



UNIVERSIDADE TÉCNICA DE LISBOA  
INSTITUTO SUPERIOR TÉCNICO

## Revisão adaptativa de crenças

Ana Margarida de Jesus Cardoso  
(Licenciada)

Dissertação para obtenção do Grau de Mestre em  
Engenharia Electrotécnica e de Computadores

Tese realizada sob a orientação de  
João Emílio Segurado Pavão Martins  
Professor Catedrático do Instituto Superior Técnico  
da Universidade Técnica de Lisboa

Constituição do Júri:

Doutor João Emílio Segurado Pavão Martins  
Professor Catedrático, do Instituto Superior Técnico, da Universidade Técnica de Lisboa

Doutor Helder Manuel Ferreira Coelho  
Professor Catedrático, da Faculdade de Ciências, da Universidade Técnica de Lisboa

Doutora Maria dos Remédios Vaz Pereira Lopes Cravo  
Professora Auxiliar, do Instituto Superior Técnico, da Universidade Técnica de Lisboa

Agosto de 1997



UNIVERSIDADE TÉCNICA DE LISBOA  
INSTITUTO SUPERIOR TÉCNICO

## Revisão adaptativa de crenças

Ana Margarida de Jesus Cardoso  
(Licenciada)

Dissertação para obtenção do Grau de Mestre em  
Engenharia Electrotécnica e de Computadores

Tese realizada sob a orientação de  
João Emílio Segurado Pavão Martins  
Professor Catedrático do Instituto Superior Técnico  
da Universidade Técnica de Lisboa

Constituição do Júri:

Doutor João Emílio Segurado Pavão Martins  
Professor Catedrático, do Instituto Superior Técnico, da Universidade Técnica de Lisboa

Doutor Helder Manuel Ferreira Coelho  
Professor Catedrático, da Faculdade de Ciências, da Universidade Técnica de Lisboa

Doutora Maria dos Remédios Vaz Pereira Lopes Cravo  
Professora Auxiliar, do Instituto Superior Técnico, da Universidade Técnica de Lisboa

Agosto de 1997

# Resumo

Ao modelar agentes racionais, é necessário dotá-los com a capacidade de mudar as suas crenças, isto é, de fazer com que acreditem em nova informação, e/ou deixem de acreditar em informação em que acreditavam anteriormente. Isto porque os agentes normalmente estão inseridos num mundo em mudança, acerca do qual têm informação incompleta, e por isso podem saltar para conclusões que mais tarde descobrem que estão erradas. Normalmente isto acontece quando o agente descobre que as suas crenças contêm informação contraditória. Neste caso, as teorias de revisão de crenças fazem com que o agente abandone uma ou mais das crenças que deram origem à contradição.

Considerando que são necessários recursos para adquirir informação e para fazer inferências a partir dela, propomos um método de mudança das crenças que não implique o desperdício de recursos associado com o seu abandono, mas que adapte a informação às novas circunstâncias. Este método é sugerido pela filosofia utilizada na evolução de teorias científicas: quando temos uma teoria científica que já previu correctamente muitos acontecimentos e surge uma observação que a contraria, é mais fácil tratar a observação como uma excepção e manter a teoria, que abandoná-la por completo apenas devido a uma observação que a contraria.

Neste trabalho, propomos uma nova abordagem à revisão de crenças, em que, como resultado da detecção de uma contradição, as crenças do agente são alteradas de forma a deixar de ter a contradição, mas mantendo também as conclusões que ainda são consistentes com o resto das suas crenças. Para isso, alteramos uma teoria de revisão de crenças e um sistema de revisão de crenças existentes. Descrevemos também a implementação de um sistema de revisão de crenças que efectua revisão adaptativa de crenças e apresentamos alguns exemplos deste tipo de revisão.

# Abstract

When modeling rational agents, we must provide them with the capability of changing their beliefs, that is, to believe new information or to disbelieve something that they formerly believed. This is because agents exist in a changing world, about which they have incomplete information; for this reason, they can jump to a conclusion that later on proves to be wrong. Usually, this happens when the agent finds out that there is a contradiction among its beliefs. In this case, belief revision theories lead the agent to give up its belief in one of the beliefs that caused the contradiction.

If we consider that we spend resources in acquiring information and in making inferences from them, we propose a way of changing the agent's beliefs without abandoning them, but adapting them to the new situation. This method is suggested by the way scientific theories evolve: when we have a scientific theory that has already correctly predicted many events but we find some contrary evidence, it is easier to treat this evidence as an exception and keep the theory, than to abandon the theory altogether, because of a single contrary evidence.

In this thesis we change a belief revision theory and a truth maintenance system so that, upon detection of a contradiction, the agent's beliefs are changed in order to resolve the contradiction, but keeping all the conclusions that still comply with the rest of the agent's knowledge. In order to do this, we change an existing belief revision theory and a truth maintenance system. We also describe the implementation of a truth maintenance system which adaptatively revises its beliefs and show some examples of this kind of revision.

## **Palavras Chave**

Representação do conhecimento

Lógicas não-monótonas

Raciocínio não-monótono

Teorias de revisão de crenças

Sistemas de revisão de crenças

Revisão adaptativa de crenças

## **Keywords**

Knowledge Representation

Non-monotonic logics

Non-monotonic reasoning

Belief revision theories

Truth maintenance systems

Adaptative belief revision

# Agradecimentos

Em primeiro lugar, gostaria de agradecer ao Professor João Pavão Martins, meu orientador, por me ter criado o interesse nesta área que levou ao desenvolvimento desta tese.

Aos membros do Grupo de Inteligência Artificial em geral, por me proporcionarem um ambiente onde é agradável trabalhar. Em especial à Professora Maria dos Remédios por, ao longo da orientação do meu Trabalho Final de Curso, me ter ajudado a dar os primeiros passos nestas áreas. Um obrigado especial também à Sofia Pinto Leitão e ao António Leitão pelo incentivo e pelas conversas acerca deste assunto.

Aos meus pais e irmã, por muitas vezes terem abdicado da minha companhia ao longo destes anos e pelo apoio moral que me têm dado.

E, “the last but not the least”, ao João, meu namorado e agora também marido, por toda a compreensão, apoio, discussões, e um sem fim de outras coisas que tem sido a nossa vida ao longo destes anos.

Esta tese foi parcialmente apoiada pelos projectos PBIC/TIT/1243/92 da Junta Nacional de Investigação Científica e Tecnológica (JNICT) e PRAXIS XXI 2/2.1/TIT/1568/95.

