoável deduzir que normalmente “vive na terra”, pois nada parece indicar que ter “pelo marrom” seja relevante com relação ao fato de “viver na terra”.

O problema da relevância, está presente na maioria dos sistemas de raciocínio default que realizam uma interpretação condicional dos defaults.

Para solucionar este problema foram propostos alguns sistemas tais como o sistema Z [Pea90] que adicionam características extra lógicas para poder tratar o problema da irrelevância.

Boutilier, por sua vez, faz uma proposta para solucionar o problema no contexto da família de lógicas condicionais por ele proposta. Ele introduz um novo conectivo, o conectivo " $>$ ", cuja definição é a seguinte:

Definição 2.48 $\alpha > \beta \equiv_{def} \Box (\alpha \supset (\Box (\alpha \supset \neg \beta) \wedge \Diamond (\alpha \wedge \beta))) \wedge \Diamond \alpha$.

O conectivo $>$, considerado no contexto das lógicas CO, possui as seguintes propriedades [?]:

Teorema 2.13 *Seja $M = \langle W, \geq \rangle$ um CO-modelo e α e β fbf's quaisquer, então:*

1. *para todo $w \in W$ $M, w \models \alpha > \beta$ se e somente se $M \models \alpha > \beta$.*
2. *Se $M \models \alpha > \beta$ então existe um grupo C em W tal que $\|\alpha \wedge \beta\| \subseteq C$ e não existe nenhum α -mundo estritamente mais normal do que C . Além disso, $\|\alpha \wedge \beta\|$ deve ser não vazio.*
3. *Se $M \models \alpha > \beta$ e $M \models \alpha \Rightarrow \beta$, então existe um grupo C em W tal que $\|\alpha \wedge \beta\| \cup N = C$, onde cada mundo em N satisfaz $\neg \alpha$ e nenhum α -mundo é estritamente mais normal que C .*

Assim, segundo a proposição anterior, $\alpha > \beta$ garante que todos os $\alpha \wedge \beta$ -mundos estão entre os α -mundos mais normais; se também se tem que $\alpha \Rightarrow \beta$, então os α -mundos mais normais satisfazem $\alpha \wedge \beta$. Isto permite garantir que se uma proposição γ é consistente com $\alpha \wedge \beta$ então $(\alpha \wedge \gamma) \Rightarrow \beta$ é dedutível. Formalmente:

Teorema 2.14

$$\models_{CO} (\alpha \Rightarrow \beta \wedge \alpha > \beta \wedge \Diamond (\alpha \wedge \beta \wedge \gamma)) \supset ((\alpha \wedge \gamma) \Rightarrow \beta).$$

Teorema 2.15 *Se $\alpha \wedge \beta \wedge \gamma$ não é contraditória então: $\models_{CO} (\alpha \Rightarrow \beta \wedge \alpha > \beta) \supset ((\alpha \wedge \gamma) \Rightarrow \beta)$.*

Como é mencionado por Boutilier [Bou92], a lógica CO* junto com o conectivo $>$ permite solucionar o problema de irrelevância em determinados casos, a partir de um conjunto de fatos presentes na base de conhecimento, embora não seja claro como raciocinar, em geral, nesta lógica. Em [Bou92] simplesmente é proposto agregar uma fórmula $\alpha > \beta$ para cada fórmula $\alpha \Rightarrow \beta$ presente na base de conhecimento, na medida que o

resultado seja consistente.

Assim, se $BC = \{\text{mamífero} \Rightarrow \text{vive_terra}\}$ e se a fórmula $\text{mamífero} > \text{vive_terra}$ é agregada a BC , então pode-se deduzir $(\text{mamífero} \wedge \text{marrom}) \Rightarrow \text{vive_terra}$, ou seja, tem-se o seguinte resultado:

$$\models BC \supset ((\text{mamífero} \wedge \text{marrom}) \Rightarrow \text{vive_terra})$$

Se $BC = \{\text{mamífero} \Rightarrow \text{vive_terra}, \Box(\text{baleia} \supset \text{mamífero}), \text{baleia} \Rightarrow \neg \text{vive_terra}, \text{baleia}\}$, então, neste exemplo, já não se pode agregar consistentemente $\text{mamífero} > \text{vive_terra}$. Mas as fórmulas

$$\begin{aligned} (\text{mamífero} \wedge \neg \text{baleia}) &> \text{vive_terra} \text{ e} \\ (\text{mamífero} \wedge \text{baleia}) &> \neg \text{vive_terra} \end{aligned}$$

podem ser agregadas a BC . Como consequência, as fórmulas

$$\begin{aligned} (BC \wedge \text{mamífero}) &\Rightarrow \text{vive_terra} \text{ e} \\ (BC \wedge \text{mamífero} \wedge \text{marrom}) &\Rightarrow \text{vive_terra} \end{aligned}$$

são válidas.

Mas há exemplos simples onde a proposta de Boutilier não permite solucionar o problema da irrelevância. Assim, por exemplo, se $BC = \{\alpha \Rightarrow \beta, \gamma \Rightarrow \delta\}$, onde α, β, γ e δ são fbf atômicas diferentes, então $\alpha > \beta$ e $\gamma > \delta$ não podem ser agregadas consistentemente a BC .

2.5 Lógicas Preferenciais

Nesta seção apresentamos um conjunto de lógicas baseadas na escolha de um determinado conjunto de modelos “mínimos”. Apresentamos primeiro a proposta de Shoham [Sho88] e a seguir a proposta de McCarthy [McC80, McC86].

2.5.1 Lógica preferencial

Uma abordagem, baseada em modelos, que permite a construção de lógicas não-monotônicas, é a lógica preferencial proposta por Shoham [Sho88]. A idéia principal da lógica preferencial é que o significado de uma fórmula não é dado, como na lógica clássica, pelo conjunto

de todos seus modelos, mas por um subconjunto “preferido” desses modelos, denominados *modelos mínimos*. Podem ser adotados diferentes critérios de preferência e cada um desses critérios resulta numa lógica diferente.

Formalmente, seja $\mathcal{L}$ uma lógica clássica¹⁴ (monotônica) e $\leq$ uma pre-ordem¹⁵ no conjunto de interpretações¹⁶ de $\mathcal{L}$ ¹⁷. Se M_1 e M_2 são duas interpretações de $\mathcal{L}$, então $M_1 \leq M_2$ significa que a interpretação M_1 é tão preferível quanto M_2 e $M_1 < M_2$ ¹⁸ significa que M_1 é preferida com relação a M_2 . Assim, $\mathcal{L}$ e $\leq$ definem uma lógica preferencial $\mathcal{L}_{\leq}$, que tem a mesma linguagem que $\mathcal{L}$ e cuja semântica é definida como segue:

Definição 2.49 *Dada uma interpretação M de $\mathcal{L}$ e uma fórmula α de $\mathcal{L}$, então M satisfaz preferencialmente α (denotado por $M \models_{\leq} \alpha$) se $M \models_{\mathcal{L}} \alpha$ e não existe uma outra interpretação M' tal que $M' < M$ e $M' \models \alpha$. Nesse caso M é chamado de modelo preferido de α .*

Definição 2.50 *α é satisfeita preferencialmente, se existe uma interpretação M tal que $M \models_{\leq} \alpha$.*

Um outro conceito que pode ser apresentado está relacionado com as fórmulas preferencialmente válidas, embora não seja muito intuitivo.

Definição 2.51 *Uma fórmula α é preferencialmente válida (denotado por $\models_{\leq} \alpha$) se e somente se $\neg\alpha$ não é preferencialmente satisfatível.*

Lema 2.1 *Uma fórmula α é preferencialmente válida se para todo M , ou*

1. $M \models \alpha$, ou
2. existe um M' tal que $M' \leq M$ e $M' \not\models \alpha$.

¹⁴Pode ser a lógica proposicional clássica (LPC), a lógica de primeira ordem (LPO) ou uma lógica modal (LM)

¹⁵Uma relação binária numa conjunto A é uma pré-ordem se ela é reflexiva e transitiva.

¹⁶Aqui interpretação é uma atribuição de valores de verdade as variáveis proposicionais na LPC, uma interpretação de primeira ordem para a LPO ou uma estrutura de Kripke e um mundo dessa estrutura, no caso da LM.

¹⁷Esta é uma formulação um pouco diferente da apresentada em [Sho88], já que Shoham pede que $\leq$ seja uma relação de ordem parcial estrita, ou seja, que a relação $\leq$ seja transitiva, anti-simétrica e não reflexiva.

¹⁸Aqui $M_1 < M_2$ significa que $M_1 \leq M_2$ e $M_2 \not\leq M_1$

As fórmulas preferencialmente válidas têm algumas propriedades estranhas, já que, por exemplo, se $\leq$ é uma relação não limitada e se α é válida, então α e $\neg\alpha$ são preferencialmente válidas. Assim, para não ter problemas com a definição de fórmula preferencialmente válida, podem ser consideradas só as relações limitadas.

Definição 2.52 *Um relação $\leq$, definida num conjunto de modelos, é limitada se não existe uma seqüência infinita de modelos $M_1, M_2, M_3, \dots$, tal que $M_1 > M_2 > M_3 > \dots$.*

Teorema 2.16 *Seja $\mathcal{L}_{\leq}$ uma lógica preferencial baseada numa relação limitada e seja α uma fórmula de $\mathcal{L}$, então α é satisfeita em $\mathcal{L}$ se e somente se α é preferencialmente satisfatível em $\mathcal{L}_{\leq}$.*

Corolário 2.1 *Com as mesmas hipóteses, α é válida em $\mathcal{L}$ se e somente se α é preferencialmente válida em $\mathcal{L}_{\leq}$.*

A seguir será definida a noção central da lógica preferencial.

Definição 2.53 *α implica preferencialmente β (denotado por $\alpha \models_{\leq} \beta$) se para todo M , se $M \models_{\leq} \alpha$ então $M \models \beta$, ou equivalentemente, se os modelos de β contém os modelos preferidos de α .*

Neste formalismo, são válidos os seguintes resultados:

Lema 2.2 *Seja $\mathcal{L}_{\leq}$ uma lógica preferencial e α uma fórmula, então α é preferencialmente satisfatível se e somente se não existe uma fórmula β em $\mathcal{L}_{\leq}$, tal que $\alpha \models_{\leq} \beta$ e $\alpha \models_{\leq} \neg\beta$.*

Lema 2.3 *Seja $\mathcal{L}_{\leq}$ uma lógica preferencial, α e β duas fórmulas e M um modelo. Se $M \models \beta$ e $M \models_{\leq} \alpha$, então $M \models_{\leq} \alpha \wedge \beta$.*

Teorema 2.17 *Seja $\mathcal{L}_{\leq}$ uma lógica preferencial e α, β e γ três fórmulas. Se $\alpha \wedge \beta \models_{\leq} \gamma$ então $\alpha \models_{\leq} \beta \supset \gamma$.*

Não-monotonicidade

Esta abordagem é, em geral, não-monotônica, já que, $\alpha \models_{\leq} \gamma$ se γ é verdadeira em todos os modelos preferidos de α , e $(\alpha \wedge \beta) \models_{\leq} \gamma$ se γ é verdadeira em todos os modelos preferidos de $\alpha \wedge \beta$; mas os modelos preferidos de $\alpha \wedge \beta$ podem não ser modelos preferidos de α (eles até podem ser disjuntos) portanto, em geral, não é válido que se $\alpha \models_{\leq} \gamma$ então $\alpha \wedge \beta \models \gamma$.

Mas pode acontecer que uma relação $\leq$ seja tal que a lógica preferencial resultante seja monotônica.

Teorema 2.18 *Seja $\mathcal{L}_{\leq}$ uma lógica preferencial. Então são equivalentes:*

1. *Para toda α, β e $\gamma \in \mathcal{L}$, se $\alpha \models_{\leq} \beta \supset \gamma$ então $\alpha \wedge \beta \models_{\leq} \gamma$.*
2. *$\mathcal{L}_{\leq}$ é monotônica.*

A demonstração dos resultados referentes à lógica preferencial está em [Sho88].

Esta abordagem é suficientemente geral para capturar os detalhes de muitos dos sistemas de raciocínio não-monotônico existentes e pode servir como base para novos enfoques de raciocínio default. Uma lógica não-monotônica é o resultado de associar com uma lógica clássica uma relação de preferência no conjunto de seus modelos. Entre os enfoques baseados em modelos mínimos, podem ser mencionados circunscrição [McC80, McC86], o trabalho proposto por Bossu e Siegel [BS85], assim como várias propostas que permitem raciocinar sobre tempo e mudanças¹⁹ [Sho88, Bak91, Sha95].

2.5.2 Circunscrição

Circunscrição é uma formalização do raciocínio não-monotônico introduzido por McCarthy [McC80] e desenvolvido com posterioridade, principalmente, por Lifschitz [Lif86]. Circunscrição é, na realidade, uma família de lógicas que podem ser consideradas como um caso particular de lógica preferencial.

Nesta proposta, em lugar de estender a linguagem ou agregar regras de inferência não-monotônicas, são definidas as fórmulas não-monotonicamente dedutíveis a partir de um conjunto de premissas como o conjunto de fórmulas dedutíveis monotonicamente, a partir de um determinado conjunto que contém as premissas. A não-monotonia surge do fato

¹⁹O trabalho original de Shoham sobre lógica preferencial foi motivado pela necessidade de uma teoria para raciocínio temporal.

de que o conjunto de fórmulas agregadas as premissas depende dessas premissas, ou seja, as fórmulas agregadas a um conjunto Γ de premissas não tem porque ser um subconjunto do conjunto de fórmulas agregadas a um conjunto Γ' , tal que $\Gamma \subset \Gamma'$.

McCarthy [McC80] observou que muitas formas de raciocínio não-monotônico podem ser consideradas como uma minimização da extensão de vários predicados. Do ponto de vista semântico, isso corresponde ao fato de não considerar todos os modelos.

Diferentes versões de circunscrição foram definidas, mas todas essas versões têm em comum a existência de um esquema de axioma de circunscrição, de primeira ordem [McC80] ou de segunda ordem [McC86]. Entre as diferentes versões podem ser mencionadas: circunscrição de predicados [McC80], circunscrição de domínio [McC80], circunscrição variável [McC86, Lif86], circunscrição paralela [McC86, Lif86], e circunscrição com prioridades [McC86]. As características fundamentais de cada uma delas são apresentadas a seguir.

Circunscrição de Predicados

Existem predicados que tipicamente não são válidos para os objetos em geral. Assim, por exemplo, o predicado "albino" não é válido para muitos seres vivos, então, se não existe informação em contrario, é razoável pensar que os seres vivos não pertencem a classe dos objetos albinos. Por este motivo é necessária uma técnica que permita deduzir que os únicos objetos albinos são aqueles dos quais pode ser demonstrado seu albinismo. Semanticamente, a minimização da extensão de um predicado pode ser feita considerando só alguns modelos. Assim, no exemplo, só seriam considerados aqueles modelos nos quais a extensão para o predicado albino é mínima.

Uma ordem parcial pode ser definida nos modelos²⁰ de um conjunto Γ de premissas de primeira ordem, que dependem de um predicado P que está sendo minimizado. A ordenação pode ser definida como segue:

Definição 2.54 *Seja Γ um conjunto de fórmulas e M_1 e M_2 dois modelos de Γ , $M_1 \leq_P M_2$ se*

1. M_1 e M_2 têm o mesmo domínio,
2. todos os símbolos de predicado de Γ , com exceção de P , tem as mesmas extensões em M_1 e M_2 ,

²⁰Os modelos considerados são os modelos de primeira ordem nos quais os predicados são representados pela sua extensão.

3. a extensão de P em M_1 está contida na extensão de P em M_2 .

Definição 2.55 M é $\leq_P$ -mínimo se não existe M' tal que $M' <_P M$.

A noção de implicação é definida em função dos modelos mínimos.

Definição 2.56 Um conjunto de premissas Γ implica minimamente uma fórmula α , com relação a P , se α é verdadeira em todos os modelos de $\Gamma \leq_P$ -mínimos.

Assim, se o conjunto de premissas Γ é

$$\Gamma = \{Pedro \neq Ana, albino(Pedro)\}$$

então $\neg albino(Ana)$ é implicado minimamente pelo conjunto de premissas, com relação ao predicado albino.

Do ponto de vista sintático, as fórmulas minimamente implicadas por um conjunto de premissas podem ser caracterizadas estendendo o conjunto de premissas. Assim, os teoremas não-monotônicos de um conjunto de premissas Γ são o conjunto de teoremas (monotônicos) de $\Gamma \cup X$, onde uma escolha apropriada de X garante que só os modelos mínimos de Γ são modelos de $\Gamma \cup X$. McCarthy descreve a circunscrição de P em Γ (denotado por $Circ(\Gamma, P)$) como:

Definição 2.57 Seja Γ uma fórmula de primeira ordem²¹ que tem um símbolo de predicado n -ário P . Seja $\Gamma(\phi)$ o resultado de substituir todas as ocorrências do predicado P , em Γ , pelo predicado ϕ , que é considerado um parâmetro. A circunscrição do predicado P em Γ é um esquema da forma

$$\Gamma(\phi) \wedge (\forall x)(\phi(x) \supset P(x)) \supset (\forall x)(P(x) \supset \phi(x)),$$

onde $x = x_1, \dots, x_n$.

Este esquema expressa o fato de que se um predicado ϕ satisfaz Γ e é menor que o predicado P (o antecedente), então o predicado P e o predicado ϕ têm as mesmas extensões (o conseqüente). Como ϕ varia no conjunto de símbolos de predicado, consegue-se com isto que a extensão de P seja a menor possível, de tal forma que Γ se satisfaça. Todas as instâncias deste esquema, junto com o conjunto de premissas Γ são usados nas deduções.

²¹ Γ é a conjunção de todas as fórmulas de um conjunto de premissas.

Também é possível, circunscrever vários predicados em paralelo (*circunscrição paralela*). Nesse caso é preciso um parâmetro para cada predicado que vai ser circunscrito e repetir o antecedente e o conseqüente para cada um dos parâmetros.

A circunscrição de predicados é correta com relação à implicação mínima [McC80], mas não é completa, ou seja, existem casos nos quais uma fórmula é verdadeira em todos os modelos mínimos mas não pode ser deduzida usando circunscrição de predicados [Dav80]. Uma versão de segunda ordem foi proposta por McCarthy [McC86] para solucionar este problema. Neste caso $Circ(\Gamma, P)$ faz referência a uma fórmula de segunda ordem

$$\Gamma(P) \wedge \neg(\exists p)(\Gamma(p) \wedge (\forall x)(p(x) \supset P(x)) \wedge \neg(\forall x)(P(x) \supset p(x))),$$

onde p é uma tupla de variáveis de predicados similares²² a P .

Circunscrição paralela de segunda ordem é completa e correta [Lif85, Eth88].

Representação de Defaults

Este formalismo pode ser utilizado para representar defaults utilizando um predicado “distinguido” que é válido para determinados indivíduos “anormais” em determinados aspectos. Assim, por exemplo, o fato que “os mamíferos vivem normalmente na terra” é representado aqui como

$$(\forall x)((mamifero(x) \wedge \neg abn(x)) \supset vive.terra(x)).$$

Intuitivamente, “os mamíferos que não são anormais vivem na terra”. O predicado abn é circunscrito de maneira tal de ter tão poucos objetos anormais quanto possível. Assim, um mamífero pode não ser anormal com relação o fato de viver na terra, mas pode ser anormal com relação a outros aspectos, por conseguinte são necessários predicados distintos abn_i para os aspectos diferentes de defaults diferentes.