Agosto de 1997

Ana Margarida de Jesus Cardoso

# Índice

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Inteligência Artificial . . . . .	1
1.2	Raciocínio não-monótono . . . . .	2
1.3	Lógicas não-monótonas . . . . .	4
1.4	Teorias de revisão de crenças . . . . .	4
1.5	Sistemas de revisão de crenças . . . . .	6
1.6	Revisão adaptativa de crenças . . . . .	6
<b>2</b>	<b>Perspectiva da área</b>	<b>9</b>
2.1	Teorias de revisão de crenças . . . . .	9
2.1.1	A teoria de Gärdenfors . . . . .	10
2.1.1.1	Os estados de crença na AGM . . . . .	10
2.1.1.2	As operações sobre as crenças . . . . .	11
2.1.1.3	Problemas com a teoria AGM . . . . .	17
2.1.2	A teoria de Nebel . . . . .	18
2.1.2.1	Bases de crenças finitas . . . . .	18
2.1.2.2	Operações sobre bases de crenças . . . . .	19
2.1.2.3	Problemas com a teoria de Nebel . . . . .	21
2.1.3	A teoria de Fuhrmann . . . . .	21
2.1.4	Teorias de coerência e teorias de fundamentos . . . . .	22

2.1.4.1	Críticas às teorias dos fundamentos . . . . .	22
2.1.4.2	As teorias de fundamentos em Inteligência Artificial . . . . .	24
2.2	Lógicas Não Monótonas e Raciocínio Não Monótono . . . . .	25
2.2.1	A lógica da omissão de Reiter . . . . .	26
2.3	Sistemas de revisão de crenças . . . . .	28
2.3.1	Sistemas baseados em justificações . . . . .	29
2.3.2	Sistemas baseados em suposições . . . . .	29
<b>3</b>	<b>Formalismos usados neste trabalho</b>	<b>31</b>
3.1	A lógica SWMC . . . . .	31
3.1.1	A linguagem da SWMC — fórmulas bem formadas . . . . .	31
3.1.2	As fbfs suportadas — registo de dependências . . . . .	32
3.1.3	As regras de inferência . . . . .	33
3.1.4	A noção de consequência — contextos e espaços de crenças . . . . .	34
3.2	A teoria de revisão de crenças baseada na SWMC . . . . .	36
3.2.1	Adição de uma fbfs a um contexto . . . . .	37
3.2.2	Remoção de uma fbfs de um contexto . . . . .	38
3.2.3	Revisão de um contexto com uma fbfs . . . . .	39
3.3	O sistema de revisão de crenças SRC . . . . .	40
3.3.1	Estruturas do SRC . . . . .	41
3.3.2	O motor de inferência . . . . .	41
3.3.3	O revisor de crenças . . . . .	42
<b>4</b>	<b>Revisão Adaptativa de Crenças</b>	<b>45</b>
4.1	Motivação . . . . .	45
4.2	Critérios para a selecção de uma regra para enfraquecer . . . . .	48
4.2.1	Utilização de preferências . . . . .	48
4.2.2	Preferências como número de excepções . . . . .	51

<b>5</b>	<b>Alterações aos formalismos usados neste trabalho</b>	<b>54</b>
5.1	Alterações à teoria de revisão de crenças . . . . .	54
5.1.1	Revisão adaptativa de um contexto com uma fbf . . . . .	55
5.2	Alterações ao sistema de revisão de crenças . . . . .	56
5.2.1	Alterações ao motor de inferência . . . . .	56
5.2.1.1	Transformação de regras universais em regras de omissão	57
5.2.1.2	Como determinar a suposição correcta? . . . . .	59
5.2.1.3	Enfraquecimento do quantificador universal . . . . .	60
5.2.2	Alterações ao revisor de crenças . . . . .	61
5.2.2.1	Remoção de uma fbf devido à detecção de uma contradição	62
<b>6</b>	<b>Um src capaz de rever adaptativamente as crenças</b>	<b>66</b>
6.1	O SNePS . . . . .	66
6.2	O SNePSwD . . . . .	68
6.2.1	Exemplo de utilização do SNePSwD . . . . .	69
6.3	O novo sistema de revisão de crenças — SNePSwDaAR . . . . .	76
6.3.1	O comando weaken-wff . . . . .	76
6.3.2	Exemplo de utilização do novo sistema SNePSwDaAR . . . . .	77
6.4	Aplicações do novo sistema de revisão de crenças . . . . .	85
<b>7</b>	<b>Trabalho futuro</b>	<b>87</b>
7.1	Evolução da lógica SWMC . . . . .	87
7.2	Tópicos em aberto na área da revisão de crenças . . . . .	88
7.3	Criação de um módulo de aprendizagem para o SRC . . . . .	89
7.3.1	Tratamento de excepções: indução de regras . . . . .	90
7.3.2	Fortalecimento de regras: passagem do particular para o geral . . .	90
7.3.3	Enfraquecimento de regras: passagem do geral para o particular .	91
7.4	Aumento da funcionalidade do SNePSwDaAR . . . . .	92

<i>ÍNDICE</i>	iv
7.5 Relação com a evolução de teorias científicas . . . . .	92
<b>A Notação utilizada</b>	<b>94</b>
<b>B Abreviaturas utilizadas</b>	<b>95</b>
<b>C Regras de inferência da lógica SWMC</b>	<b>96</b>
<b>D Regras de inferência do novo src</b>	<b>98</b>

# Capítulo 1

## Introdução

Neste capítulo fazemos o enquadramento deste trabalho na área da Inteligência Artificial. Introduzimos os conceitos de *raciocínio não-monótono*, *lógica não-monótona*, *teoria de revisão de crenças* e *sistema de revisão de crenças*, tal como serão utilizados no resto deste documento. Descrevemos também, em traços gerais, as principais contribuições deste trabalho para a área da Inteligência Artificial.

### 1.1 Inteligência Artificial

Existem muitas definições para o que é a Inteligência Artificial (IA), e nem todas estão de acordo. Por esta razão, vamos começar por explicitar o que entendemos ser a IA, uma vez que esta noção vai condicionar todo este trabalho.

Do nosso ponto de vista, a IA é o ramo da informática que se preocupa com o estudo e com a criação de *agentes racionais*. Neste trabalho usamos a palavra “agente” com o significado de um sistema computacional que utilize conhecimento para exibir comportamento inteligente, não nos comprometendo com nenhum modelo de agente em particular. Este sistema deve ter a capacidade de receber informação do exterior e de fornecer informação ao exterior quando solicitado.

Embora o Ser Humano possa servir de modelo para o agente que se pretende construir, não é obrigatório que este se lhe assemelhe em todas as características. Se possível, pretendemos construir agentes que sejam ainda melhores do que o Ser Humano a desempenhar a tarefa para a qual forem construídos.

Esta opinião está de acordo com a que é expressa em [Russell & Norvig 1995], embora as nossas preocupações sejam mais limitadas do que as destes autores: Russell e Norvig preocupam-se com o raciocínio e com as acções dos agentes racionais; neste trabalho, a preocupação fundamental é apenas com o raciocínio de agentes racionais.

Um agente com capacidade de raciocínio precisa de representar de alguma forma a informação (conhecimento) acerca da qual vai raciocinar. Consideramos que o agente tem uma *base de conhecimento* onde representa o seu conhecimento acerca do mundo. As estruturas que o agente utiliza para representar o seu conhecimento são chamadas as *crenças* do agente, isto é, aquilo que o agente acredita ser verdade acerca do mundo em que está inserido. As crenças podem ser atómicas (por exemplo, *O Piupiu é uma ave*) ou ser regras com símbolos lógicos,<sup>1</sup> que permitem ao agente derivar nova informação a partir da informação que tem na sua base de conhecimento (por exemplo, *Todas as aves voam*).

O raciocínio do agente é efectuado por um motor de inferência, que a partir da informação representada na sua base de conhecimento pode inferir nova informação, que será também representada, para depois poder ser usada em novas inferências. Por exemplo, a partir das crenças *o Piupiu é uma ave* e *todas as aves voam*, o motor de inferência deveria inferir que *o Piupiu voa*.

Nas restantes secções deste capítulo introduzimos alguns dos conceitos que vão ser mais usados ao longo deste trabalho: na secção 1.2 introduzimos o conceito de raciocínio não-monótono e explicamos porque é que este tipo de raciocínio é necessário; na secção 1.3 introduzimos as lógicas não-monótonas, que permitem representar conhecimento geral acerca do mundo e efectuar raciocínio não-monótono; na secção 1.4 introduzimos as teorias de revisão de crenças, que vão ser usadas para guiar as mudanças nas crenças do agente; na secção 1.5 introduzimos os sistemas de revisão de crenças, que são sistemas computacionais que servem para propagar as mudanças feitas nas crenças do agente à totalidade das suas crenças; finalmente, na secção 1.6, apresentamos um resumo da contribuição deste trabalho, a revisão adaptativa de crenças, que consideramos mais útil do que a revisão efectuada pelas teorias descritas na secção 1.4.

## 1.2 Raciocínio não-monótono

Para motivar a necessidade da existência do *raciocínio não-monótono* [Minsky 1975] vamos utilizar um exemplo clássico desta área. Suponhamos que o agente tinha na sua base de conhecimento a seguinte informação:

*O Piupiu é uma ave*

*As aves voam*

*Os pinguins são aves*

*Os pinguins não voam*

---

<sup>1</sup>Durante a descrição do agente, vamos considerar que o seu conhecimento é representado usando uma lógica.

Quando questionássemos o agente acerca da capacidade de voar do Piupiu, gostaríamos que ele respondesse que *o Piupiu voa*. No entanto, se mais tarde acrescentarmos à base de conhecimento do agente a informação de que *o Piupiu é um pinguim*, o agente deveria ser capaz de mudar as suas crenças, deixando de acreditar que *o Piupiu voa* e passando a acreditar que *o Piupiu não voa*.

O raciocínio não-monótono é caracterizado precisamente pelo facto de a adição de nova informação à base de conhecimento do agente poder fazer com que o agente deixe de acreditar em algumas das suas crenças anteriores.

Este tipo de raciocínio é particularmente útil quando o mundo acerca do qual o agente está a raciocinar não é estático, ou quando a informação que o agente tem acerca do mundo não é completa ou pode estar errada. Nestes casos, é importante que o agente possa deixar de acreditar em algumas das suas crenças anteriores quando adquire nova informação.

O raciocínio não-monótono também é essencial quando se pretende que o agente possa saltar para conclusões com base em informação incompleta, mesmo que algumas dessas conclusões possam estar erradas. Por exemplo, se o agente apenas souber que *o Piupiu é uma ave*, faz sentido que possa saltar para a conclusão que *o Piupiu voa*, mesmo sem ter a certeza de que ele não é um pinguim, uma vez que geralmente as aves voam e que o agente não tem informação que indique que o Piupiu não é uma ave como as outras. A este tipo de raciocínio dá-se o nome de *raciocínio de senso comum*, e é o que é utilizado por quase todos os agentes que tentam modelar e raciocinar acerca de fragmentos do mundo real.

Segundo [Lukaszewicz 1990], o raciocínio de senso comum baseia-se no conceito de racionalidade. Assim, o agente pode saltar para conclusões que sejam racionais dadas as crenças que tem na sua base de conhecimento, mesmo que mais tarde algumas dessas conclusões possam vir a ser invalidadas.

Este tipo de raciocínio coloca alguns problemas, nomeadamente:

1. Como deve ser representado o conhecimento que pode não estar certo;
2. Como é que o agente deve rever as suas crenças quando recebe nova informação, nomeadamente quando descobre que algumas das suas crenças estão erradas;
3. Como é que o agente deve propagar as alterações nas suas crenças, de uma forma eficiente, ao resto das suas crenças.

As lógicas não-monótonas surgiram para resolver o primeiro problema; as teorias de revisão de crenças para resolver o segundo e os sistemas de revisão de crenças para resolver o terceiro.

### 1.3 Lógicas não-monótonas

As lógicas não-monótonas surgiram para permitir a representação do conhecimento e para modelar o raciocínio de agentes que efectuem raciocínio de senso comum.

Por um lado, estas lógicas permitem representar crenças que indicam propriedades gerais do mundo, por exemplo, *geralmente as aves voam*, em vez de *todas as aves voam*; por outro, permitem fazer raciocínio com base em crenças que não são absolutamente certas, mas cujas consequências são racionais, se tivermos em linha de conta o resto do conhecimento do agente.

Neste trabalho vamos utilizar uma lógica não-monótona para representar o conhecimento e modelar o raciocínio de agentes racionais.

Por usarmos uma lógica não-monótona, o agente vai passar a ter a possibilidade de ter crenças que podem entrar em contradição com as suas outras crenças. Assim, de cada vez que o agente fizer uma alteração nas suas crenças, não tem forma que tem de garantir que elas continuam consistentes.

Voltando ao exemplo do Piupiu apresentado na secção 1.2, dada essa informação, se chegar a informação que *o Piupiu é um pinguim*, a base de conhecimento do agente passa a ter informação contraditória, isto é, vai ter a informação que *o Piupiu voa* e que *o Piupiu não voa*. As teorias de revisão de crenças surgiram para ajudar o agente a decidir o que fazer com as suas crenças, de modo a que a sua base de conhecimento volte a ser consistente.

### 1.4 Teorias de revisão de crenças

As teorias de revisão de crenças preocupam-se com o que deve acontecer às crenças do agente quando elas sofrem alguma alteração, nomeadamente, como é que o agente deve alterar as suas crenças face a nova informação que contrarie a informação que já se encontrava na sua base de conhecimento.

Inicialmente, as teorias de revisão de crenças foram estudadas no âmbito da Filosofia. Existem fundamentalmente dois tipos de teorias rivais, que podem ser usados na construção das teorias de revisão de crenças, as *teorias dos fundamentos* e as *teorias da coerência*, que foram distinguidas da seguinte forma em [Harman 1986]:

- Segundo as *teorias dos fundamentos*, quando se abandona uma crença devem também ser abandonadas todas as crenças que deixaram de ter razões para ser acreditadas;
- Segundo as *teorias da coerência*, devem abandonar-se apenas as crenças que estavam

envolvidas numa contradição, mantendo todas as outras que ainda forem consistentes com o restante conhecimento do agente.

Harman diz também que se considera que as teorias dos fundamentos são intuitivamente mais bem aceites, embora se reconheça que o ser humano tem mais tendência para usar uma teoria da coerência. Por estarem mais correctas do ponto de vista puramente lógico, são teorias dos fundamentos que são normalmente usadas em IA, uma vez que desta forma o agente apenas tem na sua base de conhecimento crenças nas quais tem razões para acreditar.

Neste trabalho vamos usar uma teoria dos fundamentos e consideramos que as crenças não são entidades isoladas, mas estão relacionadas umas com as outras. Por isso, não basta acrescentar novas crenças à base de conhecimento ou abandonar crenças erradas, é também necessário decidir o que deve acontecer com todas as que foram ou podem vir a ser derivadas a partir delas. A partir das relações existentes entre as várias crenças do agente podemos considerar que existem crenças de dois tipos, e que devem ser tratadas de modos diferentes quando é feita uma alteração na base de conhecimento do agente:

- *Crenças básicas* são crenças que não dependem de nenhuma outra, isto é, correspondem a informação que foi directamente adquirida pelo agente.
- *Crenças derivadas* são crenças que foram derivadas a partir de outras crenças do agente, sejam elas básicas ou derivadas.

As várias teorias de revisão de crenças indicam exactamente o que é que acontece com as crenças de um agente quando elas sofrem alguma alteração. As alterações são feitas de forma a que: o resultado de qualquer alteração a uma base de conhecimento consistente tenha como resultado outra base de conhecimento, também consistente; a mudança a fazer na base de conhecimento seja mínima, segundo alguns critérios que iremos discutir mais tarde.