Apesar de ser um formalismo muito útil, circunscrição paralela tem certos problemas. Assim, se o conjunto de premissas é

$$\Gamma = \{(\forall x)(mamifero(x) \wedge \neg abn(x) \supset vive.terra(x)), (\forall x)(orca(x) \supset abn(x)), (\forall x)(golfinho(x) \supset abn(x)), mamifero(Toto)\}$$

e o predicado abn é circunscrito, então não se pode deduzir $vive.terra(Toto)$, já que, a interpretação dos predicados diferentes do predicado abn deve ficar fixa, portanto existem alguns modelos minimais nos quais $orca(Toto)$ é verdadeiro. Logo não pode ser deduzido

²²Predicados similares são predicados que têm a mesma aridade.

vive_terra(Toto).

Para solucionar este problema foi proposta a circunscrição variável [McC86, Lif85].

Circunscrição variável e com prioridades

A circunscrição variável é uma generalização da circunscrição de predicados que permite que determinados predicados variem durante a minimização. Esta circunscrição está definida como:

Definição 2.58 *Seja $\Gamma(P; Q)$ uma fórmula²³ que tem o símbolo de predicado n -ário P e os símbolos de predicado do conjunto $Q = \{Q_1, \dots, Q_n\}$. Seja $\Gamma(\phi, \theta_1, \dots, \theta_M)$ a fórmula resultante de substituir todas as ocorrências de $P, Q_1, \dots, Q_n$ pelas variáveis de predicado $(\phi, \theta_1, \dots, \theta_M)$. Então a circunscrição variável de P em Γ com predicados variáveis $Q_1, \dots, Q_n$ (denotado por $Circ(\Gamma; P; Q)$) é a fórmula de segundo ordem*

$$\Gamma(P; Q) \wedge (\forall \phi, \theta_1, \dots, \theta_M)(\Gamma(\phi, \theta_1, \dots, \theta_M) \wedge (\forall x)(\phi(x) \supset P(x)) \supset (\forall x)(P(x) \supset \phi(x)),$$

onde $x = x_1, \dots, x_n$.

Em $Circ(\Gamma; P; Q)$, Q é o conjunto de símbolos de predicado cuja extensão pode variar.

A semântica desta circunscrição é uma modificação da semântica da circunscrição de predicados. A condição 2 da Definição 2.54 deve ser substituída por

“todos os símbolos de predicado de Γ exceto $P, Q_1, \dots, Q_n$ têm as mesmas extensões”.

O problema com este tipo de circunscrição é determinar que predicados devem ser circunscriptos e quais devem variar. McCarthy sugere um enfoque uniforme, utilizando os predicados *abn*, que são os predicados que vão ser minimizados, e a maioria dos outros (ou todos os outros) podem variar.

Uma outra proposta de circunscrição, é a circunscrição com prioridades, proposta por McCarthy que permite que determinados predicados sejam minimizados com preferência, em relação aos outros predicados. Lifschitz demonstrou que esta circunscrição é equivalente à conjunção de circunscrição variável.

²³A conjunção das fórmulas do conjunto de premissas.

Circunscrição Proposicional

Uma simplificação óbvia da circunscrição permite representar e raciocinar sobre defaults numa linguagem proposicional. As fórmulas defaults do tipo “ α então normalmente β ” são representadas nesta circunscrição proposicional como

$$\alpha \wedge \neg abn \supset \beta.$$

O critério de preferência é o critério que tenta falsificar abn , ou seja, $M_1 \leq M_2$ se e somente se

- para todos os símbolos proposicionais $p \in \mathbf{P} - \{abn\}$ $M_1 \models p$ se e somente se $M_2 \models p$ e
- $M_1 \not\models abn$.

Todas as variantes de circunscrição, tais como a que estabelece uma prioridade entre os predicados abn , circunscrição de fórmulas e circunscrição variável podem ser implementados numa linguagem proposicional. Este tipo de circunscrição será utilizada em alguns exemplos no capítulo 5.

A lógica Preferencial como uma generalização de Circunscrição

A lógica preferencial apresentada anteriormente pode ser considerada uma generalização de circunscrição, já que a lógica preferencial a partir de qualquer lógica padrão permite qualquer ordem nas interpretações, não só a ordem determinada por um axioma particular de circunscrição.

Capítulo 3

Lógicas Epistemicamente Estendidas

3.1 Introdução

Nos últimos vinte anos, a comunidade de representação de conhecimento vem realizando pesquisas em temas relacionados com vários tipos de raciocínio não-monotônico. Os trabalhos desenvolvidos tentam modelar distintos aspectos do raciocínio de senso comum, e permitem modelar, quase que exclusivamente, um único agente raciocinando sobre seu ambiente. No entanto, em muitas aplicações, esse ambiente inclui outros agentes, logo o agente necessita raciocinar sobre eles. Assim, este agente deverá raciocinar sobre o conhecimento, crenças ou a visão de mundo que os demais agentes possuem, para que a interação seja efetiva. Isso é importante, não só na área de representação de conhecimento, mas também em outras áreas relacionadas, por exemplo, a sistemas distribuídos ou banco de dados.

Por outro lado, muitas das lógicas multiagentes propostas ou são monotônicas [FHV91, GD93, HM85, HM92], e por isto possuem uma capacidade limitada para modelar o comportamento dos agentes, ou estão baseadas na lógica autoepistêmica [Mor90, Lak93, HM93, PJ95a], que modela somente alguns dos aspectos do raciocínio default.

Então, considerando-se as limitações das propostas existentes para modelar adequadamente um sistema no qual interagem vários agentes que raciocinam não-monotonicamente, ou novos formalismos devem ser propostos ou os formalismos existentes para o caso de um único agente devem que ser adaptados para poder modelar adequadamente essa interação.

Nesta dissertação são propostas as extensões de dois formalismo já existentes que possibilitam raciocinar não-monotonicamente. O trabalho está baseado no conceito de

extensões epistêmicas [Wai93] que é apresentado a seguir.

3.2 Extensões Epistêmicas de uma Lógica

Uma *extensão epistêmica* [Wai93] é uma extensão sintática e semântica de uma lógica não-monotônica que permite representar fórmulas que têm um operador de crença (denotado $\mathcal{B}$) e permite raciocinar com essas fórmulas. A lógica resultante da extensão, denominada lógica epistemicamente estendida, representa o ponto de vista de um agente particular (denominado “o sistema”) que raciocina sobre suas próprias crenças e sobre o que ele considera que são as crenças de um outro agente. Este formalismo enxerga as crenças do outro agente a partir da perspectiva do sistema.

Para obter uma extensão epistêmica de uma lógica não-monotônica, deve-se estender a linguagem da lógica original (denominada *lógica base*) e a semântica associada a essa lógica. Assim, se a lógica base tem uma linguagem $\mathcal{L}$, então a lógica estendida terá uma linguagem $\mathcal{L}^*$ que inclui um operador $\mathcal{B}$ para poder referir-se as crenças do outro agente, segundo o ponto de vista do sistema que deve interagir com ele. Uma fórmula do tipo $\alpha \wedge \mathcal{B} \gamma$ é interpretada como “o sistema acredita em α e acredita que o outro agente acredita em γ ”.

O enfoque adotado não será simétrico com relação aos agentes, já que somente poderá ser representado o conhecimento que o sistema tem das crenças do outro, mas não o conhecimento que o outro agente tem das crenças do sistema.

A semântica, por sua vez, é definida de tal forma que a lógica resultante tenha os requisitos apresentados na próxima seção.

3.2.1 Requisitos para uma Lógica Epistemicamente Estendida

No que se segue, os símbolos $\mathcal{L}$ e $\mathcal{L}^*$ são utilizados para representar a lógica base e a lógica estendida, respectivamente. O símbolo $\leadsto$ é usado para representar as regras default em ambas as lógicas¹.

Entre os requisitos que uma lógica epistemicamente estendida deve possuir para modelar o conhecimento de agentes não-monotônicos, podem ser mencionados os seguintes [Wai93]:

¹A expressão $\alpha \leadsto \beta$ será interpretada como “se α acontece então, em geral, β também”.

- O primeiro requisito está relacionado ao fato de que a lógica $\mathcal{L}^*$ deve ser uma extensão da lógica base. Assim, a lógica $\mathcal{L}^*$, restringida à fórmulas sem o operador $\mathcal{B}$, deve ter o mesmo poder dedutivo (não-monotônico) que a lógica base. Ou seja, dadas α e $\beta \in \mathcal{L}$:

$$\alpha \models_{\mathcal{L}} \beta \quad \text{se e somente se} \quad \alpha \models_{\mathcal{L}^*} \beta \quad (3.1)$$

- O sistema e o outro agente que a lógica $\mathcal{L}^*$ permite modelar devem possuir a mesma capacidade de raciocínio não-monotônico. Formalmente:

$$\alpha \models_{\mathcal{L}} \beta \quad \text{se e somente se} \quad \mathcal{B}\alpha \models_{\mathcal{L}^*} \mathcal{B}\beta \quad (3.2)$$

- A lógica $\mathcal{L}^*$ também deve modelar o comportamento do operador $\mathcal{B}$. Uma alternativa possível é a lógica $KD45$. Então, tem-se o seguinte requisito:

$$\text{Se } \models_{KD45} \alpha \quad \text{então} \quad \models_{\mathcal{L}^*} \alpha \quad (3.3)$$

- Outras propriedades que a lógica $\mathcal{L}^*$ deve possuir estão relacionadas com a inter-relação entre as crenças dos dois agentes e as regras default. Assim, $\mathcal{L}^*$ deve permitir que, a partir de fórmulas default que combinam subfórmulas com o operador $\mathcal{B}$ com subfórmulas sem o operador $\mathcal{B}$, sejam derivadas as conclusões corretas, por exemplo:

$$\begin{array}{c} \mathcal{B}\alpha \wedge (\mathcal{B}\alpha \rightsquigarrow \beta) \models_{\mathcal{L}^*} \beta \\ \text{e} \\ \alpha \wedge (\alpha \rightsquigarrow \mathcal{B}\beta) \models_{\mathcal{L}^*} \mathcal{B}\beta \end{array} \quad (3.4)$$

No capítulo 4 apresenta-se a extensão epistêmica de duas lógicas da família de lógicas condicionais propostas por [Bou94a]. No capítulo 5, por sua vez, é apresentada uma extensão epistêmica da lógica preferencial. Para desenvolver esta extensão utiliza-se uma generalização do conceito de extensão epistêmica, que permite que mais de um agente interaja com o sistema.

Capítulo 4

Extensões Epistêmicas das Lógicas Condicionais CT4O e CO

Neste capítulo apresentamos a extensão epistêmica de duas das lógicas da família de lógicas condicionais de normalidade proposta por Boutilier [Bou94b]. Também são analisadas as propriedades dessas extensões e são apresentados alguns exemplos.

4.1 Introdução

A maioria das lógicas não-monotônicas propostas não podem modelar um agente raciocinando acerca do conhecimento de outros agentes. Em geral, o estudo do raciocínio não-monotônico está limitado ao caso de um único agente; isso limita sua utilidade para modelar a comunicação entre eles. Então, para poder modelar o conhecimento de outros agentes, devem ser propostos novos formalismos ou adaptar os já existentes.

Se a direção escolhida é adaptar os formalismos já existentes e se a lógica vai modelar possíveis situações onde só interagem dois agentes, denominados *S* (o sistema) e *O* (o outro agente que interage com o sistema), então dois enfoques podem ser adotados:

- um enfoque no qual só há uma dimensão para raciocinar, aquela onde raciocina *S*, que deve importar as crenças de *O* no seu próprio espaço e raciocinar nele como se as crenças do outro fossem suas próprias.
- um outro enfoque onde ambos os agentes têm sua própria dimensão para raciocinar e ambas as crenças estão perfeitamente diferenciadas [FHV91, GD93].

Neste trabalho adotamos o segundo enfoque para a obtenção de duas lógicas denominadas *lógicas condicionais epistemicamente estendidas*. O enfoque não é simétrico com

relação ao ponto de vista dos agentes, já que, o ponto de vista adotado é o do sistema: só é representado o conhecimento que **S** tem das crenças de **O**, mas não o conhecimento que **O** tem das crenças de **S**.

Como base para a extensão, foi escolhida a família de lógicas condicionais de normalidade propostas por Boutilier, já que essas lógicas fornecem um conjunto mínimo de propriedades, desejáveis em toda lógica não-monotônica e permitem expressar as regras defaults na própria linguagem.

Apresentamos inicialmente a *extensão da lógica condicional de normalidade CT4O*, que denominaremos $CT4O^e$, e a seguir a *extensão da lógica CO*, denominada CO^e .

4.2 A Extensão de CT4O

Como já foi mencionado, serão considerados só dois agentes: **S** e **O**. Para obter a extensão de $CT4O$, primeiro precisamos definir a linguagem, a partir da linguagem condicional $\mathcal{L}_C$ definida no capítulo 2. Na extensão, as crenças de **S** estão implicitamente representadas pelas fórmulas da lógica condicional. Introduzimos um novo operador, o operador B para fazer referência às crenças do outro agente. Assim, por exemplo, a fórmula $\alpha \vee B\alpha$ é interpretada como "o sistema acredita em α ou o sistema acredita que o outro agente acredita em α ".

Mais formalmente, a linguagem $\mathcal{L}_{C^e}$ da extensão epistêmica $CT4O^e$ é definida como:

Definição 4.1 Dado um conjunto **P** de variáveis proposicionais $\{p_0, p_1, \dots\}$, as fórmulas bem formadas (fbf) de $\mathcal{L}_{C^e}$ são o menor conjunto tal que:

1. Se $\alpha \in \mathbf{P}$, então α é uma fbf (denominada fbf atômica).
2. Se α e β são fbf, então $\neg(\alpha)$, $(\alpha \supset \beta)$, $\Box(\alpha)$, $\Box(\alpha)$ são fbf.
3. Se α é uma fbf de $\mathcal{L}_C$, então $B(\alpha)$ é uma fbf.

Conjunção ($\wedge$), disjunção ($\vee$), bicondicional ($\leftrightarrow$), possibilidade ($\Diamond$), os operadores $\Diamond$, $\Box$, $\Box$ e o operador condicional ($\Rightarrow$) são definidos como na lógica $CT4O$.

Para definir a semântica de $CT4O^e$, adaptamos as estruturas clássicas de Kripke para representar o conhecimento dos dois agentes. Essencialmente, é usado o conceito de mundos possíveis, mas de uma maneira um pouco diferente. Os mundos possíveis que usamos têm duas componentes:



- uma para modelar as crenças do sistema,
- uma outra para modelar as possíveis crenças que o sistema tem, com relação às crenças do outro agente.

Inicialmente são definidos os *mundos aumentados*.

Definição 4.2 *Um mundo aumentado (denominado A-mundo) é um par $\langle w, \langle W, \geq \rangle \rangle$, onde W é um conjunto de mundos possíveis proposicionais, $w \in W$, e $\geq$ é uma relação reflexiva e transitiva definida em W .*

Um A-mundo é um CT4O-modelo com um mundo distinguido. Os conjuntos de A-mundos permitirão representar as crenças do agente O. A seguir definimos uma relação auxiliar, $\geq_\Delta$, num conjunto de A-mundos.

Definição 4.3 *Seja $\mathcal{A}$ um conjunto de A-mundos e $\langle w_1, M_1 \rangle, \langle w_2, M_2 \rangle \in \mathcal{A}$ então $\langle w_1, M_1 \rangle \geq_\Delta \langle w_2, M_2 \rangle$ se e somente se*

- $M_1 = M_2$ e
- $w_1 \geq_{M_1} w_2$, onde $\geq_{M_1}$ é a relação de normalidade do CT4O-modelo M_1 (ou do modelo M_2 , já que $\geq_{M_1}$ e $\geq_{M_2}$ são iguais pois $M_1 = M_2$).

Ou seja, dois A-mundos podem ser comparados com a relação $\geq_\Delta$ se seus CT4O-modelos são iguais e seus mundos distinguidos podem ser comparados pela relação de normalidade desses CT4O-modelos.

Teorema 4.1 *Seja $\mathcal{A}$ um conjunto de A-mundos. A relação $\geq_\Delta$ definida em $\mathcal{A}$ é uma relação reflexiva e transitiva.*

São definidos agora os *mundos de crenças*, que são os “mundos possíveis” utilizados para definir a semântica da extensão.

Definição 4.4 *Um mundo de crenças (B-mundo) é um par $\langle w, \mathcal{A} \rangle$, onde w é um mundo possível proposicional e $\mathcal{A}$ é um conjunto de A-mundos.*

Um mundo de crenças é a estrutura que permite interpretar as fórmulas da forma α e $B\alpha$. Assim, o que falta definir é a relação de “normalidade” $\geq_\square$ entre B -mundos, de tal forma que uma fórmula do tipo $B\alpha \Rightarrow B\beta$ possa ser interpretada e a lógica resultante tenha as propriedades enunciadas no capítulo 3.

$\geq_\square$ é definida com base numa relação $\geq$, definida, por sua vez, num conjunto de mundos possíveis proposicionais.

Definição 4.5 *Seja $\mathcal{X}$ um conjunto de B -mundos, e seja $\geq$ uma relação reflexiva e transitiva definida em $\Pi_1(\mathcal{X})^1$. Então, define-se em $\mathcal{X}$ a relação $\geq_\square$ com base numa relação $\geq$ como segue:*

Dados $\langle w_1, A_1 \rangle$ e $\langle w_2, A_2 \rangle \in \mathcal{X}$, então $\langle w_1, A_1 \rangle \geq_\square \langle w_2, A_2 \rangle$ se e somente se

- $w_1 \geq w_2$ e
- *para todo $a \in A_1$ existe um $b \in A_2$ tal que $a \geq_\Delta b$, e para todo $b \in A_2$ existe um $a \in A_1$ tal que $a \geq_\Delta b$.*

Teorema 4.2 *A relação $\geq_\square$, com base numa relação $\geq$ reflexiva e transitiva, também é reflexiva e transitiva.*

Finalmente, definimos o conceito central que captura a semântica de nossa lógica. A definição de $CT4O^e$ -modelo é a seguinte:

Definição 4.6 *Um $CT4O^e$ -modelo é um par $\langle \mathcal{X}, \geq \rangle$, onde $\mathcal{X}$ é um conjunto de B -mundos, $\geq$ é uma relação reflexiva e transitiva definida no conjunto $\Pi_1(\mathcal{X})$, e a relação $\geq_\square$ com base em $\geq$, definida em 4.5, é a relação de normalidade entre os B -mundos de $\mathcal{X}$.*

Além disso, pede-se também que $\mathcal{X}$ não tenha “buracos” no sentido que: se $\langle w, \{\langle w_i, M_i \rangle\}_{i \in I} \rangle \in \mathcal{X}$ e se $\{\langle w'_i, M_i \rangle\}_{i \in I}$ é um conjunto de mundos aumentados, tais que para todo $i \in I$, $\langle w_i, M_i \rangle \geq_\Delta \langle w'_i, M_i \rangle$, então existe um w_0 tal que $w \geq w_0$ e $\langle w_0, \{\langle w'_i, M_i \rangle\}_{i \in I} \rangle \in \mathcal{X}$.