Voltando uma vez mais ao exemplo do Piupiu da secção 1.2, vai ser necessário abandonar pelo menos uma das crenças que deram origem à contradição, para que apenas se possa tirar uma conclusão acerca da capacidade de voar do Piupiu. Podemos supor, por exemplo, que deixamos de acreditar que *o Piupiu é um pinguim*. Mas isto não é suficiente para que a base de conhecimento volte a ser consistente: é também necessário abandonar todas as crenças que foram derivadas a partir desta, nomeadamente, que *o Piupiu não voa*. Uma forma de fazer isso é eliminar todas as conclusões que o agente tiver derivado anteriormente e voltar a tirar todas as conclusões possíveis a partir da sua nova base de conhecimento.

Se, de um ponto de vista meramente teórico, esta abordagem está correcta, de um ponto de vista computacional não faz sentido repetir inferências que já foram efectuadas

antes, quando a maior parte da base de conhecimento se mantém inalterada. Os sistemas de revisão de crenças surgiram precisamente para resolver este problema.

## 1.5 Sistemas de revisão de crenças

Enquanto que as teorias de revisão de crenças se preocupam com o que deve acontecer com as crenças de um agente quando elas sofrem alguma alteração, os sistemas de revisão de crenças preocupam-se com o modo como essa alteração pode ser propagada à totalidade das crenças do agente de um modo eficiente.

Nos sistemas de revisão de crenças, as crenças do agente são representadas através de nós de uma rede de dependências [Doyle 1983], e as dependências entre as várias crenças são representadas como justificações entre os vários nós. A propagação das alterações à totalidade das crenças do agente é feita usando a rede construída desta forma.

Os sistemas de revisão de crenças podem ser classificados segundo várias perspectivas: quanto à monotonicidade, quanto ao tipo de dependências que consideram para os nós e quanto à capacidade de fazer inferência.

1. Quanto à monotonicidade, os sistemas podem ser monótonos ou não-monótonos, isto é, podem considerar (não monótonos) ou não considerar (monótonos) justificações para os nós baseadas na ausência de informação.
2. Quanto ao tipo de suporte que consideram para os nós podem ser baseados em justificações (associa com cada nó os nós que lhe deram origem) ou em suposições (associa com cada nó os nós que correspondem às crenças básicas que lhe deram origem).
3. Quanto à capacidade de inferência podem ou não ter capacidades dedutivas.

## 1.6 Revisão adaptativa de crenças

As teorias de revisão de crenças desenvolvidas até agora (por exemplo, [Gärdenfors 1988, Nebel 1989, Fuhrmann 1991, Cravo 1993a]) apenas consideram a adição ou a remoção de crenças da base de conhecimento do agente. No entanto, o abandono de uma crença pode implicar o abandono de muitas outras que ainda são consistentes com a base de conhecimento do agente e que poderiam continuar a ser acreditadas sem problemas.

Consideremos mais uma vez que o agente tinha a seguinte informação:

*O Piupiu é uma ave*

*Todas as aves voam*

e que recebia a informação de que *o Piupiu não voa*. A nova crença que representa esta informação vai provocar uma contradição entre as crenças do agente, uma vez que o agente vai poder derivar que *o Piupiu voa* e tem a informação que *o Piupiu não voa*.

Se considerarmos que a informação específica acerca do Piupiu está correcta, por exemplo, por corresponder a uma observação directa que o agente fez do Piupiu, o agente deveria abandonar a regra que diz que *todas as aves voam*. Mas assim, uma vez que vamos usar uma teoria dos fundamentos, o agente deixaria de acreditar que todas as outras aves que existissem na sua base de conhecimento voavam. Neste caso, faria sentido seguir uma teoria da coerência e manter as crenças que representam o facto de as outras aves voarem, mesmo sem ter justificação para elas.

Mas podemos considerar uma outra abordagem, que consiste, não em abandonar a regra que diz que *todas as aves voam*, mas em substituí-la por uma mais fraca, que diz que *normalmente as aves voam*, e que admite a existência de excepções. Neste caso, as proposições acerca da capacidade de voar das outras aves continuariam a ser acreditadas, só que agora com base numa regra mais fraca, e apenas enquanto não fosse detectada uma contradição que as envolvesse.

Com esta abordagem, conseguimos manter tanto conhecimento como se estivéssemos a usar uma teoria da coerência, mas continuando a usar uma teoria dos fundamentos, o que nos permite ter sempre uma justificação, e conseqüentemente uma explicação, para todas as crenças existentes na base de conhecimento do agente.

Na realidade, este tipo de comportamento pode também ser observado no Ser Humano: posto perante o mesmo problema, um Ser Humano iria adaptar as suas crenças de forma a deixar de ter a contradição, mas continuaria a acreditar que as outras aves voam, e que o Piupiu é “a excepção que confirma a regra”. Nas teorias científicas observa-se um fenómeno semelhante: uma teoria que preveja correctamente as observações efectuadas durante um longo período de tempo é confirmada à custa dessas previsões; por isso, é natural que, quando surgirem as primeiras observações que contrariem a teoria elas sejam consideradas como excepções, mas que a teoria continue a ser utilizada tal como estava (a este respeito ver, por exemplo, [Quine & Ullian 1978, pp. 20–34]). Foi o que aconteceu, por exemplo, com a mecânica clássica (modelada pelas Leis de Newton), que só algumas décadas depois das primeiras observações que contrariavam é que foi substituída pela teoria da relatividade de Einstein, que a inclui como caso particular, mas que revolucionou completamente os conceitos de tempo e de espaço que eram considerados até essa altura [Kuhn 1970, pp. 98–99, 101–102].

A capacidade de adaptação das crenças em vez do seu abandono também pode ser útil quando o agente tem várias fontes de informação, nas quais pode ter diferentes graus de confiança, e que podem fornecer informação contraditória entre si. Neste caso, é im-

portante que o agente consiga decidir em que informação é que vai acreditar e adaptar o resto das suas crenças de modo a que a sua base de conhecimento seja consistente. Mais uma vez, este tipo de comportamento também pode ser observado no Ser Humano: se um amigo de longa data e um outro, conhecido por ser mentiroso, relatarem um acontecimento de formas contraditórias, geralmente vamos acreditar no amigo de longa data (mesmo que numa dada situação seja o mentiroso a dizer a verdade...).<sup>2</sup> Eventualmente, podemos acreditar em alguns dos pormenores relatados pelo mentiroso, pelo menos enquanto não surgir um novo relato que os contrarie, o que poderia acontecer facilmente se ouvirmos primeiro o mentiroso e só depois encontrarmos o nosso amigo, que conta a outra versão. Neste caso, vamos ter que rever as nossas crenças, possivelmente adaptando-as aos novos dados, de modo a que o nosso conhecimento volte a ser consistente.

Neste trabalho propomos a revisão adaptativa de crenças e estudamos a sua incorporação num sistema computacional. Para isso, estudamos quais os mecanismos de que este necessita e quais as alterações que devem ser introduzidas nas teorias de revisão de crenças tradicionais. Para servir de base ao desenvolvimento do sistema, expandimos a teoria de revisão de crenças descrita em [Cravo 1992, Cravo 1993a], para que passe a considerar também o enfraquecimento de crenças. Alteramos também a implementação de um sistema de revisão de crenças que utiliza essa teoria [Cravo 1992, Cravo 1995], para que a nova teoria possa ser testada.

Esta extensão foi feita com dois objectivos em vista:

- Modelar parte do raciocínio de senso comum;
- Tornar mais robustos os sistemas computacionais que utilizem a nova teoria de revisão de crenças.

O resultado desta tese pode ser utilizado, por exemplo, para auxiliar a construção incremental de bases de conhecimento, detectando contradições e sugerindo formas de as resolver. Por exemplo, se tivermos uma base de conhecimento acerca de algumas aves marítimas que voam (gaiotas, pelicanos, etc.), faz sentido ter uma regra que represente o facto de que *todas as aves voam*, o que facilita o raciocínio ao nível da lógica. Se essa base de conhecimento for estendida para incluir mais aves marítimas que não voam (pinguins), a regra anterior deve ser adaptada para passar a incluir também aves que não voam, e por isso deve passar a ser uma regra geral, que representa o facto de que *normalmente as aves voam*, mas que pode ter excepções.

---

<sup>2</sup>Quando existem várias fontes de informação, pode fazer sentido dizer que se preferem as crenças que provêm de uma determinada fonte em detrimento das que provêm de outra. A especificação de preferências entre crenças vai ser discutida na secção 4.2.1.

## Capítulo 2

# Perspectiva da área

Neste capítulo fazemos uma perspectiva da área em que se enquadra esta tese, a área das teorias de revisão de crenças. Descrevemos também uma lógica não-monótona e os sistemas de revisão de crenças em geral, pois alguns dos conceitos aqui introduzidos serão necessários para compreender os restantes capítulos desta tese.

### 2.1 Teorias de revisão de crenças

*As teorias de revisão de crenças* descrevem de que forma se devem processar as alterações nas crenças de um agente racional. Estas teorias, desenvolvidas inicialmente por filósofos, têm merecido recentemente maior atenção em IA.

Tal como já foi dito, os investigadores em IA preocupam-se com a realização de sistemas computacionais práticos, que possam demonstrar inteligência. Estes sistemas poderão depois ser utilizados para ajudar o ser humano a desempenhar algumas das suas tarefas, raciocinando acerca de alguns fragmentos do mundo real. Estas tarefas podem ser tão simples como automatizar a classificação taxonómica de animais ou tão complexas como o planeamento do percurso dos materiais numa linha de montagem, ou o aconselhamento de percursos entre dois locais à escolha. Mas o mundo real é demasiado complexo para poder ser totalmente apreendido, e qualquer agente real tem limitações na sua capacidade de percepção do mundo.

Estas características exigem que um agente inteligente tenha a capacidade de rever as suas crenças. Para além disso, impõem certos constrangimentos às teorias de revisão de crenças, que tornam impraticáveis algumas das teorias desenvolvidas por filósofos. Essas teorias, no entanto, serviram de ponto de partida para as teorias desenvolvidas posteriormente em IA.

Enquanto que num Ser Humano é concebível que as crenças possam sofrer alterações

devido a processos subconscientes, não tendo necessariamente uma explicação racional, um aspecto comum a todas as teorias de revisão de crenças que vamos apresentar é o facto de se preocuparem apenas com o estudo da revisão de crenças *racional*.

Da mesma forma que a lógica caracteriza as inferências que um agente perfeitamente racional pode fazer com o conhecimento que tem, as teorias de revisão de crenças tentam caracterizar as alterações que um agente racional deve fazer nas suas crenças face a nova informação.

De seguida vamos apresentar uma perspectiva do trabalho feito na área das teorias de revisão de crenças. A primeira formalização da revisão de crenças racional foi efectuada em [Gärdenfors 1988], onde são apresentadas as várias operações que um agente racional deve poder efectuar sobre as suas crenças e se introduz o conceito de valor epistémico de uma crença. Após a apresentação desta teoria, descrevemos alguns dos seus problemas quando aplicada na construção de um sistema computacional, e apresentamos alguns dos trabalhos feitos posteriormente para resolver parte desses problemas. A teoria de [Nebel 1990] define as mesmas operações, mas sobre bases de crenças finitas, tendo em conta que se pretende implementar um sistema que siga a teoria nas alterações que faz às suas crenças. [Fuhrmann 1991] apresenta uma outra teoria de revisão de crenças, que também lida com bases finitas, mas que não exige que o valor epistémico das crenças seja representado usando uma ordem total.

### 2.1.1 A teoria de Gärdenfors

A primeira Teoria de Revisão de Crenças que vamos descrever foi desenvolvida por filósofos [Alchourrón, Gärdenfors & Makinson 1985] e serviu de base para todos os outros trabalhos que vamos apresentar. É normalmente designada por teoria AGM, devido aos seus autores.

A descrição mais completa da teoria AGM, incluindo motivação e aplicações, foi feita em [Gärdenfors 1988]. A teoria AGM descreve o processo de revisão de crenças ao nível do conhecimento,<sup>1</sup> não se comprometendo com uma representação simbólica particular das crenças do agente. A teoria define critérios de racionalidade que devem guiar as mudanças nas crenças de um agente racional ideal.

#### 2.1.1.1 Os estados de crença na AGM

O *estado de crença* de um agente designa o estado cognitivo em que o agente se encontra num determinado instante, isto é, o que é que ele sabe e em que é que acredita.

A teoria AGM estuda as operações que alteram as crenças de um agente, utilizando

---

<sup>1</sup>No sentido de [Newell 1982].

a lógica proposicional para representar os seus estados de crença. Um estado de crença é representado por um conjunto de fórmulas: as fórmulas pertencentes ao conjunto correspondem àquilo em que o agente acredita, isto é, à sua base de conhecimento.

A teoria AGM estuda as operações de mudança nas crenças, não de um agente real, mas sim de um agente ideal perfeitamente racional. Por isso, é necessário impor certos critérios de racionalidade que os conjuntos de fórmulas devem satisfazer para que possam representar um estado de crença racional:

- o conjunto de fórmulas deve ser consistente, isto é, não deve conter fórmulas contraditórias entre si;
- o conjunto deve ser fechado dedutivamente, isto é, todas as consequências lógicas das fórmulas existentes no conjunto devem também pertencer ao conjunto.

Aos conjuntos de fórmulas proposicionais que satisfazem estes critérios de racionalidade e que, portanto, representam estados de crença de um agente, chamamos *conjuntos de crenças*.

### 2.1.1.2 As operações sobre as crenças

A representação escolhida para modelar as crenças de um agente determina as operações de mudança de crenças que podem ser definidas. A utilização de conjuntos de crenças para a representação das crenças faz com que só existam três operações possíveis.

Na realidade, dado um conjunto de crenças  $\Delta$ , e uma fórmula  $A$  que representa uma crença, ou  $A \in \Delta$ , ou  $A \notin \Delta$ .<sup>2</sup> Neste último caso, podemos ter ainda  $\neg A \in \Delta$  ou  $\neg A \notin \Delta$ . A estas três possibilidades correspondem três *atitudes*<sup>3</sup> possíveis para a crença representada por  $A$ :

- a crença é *aceite*, se  $A \in \Delta$ ;
- a crença é *rejeitada*, se  $\neg A \in \Delta$ ;
- a crença é *indeterminada*, se  $A \notin \Delta$  e  $\neg A \notin \Delta$ .

A utilização de conjuntos para representar as crenças de um agente permite apenas como alterações a adição de uma nova crença e a remoção de uma crença existente. No caso da adição de uma crença, existem ainda duas situações possíveis que interessa distinguir: se a nova crença entra em contradição com o conjunto de crenças do agente ou não, isto é, se a crença era rejeitada ou indeterminada. São estas três situações que dão origem às três operações definidas na teoria AGM:

<sup>2</sup>A notação utilizada ao longo deste trabalho está descrita no apêndice A.