A *satisfatibilidade* nos $CT4O^e$ -modelos é definida da forma usual. Dado um $CT4O^e$ -modelo $M = \langle \mathcal{X}, \geq \rangle$, $\geq_\square$ a relação, com base na relação $\geq$, definida no conjunto $\mathcal{X}$, x um B -mundo de $\mathcal{X}$ e Π_1 e Π_2 a primeira e segunda projeção estendidas para conjuntos de pares², então definimos satisfatibilidade de um fórmula α num B -mundo x do $CT4O^e$ -modelo M , denotado por $M, x \models_{CT4O^e} \alpha$ como:

¹ $\Pi_1(\mathcal{X}) = \{x : \langle x, y \rangle \in \mathcal{X}\}$

²Se A é um conjunto de pares ordenados, então $\Pi_1(A) = \{x : \langle x, y \rangle \in A\}$, $\Pi_2(A) = \{y : \langle x, y \rangle \in A\}$

Definição 4.7

- $M, x \models_{CT4O^e} \alpha$ se e somente se $\alpha \in \Pi_1(x)$, para α fbf atômica.
- $M, x \models_{CT4O^e} \neg \alpha$ se e somente se $M, x \not\models_{CT4O^e} \alpha$.
- $M, x \models_{CT4O^e} \alpha \supset \beta$ se e somente se $M, x \models_{CT4O^e} \beta$ ou $M, x \not\models_{CT4O^e} \alpha$.
- $M, x \models_{CT4O^e} \Box \alpha$ se e somente se $M, y \models_{CT4O^e} \alpha$, para todo y tal que $x \geq_\Box y$.
- $M, x \models_{CT4O^e} \Box \alpha$ se e somente se $M, y \models_{CT4O^e} \alpha$, para todo y tal que $x \not\geq_\Box y$.
- $M, x \models_{CT4O^e} B\alpha$ se e somente se $\Pi_2(a), \Pi_1(a) \models_{CT4O} \alpha$, para todo $a \in \Pi_2(x)$.

Quando estiver claro pelo contexto qual é a lógica com que se está trabalhando $M, x \models_{CT4O^e} \alpha$ se denotará simplesmente $M, x \models \alpha$.

Finalmente, a *validade* de uma fórmula num modelo e o conceito de *fórmula válida* são definidos da maneira usual:

Definição 4.8 Uma fbf α de $\mathcal{L}_C$ é válida num $CT4O^e$ -modelo $M = \langle \mathcal{X}, \geq \rangle$, denotado por $M \models_{CT4O^e} \alpha$, se e somente se $M, x \models_{CT4O^e} \alpha$, para todo $x \in \mathcal{X}$.

Definição 4.9 α é $CT4O^e$ -válida, denotado por $\models_{CT4O^e} \alpha$, quando $M \models_{CT4O^e} \alpha$, para todo $CT4O^e$ -modelo M .

Definição 4.10 A fórmula α é satisfeita se e somente se $\neg \alpha$ não é válida.

4.2.1 Resultados

Com semântica definida anteriormente à lógica $CT4O^e$ satisfaz os requisitos mencionados no capítulo 3. Temos portanto os seguintes resultados:

Teorema 4.3 Se $\alpha \in \mathcal{L}_C$ então: $\models_{CT4O} \alpha$ se e somente se $\models_{CT4O^e} \alpha$.

O teorema anterior estabelece que a lógica $CT4O^e$ é uma extensão conservativa de $CT4O$. Na realidade este teorema é mais forte que o primeiro requisito exigido no capítulo 3, mas é óbvio que se deduz do teorema anterior.

Corolário 4.1 *Se $\alpha, \beta \in \mathcal{L}_C$, então $\models_{CT4O} \alpha \Rightarrow \beta$ se e somente se $\models_{CT4O^e} \alpha \Rightarrow \beta$.*

O segundo requisito apresentado no capítulo 3 tem a seguinte formulação na lógica $CT4O^e$:

Teorema 4.4 *Se $\alpha, \beta \in \mathcal{L}_C$, então $\models_{CT4O} \alpha \Rightarrow \beta$ se e somente se $\models_{CT4O^e} B\alpha \Rightarrow B\beta$.*

A lógica $CT4O^e$ modela o comportamento do operador B :

Teorema 4.5 *Se $\alpha \in \mathcal{L}_C$, então $\models_{KD} \alpha$ então $\models_{CT4O^e} \alpha^3$.*

Além das propriedades mencionadas anteriormente em $CT4O^e$ também são válidas:

Teorema 4.6 *Os seguintes são resultados válidos em $CT4O^e$:*

1. $\models_{CT4O^e} ((\alpha \Rightarrow B\beta) \wedge \alpha) \Rightarrow B\beta$.
2. $\models_{CT4O^e} ((B\alpha \Rightarrow \beta) \wedge B\alpha) \Rightarrow \beta$.

Teorema 4.7 *Seja M um $CT4O^e$ -modelo; então*

1. $M \models_{CT4O^e} B(\alpha \wedge \beta)$ se e somente se $M \models_{CT4O^e} B(\alpha) \wedge B(\beta)$.
2. $M \models_{CT4O^e} B(\neg\alpha)$ então $M \models_{CT4O^e} \neg B(\alpha)$.
3. $\models_{CT4O^e} B(\alpha \wedge (\alpha \supset \beta)) \supset B(\beta)$.
4. $\models_{CT4O^e} B(\alpha \wedge (\alpha \Rightarrow \beta)) \Rightarrow B(\beta)$.

A demonstração destes resultados encontra-se no apêndice A.

Na lógica que acabamos de definir, pode-se representar as crenças de um outro agente e raciocinar sobre essas crenças. Nesta lógica, as propriedades essenciais da lógica $CT4O$ continuam sendo verdadeiras quando o sistema raciocina sobre as crenças do outro agente. Assim, segundo o teorema 4.6, se S inicialmente acredita que $mamífero \Rightarrow B(vive_terra)$ e $mamífero$, então, por default, o sistema acredita que o agente O acredita em $vive_terra$.

³Este requisito é equivalente ao requisito (3.3) do capítulo 3, já que a linguagem de $CT4O^e$ não permite o aninhamento do operador B .

Ou seja temos o seguinte resultado:

$$\models_{CT4O^c} \left(\begin{array}{c} \text{mamífero} \Rightarrow \mathcal{B}(\text{vive_terra}) \wedge \\ \text{mamífero} \end{array} \right) \Rightarrow \mathcal{B}(\text{vive_terra}). \quad (4.1)$$

Mas a lógica vai além dos resultados anteriores, levando em consideração a lógica do operador $\mathcal{B}$. Se às duas afirmações anteriores agregamos $\mathcal{B}(\neg \text{mamífero})$, então temos

$$\not\models_{CT4O^c} \left(\begin{array}{c} \text{mamífero} \Rightarrow \mathcal{B}(\text{vive_terra}) \wedge \\ \text{mamífero} \wedge \\ \neg \mathcal{B}(\text{vive_terra}) \end{array} \right) \Rightarrow \mathcal{B}(\text{vive_terra}) \quad (4.2)$$

O exemplo anterior pode ser generalizado da seguinte forma:

Teorema 4.8 *Os seguintes resultados são válidos em $\mathcal{L}_{C^c}$:*

1. $\not\models_{CT4O^c} (\alpha \Rightarrow \mathcal{B}(\beta) \wedge \alpha \wedge \neg \mathcal{B}(\beta)) \Rightarrow \mathcal{B}(\beta).$
2. $\not\models_{CT4O^c} (\alpha \Rightarrow \mathcal{B}(\beta) \wedge \alpha \wedge \mathcal{B}(\neg \beta)) \Rightarrow \mathcal{B}(\beta).$

Um outro tipo de raciocínio que pode ser feito em $CT4O^c$ é o seguinte. Seja o seguinte conjunto de premissas:

$$KB = \left\{ \begin{array}{l} \mathcal{B}(\text{mamífero}) \Rightarrow \mathcal{B}(\text{vive_terra}) \\ \Box(\mathcal{B}(\text{orca}) \supset \mathcal{B}(\text{mamífero})) \\ \mathcal{B}(\text{orca}) \Rightarrow \mathcal{B}(\neg \text{vive_terra}) \end{array} \right\}. \quad (4.3)$$

Do fato de ter-se $\mathcal{B}(\text{orca})$, infere-se, por default, que $\mathcal{B}(\neg \text{vive_terra})$, já que, o default mais específico se aplica e cancela o default mais genérico. Pode-se ver que este raciocínio não é semelhante ao exemplo 2.2 do capítulo 2, já que nesse caso, as fórmulas vive_terra e $\neg \text{vive_terra}$ estão dentro do escopo do operador de crença.

Também a partir dessas afirmações contidas em KB , podemos deduzir, por default, $\mathcal{B}(\text{mamífero}) \Rightarrow \neg \mathcal{B}(\text{orca})$, mas não $\mathcal{B}(\text{mamífero}) \Rightarrow \neg \text{orca}$, já que, esta última inferência seria uma atribuição de crenças ao outro agente, baseada nas crenças do sistema,

o que é um raciocínio incorreto.

Um outro modo de inferência permitida em $CT4O^e$ é o seguinte: a partir de $mamífero \Rightarrow B(mamífero)$, $B(mamífero \Rightarrow vive_terra)$ e $\Diamond\Box(mamífero)$, pode-se deduzir, por default, que $B(vive_terra)$, ou seja

$$\models_{CT4O^e} \left(\begin{array}{l} mamífero \Rightarrow B(mamífero) \wedge \\ B(mamífero \Rightarrow vive_terra) \wedge \\ \Diamond\Box(mamífero) \end{array} \right) \Rightarrow B(vive_terra). \quad (4.4)$$

Esta inferência mostra que $CT4O^e$ permite um encadeamento de inferências entre as crenças de **S** e as de **O**; os defaults do sistema podem ser compostos, em alguns casos com os defaults do outro agente.

4.3 A Extensão de CO

A lógica CO pode ser estendida da mesma forma que $CT4O$. Lembremos que a diferença entre $CT4O$ e CO é que a relação de normalidade em CO é totalmente conexa.

Assim, para estender a lógica CO , na definição de um A -mundo $\langle w, \langle W, \geq \rangle \rangle$ é necessário que $\geq$ seja totalmente conexa. As relação $\geq_\Delta$ e os B -mundos são definidos de maneira análoga e $\geq_\Box$ tem que ser totalmente conexa.

Para definir um CO^e -modelo é necessário somente adicionar mais restrições à relação $\geq_\Box$ e aos A -mundos:

Definição 4.11 *Um CO^e -modelo é um par $\langle \mathcal{X}, \geq \rangle$, onde $\mathcal{X}$ é um conjunto de B -mundos, $\geq$ é uma relação reflexiva, transitiva definida no conjunto $\Pi_1(\mathcal{X})$, e a relação $\geq_\Box$, como foi definida em $CT4O^e$, é totalmente conexa.*

Além disso, se $\langle w, \{\langle w_i, M_i \rangle\}_{i \in I} \rangle \in \mathcal{X}$ e se $\{\langle w'_i, M_i \rangle\}_{i \in I}$ é um conjunto de mundos aumentados tais que, para todo $i \in I$, $\langle w_i, M_i \rangle \geq_\Delta \langle w'_i, M_i \rangle$, então existe um w_0 , tal que $w \geq w_0$ e $\langle w_0, \{\langle w'_i, M_i \rangle\}_{i \in I} \rangle \in \mathcal{X}$ e os modelos dos mundos aumentados são CO -modelos.

A lógica CO^e mantém todas as propriedades de $CT4O^e$ e tem outras novas. Por exemplo, a partir de $B(mamífero) \not\Rightarrow B(orca)$ e $B(mamífero) \Rightarrow B(vive_terra)$ podemos

deduzir em CO^e $(\mathcal{B}(\text{mamífero}) \wedge \mathcal{B}(\neg \text{orca})) \Rightarrow \mathcal{B}(\text{vive terra})$. Ou seja, a seguinte fórmula é válida em CO^e , mas não em $CT4O^e$:

$$\models_{CO^e} \left(\frac{\mathcal{B}(\text{mamífero}) \not\Rightarrow \mathcal{B}(\text{orca}) \wedge}{\mathcal{B}(\text{mamífero}) \Rightarrow \mathcal{B}(\text{vive terra})} \right) \supset ((\mathcal{B}(\text{mamífero}) \wedge \mathcal{B}(\neg \text{orca})) \Rightarrow \mathcal{B}(\text{vive terra})) \quad (4.5)$$

Capítulo 5

Uma Extensão Multiagentes de uma Lógica Preferencial

Neste capítulo apresentamos uma lógica multiagentes não-monotônica baseada numa lógica preferencial. Nesta nova lógica, cada agente raciocina de forma independente utilizando a lógica preferencial. As propriedades da nova lógica são analisadas e são apresentados alguns exemplos.

5.1 Introdução

De forma genérica, a lógica multiagentes proposta foi desenvolvida da seguinte forma: com base numa lógica proposicional $\mathcal{L}$ (monotônica), é definida uma lógica multiagentes $\mathcal{L}^*$ (monotônica), que permite representar as crenças de um número finito de agentes. Em seguida, a partir de uma relação de preferência $\leq$ nas interpretações de $\mathcal{L}$, é definida uma relação de preferência $\preceq$ entre as interpretações de $\mathcal{L}^*$. Em função da relação $\preceq$ é definida, finalmente, uma lógica preferencial multiagentes $\mathcal{L}_{\preceq}^*$, seguindo o enfoque proposto por Shoham [Sho87, Sho88].

Apresentamos inicialmente a lógica $\mathcal{L}^*$ e em seguida a lógica $\mathcal{L}_{\preceq}^*$.

5.2 A Lógica $\mathcal{L}^*$

A lógica multiagentes $\mathcal{L}^*$ (monotônica) é obtida estendendo-se uma lógica base proposicional $\mathcal{L}$ (monotônica). A linguagem de $\mathcal{L}^*$ é a mesma de $\mathcal{L}$, aumentada com um conjunto de novos operadores $\{\mathcal{B}_0, \dots, \mathcal{B}_n\}$ para representar as crenças de cada um dos agentes $0, 1, \dots, m$. Formalmente:

Definição 5.1 Dado um conjunto P de variáveis proposicionais $\{p_0, p_1, \dots, p_m\}$, as fórmulas bem formadas (fbf) de $\mathcal{L}^*$ são o menor conjunto tal que:

1. Se $\alpha \in P$ então α é uma fbf (denominada fbf atômica).
2. Se α e β são fbf, então $\neg(\alpha)$, $(\alpha \supset \beta)$ e $\mathcal{B}_i(\alpha)$ são fbf.

A conjunção ($\wedge$), a disjunção ($\vee$) e o bicondicional ($\leftrightarrow$) estão definidos da maneira usual na lógica proposicional.

A profundidade de uma fórmula α (denotada $\text{prof}(\alpha)$) é definida por indução na estrutura da fórmula, e determina a quantidade máxima de operadores $\mathcal{B}_i$ aninhados. Formalmente:

Definição 5.2 A profundidade de uma fbf é definida como:

1. $\text{prof}(\alpha) = 0$ se $\alpha \in P$;
2. $\text{prof}(\neg\alpha) = \text{prof}(\alpha)$;
3. $\text{prof}(\alpha \supset \beta) = \max\{\text{prof}(\alpha), \text{prof}(\beta)\}$;
4. $\text{prof}(\mathcal{B}_i\alpha) = \text{prof}(\alpha) + 1$, para $i \in \{0, 1, \dots, n\}$

Para modelar as crenças de cada um dos agentes usamos as estruturas de crenças propostas por Fagin, Halpern e Vardi [FHV91]. Em nosso trabalho, a noção de mundo é um pouco diferente da noção de mundos das estruturas de Kripke. Os mundos estão definidos indutivamente¹:

Definição 5.3 Uma atribuição de crenças de ordem 0 (ou 0-atribuição) f_0 , é uma atribuição de valores de verdade às variáveis proposicionais da linguagem $\mathcal{L}$. Chama-se a $\langle f_0 \rangle$ de mundo 1-ário. Assuma que os mundos k -ários (ou, para abreviar, k -mundos) tenham sido definidos. Seja W_k o conjunto de todos os k -mundos. Uma atribuição de crenças de ordem k (ou k -atribuição) é uma função $f_k : \{0, 1, \dots, n\} \rightarrow 2^{W_k}$. Uma $(k+1)$ -seqüência de atribuição de crenças é uma seqüência $\langle f_0, \dots, f_k \rangle$, onde cada f_i , $i \in \{0, \dots, k\}$, é uma i -ésima-atribuição. Um $(k+1)$ -mundo é uma $(k+1)$ -seqüência de atribuição de crenças, que satisfaz as seguintes restrições, para cada agente i , $i \in \{0, 1, \dots, n\}$:

¹A definição de mundo é similar à definição apresentada em [FHV91] alterando-se apenas a restrição 1 da definição. Esta modificação foi feita para que o comportamento dos operadores $\mathcal{B}_i$ seja correspondente ao da lógica KD45 (crenças) em vez da lógica S5 (conhecimento).

1. $f_k(i)$ é não vazia se $k \geq 1$.
2. Se $\langle g_0, \dots, g_{k-1} \rangle \in f_k(i)$ e $k > 1$, então $g_{k-1}(i) = f_{k-1}(i)$.
3. $\langle g_0, \dots, g_{k-2} \rangle \in f_{k-1}(i)$ se e somente se existe uma $(k-1)$ -atribuição g_{k-1} tal que $\langle g_0, \dots, g_{k-1} \rangle \in f_k(i)$, se $k > 1$.

Intuitivamente, um 1-mundo é uma descrição da “realidade” e uma atribuição de crenças f_k associa a cada agente um conjunto de “mundos possíveis”. Assim, $f_k(i)$ é o conjunto de $(k-1)$ -mundos que o agente i considera “possíveis”.

Note-se que as restrições na definição de k -mundos implicam que:

- Se $\langle g_0, \dots, g_{k-1} \rangle \in f_k(i)$ e $k > 1$, então $g_j(i) = f_j(i)$, se $1 \leq j < k$.
- De maneira análoga, a “compatibilidade” entre f_k e f_{k-1} , conforme expressada pela restrição 3 anterior, também é válida entre f_{k-1} e f_j , se $0 < j < k - 1$.

As *estruturas de crenças* são definidas em função dos k -mundos, e representam, nesta abordagem, o conceito equivalente ao de estruturas de Kripke da lógica modal.

Definição 5.4 Uma seqüência $\langle f_0, f_1, \dots \rangle$ é chamada de *estrutura de crenças* se cada prefixo $\langle f_0, \dots, f_k \rangle$ é um k -mundo, para todo $k \geq 0$.

Assim, um k -mundo descreve crenças de profundidade $k-1$ e uma estrutura de crenças descreve crenças de profundidade arbitrária.

Para uma fórmula α , os mundos que têm a informação necessária para avaliá-la são os k -mundos tais que $k \geq \text{prof}(\alpha)$.