<sup>3</sup>Gärdenfors utiliza o termo *atitude epistémica*.

1. *Expansão*: passar a aceitar (ou rejeitar) uma crença que anteriormente era indeterminada;
2. *Contração*: deixar de acreditar numa crença (ou na sua negação) que era aceite (rejeitada) anteriormente, passando esta a ser indeterminada;
3. *Revisão*: passar a aceitar uma crença que anteriormente era rejeitada (ou vice-versa).

Embora Gärdenfors não tenha considerado este assunto, devemos notar que, se utilizarmos outro tipo de representação para as crenças de um agente, pode ser possível a existência de outras operações. Por exemplo, se a representação fizesse distinção entre crenças plausíveis e crenças certas, talvez fizesse sentido uma operação que alterasse uma crença certa para plausível, ou vice-versa.

A formalização das operações sobre as crenças do agente é feita impondo, para cada tipo de operação, uma série de postulados de racionalidade que devem ser satisfeitos pelas possíveis definições dessa operação. Esses postulados podem servir para caracterizar completamente a operação, como no caso da expansão, ou apenas para impor restrições a essas definições, como nos casos da contração e da revisão.

Alguns dos critérios de racionalidade baseiam-se no facto de estarmos a utilizar conjuntos de crenças e de termos uma lógica subjacente. Por exemplo, a exigência que todas as consequências de uma crença sejam aceites quando ela é aceite.

Para além destes critérios, existem outros que são independentes da representação escolhida para as crenças do agente e que, normalmente, são adoptados como princípios desejáveis em qualquer teoria de revisão de crenças. Um desses critérios, e o mais importante em todas as operações da teoria AGM, é o critério da *economia de informação*, ou da *mudança mínima* (introduzido em [Quine & Ullian 1978] e mais tarde em [Harman 1986, pp. 46]). De acordo com este critério, as alterações efectuadas nas crenças de um agente devem ser mínimas. Subjacente a este critério está o conceito de que a informação é preciosa e difícil de obter. Assim sendo, não devem ser abandonadas crenças de uma forma gratuita, mas apenas nos casos em que temos razões para as abandonar.

Apesar de universalmente aceite como princípio, o conceito de mudança mínima depende da representação utilizada. No caso dos conjuntos de crenças, isso corresponde a manter o maior número possível de fórmulas após a mudança.

Nas secções seguintes apresentamos cada operação considerada na AGM em mais detalhe. Para uniformizar a notação ao longo da descrição dos vários formalismos descritos neste trabalho, vamos usar a notação que está descrita no apêndice A. A notação relevante para a teoria AGM é a seguinte:

- os conjuntos de crenças, fechados dedutivamente, são representados por letras greegas maiúsculas:  $\Delta, \Sigma, \dots$

- as fórmulas são representadas por letras romanas maiúsculas:  $A, B, \dots$ ;
- as operações de expansão, contracção e revisão de um conjunto de crenças  $\Delta$  com uma fórmula  $A$ , são representadas, respectivamente, por  $(\Delta + A)$ ,  $(\Delta - A)$ ,  $(\Delta * A)$ ;
- o conjunto de todas as consequências de um conjunto finito de fórmulas  $\alpha$  é representado por  $Cn(\alpha)$ ;
- o conjunto  $\Delta_{\perp}$  representa o conjunto de crenças inconsistente, isto é, o conjunto de todas as fórmulas da linguagem.

**A operação de expansão** A *expansão* é a mais simples das três operações que alteram um conjunto de crenças: uma crença que era indeterminada passa a ser aceite pelo agente.

Gärdenfors apresenta seis postulados de racionalidade que uma operação de expansão deve satisfazer. Os postulados têm, fundamentalmente, a ver com:

- o sucesso da operação— a crença com a qual se faz a expansão de um conjunto de crenças deve ser aceite no resultado da expansão, que deve continuar a ser um conjunto de crenças (isto é, fechado dedutivamente);
- o critério da mudança mínima— a operação de expansão não deve abandonar crenças mantidas anteriormente, nem adicionar crenças para as quais não existam razões para serem acreditadas.

Tendo em conta os seis postulados para a expansão, prova-se que a única operação que satisfaz esses postulados é a operação que adiciona a crença ao conjunto de crenças e faz o fecho dedutivo desse conjunto. Formalmente, a operação de expansão é definida na AGM por

$$\Delta + A = Cn(\Delta \cup \{A\})$$

Repare-se que, se  $\neg A \in \Delta$ , então  $\Delta + A = \Delta_{\perp}$ .

**A operação de revisão** A revisão corresponde à passagem de uma situação em que uma crença é rejeitada pelo agente a uma situação em que essa mesma crença é aceite. Este tipo de operação é muito importante no raciocínio de senso comum (e no raciocínio não-monótono), em que grande parte das crenças do agente são mantidas com base em informação incompleta ou suposições. Neste tipo de raciocínio é natural que, com a chegada de nova informação, o agente tenha que aceitar crenças que entram em contradição com as anteriormente mantidas.

Naturalmente, a revisão, ao contrário da expansão, é não-monótona, uma vez que não é possível manter todas as crenças anteriores quando se adiciona nova informação.<sup>4</sup>

Os critérios de racionalidade impostos na AGM para as operações são critérios lógicos, que as operações devem satisfazer de acordo com a lógica subjacente aos conjuntos de crenças. No caso da revisão, em que é necessário escolher que crenças devem ser abandonadas para manter a consistência, estes critérios limitam as alternativas possíveis, mas não são suficientes para determinar de forma única a escolha a fazer. Essa escolha só pode ser feita com base em informação que não está representada nos conjuntos de crenças: o *valor epistémico*<sup>5</sup> das várias crenças, as relações causais existentes entre elas, etc.

**A operação de contracção** A operação de contracção corresponde à remoção de uma crença do conjunto de crenças do agente, isto é, uma crença aceite passa a ser indeterminada. Esta operação permite que um agente racional possa abandonar crenças para as quais deixaram de existir razões que as suportassem, por exemplo.

A contracção de um conjunto de crenças com uma determinada crença tem como objectivo remover essa crença do conjunto de crenças do agente, isto é, a crença contraída não deve pertencer ao conjunto de crenças que resulta da contracção. No entanto, para obter isto não basta remover a crença do conjunto de crenças. É necessário remover também crenças que permitam derivar a crença que se pretende remover, uma vez que os conjuntos de crenças são fechados dedutivamente. E aqui, de novo, temos o mesmo problema que na revisão: não existe uma forma única de invalidar as derivações da crença a remover e os constrangimentos lógicos impostos pelos postulados à operação de contracção não são suficientes para fazer a escolha; logo, não vai ser possível determinar o seu resultado de forma única.

**Relação entre as operações de revisão e contracção** As operações de revisão e contracção, como vimos nas duas últimas secções, não ficam completamente determinadas a partir dos seus postulados. Os motivos para que isso aconteça são semelhantes em ambos os casos: não existir uma única alternativa para a remoção de fórmulas que, no caso da revisão evitem a contradição com a fórmula a adicionar, ou, no caso da contracção permitam derivar a crença que pretendemos remover.

Na realidade, a operação de revisão, de modo a manter a consistência do resultado final, tem necessidade de remover as crenças que entrem em conflito com a que se pretende adicionar, isto é, necessita de remover a negação da fórmula a adicionar. Esta observação sugere a existência de uma relação entre as operações de revisão e de contracção.

---

<sup>4</sup>A menos que a nova informação não entre em contradição com as crenças do agente, caso em que a operação de revisão corresponde a uma expansão.

<sup>5</sup>A importância e utilização dos valores epistémicos na escolha das alternativas será considerada na secção 2.1.1.2.

Essa relação, estabelecida por Levi em [Levi 1977], permite definir a operação de revisão à custa das operações de contracção e expansão. A definição da revisão fica assim estabelecida pela seguinte *identidade de Levi*:

$$(\Delta * A) = (\Delta - \neg A) + A$$

Esta definição expressa aquilo que tínhamos sugerido anteriormente: para efectuarmos a revisão de um conjunto de crenças  $\Delta$  com uma fórmula  $A$ , devemos primeiro remover  $\neg A$  do conjunto de crenças e depois adicionar  $A$  ao resultado. Repare-se que ao contrairmos  $\Delta$  com  $\neg A$  eliminamos todas as crenças que permitem derivar  $\neg A$  e assim é possível adicionar  $A$  sem que isso entre em contradição com as restantes crenças.

Existe uma relação idêntica que define a operação de contracção em termos das operações de revisão e expansão. Essa definição, proposta por Harper em [Harper 1987], é dada pela seguinte *identidade de Harper*:

$$(\Delta - A) = \Delta \cap (\Delta * \neg A)$$

Também neste caso é intuitivamente aceitável a definição. A revisão de  $\Delta$  com  $\neg A$  remove o menor número de crenças de  $\Delta$  de modo a poder adicionar  $\neg A$  sem obter uma contradição, ou seja,  $A$  e tudo o que permita derivar  $A$ . É precisamente isto que pretendemos que a operação de contracção faça. No entanto, a operação de revisão também adiciona  $\neg A$  ao conjunto de crenças, o que não desejamos no caso da contracção. A intersecção de  $\Delta$  com o resultado da revisão permite manter todas as crenças de  $\Delta$  que não tenham sido removidas pela revisão, isto é, todas as crenças de  $\Delta$ , excepto as que permitiam derivar  $A$ , obtendo assim o resultado desejado para a contracção.

A adequação destas duas definições é estabelecida em [Gärdenfors 1988, teoremas 3.2–3.5], onde Gärdenfors prova que a operação de revisão definida à custa de uma operação de contracção que satisfaça os postulados da contracção satisfaz todos os postulados da revisão, e, conversamente, a operação de contracção definida à custa de uma operação de revisão que satisfaça os postulados da revisão satisfaz todos os postulados para a contracção.

**Valores epistémicos** Os postulados para a revisão e contracção impõem constrangimentos que eventuais definições dessas operações têm que satisfazer, mas não nos permitem obter uma definição para as operações.

De modo a podermos operacionalizar a operação de contracção, precisamos de fornecer uma definição que, dado um conjunto de crenças  $\Delta$  e uma fórmula  $A$ , nos diga que crenças devemos abandonar de modo a remover  $A$  de  $\Delta$ .

Existem na realidade várias definições possíveis para esta operação, todas elas logi-

camente aceitáveis (isto é, satisfazendo os postulados). É necessário portanto, utilizar informação adicional para escolher entre as diversas alternativas. Designamos essa informação adicional por *valor epistémico*<sup>6</sup> das crenças.

O *valor epistémico* de uma crença mede a “força” que essa crença tem. Não é um valor absoluto, mas sim relativo—o valor epistémico de uma crença é dado relativamente ao de outra. Dizemos que o valor epistémico de uma crença  $A$  é maior do que o de uma crença  $B$ , se a crença em  $A$  é mais importante para o agente do que a crença em  $B$ . Do ponto de vista da contracção isto significa que, se o agente tiver que decidir que crença abandonar, prefere abandonar  $B$  a abandonar  $A$ .

O valor epistémico de uma crença pode ser visto como uma medida (qualitativa) do valor informativo, ou poder explicativo, que essa crença tem no conjunto de crenças do agente. Por isso mesmo, o valor epistémico depende fortemente do conjunto de crenças particular que o agente tenha, sofrendo alterações quando o conjunto de crenças é alterado.

É com base nessa informação que Gärdenfors propõe a definição de uma função de selecção que permite escolher a melhor de entre as várias alternativas possíveis para remover uma fórmula de um conjunto, e, conseqüentemente, da operação de contracção.

Os valores epistémicos entre as crenças definem uma ordem entre todas as fórmulas da linguagem, ordem essa que deve satisfazer certos postulados de racionalidade, explicados em [Gärdenfors 1988].

O valor epistémico que temos estado a considerar até agora apenas relaciona crenças individuais e não conjuntos de crenças. A função de selecção necessita é de uma ordem entre conjuntos de crenças para poder fazer a escolha: dada uma ordem entre os vários conjuntos de fórmulas que podem corresponder a contrair o conjunto de fórmulas  $\Delta$  com  $A$ , a função de selecção utiliza essa ordem para seleccionar os elementos que correspondem a máximos dessa ordem, que podem ser vários, uma vez que a ordem entre os conjuntos de fórmulas, ao contrário da ordem entre as crenças, não precisa de ser total. Representamos por  $\Delta \perp A$  o conjunto de subconjuntos máximos de  $\Delta$  que não contêm  $A$ .

Intuitivamente, a ordem entre os vários elementos de  $\Delta \perp A$  é obtida a partir dos valores epistémicos para as várias crenças da seguinte forma: dizemos que um conjunto de crenças é preferido em relação a outro se as crenças mantidas por ele são mais importantes do que as mantidas pelo outro. Utilizando o valor epistémico das crenças, Gärdenfors consegue assim definir a operação de contracção.

---

<sup>6</sup>Do inglês “epistemic entrenchment”.

### 2.1.1.3 Problemas com a teoria AGM

Para finalizar a apresentação da teoria AGM, vamos apresentar alguns dos problemas que se põem na sua utilização em IA, e, em particular, na construção de um agente racional real. As tentativas de resolução de alguns desses problemas serão consideradas nas secções seguintes.

Grande parte dos problemas apontados à AGM deve-se à escolha feita na teoria para a representação dos estados epistémicos dos agentes. Na realidade, a representação do estado epistémico do agente condiciona as atitudes que ele pode ter para com as suas crenças, bem como as operações sobre elas.

Um dos problemas abordados por Nebel em [Nebel 1989] é o de a utilização de teorias proposicionais fechadas tornar esta teoria pouco adequada na modelação de um agente racional em IA. Não é possível lidar com conjuntos de fórmulas infinitos num sistema computacional. Mesmo considerando representações finitas para os conjuntos de crenças não é fácil caracterizar as operações nessas representações. Na secção 2.1.2.1 apresentamos a proposta de Nebel para resolver este problema.

O outro problema abordado por Nebel diz respeito à forma como as operações definidas na AGM alteram as crenças. De acordo com Nebel, ao removermos uma crença de um conjunto de crenças, é natural supor que, sendo essa crença a única justificação existente para uma outra crença, esta última seja também abandonada. No entanto, nos conjuntos de crenças não é mantida informação acerca das justificações para as várias crenças, não permitindo abandonar as crenças derivadas.

Embora do ponto de vista da IA, enquanto disciplina que estuda agentes racionais, o registo das dependências entre as várias crenças seja desejável e essencial, essa posição não é adoptada por todas as abordagens existentes para a revisão de crenças. Na secção 2.1.4 vamos considerar os dois tipos de abordagens existentes nesta área apresentando os argumentos de cada uma delas.

Para além destes problemas apontados por Nebel, julgamos que o modelo escolhido na AGM para representar o estado epistémico de um agente tem outros problemas que dificultam a sua aplicação em IA.