Definimos, a seguir, a noção de *satisfatibilidade* de uma fórmula α num $k+1$ -mundo $\langle f_0, \dots, f_k \rangle$ (denotado por $\langle f_0, \dots, f_k \rangle \models_{\mathcal{L}^*} \alpha$) e em seguida numa estrutura de crenças.

Definição 5.5 Um $(k+1)$ -mundo $\langle f_0, \dots, f_k \rangle$ satisfaz uma fórmula de profundidade menor ou igual a k nos seguintes casos:

- $\langle f_0, \dots, f_k \rangle \models_{\mathcal{L}^*} \alpha$ se e somente se $f_0 \models_{\mathcal{L}} \alpha$, quando $\alpha \in \mathcal{L}$
- $\langle f_0, \dots, f_k \rangle \models_{\mathcal{L}^*} \neg \alpha$ se e somente se $\langle f_0, \dots, f_k \rangle \not\models_{\mathcal{L}^*} \alpha$;
- $\langle f_0, \dots, f_k \rangle \models_{\mathcal{L}^*} \alpha \supset \beta$ se e somente se $\langle f_0, \dots, f_k \rangle \not\models_{\mathcal{L}^*} \alpha$ ou $\langle f_0, \dots, f_k \rangle \models_{\mathcal{L}^*} \beta$;

- $\langle f_0, \dots, f_k \rangle \models_{\mathcal{L}^*} B_i \alpha$ se e somente se $\langle g_0, \dots, g_{k-1} \rangle \models_{\mathcal{L}^*} \alpha$, para todo $\langle g_0, \dots, g_{k-1} \rangle \in f_k(i)$.

Pode-se demonstrar que para determinar se um k -mundo $\langle f_0, \dots, f_k \rangle$ satisfaz uma fórmula α , basta considerar o prefixo $\langle f_0, \dots, f_{\text{prof}(\alpha)} \rangle$. Formalmente:

Teorema 5.1 *Suponha que $\text{prof}(\alpha) = r$ e $k \geq r$. Então, $\langle f_0, \dots, f_k \rangle \models_{\mathcal{L}^*} \alpha$ se e somente se $\langle f_0, \dots, f_r \rangle \models_{\mathcal{L}^*} \alpha$.²*

Considerando-se o último resultado, a noção de satisfatibilidade de uma fórmula numa estrutura de crenças é definida como:

Definição 5.6 *Uma estrutura de crenças $f = \langle f_0, f_1, \dots \rangle$ satisfaz α (denotado por $f \models_{\mathcal{L}^*} \alpha$) se $\langle f_0, \dots, f_r \rangle \models_{\mathcal{L}^*} \alpha$, onde $r = \text{prof}(\alpha)$.*

Definição 5.7 *Uma fbf α é válida em $\mathcal{L}^*$ se $f \models_{\mathcal{L}^*} \alpha$, para toda estrutura de crenças f .*

Para um melhor entendimento dos conceitos anteriormente apresentados, consideremos o seguinte exemplo: suponha que só existe uma variável proposicional, ou seja, $P = \{p\}$, e que só existem dois agentes, o agente 1 e o agente 2. Suponha também que “ p é verdadeira”, que “o agente 1 não sabe se p é verdadeira ou falsa” e que “o agente 2 acredita que p é verdadeira”. Além disso, suponha que “o agente 1 acredita que o agente 2 sabe se p é verdadeira ou falsa”.

Podemos representar o exemplo anterior assim:

$$\begin{aligned} & p \wedge \\ & B_1(p \vee \neg p) \wedge B_2(p) \wedge \\ & (B_1(p \wedge B_2(p)) \vee B_1(\neg p \wedge B_2(\neg p))). \end{aligned}$$

Então, uma possível estrutura de crenças onde a fórmula anterior é verdadeira poderia ser $f = \langle f_0, f_1, f_2, \dots \rangle$, onde

$$\begin{aligned} f_0 &= \langle p \rangle \\ f_1(1) &= \{\langle p \rangle, \langle \neg p \rangle\} \quad \text{e} \quad f_1(2) = \{\langle p \rangle\} \\ f_2(1) &= \{w_1, w_2\} \quad \text{e} \quad f_2(2) = \{w_1, w_3\} \\ &\dots \end{aligned}$$

²A demonstração deste teorema encontrar-se em [FHV91].

e w_1, w_2, w_3 são os 2-mundos definidos como segue:

$$\begin{aligned} w_1 &= \langle \langle p \rangle, (f_1(1) = \{\langle p \rangle, \langle \neg p \rangle\}, f_1(2) = \{\langle p \rangle\}) \rangle, \\ w_2 &= \langle \langle \neg p \rangle, (f_1(1) = \{\langle p \rangle, \langle \neg p \rangle\}, f_1(2) = \{\langle \neg p \rangle\}) \rangle, \\ w_3 &= \langle \langle p \rangle, (f_1(1) = \{\langle p \rangle\}, f_1(2) = \{\langle p \rangle\}) \rangle. \end{aligned}$$

5.3 A Lógica $\mathcal{L}_\preceq^*$

Apresentamos neste momento uma lógica preferencial $\mathcal{L}^*$, cuja relação de preferência $\preceq$ é definida em função de uma relação de preferência, $\leq^3$, definida na lógica $\mathcal{L}$.

5.3.1 A Relação de Preferência $\preceq$

Inicialmente definiremos a extensão de uma pré-ordem $\leq$ num conjunto A , ao conjunto de partes de A (2^A).

Definição 5.8 *Seja $\leq$ uma pré-ordem definida num conjunto A . Então, $\sqsubseteq$ é uma relação em 2^A definida como segue: $A_1 \sqsubseteq A_2$ se para todo $a \in A_1$ existe $b \in A_2$ tal que $a \leq b$ e para todo $b \in A_2$ existe $a \in A_1$ tal que $a \leq b$.*

Teorema 5.2 *Se $\leq$ é uma pré-ordem definida num conjunto A , então $\sqsubseteq$ é uma pré-ordem no conjunto 2^A .*

A seguir será definida uma relação de preferência no conjunto de estruturas de crença.

Definição 5.9 *Seja $\leq$ uma pré-ordem definida no conjunto de interpretações de $\mathcal{L}$. Se f e f' são duas estruturas de crença, então $f \preceq f'$ se e somente se para cada prefixo $\langle f_0, \dots, f_{k-1} \rangle$ de f , e para cada prefixo $\langle f'_0, \dots, f'_{k-1} \rangle$ de f' se verifica que $\langle f_0, \dots, f_{k-1} \rangle \preceq_k \langle f'_0, \dots, f'_{k-1} \rangle$, onde $\preceq_k$ é definida no conjunto de k -mundos como:*

- $\langle f_0 \rangle \preceq_1 \langle f'_0 \rangle$ se e somente se $f_0 \leq f'_0$.

³ $\leq$ é uma pré-ordem limitada.

- Assuma que $\preceq_k$ foi definida indutivamente no conjunto de k -mundos, então $\preceq_{k+1}$ é definida no conjunto de $(k+1)$ -mundos como:

$\langle f_0, \dots, f_k \rangle \preceq_{k+1} \langle f'_0, \dots, f'_k \rangle$ se e somente se $\langle f_0, \dots, f_{k-1} \rangle \preceq_k \langle f'_0, \dots, f'_{k-1} \rangle$, e para cada i tal que $1 \leq i \leq n$, $f_k(i) \sqsubseteq_k f'_k(i)$, onde $\sqsubseteq_k$ é definida a partir de $\preceq_k$, como na Definição 5.8.

Teorema 5.3 Se $\leq$ é uma pré-ordem definida num conjunto de interpretações de $\mathcal{L}$, então $\preceq$ é uma pré-ordem no conjunto de estruturas de crenças de $\mathcal{L}^*$.

$\mathcal{L}^*$ com a pré-ordem $\preceq$ define uma lógica multiagentes $\mathcal{L}_\preceq^*$ que permite modelar as crenças de agentes que raciocinam não-monotonicamente.

5.4 Interpretação Epistêmica da Lógica $\mathcal{L}_\preceq^*$

Os conceitos apresentados no capítulo 3 podem ser generalizados para modelar mais de dois agentes. Assim, a lógica multiagentes que acabamos de apresentar pode ser interpretada no contexto das lógicas epistemicamente estendidas. Podemos pensar que a lógica base $\mathcal{L}_\preceq$ permite representar as crenças de um agente que, segundo seu ponto de vista, define uma relação de preferência no conjunto de interpretações da lógica $\mathcal{L}$. Então as fórmulas de $\mathcal{L}$ representam a “realidade”, segundo a visão de um agente “distinguido” que chamamos de sistema ou agente 0. A partir dessa relação, é definida a extensão epistêmica de $\mathcal{L}_\preceq$, que continua representando o ponto de vista particular de um agente. Assim, a fórmula

$$\alpha \wedge B_1(\gamma \supset B_2\gamma)$$

pode ser lida como “o sistema acredita em α e acredita que o agente 1 acredita que se γ é verdadeira então o agente 2 acredita nela”.

Portanto, a lógica multiagentes representa o ponto de vista do sistema. Neste contexto o operador modal B_0 pode:

- representar a atribuição de crenças feitas ao sistema por um outro agente. Logo, a fórmula $B_i B_0 \alpha$ é interpretada como “o sistema acredita que o agente i acredita que o sistema (ou agente 0) acredita em α ”.
- representar as crenças introspectivas do sistema. Para que isto seja feito corretamente, deve-se pedir que a fórmula

$$\alpha \leftrightarrow B_0 \alpha \tag{5.1}$$

seja uma fórmula válida, fora do escopo de qualquer operador modal B_i , com $i \neq 0$. Assim, por exemplo, não desejamos que as seguinte fórmulas sejam válidas em nossa lógica:

$$B_i(\alpha \leftrightarrow B_0\alpha) \quad (5.2)$$

$$B_i(B_0\alpha) \leftrightarrow B_i\alpha \quad (5.3)$$

$$(5.4)$$

Para que 5.1 seja válida devemos redefinir as estruturas de crenças.

Considerando-se as justificativas anteriores, a Definição 5.4 é substituída pela seguinte definição:

Definição 5.10 *A seqüência $\langle f_0, f_1, \dots \rangle$ é uma estrutura de crenças se cada prefixo $\langle f_0, \dots, f_k \rangle$ é um k -mundo, para cada k , e $f_k(0) = \{\langle f_0, f_1, \dots, f_{k-1} \rangle\}$, para $k > 0$.*

Com esta mudança, serão analisadas a seguir as propriedades da lógica.

5.5 Propriedades da Lógica $\mathcal{L}_\leq^*$

A lógica $\mathcal{L}_\leq^*$ que acabamos de definir tem as propriedades exigidas no capítulo 3 para uma lógica epistemicamente estendida.

A primeira propriedade de $\mathcal{L}_\leq^*$ considera o fato de $\mathcal{L}_\leq^*$ ser uma extensão conservativa de $\mathcal{L}_\leq$:

Teorema 5.4 *Se α e β são duas fórmulas $\mathcal{L}_\leq$, então $\alpha \models_\leq \beta$ se e somente se $\alpha \models_\leq^* \beta$.*

Como as fórmulas sem o operador de crença representam as crenças do sistema sobre o mundo, esta propriedade estabelece que o sistema continua raciocinando sobre o mundo usando a lógica $\mathcal{L}_\leq$.

Nesta lógica o segundo requisito apresentado no capítulo 3 é formulado na lógica $\mathcal{L}_\leq^*$ como segue:

Teorema 5.5 *$\alpha \models_\leq \beta$ se e somente se $B_{i_1} \dots B_{i_r} \alpha \models_\leq B_{i_1} \dots B_{i_r} \beta$, quando $r > 0$.*

Uma terceira propriedade trata do comportamento dos operadores B_i . Cada B_i satisfaz os requisitos da lógica $KD45$, comumente utilizada para caracterizar crenças.

Teorema 5.6 *Em $\mathcal{L}_\leq^*$ é válido:*

1. *Todas as instâncias das tautologias.*
2. $B_i(\alpha) \supset B_i B_i(\alpha)$.
3. $\neg B_i(\alpha) \supset B_i \neg B_i(\alpha)$.
4. $(B_i(\alpha) \wedge B_i(\alpha \supset \beta)) \supset B_i(\beta)$.

Um outro resultado estabelece a independência do raciocínio realizado exclusivamente fora do escopo dos operadores de crença, e do raciocínio feito dentro do escopo dos operadores de crença:

Teorema 5.7 *Se $\alpha \models_{\leq} \beta$ e $\gamma \models_{\leq} \delta$ então $\alpha \wedge B_{i_1} \dots B_{i_k} \gamma \models_{\leq} \beta \wedge B_{i_1} \dots B_{i_k} \delta$, para $k > 0$ e $i_1 \neq 0$.*

Isto estabelece que as inferências default, realizadas de forma independente em $\mathcal{L}_\leq$, podem ainda ser realizadas mesmo se o antecedente de uma delas está dentro do escopo dos operadores de crença e o antecedente da outra não. O aspecto interessante desta propriedade é que, embora α e γ interfiram uma com a outra em $\mathcal{L}_\leq$, a interferência não ocorre se uma delas estiver dentro do escopo de operadores de crença em $\mathcal{L}_\leq$.

Os teoremas anteriores estabelecem propriedades genéricas da lógica $\mathcal{L}_\leq^*$, independentemente de como é definida a relação de preferência base $\leq$. Mostraremos, a seguir, que se a lógica base é circunscrição proposicional, então as regras defaults que combinam subfórmulas, com e sem o operador de crença, são corretamente tratadas.

Se $\leq_x$ é uma relação de preferência que minimiza a variável proposicional abn_1 , e se $\preceq_x$ é a relação de preferência correspondente para as interpretações de $\mathcal{L}^*$, então os seguintes resultados são verdadeiros:

- $B_i \alpha \wedge (B_i \alpha \wedge \neg abn_1 \supset \beta) \models_{\preceq_x} \beta$,
- $\alpha \wedge (\alpha \wedge \neg abn_1 \supset B_i \beta) \models_{\preceq_x} B_i \beta$,
- $\alpha \wedge (\alpha \wedge \neg abn_1 \supset B_i \beta) \wedge B_i(\neg \beta) \not\models_{\preceq_x} B_i \beta$.

Os resultados anteriores são válidos também se B_i é substituído por qualquer seqüência $B_{i_1}B_{i_2}\dots B_{i_r}$, para $r > 0$ e $i_1 \neq 0$. Os dois primeiros exemplos mostram que se a subfórmula com o operador B_i é o antecedente ou o conseqüente da regra default, e o antecedente é verdadeiro, então a conseqüência default será inferida. O último exemplo mostra que a lógica, quando determina se uma conseqüência default pode ser consistentemente afirmada, considera a lógica dos operadores B_i .

5.6 Exemplos

Para exemplificar as formas de raciocínio desta lógica preferencial multiagentes podemos considerar a seguinte situação: Você (agente 0) está com um amigo (agente 1) num clube aguardando um outro amigo (agente 2). Se o agente 2 sai do trabalho as 17 horas (*sai_17*) então ele chega no clube, geralmente, as 18 horas (*chega_18*), mas isso não acontece se há um engarrafamento na estrada (*eng*). Podemos supor, também, que todos os agentes conhecem estes fatos.

Começemos resolvendo o problema da representação da situação descrita anteriormente. O default “se alguém sai do trabalho as 17 horas então, geralmente, chega no clube as 18 horas” pode ser representado simplesmente como:

$$sai_17 \wedge \neg abn_1 \supset chega_18 \quad (5.5)$$

numa lógica preferencial que circunscreve a variável proposicional abn_1 . Temos agora que determinar como representar que o agente 1 acredita em (5.5), de maneira tal que o default se comporte corretamente. Mas, como a lógica preferencial multiagentes foi desenvolvida com base numa relação de preferência proposicional, de tal forma que cada agente da sociedade raciocina (sobre o mundo) usando essa relação de preferência, então a representação do fato que o agente 1 acredita em (5.5) é:

$$B_1(sai_17 \wedge \neg abn_1 \supset chega_18) \quad (5.6)$$

pois a lógica onde serão feitos os raciocínios é uma lógica preferencial multiagentes, que usa a relação $\preceq_x$ derivada da circunscrição proposicional de abn_1 ($\leq_x$).

Esta forma de representação é adequada, pois se $\leq_x$ é a relação que prefere os modelos onde abn_1 é falso, então podemos obter aplicando o Teorema 5.5:

$$B_1(sai_17) \wedge B_1(sai_17 \wedge \neg abn_1 \supset chega_18) \models_{\preceq_x} B_1(chega_18). \quad (5.7)$$

Continuando o exemplo, suponha que você saiba que “o agente 2 deixou o trabalho às 17 horas”. Suponha também que o agente 1 fale para você que “há um engarrafamento na

estrada e que ele acredita que o agente 2 não saiba disso". Por outro lado, você acredita que "não há engarrafamento" porque acaba de chegar no clube e a estrada tinha o tráfego normal nessa hora. Então, segundo as crenças de cada agente, você acredita que "o agente 2 vai chegar às 18 horas", o agente 1 acredita que "o agente 2 não vai chegar na hora mas acredita que o agente 2 pensa que vai chegar a tempo". Este tipo de raciocínio é modelado pela lógica $\mathcal{L}_{\preceq_x}$, onde $\preceq_x$ é a relação de preferência entre estruturas de crenças derivada da relação $\leq_x$.

Assim, o raciocínio anterior pode ser expressado como:

$$\left(\begin{array}{l} sai_17 \wedge \\ \neg eng \wedge \\ \alpha \end{array} \right) \wedge \left(\begin{array}{l} B_1(sai_17) \wedge \\ B_1(eng) \wedge \\ B_1(\alpha) \end{array} \right) \wedge \left(\begin{array}{l} B_1 B_2(sai_17) \wedge \\ B_1 B_2(\alpha) \end{array} \right) \models_{\preceq_x} \begin{array}{l} chega_18 \wedge \\ B_1(\neg chega_18) \wedge \\ B_1 B_2(chega_18) \end{array}$$

onde α é a abreviatura de:

$$(sai_17 \wedge \neg abn_1 \supset chega_18) \wedge (sai_17 \wedge eng \supset \neg chega_18)$$

O resultado anterior é deduzido aplicando-se os Teoremas 5.5 e 5.7.

O exemplo anterior ilustra o fato de que os defaults que estão dentro do escopo do operador B_i comportam-se conforme o esperado.

Considere agora que os três agentes acreditam que "se chove então geralmente há um engarrafamento na estrada", e que "se está chovendo todo mundo pode perceber isso". Introduzimos as seguintes abreviaturas:

$$\begin{aligned} \beta &\equiv \alpha \wedge (chove \wedge \neg abn_2 \supset eng) \\ \gamma &\equiv (chove \wedge \neg abn_3 \supset B_1 chove) \end{aligned}$$

Suponha também que temos definida uma relação de preferência proposicional $\leq_y$, onde as variáveis proposicionais abn_2 e abn_3 são minimizadas com maior prioridade que abn_1 . Então, se você sabe que "está chovendo", então você pode concluir que "o agente 1 acredita que o agente 2 não vá chegar às 18 horas".

$$chove \wedge \gamma \wedge B_1(sai_17) \wedge B_1(\beta) \models_{\preceq_y} B\neg chega_18$$

Este exemplo ilustra uma forma de encadeamento através dos contextos de crença do agente 0 e os outros agentes que interagem com ele. Pode-se notar neste exemplo, que

no espaço de crenças do agente 1 o default “sair do trabalho às 17 horas” normalmente implica “chegar no clube às 18 horas”, mas é desabilitado pelo fato de que “se chove há um engarrafamento e não chega às 18 horas”.

Capítulo 6

Conclusões e Extensões

Neste trabalho foram apresentadas as extensões epistêmicas de duas famílias de lógicas não-monotônicas.

Com relação às extensões epistêmicas de $CT4O$ e CO , ambas estendem suas lógicas base para poder tratar fórmulas que possuem um operador de crença, de tal maneira que todas as propriedades das lógicas originais sejam mantidas e os requisitos, apresentadas no capítulo 3, das lógicas epistemicamente estendidas sejam satisfeitos. Além disso, acreditamos que a metodologia aplicada para desenvolver as lógicas $CT4O^e$ e CO^e (ou seja, o uso A -mundos, B -mundos, e a definição de $\geq_\square$ baseada na relação $\geq$) pode ser usada para estender algumas das outras lógicas condicionais.

Este enfoque tem, porém, algumas limitações. Como foi mencionado no capítulo 2, [Bou94b] propõe a lógica CO^* , que oferece uma solução limitada para o problema da irrelevância. Na lógica CO^* , a solução do problema da irrelevância depende do fato de que a relação de normalidade seja totalmente conexa e do fato de que todos os mundos possíveis estejam presentes. Mas, seguindo a metodologia proposta, não é possível desenvolver uma extensão epistêmica de CO^* , já que não é possível construir uma estrutura onde todos os possíveis B -mundos estejam presentes e $\geq_\square$ seja uma relação totalmente conexa.

Para resolver o problema da informação irrelevante, o que poderia ser feito é fortalecer a lógica base com características extra lógicas, tal como é feito na proposta de Delgrande [Del88] apresentada no capítulo 2. Ou, melhor ainda, poderia-se definir uma relação de preferência no conjunto de modelos, tal como é feito no trabalho mais recente de Delgrande [Del94] e, a partir dessa lógica, definir uma extensão epistêmica, onde a relação de preferência leve em conta tanto as crenças do sistema como as crenças do outro agente.

Com relação à linguagem utilizada nas extensões, devemos lembrar que ela não permite

representar fórmulas com ocorrências aninhadas do operador de crenças. Uma possível extensão do trabalho realizado seria considerar uma linguagem que incluía esse tipo de fórmula e estender a semântica de forma que o operador de crenças tenha o comportamento correspondente à lógica *KD45*. Também deveria ser considerada uma extensão da linguagem para poder representar as crenças de mais de dois agentes.

Por outro lado, como as lógicas *CT40* e *CO* têm caracterizações sintáticas corretas e completas, deixamos como um trabalho futuro a caracterização sintática das lógicas estendidas.

Com relação à lógica preferencial multiagentes apresentada, ela é uma extensão do trabalho apresentado em [Wai93]. Na nova lógica, cada agente raciocina independentemente, utilizando a lógica base não-monotônica. Neste trabalho foram utilizadas estruturas de crenças para definir a semântica, em lugar das tradicionais estruturas de Kripke.

Até onde é conhecido, esta é a primeira proposta de uma lógica multiagentes não-monotônica baseada num enfoque preferencial, já que as lógicas multiagentes propostas na literatura estão baseadas na lógica autoepistêmica, que só modela alguns aspectos do raciocínio default.

A extensão multiagentes da lógica preferencial proposta permite representar, em princípio, as crenças de um agente distinguido (o sistema) e as crenças de outros agentes que interagem com ele. Mas essas crenças (dos outros agentes) são enxergadas do ponto de vista do sistema. A relação de preferência, por sua vez, corresponde a uma relação de preferência compatível com as preferências do sistema. Portanto, o enfoque adotado não é um enfoque simétrico (com relação as crenças dos agentes). Um trabalho futuro, relacionado com a lógica aqui desenvolvida, seria analisar a possibilidade de considerar uma relação de preferência diferente para cada agente.

Um outro trabalho futuro, associado com a lógica preferencial multiagentes, está relacionado à utilização de uma lógica de primeira ordem como lógica base. As estruturas de crenças deveriam ser modificadas adequadamente e deveria-se analisar as propriedades dessa extensão.

Apêndice A

Demonstrações dos Teoremas do Capítulo 4

Teorema 4.2 *Seja A um conjunto de A -mundos. A relação $\geq_\Delta$ definida em A é uma relação reflexiva e transitiva.*

Demonstração: A demonstração é óbvia, a partir da definição da relação e do fato que a relação de normalidade dos A -mundos é reflexiva e transitiva.

Teorema 4.3 *A relação $\geq_\square$, com base numa relação $\geq$ reflexiva e transitiva, também é reflexiva e transitiva.*

Demonstração:

A demonstração é óbvia, a partir da definição da relação e do fato que a relação de normalidade na qual está baseada $\geq_\square$ é reflexiva e transitiva.

Teorema 4.4 *Se $\alpha \in \mathcal{L}_c$ então: $\models_{CT4O} \alpha$ se e somente se $\models_{CT4O^c} \alpha$.*

Demonstração:

$\Rightarrow$)

Primeiro vamos demonstrar que $\models_{CT4O} \alpha$ então $\models_{CT4O^c} \alpha$. Na realidade, o que vai ser demonstrado é um enunciado equivalente

$$\not\models_{CT4O^c} \alpha \text{ então } \not\models_{CT4O} \alpha.$$

Suponha então que $\not\models_{CT4O^c} \alpha$, então existe um $CT4O^c$ -modelo $M^c = \langle \mathcal{X}, \geq \rangle$ tal que $M^c \not\models \alpha$.

Se $\mathcal{X} = \{x_1, x_2, \dots\}$ é o conjunto de mundos de crenças e $\mathbf{P}$ o conjunto de variáveis proposicionais, então introduzimos um conjunto $\mathbf{P}_{\mathcal{X}}$ de novas variáveis proposicionais, tal que $\mathcal{X}$ e $\mathbf{P}_{\mathcal{X}}$ tem a mesma quantidade de elementos. Então podemos pensar que associado a cada B -mundo $x \in \mathcal{X}$ existe um único $p_x \in \mathbf{P}_{\mathcal{X}}$. Também pode-se associar a cada $x \in \mathcal{X}$ um mundo possível proposicional $w_x = \Pi_1(x) \cup \{p_x\}$

Considere então um novo conjunto de variáveis proposicionais $\mathbf{P}' = \mathbf{P} \cup \mathbf{P}_{\mathcal{X}}$ e um conjunto de mundos possíveis proposicionais $W' = \{w_x : x \in \mathcal{X}\}$.

Em W' definimos uma relação binária, $\geq'$ da seguinte forma:

$$w_x \geq' w_{x'} \text{ se e somente se } x \geq_{\square} x', \text{ com } w_x \text{ e } w_{x'} \in W'.$$

A relação $\geq'$ assim definida é reflexiva e transitiva, já que $\geq_{\square}$ tem essas propriedades.

Então $M' = \langle W', \geq' \rangle$, é um $CT4O$ -modelo, isomorfo a M .

Portanto, como $M \not\models_{CT4O^e} \alpha$ então $M' \not\models_{CT4O} \alpha$, já que determinar a validade da fórmula α no modelo M' é equivalente a determinar a validade no modelo M .

Logo, pelo resultado anterior, temos que $\not\models_{CT4O} \alpha$.

$\Leftarrow$)

Vamos demonstrar agora que se $\models_{CT4O^e} \alpha$ então $\models_{CT4O} \alpha$.

Vamos supor que $\models_{CT4O^e} \alpha$.

Seja $M = \langle W, \geq \rangle$ um $CT4O$ -modelo. Vamos demonstrar que $M \models_{CT4O} \alpha$.

A partir de M vamos a construir um $CT4O^e$ -modelo.

Seja $M^e = \langle \mathcal{X}, \geq \rangle$, onde $\mathcal{X} = \{\langle w, \{\langle w, M \rangle\} \rangle : w \in W\}$. Portanto M^e é um $CT4O^e$ -modelo, tal que M^e com a relação $\geq_{\square}$ é isomorfo ao modelo M .

Como $\models_{CT4O^e} \alpha$ e M^e é um $CT4O^e$ -modelo, então $M^e \models_{CT4O^e} \alpha$. Logo, como M^e e M são isomorfos, e levando em conta a definição de M^e , resulta que $M \models_{CT4O} \alpha$.

Portanto $\models_{CT4O} \alpha$ ■

Corolário 4.1 Se $\alpha, \beta \in \mathcal{L}_C$ então: $\models_{CT4O} \alpha \Rightarrow \beta$ se e somente se $\models_{CT4O^c} \alpha \Rightarrow \beta$.

Demonstração:

A demonstração é óbvia a partir do teorema anterior.

Teorema 4.5 Se α e β são fbf de $\mathcal{L}_C$, então

$$\models_{CT4O} \alpha \Rightarrow \beta \text{ se e somente se } \models_{CT4O^c} \mathcal{B}(\alpha) \Rightarrow \mathcal{B}(\beta).$$

Demonstração:

$\Rightarrow$)

Suponha que $\models_{CT4O} \alpha \Rightarrow \beta$.

Seja $M = \langle \mathcal{X}, \geq \rangle$ um $CT4O^c$ -modelo e seja $x \in \mathcal{X}$ um B -mundo. Será demonstrado que $M, x \models_{CT4O^c} \mathcal{B}(\alpha) \Rightarrow \mathcal{B}(\beta)$.

Mas, para demonstrar $M, x \models_{CT4O^c} \mathcal{B}(\alpha) \Rightarrow \mathcal{B}(\beta)$, segundo a definição de $\Rightarrow$, deve ser demonstrado que $M, x \models \Box (\Box \neg \mathcal{B}(\alpha) \vee \Diamond (\mathcal{B}(\alpha) \wedge \Box (\mathcal{B}(\alpha) \supset \mathcal{B}(\beta))))$.

Então, pela definição de satisfatibilidade do operador $\Box$, deve ser demonstrado que $M, x' \models \Box \neg \mathcal{B}(\alpha) \vee \Diamond (\mathcal{B}(\alpha) \wedge \Box (\mathcal{B}(\alpha) \supset \mathcal{B}(\beta)))$, para todo $x' \in \mathcal{X}$.

Seja $x' \in \mathcal{X}$, se $M, x' \models \Box \neg \mathcal{B}(\alpha)$ então $M, x' \models \Box \neg \mathcal{B}(\alpha) \vee \Diamond (\mathcal{B}(\alpha) \wedge \Box (\mathcal{B}(\alpha) \supset \mathcal{B}(\beta)))$.

Se $M, x' \not\models \Box \neg \mathcal{B}(\alpha)$, então demonstraremos que $M, x' \models \Diamond (\mathcal{B}(\alpha) \wedge \Box (\mathcal{B}(\alpha) \supset \mathcal{B}(\beta)))$.

Mas como $M, x' \not\models \Box \neg \mathcal{B}(\alpha)$, então existe x'' tal que $x' \geq_\Box x''$ e $M, x'' \models \mathcal{B}(\alpha)$. Então, pela definição de satisfatibilidade de uma fórmula do tipo $\mathcal{B}(\alpha)$, temos que $\Pi_2(c), \Pi_1(c) \models \alpha$, para todo $c \in \Pi_2(x'')$.

Levando em consideração que $\models_{CT4O} \alpha \Rightarrow \beta$, que $\Pi_2(c), \Pi_1(c) \models \alpha$, para todo $c \in \Pi_2(x'')$ e que $\Pi_2(c)$ é um $CT4O$ -modelo, então para todo $c \in \Pi_2(x'')$ existe w_c tal que $\Pi_1(c) \geq w_c$ e $\Pi_2(c), w_c \models \alpha \wedge \Box (\alpha \supset \beta)$.

Mas, pela propriedade dos $CT4O^c$ -modelos, existe w_0 tal que $\langle w_0, \{ \langle w_c, \Pi_2(c) \rangle \}_{c \in \Pi_2(x'')} \rangle \in \mathcal{X}$. Seja então $x''' = \langle w_0, \{ \langle w_c, \Pi_2(c) \rangle \}_{c \in \Pi_2(x'')} \rangle$, x''' é tal que $x'' \geq_\Box x'''$ e $M, x''' \models \mathcal{B}(\alpha)$, pois temos que $\Pi_2(c), w_c \models_{CT4O} \alpha \wedge \Box (\alpha \supset \beta)$.

Além disso para todo x^{iv} tal que $x''' \geq_\square x^{iv}$ e $M, x^{iv} \models_{CT4O^*} B(\alpha)$, temos que $M, x^{iv} \models_{CT4O^*} B(\beta)$, já que $\Pi_2(c), w_c \models_{CT4O} \alpha \wedge \Box(\alpha \supset \beta)$.

Então $M, x''' \models B(\alpha) \wedge \Box(B(\alpha) \supset B(\beta))$.

Logo, como $x'' \geq_\square x'''$ e $x' \geq_\square x''$, então $M, x' \models_{CT4O^*} \Diamond(B(\alpha) \wedge \Box(B(\alpha) \supset B(\beta)))$.

Então $M, x \models_{CT4O^*} B(\alpha) \Rightarrow B(\beta)$.

Com o qual $M \models_{CT4O^*} B(\alpha) \Rightarrow B(\beta)$.

Portanto temos $\models_{CT4O^*} B(\alpha) \Rightarrow B(\beta)$

$\Leftarrow$)

Reciprocamente demonstraremos que se $\models_{CT4O^*} B(\alpha) \Rightarrow B(\beta)$, então $\models_{CT4O} \alpha \Rightarrow \beta$.

Mas, ao invés de demonstrar isto, será demonstrado um enunciado equivalente:

$\not\models_{CT4O} \alpha \Rightarrow \beta$ então $\not\models_{CT4O^*} B(\alpha) \Rightarrow B(\beta)$.

Suponha que $\not\models_{CT4O} \alpha \Rightarrow \beta$. Logo, existe $M = \langle W, \geq \rangle$, $CT4O$ -modelo, tal que $M \not\models_{CT4O} \alpha \Rightarrow \beta$.

Então existe $w_0 \in W$ tal que $M, w_0 \not\models_{CT4O} \alpha \Rightarrow \beta$.

Seja $M' = \langle \mathcal{X}, \geq \rangle$, onde $\geq$ é a relação do modelo M e $\mathcal{X} = \{ \langle w_1, \{ \langle w_1, M \rangle \} \rangle : w_1 \in W \}$.

Segundo a definição M' é um $CT4O^*$ -modelo. Além disso, M' com a relação $\geq_\square$ é isomorfo a M . Como a segunda componente dos B -mundos de M' são conjuntos unitários tais que a primeira componente do A -mundo, coincide com a primeira componente do B -mundo correspondente e pelo isomorfismo temos então que:

$$M', \langle w_0, \{ \langle w_0, M \rangle \} \rangle \not\models_{CT4O^*} B(\alpha) \Rightarrow B(\beta).$$

Portanto $M' \not\models_{CT4O^*} B(\alpha) \Rightarrow B(\beta)$.

Então foi demonstrado que:

$$\not\models_{CT4O} \alpha \Rightarrow \beta \text{ então } \not\models_{CT4O^*} B(\alpha) \Rightarrow B(\beta) \blacksquare$$

Teorema 4.6 *O seguintes resultados são válidos em $CT4O^e$:*

1. $\models_{CT4O^e} ((\alpha \Rightarrow B(\beta)) \wedge \alpha) \Rightarrow B(\beta)$.
2. $\models_{CT4O^e} ((B(\alpha) \Rightarrow \beta) \wedge B(\alpha)) \Rightarrow \beta$.

Demonstração:

1) Seja $M = \langle \mathcal{X}, \geq \rangle$ um $CT4O^e$ -modelo e x um B -mundo em $\mathcal{X}$.

Se $M, x \models \Box \neg((\alpha \Rightarrow B(\beta)) \wedge \alpha)$ então $M, x \models_{CT4O^e} \Box \neg((\alpha \Rightarrow B(\beta)) \wedge \alpha) \vee \Diamond((\alpha \Rightarrow B(\beta)) \wedge \alpha \wedge \Box((\alpha \wedge \alpha \Rightarrow B(\beta)) \supset B(\beta)))$.

Caso contrario, se $M, x \not\models \Box \neg((\alpha \Rightarrow B(\beta)) \wedge \alpha)$, então $M, x \models \Diamond((\alpha \Rightarrow B(\beta)) \wedge \alpha)$. Então existe $x' \in \mathcal{X}$, tal que $x \geq x'$ e $M, x' \models_{CT4O^e} (\alpha \Rightarrow B(\beta)) \wedge \alpha$. Portanto $M, x' \models_{CT4O^e} \alpha \Rightarrow B(\beta)$ e $M, x' \models_{CT4O^e} \alpha$

Então pela definição de $\Rightarrow$, existe x'' tal que $x' \geq x''$ e $M, x'' \models_{CT4O^e} \alpha \wedge \Box(\alpha \supset B(\beta))$.

Então $M, x'' \models_{CT4O^e} \alpha \wedge (\alpha \Rightarrow B(\beta)) \wedge B(\beta)$.

Logo $M, x'' \models_{CT4O^e} (\alpha \Rightarrow B(\beta)) \wedge \alpha \wedge \Box((\alpha \wedge \alpha \Rightarrow B(\beta)) \supset B(\beta))$.

Portanto $M, x' \models_{CT4O^e} \Diamond((\alpha \Rightarrow B(\beta)) \wedge \alpha \wedge \Box((\alpha \wedge \alpha \Rightarrow B(\beta)) \supset B(\beta)))$.

Então temos que $M, x \models_{CT4O^e} \Box \neg((\alpha \Rightarrow B(\beta)) \wedge \alpha) \vee \Diamond((\alpha \Rightarrow B(\beta)) \wedge \alpha \wedge \Box((\alpha \wedge \alpha \Rightarrow B(\beta)) \supset B(\beta)))$.

Com o qual temos o resultado desejado:

$$M \models_{CT4O^e} ((\alpha \Rightarrow B(\beta)) \wedge \alpha) \Rightarrow B(\beta)$$

A demonstração da parte 2) é similar à anterior ■

Teorema 4.7 *Seja M um $CT4O^e$ -modelo; então:*

1. $M \models_{CT4O^e} B(\alpha \wedge \beta)$ se e somente se $M \models_{CT4O^e} B(\alpha) \wedge B(\beta)$.
2. $M \models_{CT4O^e} B(\neg\alpha)$ então $M \models_{CT4O^e} \neg B(\alpha)$.
3. $\models_{CT4O^e} B(\alpha \wedge (\alpha \supset \beta)) \supset B(\beta)$.
4. $\models_{CT4O^e} B(\alpha \wedge (\alpha \Rightarrow \beta)) \Rightarrow B(\beta)$.

Demonstração:

1)

$M \models_{CT4O^e} B(\alpha \wedge \beta)$ se e somente se $M, x \models_{CT4O^e} B(\alpha \wedge \beta)$, para todo $x \in \mathcal{X}$
 se e somente se $\Pi_2(c), \Pi_1(c) \models_{CT4O} \alpha \wedge \beta$, para todo c tal que $c \in \Pi_2(x)$, para todo $x \in \mathcal{X}$
 se e somente se $\Pi_2(c), \Pi_1(c) \models_{CT4O} \alpha$ e $\Pi_2(c), \Pi_1(c) \models_{CT4O} \beta$, para todo c tal que $c \in \Pi_2(x)$, para todo $x \in \mathcal{X}$
 se e somente se $M, x \models_{CT4O^e} B(\alpha)$ e $M, x \models_{CT4O^e} B(\beta)$, para todo $x \in \mathcal{X}$
 se e somente se $M, x \models_{CT4O^e} B(\alpha) \wedge B(\beta)$, para todo $x \in \mathcal{X}$
 se e somente se $M \models_{CT4O^e} B(\alpha)$ e $M \models_{CT4O^e} B(\beta)$

2)

Se $M \models_{CT4O^e} B(\neg\alpha)$ então $M, x \models_{CT4O^e} B(\neg\alpha)$, para todo $x \in \mathcal{X}$.

Portanto $\Pi_2(c), \Pi_1(c) \models_{CT4O} \neg\alpha$, para todo c tal que $c \in \Pi_2(x)$, para todo $x \in \mathcal{X}$.

Então $M, x \not\models_{CT4O^e} B(\alpha)$. Portanto $M, x \models_{CT4O^e} \neg B(\alpha)$.

Logo $M \models_{CT4O^e} \neg B(\alpha)$.

3)

Seja $x \in \mathcal{X}$.

Se $M, x \models_{CT4O^e} B(\alpha \wedge (\alpha \supset \beta))$ então $\Pi_2(c), \Pi_1(c) \models_{CT4O} \alpha \wedge (\alpha \supset \beta)$, para todo c tal que $c \in \Pi_2(x)$.

Então $\Pi_2(c), \Pi_1(c) \models_{CT4O} \beta$, para todo c tal que $c \in \Pi_2(x)$, para todo $x \in \mathcal{X}$.

Então $M, x \models_{CT4O^e} B(\beta)$, para todo $x \in \mathcal{X}$.

Logo $\models_{CT4O^e} B(\alpha \wedge (\alpha \supset \beta)) \supset B(\beta)$.

4)

Sai a partir do teorema 4.5.

Teorema 4.8 *O seguintes resultados são válidos em $CT4O^e$:*

1. $\nVdash_{CT4O^e} (\alpha \Rightarrow B(\beta) \wedge \alpha \wedge \neg B(\beta)) \Rightarrow B(\beta)$.
2. $\nVdash_{CT4O^e} (\alpha \Rightarrow B(\beta) \wedge \alpha \wedge B(\neg\beta)) \Rightarrow B(\beta)$.

Demonstração:

Para demonstrar o ponto 1 basta definir o $CT4O^e$ -modelo que segue.

Seja $P = \{p, q\}$.

Sejam $w_0 = \{\}$, $w_1 = \{q\}$, $w_2 = \{p\}$ e $w_3 = \{p, q\}$ e $W = \{w_0, w_1, w_2, w_3\}$

Os A -mundos são:

$a_0 = (w_1, E)$, $a_1 = (w_2, E)$, $a_2 = (w_3, E)$, onde $E = \langle W, \geq \rangle$ e $\geq$ está definida por $w_0 \geq w_1$, $w_1 \geq w_2$, $w_2 \geq w_3$.

Seja $\mathcal{X} = \{x_0, x_1, x_2, x_3\}$, onde os B -mundos são:

$$x_0 = (w_0, \{a_0\}), x_1 = (w_1, \{a_1\}), x_2 = (w_2, \{a_1\}), x_3 = (w_3, \{a_2\}),$$

Então, considerado as fórmulas $\alpha \equiv \neg p$ e $\beta \equiv q$, $M = \langle \mathcal{X}, \geq \rangle$ é um $CT4O$ -modelo tal que:

$$M \models_{CT4O^e} \alpha \Rightarrow B(\beta), M, x_0 \models_{CT4O^e} \alpha \text{ e } M, x_0 \models_{CT4O^e} \neg B(\beta)$$

e $M, x \nVdash_{CT4O^e} \alpha \wedge \neg B(\beta)$, em todo outro mundo x diferente de x_0 . Portanto como $M, x_0 \nVdash_{CT4O^e} B\beta$ então $\nVdash_{CT4O^e} ((\alpha \Rightarrow B\beta) \wedge \alpha \wedge \neg B(\beta)) \Rightarrow B(\beta)$.

O ponto 2 é demonstrado de maneira semelhante ao 1 ■

Apêndice B

Demonstrações dos Teoremas do Capítulo 5

Teorema 5.2 *Se $\leq$ é uma pré-ordem definida num conjunto A , então $\sqsubseteq$ é uma pré-ordem no conjunto 2^A .*

Demonstração: A demonstração é óbvia, a partir da definição da relação $\sqsubseteq$ e do fato que a relação $\leq$ é reflexiva e transitiva.

Teorema 5.3 *Se $\leq$ é uma pré-ordem definida num conjunto de interpretações de $\mathcal{L}$, então $\preceq$ é uma pré-ordem no conjunto de estruturas de crenças de $\mathcal{L}^*$.*

Demonstração:

A demonstração é óbvia, a partir da definição da relação $\preceq$ e do teorema anterior.

Teorema 5.4 *Se α e β são duas fórmulas $\mathcal{L}_{\leq}$, então $\alpha \models_{\leq} \beta$ se e somente se $\alpha \models_{\preceq} \beta$.*

Demonstração:

$\Rightarrow$)

Se $\alpha \models_{\leq} \beta$, então para toda interpretação M , se $M \models_{\leq} \alpha$ então $M \models \beta$.

Seja $f = \langle f_0, f_1, \dots \rangle$ uma estrutura de crenças tal que $f \models_{\preceq} \alpha$. Logo, temos que $\langle f_0 \rangle \models_{\mathcal{L}^*} \alpha$, com o qual $f_0 \models_{\mathcal{L}} \alpha$.

Mas f_0 é uma interpretação tal que não existe nenhuma outra interpretação M tal que $M < f_0$ e $M \models_{\mathcal{L}} \alpha$, já que, se existisse uma tal M , pode-se definir $f' = \langle f'_0, f'_1, \dots \rangle$, tal que

$$f'_0 = M \text{ e}$$

$$h_k(i) = \begin{cases} f_k(i) & \text{para todo } i \neq 0 \\ \{\langle f_0, f_1, \dots, f_{k-1} \rangle\} & \text{para } i = 0 \end{cases}$$

para todo $k > 0$. A estrutura assim definida é tal que $\langle f'_0, f'_1, \dots \rangle \models_{\mathcal{L}} \alpha$ e $\langle f'_0, f'_1, \dots \rangle < \langle f_0, f_1, \dots \rangle$, o que não pode acontecer, já que, $f \models_{\mathcal{L}} \alpha$.

Então, como $\alpha \models_{\leq} \beta$ e $f_0 \models_{\leq} \alpha$, então $f_0 \models_{\mathcal{L}} \beta$ e portanto $f \models_{\mathcal{L}} \beta$.

Logo, temos que $\alpha \models_{\leq} \beta$.

$\Leftarrow$)

Reciprocamente, suponha que $\alpha \models_{\leq} \beta$.

Se M é uma interpretação tal que $M \models_{\leq} \alpha$, então podemos definir uma estrutura de crenças $f = \langle M, f_1, f_2, \dots \rangle$ na qual $f_1(i) = \{\langle M \rangle\}$, para todo i tal que $1 \leq i \leq n$, e se supomos definida f_s para todo $s \leq k$, então f_{k+1} é definida como $f_{k+1}(i) = \{\langle M, f_1, \dots, f_k \rangle\}$, quando $1 \leq i \leq n$. A estrutura f assim definida é tal que $f \models_{\leq} \alpha$. Mas, como por hipótese temos que $\alpha \models_{\leq} \beta$, daí resulta que $f \models_{\mathcal{L}} \beta$. Então $M \models_{\mathcal{L}} \beta$, como era desejado.

Portanto $\alpha \models_{\leq} \beta$ ■

A seguir serão introduzidos alguns conceitos e demonstrados alguns resultados com o objetivo de demonstrar: $\alpha \models_{<} \beta$ se e somente se $B_{i_1} \dots B_{i_r} \alpha \models B_{i_1} \dots B_{i_r} \beta$, quando $r > 0$.

Seja $\varphi = B_{i_1} B_{i_2} \dots B_{i_r} \alpha$ uma fórmula tal que α não tem ocorrências do operador B_{i_i} , para todo $i \in \{1, \dots, n\}$; intuitivamente se uma estrutura de crenças f (ou um $(\text{prof}(\varphi))$ - mundo¹ vai satisfazer a fórmula φ , então a fórmula α é satisfeita por um conjunto de interpretações da lógica $\mathcal{L}$. Quais são essas interpretações, é o que tenta formalizar a seguinte definição:

¹Pelos resultados vistos no capítulo 5, basta considerar só os mundos de profundidade φ .

Definição auxiliar 1 Seja $\varphi = B_{i_1}B_{i_2}\dots B_{i_r}\alpha$ uma fórmula tal que α não tem ocorrências dos operadores B_{i_i} , para todo $i \in \{1, \dots, n\}$, e seja $f = \langle f_0, f_1, \dots, f_r \rangle$ um $(r+1)$ -mundo tal que $\langle f_0, f_1, \dots, f_r \rangle \models_{\mathcal{L}} \varphi$, então o conjunto de interpretações básicas associadas a $\langle f_0, f_1, \dots, f_r \rangle$ (denotado por $M_{f_0, \dots, f_r}$) está definido indutivamente como:

Se $r = 1$, como $\langle f_0, f_1 \rangle \models B_{i_1}\alpha$ então o conjunto de interpretações básicas é:

$$M_{f_0, \dots, f_1} = \{f'_0 : f'_0 \in f_1(i_1)\}.$$

Suponha definido o conjunto de interpretações básicas para todo s -mundo que satisfaz uma fórmula de profundidade s , para $s \leq r$.

Se define a seguir o conjunto de interpretações básicas $M_{f_0, \dots, f_r}$ associadas a um $(r+1)$ -mundo que satisfaz a fórmula $B_{i_1}B_{i_2}\dots B_{i_r}\alpha$:

Se $\langle f_0, f_1, \dots, f_r \rangle \models_{\mathcal{L}} B_{i_1}B_{i_2}\dots B_{i_r}\alpha$, então tem-se que $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle \models_{\mathcal{L}} B_{i_2}\dots B_{i_r}\alpha$ quando $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle \in f_r(i_1)$.

Portanto, pela hipótese de indução $M_{f'_0, \dots, f'_{r-1}}$ está definida para todo $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle \in f_r(i_1)$.

Então $M_{f_0, \dots, f_r}$ é definida como:

$$M_{f_0, \dots, f_r} = \left\{ \bigcup_{\langle f'_0, f'_1, \dots, f'_{r-1} \rangle \in f_r(i_1)} M_{f'_0, \dots, f'_{r-1}} \right\}.$$

Os lemas que seguem, permitem demonstrar que uma estrutura de crenças satisfaz preferencialmente um formula $\varphi = B_{i_1}\dots B_{i_r}\alpha$, onde α não tem ocorrências dos operadores B_{i_i} , para todo $i \in \{1, \dots, n\}$, se e somente se, essa mesma estrutura satisfaz preferentemente a fórmula obtida eliminando ocorrências consecutivas dos operadores B_{i_i} .

Lema 5.1 Seja $\langle f_0, f_1, \dots, f_{r+k} \rangle$ um $(r+k+1)$ -mundo, prefixo do $(r+k+2)$ -mundo $\langle f_0, f_1, \dots, f_{r+k+1} \rangle$, então $\langle f_0, f_1, \dots, f_{r+k+1} \rangle \models_{\mathcal{L}} B_{i_1}\dots B_{i_r}B_{i_{r+1}}B_{i_{r+2}}\dots B_{i_{r+k}}\alpha$ se e somente se $\langle f_0, f_1, \dots, f_{r+k} \rangle \models_{\mathcal{L}} B_{i_1}\dots B_{i_r}B_{i_{r+1}}B_{i_{r+2}}\dots B_{i_{r+k}}\alpha$, onde $B_{i_1}\dots B_{i_r}B_{i_{r+1}}$ não tem ocorrências consecutivas do mesmo operador B_{i_i} , para todo $i \in \{1, \dots, n\}$.

Demonstração:

A demonstração será feita por indução.

Suponha $s=0$.

Se $\langle f_0, \dots, f_k \rangle \models_{\mathcal{L}^*} B_{i_1} B_{i_2} B_{i_3} \dots B_{i_k} \alpha$ então $\langle f'_0, \dots, f'_{k-1} \rangle \models_{\mathcal{L}^*} B_{i_1} B_{i_3} \dots B_{i_k} \alpha$ para todo $\langle f'_0, \dots, f'_{k-1} \rangle \in f_k(i_1)$. Mas isto acontece se, para todo $\langle f''_0, \dots, f''_{k-2} \rangle \in f'_{k-1}(i_1)$, $\langle f''_0, \dots, f''_{k-2} \rangle \models_{\mathcal{L}^*} B_{i_3} \dots B_{i_k} \alpha$.

Logo, levando em conta que $f'_{k-1}(i_1) = f_{k-1}(i_1)$ então para todo $\langle f'_0, \dots, f'_{k-2} \rangle \in f_{k-1}(i_1)$, $\langle f'_0, \dots, f'_{k-2} \rangle \models_{\mathcal{L}^*} B_{i_3} \dots B_{i_k} \alpha$ (ou de maneira equivalente $\langle f'_0, \dots, f'_{k-2} \rangle \models_{\mathcal{L}^*} B_{i_3} \dots B_{i_k} \alpha$). Portanto $\langle f_0, \dots, f_{k-1} \rangle \models_{\mathcal{L}^*} B_{i_1} B_{i_3} \dots B_{i_k} \alpha$.

De maneira análoga, pode-se demonstrar que se $\langle f_0, \dots, f_{k-1} \rangle \models_{\mathcal{L}^*} B_{i_1} B_{i_3} \dots B_{i_k} \alpha$, então $\langle f_0, \dots, f_k \rangle \models_{\mathcal{L}^*} B_{i_1} B_{i_2} B_{i_3} \dots B_{i_k} \alpha$.

Suponha agora o resultado válido para todo $s < r$. Vai ser demonstrado o resultado para r .

Se $\langle f_0, \dots, f_{r+k+1} \rangle \models_{\mathcal{L}^*} B_{i_1} B_{i_2} B_{i_r} B_{i_{r+1}} B_{i_{r+1}} \dots B_{i_{r+k}} \alpha$, então $\langle f'_0, \dots, f'_{r+k} \rangle \models_{\mathcal{L}^*} B_{i_2} B_{i_r} B_{i_{r+1}} B_{i_{r+1}} \dots B_{i_{r+k}} \alpha$, para todo $\langle f'_0, \dots, f'_{r+k} \rangle \in f_{r+k+1}(i_1)$.

Mas se $\langle f'_0, \dots, f'_{r+k} \rangle \models_{\mathcal{L}^*} B_{i_2} B_{i_r} B_{i_{r+1}} B_{i_{r+1}} \dots B_{i_{r+k}} \alpha$, então aplicando a hipótese de indução $\langle f'_0, \dots, f'_{r+k-1} \rangle \models_{\mathcal{L}^*} B_{i_1} B_{i_2} B_{i_r} B_{i_{r+1}} \dots B_{i_{r+k}} \alpha$. Assim, levando em conta a definição de $(r+k+1)$ -mundos $\langle f'_0, \dots, f'_{r+k} \rangle \in f_{r+k+1}(i_1)$ se e somente se $\langle f'_0, \dots, f'_{r+k-1} \rangle \in f_{r+k+1}(i_1)$, então $\langle f'_0, \dots, f'_{r+k-1} \rangle \models_{\mathcal{L}^*} B_{i_1} B_{i_2} B_{i_r} B_{i_{r+1}} \dots B_{i_{r+k}} \alpha$, para todo $\langle f'_0, \dots, f'_{r+k-1} \rangle \in f_{r+k+1}(i_1)$. Portanto $\langle f_0, f_1, \dots, f_{r+k} \rangle \models_{\mathcal{L}^*} B_{i_1} \dots B_{i_r} B_{i_{r+1}} B_{i_{r+1}} \dots B_{i_{r+k}} \alpha$.

De maneira análoga, pode-se demonstrar que se $\langle f_0, f_1, \dots, f_{r+k} \rangle \models_{\mathcal{L}^*} B_{i_1} \dots B_{i_r} B_{i_{r+1}} B_{i_{r+1}} \dots B_{i_{r+k}} \alpha$, então $\langle f_0, \dots, f_{r+k+1} \rangle \models_{\mathcal{L}^*} B_{i_1} B_{i_2} B_{i_r} B_{i_{r+1}} B_{i_{r+1}} \dots B_{i_{r+k}} \alpha$.

Com o qual fica demonstrado o caso para r , e portanto conclui a demonstração do lema ■

Lema 5.2 *Seja $\langle f_0, f_1, \dots, f_{r+k} \rangle$ um $(r+k+1)$ -mundo prefixo do $(r+k+2)$ -mundo $\langle f_0, f_1, \dots, f_{r+k+1} \rangle$, tal que $\langle f_0, f_1, \dots, f_{r+k+1} \rangle \models_{\mathcal{L}^*} B_{i_1} \dots B_{i_r} B_{i_{r+1}} B_{i_{r+1}} B_{i_{r+3}} \dots B_{i_{r+k}} \alpha$, onde $B_{i_1} \dots B_{i_r} B_{i_{r+1}}$ não tem mais de uma ocorrência consecutiva do mesmo operador B_{i_i} , para todo $i \in \{1, \dots, n\}$. Então as interpretações básicas associadas a $\langle f_0, f_1, \dots, f_{r+k} \rangle$ e o mesmo que o conjunto de interpretações básicas associadas a $\langle f_0, f_1, \dots, f_{r+k+1} \rangle$.*

Demonstração:

A demonstração será feita por indução.

Suponha $r=0$.

Se $\langle f_0, \dots, f_k \rangle \models_{\mathcal{L}} B_{i_1} B_{i_2} B_{i_3} \dots B_{i_k} \alpha$. Então pela definição de interpretações básicas associadas:

$$M_{f_0, \dots, k} = \left\{ \bigcup_{\langle f'_0, f'_1, \dots, f'_{k-1} \rangle \in f_k(i_1)} M_{f'_0, \dots, k-1} \right\}.$$

Por sua vez $M_{f'_0, \dots, k-1}$ está definida como:

$$M_{f'_0, \dots, k-1} = \left\{ \bigcup_{\langle f''_0, f''_1, \dots, f''_{k-2} \rangle \in f'_{k-1}(i_1)} M_{f''_0, \dots, k-2} \right\}.$$

Mas como $f'_{k-1}(i_1) = f_{k-1}(i_1)$, então

$$M_{f'_0, \dots, k-1} = \left\{ \bigcup_{\langle f''_0, f''_1, \dots, f''_{k-2} \rangle \in f_{k-1}(i_1)} M_{f''_0, \dots, k-2} \right\}. \quad (5.1)$$

Mas, pelo lema anterior, se $\langle f_0, \dots, f_k \rangle \models_{\mathcal{L}} B_{i_1} B_{i_2} B_{i_3} \dots B_{i_k} \alpha$, então Se $\langle f_0, \dots, f_{k-1} \rangle \models_{\mathcal{L}} B_{i_1} B_{i_2} \dots B_{i_k} \alpha$. Portanto o conjunto de interpretações básicas associadas é:

$$M_{f_0, \dots, k-1} = \left\{ \bigcup_{\langle f'_0, f'_1, \dots, f'_{k-2} \rangle \in f_{k-1}(i_1)} M_{f'_0, \dots, k-2} \right\}.$$

Mas levando em conta que $\langle f'_0, f'_1, \dots, f'_{k-1} \rangle \in f_k(i_1)$ se e somente se $\langle f'_0, \dots, f'_{k-2} \rangle \in f_{k-1}(i_1)$, então por 5.1, os conjuntos $M_{f_0, \dots, k}$ e $M_{f'_0, \dots, k-1}$ são iguais. Portanto acaba a demonstração para o caso base.

Suponha agora o resultado válido para todo $s < r$. Vai ser demonstrado o resultado para r .

Se $\langle f_0, \dots, f_{r+k+1} \rangle \models_{\mathcal{L}} B_{i_1} B_{i_2} B_{i_r} B_{i_{r+1}} B_{i_{r+1}} \dots B_{i_{r+k}} \alpha$ então o conjunto de interpretações básicas é:

$$M_{f_0, \dots, r+k+1} = \left\{ \bigcup_{\langle f'_0, f'_1, \dots, f'_{r+k} \rangle \in f_{r+k+1}(i_1)} M_{f'_0, \dots, r+k} \right\}.$$

Mas pelo Lema 5.1 $\langle f_0, \dots, f_{r+k} \rangle \models_{\mathcal{L}} B_{i_1} B_{i_2} B_{i_r} B_{i_{r+1}} \dots B_{i_{r+k}} \alpha$ então o conjunto de interpretações básicas associadas é o conjunto

$$M_{f_0, \dots, r+k} = \left\{ \bigcup_{\langle f'_0, f'_1, \dots, f'_{r+k-1} \rangle \in f_{r+k}(i_1)} M_{f'_0, \dots, r+k-1} \right\}.$$

Logo, levando em conta que $\langle f'_0, f'_1, \dots, f'_{r+k} \rangle \models_{\mathcal{L}} B_{i_2} B_{i_r} B_{i_{r+1}} B_{i_{r+1}} \dots B_{i_{r+k}} \alpha$ se e somente se $\langle f'_0, f'_1, \dots, f'_{r+k-1} \rangle \models_{\mathcal{L}} B_{i_2} B_{i_r} B_{i_{r+1}} \dots B_{i_{r+k}} \alpha$, e aplicando a hipótese indutiva tem-se que $M_{f'_0, \dots, r+k} = M_{f'_0, \dots, r+k-1}$. Portanto $M_{f_0, \dots, r+k+1} = M_{f_0, \dots, r+k}$.

Com o qual fica demonstrado o caso para r , e portanto conclui a demonstração do lema ■

Lema 5.3 *Seja f uma estrutura de crenças, então $f \models_{\leq} B_{i_1} B_{i_2} \dots B_{i_r} B_{i_{r+1}} B_{i_{r+1}} B_{i_{r+3}} \dots B_{i_{r+k}} \alpha$ se e somente se $f \models_{\leq} B_{i_1} B_{i_2} \dots B_{i_r} B_{i_{r+1}} B_{i_{r+3}} \dots B_{i_{r+k}} \alpha$, onde $B_{i_1} \dots B_{i_r} B_{i_{r+1}}$ não tem mas de uma ocorrência consecutiva do mesmo operador B_{i_i} , para todo $i \in \{1, \dots, n\}$.*

Demonstração:

$\Rightarrow$)

Se $f \models_{\leq} B_{i_1} \dots B_{i_r} B_{i_{r+1}} B_{i_{r+1}} B_{i_{r+3}} \dots B_{i_{r+k}} \alpha$ então pelo Lema ??, $f \models_{\mathcal{L}} B_{i_1} \dots B_{i_r} B_{i_{r+1}} B_{i_{r+3}} \dots B_{i_{r+k}} \alpha$.

Se $f \not\models_{\leq} B_{i_1} \dots B_{i_r} B_{i_{r+1}} B_{i_{r+3}} \dots B_{i_{r+k}} \alpha$, então existe $f' \prec f$ tal que $f' \models_{\mathcal{L}} B_{i_1} \dots B_{i_r} B_{i_{r+1}} B_{i_{r+3}} \dots B_{i_{r+k}} \alpha$. Então pelo Lema ?? $f' \models_{\leq} B_{i_1} \dots B_{i_r} B_{i_{r+1}} B_{i_{r+1}} B_{i_{r+3}} \dots B_{i_{r+k}} \alpha$, mas isto não pode acontecer pois $f \models_{\leq} B_{i_1} \dots B_{i_r} B_{i_{r+1}} B_{i_{r+1}} B_{i_{r+3}} \dots B_{i_{r+k}} \alpha$.

Portanto, $f \models_{\leq} B_{i_1} \dots B_{i_r} B_{i_{r+1}} B_{i_{r+3}} \dots B_{i_{r+k}} \alpha$.

A outra implicação se demonstra de maneira análoga ■

Corolário 5.1 *Seja f uma estrutura de crenças f e $\varphi = B_{i_1} \dots B_{i_r} \alpha$ uma fórmula tal que α não tem ocorrências dos operadores B_{i_i} , para todo $i \in \{1, \dots, n\}$. $f \models_{\leq} \varphi$ se e somente se f satisfaz preferentemente a fórmula obtida eliminando as ocorrências consecutivas dos operadores B_{i_i} , para todo $i \in \{1, \dots, n\}$.*

Demonstração:

A demonstração é trivial a partir do lema anterior.

Lema 5.4 *Seja $\varphi = B_{i_1}B_{i_2}\dots B_{i_r}\alpha$ uma fórmula tal que α não tem ocorrências dos operadores B_i , para todo $i \in \{1, \dots, n\}$ e não tem mais de uma ocorrência consecutiva do mesmo operador B_i , para $i \in \{1, \dots, n\}$, e seja $f = \langle f_0, f_1, \dots \rangle$ uma estrutura de crenças tal que $f \models_{\mathcal{L}^*} \varphi$.*

Se as interpretações do conjunto de interpretações básicas associadas com $\langle f_0, f_1, \dots, f_r \rangle$ não satisfazem preferencialmente α , então existe um $r+1$ -mundo $\langle h_0, h_1, \dots, h_r \rangle$ associado a $\langle f_0, f_1, \dots, f_r \rangle$, tal que $\langle h_0, h_1, \dots, h_r \rangle \prec_{r+1} \langle f_0, f_1, \dots, f_r \rangle$, $h_j = f_j$, para todo j tal que $0 \leq j \leq r-1$ e $h_r(i) = f_r(i)$, para todo i tal que $i \in \{1, \dots, n\}$ e $i \neq i_1$ e as interpretações básicas associadas a $\langle h_0, h_1, \dots, h_r \rangle$ satisfazem preferencialmente α .

Além disso, para todo $k \geq 1$, pode-se definir $\langle h_0, \dots, h_r, \dots, h_{r+k} \rangle$ associado a $\langle f_0, \dots, f_r, \dots, f_{r+k} \rangle$, tal que $\langle h_0, \dots, h_r, \dots, h_{r+k} \rangle \prec_{r+k+1} \langle f_0, \dots, f_r, \dots, f_{r+k} \rangle$, $\langle h_0, \dots, h_r, \dots, h_{r+k-1} \rangle$ é um prefixo de $\langle h_0, \dots, h_r, \dots, h_{r+k} \rangle$, e $h_{r+k}(i) = f_{r+k}(i)$ para todo i tal que $i \in \{1, \dots, n\}$ e $i \neq i_1$.

Demonstração:

A prova é por indução na profundidade de φ .

Paso base:

Seja $\varphi = B_{i_1}\alpha$ e f uma estrutura de crenças tal que $f \models_{\mathcal{L}^*} \varphi$. Então $\langle f_0, f_1 \rangle \models_{\mathcal{L}^*} B_{i_1}\alpha$.

Se $\langle f_0, f_1 \rangle \models_{\mathcal{L}^*} B_{i_1}\alpha$ então para todo $\langle f'_0 \rangle \in f_1(i_1)$, $\langle f'_0 \rangle \models_{\mathcal{L}^*} \alpha$.

Se $\langle f'_0 \rangle \in f_1(i_1)$, então podemos associar com ela uma interpretação h'_0 como segue:

$$h'_0 = \begin{cases} f'_0 & \text{se } f'_0 \models_{\leq} \alpha \\ I & \text{onde } I \text{ é uma interpretação tal que } I < f'_0 \text{ e } I \models_{\leq} \alpha, \text{ caso contrario} \end{cases}$$

Para cada $\langle f'_0 \rangle \in f_1(i_1)$, $\langle h'_0 \rangle_{\langle f'_0 \rangle}$ denotará o 1-mundo no qual h'_0 é a interpretação associada com f'_0 .

Pode-se definir então $\langle h_0, h_1 \rangle$ como segue:

$$\begin{aligned} h_0 &= f_0 \\ h_1(i) &= \begin{cases} f_1(i) & \text{para todo } i \neq i_1 \\ \{ \langle h'_0 \rangle_{f'_0} : \langle f'_0 \rangle \in f_1(i_1) \} & i = i_1 \end{cases} \end{aligned}$$

$\langle h_0, h_1 \rangle$ assim definido é um 2-mundo, tal que $\langle h_0, h_1 \rangle \prec \langle f_0, f_1 \rangle$, e $h_1(i) = f_1(i)$, para todo i tal que $i \in \{1, \dots, n\}$ e $i \neq i_1$. Além disso, pela definição de $h_1(i_1)$ as interpretações associadas a $\langle h_0, h_1 \rangle$ satisfazem preferencialmente α .

Veremos agora que para todo j , tal que $j > 0$, pode-se definir um $(1+j+1)$ -mundo $\langle h_0, \dots, h_r, \dots, h_{1+j} \rangle$ tal que $\langle h_0, \dots, h_{1+j} \rangle \prec_{1+j+1} \langle f_0, f_1, \dots, f_{1+j} \rangle$, $\langle h_0, \dots, h_{1+j-1} \rangle$ é um prefixo de $\langle h_0, \dots, h_{1+j} \rangle$ e $h_{1+j}(i) = f_{1+j}(i)$ para todo i tal que $i \in \{1, \dots, n\}$ e $i \neq i_1$.

Consideremos primeiro o caso $j = 1$.

Só é preciso definir h_2 , já que h_0 e h_1 já estão definidos.

Agora bem, se $\langle f'_0, f'_1 \rangle \in f_2(i_1)$, então pelas propriedades dos 3-mundos $\langle f'_0 \rangle \in f_1(i_1)$ e $f'_1(i_1) = f_1(i_1)$, então definimos o 2-mundo associado a $\langle f'_0, f'_1 \rangle$ e denotado $\langle h'_0, h'_1 \rangle_{\langle f'_0, f'_1 \rangle}$ ² como:

$$h'_0 \text{ é a interpretação associada a } f'_0 \text{ na definição de } h_1(i_1)$$

$$h'_1(i) = \begin{cases} f'_1(i) & \text{se } i \neq i_1 \\ h_1(i) & \text{se } i = i_1 \end{cases}$$

Então $h_2(i)$ define-se como:

$$h_2(i) = \begin{cases} f_2(i) & \text{quando } i \neq i_1 \\ \{ \langle h'_0, h'_1 \rangle_{\langle f'_0, f'_1 \rangle} : \langle f'_0, f'_1 \rangle \in f_2(i) \} & \text{quando } i = i_1 \end{cases}$$

Pela construção, $\langle h_0, h_1, h_2 \rangle$ é um 3-mundo, $\langle h_0, h_1, h_2 \rangle \prec_3 \langle f_0, f_1, f_2 \rangle$, $\langle h_0, h_1 \rangle$ é um prefixo de $\langle h_0, h_1, h_2 \rangle$, e $h_2(i) = f_2(i)$, para todo i , tal que $i \in \{1, \dots, n\}$ e $i \neq i_1$.

Suponha definidos $\langle h_0, h_1, \dots, h_{1+j} \rangle$ para $1 \leq j \leq k$, então vamos definir $\langle h_0, h_1, \dots, h_{1+k+1} \rangle$.

Só precisamos definir h_{1+k+1} . Definamos primeiro $h_{1+k+1}(i_1)$.

Se $\langle f'_0, f'_1, \dots, f'_{1+k} \rangle \in f_{1+k+1}(i_1)$, então $\langle f'_0, f'_1, \dots, f'_{1+k-1} \rangle \in f_{1+k}(i_1)$ e $f'_{1+k}(i_1) = f_{1+k}(i_1)$.

Então, se $\langle f'_0, f'_1, \dots, f'_{1+k-1} \rangle \in f_{1+k}(i_1)$, pela construção existe um $(1+k)$ -mundo $\langle h'_0, h'_1, \dots, h'_{1+k-1} \rangle_{\langle f'_0, f'_1, \dots, f'_{1+k-1} \rangle}$ associado a $\langle f'_0, f'_1, \dots, f'_{1+k-1} \rangle$. Para definir o $(k+1)$ -mundo associado a $\langle f'_0, f'_1, \dots, f'_{1+k} \rangle$, só falta definir h'_{1+k} , então:

²A seguir um s -mundo $\langle h_0, h_1, \dots, h_{s-1} \rangle$ associado a um s -mundo $\langle f_0, f_1, \dots, f_{s-1} \rangle$ será denotado $\langle h_0, h_1, \dots, h_{s-1} \rangle_{\langle f_0, f_1, \dots, f_{s-1} \rangle}$.

$$h'_{i+k}(i) = \begin{cases} f'_{i+k}(i) & i \neq i_1 \\ h_{i+k}(i) & \text{caso contrário} \end{cases}$$

Portanto h_{1+k+1} está definida como:

$$h_{1+k+1}(i) = \begin{cases} f_{1+k+1}(i) & \text{se } i \neq i_1 \\ \{ \langle h'_0, \dots, h'_{1+k} \rangle_{\langle f'_0, \dots, f'_{1+k} \rangle} : \langle f'_0, \dots, f'_{1+k-1} \rangle \in f_{1+k}(i_1) \} & \text{caso contrário} \end{cases}$$

$\langle h_0, h_1, \dots, h_{1+k+1} \rangle$ assim definido é um $(k+3)$ -mundo, $\langle h_0, h_1, \dots, h_{1+k} \rangle$ é um prefixo de $\langle h_0, h_1, \dots, h_{1+k+1} \rangle$, $\langle h_0, \dots, h_{1+k+1} \rangle \prec_{1+k+2} \langle f_0, \dots, f_{1+k+1} \rangle$ e $h_{r+k+1}(i) = f_{r+k+1}(i)$, para todo i tal que $i \in \{1, \dots, n\}$ e $i \neq i_1$.

Portanto foi demonstrado o caso base $s=0$.

Suponha agora que o resultado é válido para todo $s < r$. Vai ser demonstrado que é válido para $s = r$.

Seja $f = \langle f_0, f_1, \dots \rangle$ tal que $\langle f_0, f_1, \dots \rangle \models_{\mathcal{L}} B_{i_1} \dots B_{i_r} \alpha$ e as interpretações básicas associadas a $\langle f_0, f_1, \dots, f_r \rangle$ não satisfazem preferencialmente α .

Então, pela definição, $\langle f_0, f_1, \dots, f_r \rangle \models_{\mathcal{L}} B_{i_1} \dots B_{i_r} \alpha$. Portanto $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle \models_{\mathcal{L}} B_{i_2} \dots B_{i_r} \alpha$, para todo $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle \in f_r(i_1)$.

Como $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle$ é um r -mundo tal que $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle \models B_{i_2} \dots B_{i_r} \alpha$ então pode-se associar com cada $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle$ um r -mundo da seguinte maneira:

- $\langle h'_0, \dots, h'_{r-1} \rangle_{\langle f'_0, \dots, f'_{r-1} \rangle}$ é $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle$, quando as interpretações básicas associadas a $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle$ satisfazem preferencialmente α ,
- $\langle h'_0, \dots, h'_{r-1} \rangle_{\langle f'_0, \dots, f'_{r-1} \rangle}$ é (aplicando a hipótese indutiva) o r -mundo associado a $\langle f'_0, \dots, f'_{r-1} \rangle$, caso as interpretações básicas associadas a $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle$ não satisfaçam preferencialmente α . Neste caso as interpretações básicas associadas a $\langle h'_0, h'_1, \dots, h'_{r-1} \rangle_{\langle f'_0, \dots, f'_{r-1} \rangle}$ satisfazem preferencialmente α .

Então o $(r+1)$ -mundo $\langle h_0, h_1, \dots, h_r \rangle$ está definido como:

$$h_j = f_j \text{ quando } 0 \leq j \leq r-1$$

$$h_r(i) = \begin{cases} f_r(i) & \text{quando } i \neq i_1 \\ \{ \langle h'_0, h'_1, \dots, h'_{r-1} \rangle_{\langle f'_0, \dots, f'_{r-1} \rangle} : \langle f'_0, f'_1, \dots, f'_{r-1} \rangle \in f_r(i) \} & \text{quando } i = i_1 \end{cases}$$

É fácil ver, a partir da construção, que $\langle h_0, h_1, \dots, h_r \rangle$ é um $r + 1$ -mundo que tem as propriedades desejadas.

Definido $\langle h_0, h_1, \dots, h_r \rangle$, para definir $\langle h_0, h_1, \dots, h_{r+1} \rangle$ a partir de $\langle f_0, f_1, \dots, f_{r+1} \rangle$, basta definir h_{r+1} .

Primeiro vamos definir $h_{r+1}(i_1)$.

Se $\langle f'_0, f'_1, \dots, f'_r \rangle \in f_{r+1}(i_1)$ então $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle \in f_r(i_1)$ é $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle \models_{\mathcal{L}} B_{i_2} \dots B_{i_r} \alpha$.

Então é possível associar a cada $\langle f'_0, f'_1, \dots, f'_r \rangle \in f_{r+1}(i_1)$ um $(r+1)$ -mundo $\langle h'_0, h'_1, \dots, h'_r \rangle_{\langle f'_0, \dots, f'_r \rangle}$ da seguinte maneira:

- $\langle h'_0, h'_1, \dots, h'_r \rangle_{\langle f'_0, \dots, f'_r \rangle}$ é $\langle f'_0, f'_1, \dots, f'_r \rangle$, quando as interpretações básicas associadas a $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle$ satisfazem preferencialmente α ,
- $\langle h'_0, \dots, h'_r \rangle_{\langle f'_0, \dots, f'_r \rangle}$ é, caso contrario, o $r + 1$ -mundo associado a $\langle f'_0, \dots, f'_r \rangle$ que é obtido aplicando a hipótese indutiva, já que $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle$ é o prefixo de uma estrutura de crenças f' que satisfaz $B_{i_2} \dots B_{i_r} \alpha$

Definimos, a seguir, h_{r+1} como:

$$h_{r+1}(i) = \begin{cases} f_{r+1}(i) & \text{quando } i \neq i_1 \\ \langle h'_0, h'_1, \dots, h'_r \rangle_{\langle f'_0, \dots, f'_r \rangle} : \langle f'_0, \dots, f'_r \rangle \in f_{r+1}(i_1) \} & \text{se } i = i_1 \end{cases}$$

Então $\langle h_0, h_1, \dots, h_{r+1} \rangle$ é um $(r+2)$ -mundo tal que $\langle h_0, h_1, \dots, h_{r+1} \rangle \prec_{r+2} \langle f_0, f_1, \dots, f_{r+1} \rangle$, $\langle h_0, h_1, \dots, h_r \rangle$ é um prefixo de $\langle h_0, h_1, \dots, h_{r+1} \rangle$, e $h_{r+1}(i) = f_{r+1}(i)$, para todo i tal que $i \in \{1, \dots, n\}$ e $i \neq i_1$.

Suponha definido $\langle h_0, h_1, \dots, h_{r+j} \rangle$ para $1 \leq j \leq k$, então vamos definir $\langle h_0, h_1, \dots, h_{r+k+1} \rangle$.

Só é preciso definir h_{r+k+1} . Vamos definir inicialmente $h_{r+k+1}(i_1)$.

Se $\langle f'_0, f'_1, \dots, f'_{r+k} \rangle \in f_{r+k+1}(i_1)$, então $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle \in f_r(i_1)$ é $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle \models_{\mathcal{L}} B_{i_2} \dots B_{i_r} \alpha$.

Então é possível associar a cada $\langle f'_0, f'_1, \dots, f'_{r+k} \rangle \in f_{r+k+1}(i_1)$ um $(r+k+1)$ -mundo $\langle h'_0, h'_1, \dots, h'_{r+k} \rangle_{\langle f'_0, \dots, f'_{r+k} \rangle}$ da seguinte maneira:

- $\langle h'_0, h'_1, \dots, h'_{r+k} \rangle_{\langle f'_0, \dots, f'_{r+k} \rangle}$ é $\langle f'_0, f'_1, \dots, f'_{r+k} \rangle$, quando as interpretações básicas associadas a $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle$ satisfazem preferencialmente α ,
- $\langle h'_0, \dots, h'_{r+k} \rangle_{\langle f'_0, \dots, f'_{r+k} \rangle}$ é, caso contrario, o $(r+k+1)$ -mundo associado a $\langle f'_0, \dots, f'_{r+k} \rangle$ que é obtido aplicando a hipótese indutiva, já que $\langle f'_0, f'_1, \dots, f'_{r-1} \rangle$ é o prefixo de uma estrutura de crenças f' que satisfaz $B_{i_2} \dots B_{i_r} \alpha$.

Assim a definição de h_{r+1+k} é:

$$h_{r+k+1}(i) = \begin{cases} f_{r+k+1}(i) & \text{quando } i \neq i_1 \\ \langle h'_0, h'_1, \dots, h'_{r+k} \rangle_{\langle f'_0, \dots, f'_{r+k} \rangle} : \langle f'_0, f'_1, \dots, f'_{r+k} \rangle \in f_{r+k+1}(i_1) \end{cases} \text{ se } i = i_1$$

Então $\langle h_0, h_1, \dots, h_{r+k+1} \rangle$ é um $(r+k+2)$ -mundo, $\langle h_0, h_1, \dots, h_{r+k} \rangle$ é um prefixo de $\langle h_0, h_1, \dots, h_{r+k+1} \rangle$, $\langle f_0, f_1, \dots, f_{r+k+1} \rangle \prec_{r+k+2} \langle h_0, h_1, \dots, h_{r+k+1} \rangle$ e $h_{r+k}(i) = f_{r+k}(i)$, para todo i , tal que $i \in \{1, \dots, n\}$ e $i \neq i_1$ ■

Teorema 5.5 $\alpha \models_{\leq} \beta$ se e somente se $B_{i_1} \dots B_{i_r} \alpha \models_{\leq} B_{i_1} \dots B_{i_r} \beta$, quando $r > 0$.

Demonstração:

$\Rightarrow$) Suponha que $\alpha \models_{\leq} \beta$ e seja $f = \langle f_0, f_1, \dots \rangle$ tal que $f \models_{\leq} B_{i_1} \dots B_{i_r} \alpha$.

Se M é o conjunto das interpretações básicas associadas a $\langle f_0, f_1, \dots, f_r \rangle$. Então para todo $I \in M$ temos que $I \models_{\leq} \alpha$, já que se existisse uma interpretação $I_0 \in M$, tal que $I_0 \not\models_{\leq} \alpha$, pelo lema 5.4 e pelo Lema 5.2, existiria uma estrutura de crenças $h = \langle h_0, h_1, \dots \rangle$ tal que $h \prec f$ e $h \models_{\leq} B_{i_1} \dots B_{i_r} \alpha$, o que não pode acontecer.

Então, como $I \models_{\leq} \alpha$ quando $I \in M$ e como $\alpha \models_{\leq} \beta$, então $I \models \beta$ e então $f \models \beta$.

$\Leftarrow$)

Para demonstrar a recíproca, suponha que $B_{i_1} \dots B_{i_r} \alpha \models_{\leq} B_{i_1} \dots B_{i_r} \beta$ quando $r > 0$, e que $\alpha \not\models_{\leq} \beta$.

Se $\alpha \not\models_{\leq} \beta$ então existe uma interpretação I tal que $I \models_{\leq} \alpha$ e $I \not\models_{\leq} \beta$, então podemos definir uma estrutura de crenças $f = \langle I, f_1, \dots \rangle$ onde $f_1(i) = \{\langle I \rangle\}$ para todo i tal que $1 \leq i \leq n$ e se f_k foi definido, então f_{k+1} se define como $f_{k+1}(i) = \{\langle I, f_1, \dots, f_k \rangle\}$ quando $1 \leq i \leq n$. É fácil ver que $f \models_{\leq} \alpha$.

Mas por hipótese $B_{i_1} \dots B_{i_r} \alpha \models_{\leq} B_{i_1} \dots B_{i_r} \beta$, quando $r > 0$ e como a estrutura de crenças assim definida é tal que $f \models_{\leq} B_{i_1} \dots B_{i_r} \alpha$ então $f \models B_{i_1} \dots B_{i_r} \beta$ e então pela definição de f , $I \models_{\mathcal{L}} \beta$, o que é uma contradição.

Então temos que se $B_{i_1} \dots B_{i_r} \alpha \models_{\leq} B_{i_1} \dots B_{i_r} \beta$, quando $r > 0$, então $\alpha \models_{\leq} \beta$ ■

Teorema 5.6 Em $\mathcal{L}_{\leq}^*$ é válido:

1. Todas as instâncias das tautologias.
2. $B_i \alpha \rightarrow B_i B_i \alpha$.
3. $\neg B_i \alpha \rightarrow B_i \neg B_i \alpha$.
4. $B_i \alpha \wedge B_i (\alpha \rightarrow \beta) \rightarrow B_i \beta$.

Demonstração:

1)

Se α é uma fórmula logicamente válida e se $f = \langle f_0, f_1, \dots \rangle$ é uma estrutura de crenças, então $f \models_{\mathcal{L}^*} \alpha$, já que $f_0 \models_{\mathcal{L}} \alpha$, pois f_0 é uma interpretação e α é uma fórmula logicamente válida. Então para todas as estruturas de crenças f , $f \models \alpha$.

2)

Se deduz das propriedades de k -mundos.

3)

Se deduz das propriedades de k -mundos.

4)

Pode-se demonstrar facilmente ■

Teorema 5.7 Sejam α e γ fbf de $\mathcal{L}$. Se $\alpha \models_{\leq} \beta$ e $\gamma \models_{\leq} \delta$ então $\alpha \wedge B_{i_1} \dots B_{i_k} \gamma \models_{\leq} \beta \wedge B_{i_1} \dots B_{i_k} \delta$ para $k > 0$ e $i_1 \neq 0$.

Demonstração:

Seja $f = \langle f_0, f_1, \dots \rangle$ uma estrutura de crenças tal que $f \models_{\leq} \alpha \wedge B_{i_1} \dots B_{i_k} \gamma$, então $f_0 \models_{\leq} \alpha$ e as interpretações básicas associadas a $f = \langle f_0, f_1, \dots, f_r \rangle$ satisfazem preferencialmente γ , já que se isso não acontecesse poderia-se definir uma estrutura de crenças h tal que $h \prec f$ e $h \models_{\leq} \alpha \wedge B_{i_1} \dots B_{i_k} \gamma$.

Portanto $f \models_{\leq} \alpha$ e $f \models_{\leq} B_{i_1} \dots B_{i_k} \gamma$.

Logo $f \models_{\mathcal{L}^*} \beta \wedge B_{i_1} \dots B_{i_k} \delta$ ■

Bibliografia

- [AGM88] C. Alchourrón, P. Gärdenfors and D. Makinson. On the logic of theory of change: Partial meet contraction and revision functions. *J. Symbolic Logic*, 50:510-530. 1985.
- [Bak91] A. B. Baker. Nonmonotonic reasoning in the framework of situation calculus. *Artificial Intelligence*, 49:5-23, 1991.
- [Bes89] P. Besnard. *An Introduction to Default Logic*. Springer-Verlag. 1989.
- [Bir91] L. Birnbaum. Rigor Mortis: a reponse to Nilsson's "Logic and Artificial Intelligence". *Artificial intelligence*, 47:57-77. 1991.
- [Bou92] C. Boutilier. Conditional Logics for Default Reasoning and Belief Revision. *Tech. Report KRR-TR-92-1*. University of Toronto. 1992.
- [Bou93] C. Boutilier. On the semantics of stable inheritance reasoning. *Comput. Intell.* 9(1):73-110. 1993.
- [Bou94a] C. Boutilier. Unifying default reasoning and belief revision in a modal framework. *Artificial Intelligence*, 68:33-85. 1994.
- [Bou94b] C. Boutilier. Conditional logics on normality: a modal approach. *Artificial Intelligence*, 68:87-154. 1994.
- [Bra77] R. J. Brachman. What's in a concept: strutral foundations for semantic networks. *Int. J. Man-Machine Studies*, 9:127-152. 1977.
- [Bre94] G. Brewka. Reasoning About Priorities in Default Logic. Proceedings AAAI-94, pag. 940-945. 1994
- [BS85] G. Bossu and P. Siegel. Saturation nonmonotonic reasoning and the closed-world assumption. *Artificial Intelligence*, 25:13-63. 1985.
- [BW77] D. G. Bobrow and T. Winograd. An overview of KRL, a Knowledge Representation Language. *Cognitive Science* 1(1):3-46. 1977.

- [Che80] B. Chellas. *Modal Logic: An Introduction*. Cambridge University Press. 1980.
- [Dav80] M. Davis. The mathematics of non-monotonic reasoning. *Artificial Intelligence*, 13:73-80. 1980.
- [Del87] J.P. Delgrande. A first-order logic for prototypical properties. *Artificial Intelligence*, 33:105-130. 1987.
- [Del88] J.P. Delgrande. An approach to default reasoning based on a first-order conditional logic: revised report. *Artificial Intelligence*, 36:63-90. 1988.
- [Del90] J. P. Delgrande. A semantics for a class of inheritance networks. *Proceedings Canadian Society for Computational Studies of Intelligence Conference*, pag. 54-60. Ottawa. 1990.
- [Del94] J. P. Delgrande. A Preference-Based Approach to Default Reasoning: Preliminary Report. *Proceedings of the Eleventh National Conference on Artificial Intelligence (AAAI'94)*, pag. 902-908. 1984
- [Eth88] D.W. Etherington. Reasoning with Incomplete Information: Investigations of Non-monotonic Reasoning. Pitman. London. 1988.
- [FH88] J. Fagin, J.Y. Halpern. Belief, awareness and limited reasoning. *Artificial Intelligence* 34:39-76. 1988
- [FHV91] R. Fagin, J. Halpern and M. Vardi. A Model-theoretic Analysis of Knowledge. *Journal of the ACM*. Vol. 38/2:382-428. 1991.
- [Gar88] P. Gardenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press, Cambridge, MA. 1988.
- [GD93] P. Gmytrasiewicz and E. Durfee. A Logic of Knowledge and Belief for Recursive Modeling: Preliminary Report. *Proceedings AAAI'93*. Pag. 628-634. 1993.
- [Gef92] H. Geffner. Hight-Probabilities, Model-Preference and Defaults Arguments. *Minds and Machines* 2:51-70. 1992.
- [GP88] H. Geffner and J. Pearl. A Framework for Reasoning with Defaults. UCLA Cognitive Systems Lab., TR870058. 1988.
- [GP92] H. Geffner and J. Pearl. Conditional entailment: bridging two approaches to default reasoning. *Artificial Intelligence*, 53:209-244. 1992.

- [Hal87] J. Halpern. Using reasoning about knowledge to analyze distributed systems. *Annual Review of Computer Science* 2:37-68. 1987.
- [HC68] G. E. Hughes and M. J. Cresswell. *An introduction to Modal Logic*. London. Methuen and Co. Ltd.. 1968.
- [Hin62] J. Hintikka. *Knowledge and belief: an introduction to the two notions*. Cornell University Press. 1962.
- [HM85] J. Y. Halpern and Y. Moses. A Guide to the Modal Logics of Knowledge and Belief: Preliminary Draft. *Proceedings Thirteenth International Joint Conference on Artificial Intelligence (IJCAI'85)*. Pag. 480-490. 1985.
- [HM92] J. Y. Halpern and Y. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, 54(3):319-379. 1992.
- [HM93] J. Y. Halpern and Y. O. Moses. Reasoning about only knowing with many agents. *Proceedings IJCAI'93*, pag. 655-661. 1993.
- [KaL88] L. N. Kanal, J.F. Lemmer. *Uncertainty in Artificial Intelligence*. North-Holland. 1988.
- [KL88] S. Kraus and D. Lehmann. Knowledge, belief and time. *Theor. Comput. Sci.*, 58:155-174. 1988.
- [KLM90] S. Kraus, D. Lehmann and M. Magidor. Nonmonotonic Reasoning, Preferential Models and Cumulative Logics. *Artificial Intelligence*, 44:167-207. 1990.
- [Kon86] K. Konolige. *A deduction Model of Belief*. Morgan Kaufman. 1986.
- [Kon88] K. Konolige. On the Relation between Default and Autoepistemic Logic. *Artificial Intelligence*, 35(3):343-382. 1988.
- [Kri63] S. A. Kripke. Semantical analysis of modal logic I, normal propositions calculi. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 9:67-96. 1963.
- [Lak93] G. Lakemeyer. All they know: a study in multi-agent autoepistemic reasoning. *Proceedings IJCAI'93*. 1993.
- [Lam91] P. Lamarre. S4 as the conditional logic of nonmonotonicity. *Proceedings Second International Conference on Principles of Knowledge Representation and Reasoning*. Cambridge. MA. Pag. 402-408. 1991.

- [Lev90] J. L. Levesque. All I know: A study in autoepistemic logic. *Artificial Intelligence*, 42:263-309. 1990.
- [Lew73] D. Lewis. *On the plurality of worlds*. Blackwell. Oxford. 1973.
- [Lif85] V. Lifschitz. Computing circumscriptions. *Proceedings IJCAI'95*, pag. 121-127. 1985.
- [Lif86] V. Lifschitz. On the Satisfiability of Circumscription. *Artificial Intelligence*, 28:17-27. 1986.
- [McC80] J. McCarthy. Circumscription -a form of non-monotonic reasoning. *Artificial Intelligence*, 13:27-39. 1980.
- [McC86] J. McCarthy. Applications of circumscription to formalizing commonsense reasoning. *Artificial Intelligence*, 28:89-116. 1986.
- [McD80] D. McDermott and J. Doyle. Nonmonotonic Logic I. *Artificial Intelligence* 13:41-72. 1980.
- [Men79] E. Mendelson. *Introduction to mathematical logic*, Second Edition. D. Van Nostrand Company. 1979.
- [Min81] M. Minsky. A framework for Representing Knowledge. *Mind Design*. Edited By J. Haugeland. Pag. 95-128. The MIT Press. 1981.
- [Moo80] R.C. Moore. *Reasoning About Knowledge and Action*. Technical Note 191. SRI International. Artificial Intelligence Center. 1980.
- [Moo82] R. Moore. The Role of Logic in Knowledge representation and Commonsense Reasoning. *Proc. AAAI'82*. Pittsburgh. Pag. 428-433. 1982.
- [Moo85b] R. Moore. Semantical considerations on non-monotonic logics. *Artificial intelligence*, 25:75-94. 1985.
- [Mor90] L. Morgenstern. A theory of multiple agent nonmonotonic reasoning. *Proceedings AAAI'90*, pag. 538-544. 1990.
- [MS93] Y. Moses and Y. Shoham. Belief as defeasible Knowledge. *Artificial Intelligence*. 64. 1993.
- [Nil91] N. Nilson. Logic and Artificial Intelligence. *Artificial Intelligence*, 47:31-56.1991
- [Pea88] J. Pearl. *Probabilistic reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufman. 1988.

- [Pea89] J. Pearl. Probabilistic semantics for nonmonotonic reasoning: A survey. *Proceedings of First International Conference on Principles of Knowledge Representation and Reasoning*, pag. 505-516. 1989.
- [Pea90] J. Pearl. System Z: a natural ordering of defaults with tractable applications to default reasoning. In *M. Vardi, ed. Proceedings Third Conference on Theoretical Aspects of Reasoning About Knowledge*. Pag. 121-135. 1990.
- [PJ95a] Y. Permpoontanalrp and J.Y. Jiang. Generalized proof-theory for multi-agent autoepistemic reasoning. *Proceedings of the ICMAS*, pag. 304-311. 1995.
- [Poo88] D. Poole. A Logical Framework for Default Reasoning. *Artificial Intelligence*, 36:27-47. 1988
- [RC81] R. Reiter and G. Criscuolo. On Interacting Defaults. *Proc. IJCAI81*. 1981.
- [Rei80] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81-132. 1986.
- [Rei84] R. Reiter. Towards a logical reconstruction of relational database theory. In M. Brodie, J. Mylopoulos, and J. W. Schmidt, editors, *On Conceptual Modelling*. Pag. 163-189. Springer Verlag, New York. 1984.
- [RG77] R. B. Robert and I. Goldstein. The FRL Manual. *AI memo 409*. MIT, Cambridge, MA. 1977.
- [Sha95] M. Shanahan. A circumscriptive calculus of events. *Artificial Intelligence*, 77(2):249-284. 1995.
- [Sho87] Y. Shoham. A semantical approach to non-monotonic logics. *Proceedings IJCAI'87*. pag. 388-392. 1987.
- [Sho88] Y. Shoham. *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence*. MIT Press. Cambridge, MA. 1988.
- [Sow84] J. Sowa. *Conceptual Structures: Information Proceeding in Mind and Machine*. Addison-Wesley, Reading, MA. 1984.

- [ZK92] L. Zadeh and J. Kacprzyk. *Fuzzy Logic for the Management of Uncertainty*. John Wiley & Sons Inc. 1992.