A maior parte do trabalho desenvolvido na área de raciocínio de senso comum rejeita a utilização da lógica proposicional, e até mesmo da lógica de primeira ordem, como linguagem adequada para representar e raciocinar neste domínio. Novas lógicas foram desenvolvidas para lidar com os problemas específicos deste tipo de raciocínio, nomeadamente lógicas não-monótonas. Estando as operações sobre as crenças dependentes da representação escolhida para os estados epistémicos, é necessário reconsiderar essas operações quando consideramos outro tipo de linguagens para representar as crenças.

Um outro problema que não tem sido discutido até agora, e que de certa forma está relacionado com o anterior, é o facto de uma parte daquilo que nós julgamos ser conhecimento específico do agente—o valor epistémico das várias crenças—que é utilizado nas operações de revisão e contracção não estar sequer representado no estado epistémico do agente.

Para além destes dois problemas epistemológicos existe um outro, este de natureza mais pragmática: é essencial considerar o facto de que um agente racional real tem limitações à quantidade de raciocínio que pode efectuar; não é razoável esperar que seja possível construir um sistema computacional omnisciente.

Finalmente, para terminar esta lista (não exaustiva) de problemas que a teoria AGM levanta, convém referir o facto de os postulados introduzidos pela AGM para caracterizar as várias operações não serem adequados a todas as circunstâncias.

Katsuno e Mendelzon, em [Katsuno & Mendelzon 1992], mostram outra formalização para a operação de revisão, mais adequada do que a da AGM para lidar com mudanças no mundo.

### 2.1.2 A teoria de Nebel

Um dos principais problemas na aplicação da teoria AGM em IA é o facto de serem utilizados conjuntos de crenças fechados (necessariamente infinitos) para representar as crenças de um agente.

Na realidade, uma vez que em IA se pretende construir agentes computacionais, é necessário lidar com as alterações a fazer nas crenças de um agente estando estas representadas de uma forma finita.

Nebel, em [Nebel 1989], define as operações de revisão e contracção em bases finitas de crenças e estabelece relações entre as operações definidas desta forma e as operações da teoria AGM.

#### 2.1.2.1 Bases de crenças finitas

A característica fundamental desta nova abordagem é o facto de se considerar uma outra representação para as crenças do agente: em vez de serem representadas por conjuntos de crenças, as crenças passam a ser representadas por *bases de crenças finitas*, isto é, conjuntos de fórmulas finitas que não satisfazem o requisito de fecho dedutivo.

As fórmulas colocadas nestas bases de crenças correspondem às observações efectuadas pelo agente, às regras que caracterizam um determinado domínio e que constituem o conhecimento que o agente tem acerca desse domínio, etc. Naturalmente, a partir destas

bases de crenças, o agente pode deduzir novas crenças, usando a lógica subjacente (continuamos a assumir a lógica proposicional). As crenças do agente vão corresponder a tudo aquilo que o agente conseguir derivar a partir da sua base de crenças, com a operação de consequência lógica.

A utilização de bases de crenças permite ainda fazer uma distinção entre dois tipos de crenças: as que estão na base de crenças e que correspondem a crenças básicas do agente, e as que são derivadas a partir das outras. Esta distinção vai permitir obter um dos efeitos defendidos por Nebel: o de, ao abandonar uma crença, abandonar também as crenças derivadas a partir dela.

Um outro aspecto importante acerca da representação com bases de crenças é que existem várias bases de crenças diferentes que são logicamente equivalentes. Por exemplo,  $Cn(\{A, B\}) = Cn(\{A \wedge B\})$ . Mas, os resultados obtidos pela mesma operação em cada um destes casos podem ser diferentes. Isto não acontece no caso da AGM, uma vez que se consideram sempre conjuntos dedutivamente fechados.

### 2.1.2.2 Operações sobre bases de crenças

As operações que alteram as crenças do agente são agora definidas nestas bases de crenças: partem de uma base de crenças e produzem uma nova base de crenças que deve reflectir a alteração desejada. Todas as operações manipulam apenas as crenças básicas. Estas alterações nas crenças básicas induzem, obviamente, alterações nas crenças derivadas.

Na descrição que se segue das operações vamos utilizar  $(\alpha - A)$  e  $(\alpha * A)$  para representar, respectivamente, a contracção e a revisão da base de crenças  $\alpha$  com a crença representada por  $A$ .<sup>7</sup> A operação de expansão não é definida em [Nebel 1989], mas, como sabemos, esta operação é um caso particular da operação de revisão.

Nebel apresenta, para a operação de revisão, a seguinte definição:

$$(\alpha * A) = (\alpha - \neg A) \cup \{A\}$$

Esta definição é em tudo idêntica à identidade de Levi, agora no caso das bases de crenças. Repare-se que podemos considerar que a operação de expansão consiste em adicionar a crença à base de crenças, sem necessidade de efectuar o fecho dedutivo.

Estando a operação de revisão definida à custa da operação de contracção, apenas precisamos de nos preocupar com a operação de contracção.

Para a definição da operação de contracção  $(\alpha - A)$ , vamos utilizar, de acordo com o

---

<sup>7</sup>Nebel utiliza  $B \sim A$  e  $B \dot{\sim} A$ , respectivamente, para representar as operações, mas para mantermos a notação utilizada para a apresentação da AGM, adoptamos esta.

critério da mudança mínima, os subconjuntos máximos de  $\alpha$  que não permitem derivar  $A$ . Continuamos a designar o conjunto de todos esses conjuntos por  $\alpha \perp A$ .<sup>8</sup>

É com base nestes conjuntos que Nebel apresenta uma definição para a operação de contracção em bases de crenças. Para garantir a satisfação dos postulados para a operação de contracção definidos na AGM, incluindo o da recuperação, Nebel define a operação de contracção de modo a incluir um termo que lhe permita recuperar a informação quando  $A$  é adicionado:

$$(\alpha - A) = \begin{cases} (\bigvee_{C \in (\alpha \perp A)} C) \wedge (\alpha \vee \neg A) & \text{se } \{\} \not\vdash A \\ \alpha & \text{caso contrário} \end{cases}$$

A ideia subjacente a esta definição é a de manter tanta informação da base quanto possível, fazendo a disjunção quando houver várias alternativas.

Por exemplo, dada a base de crenças

$$\alpha = \{A, A \rightarrow B, C, D\}$$

temos

$$\alpha \perp B = \{\{A, C, D\}, \{A \rightarrow B, C, D\}\},$$

e, de acordo com a definição acima,

$$(\alpha - B) = \{(A \wedge C \wedge D) \vee ((A \rightarrow B) \wedge C \wedge D)\}.$$

O facto de a operação de contracção definida em bases satisfazer os postulados impostos na AGM é um bom indício da sua adequação. Por isto mesmo, é importante investigar que tipo de operação de contracção é a definida para as bases de crenças.

Nebel prova [Nebel 1989, teorema 14] que a operação de contracção em bases de crenças corresponde a uma operação de contracção de intersecção parcial,<sup>9</sup> em que a função de selecção escolhe, basicamente, os elementos de  $Cn(\alpha) \perp A$  que contêm os subconjuntos máximos de  $\alpha$ . A base de crenças  $\alpha$  induz assim uma ordem parcial nos elementos de  $Cn(\alpha) \perp A$ .

Para estender a ordem parcial obtida entre os elementos de  $Cn(\alpha) \perp A$  a uma ordem total, Nebel considera ainda a existência de uma ordem entre as fórmulas de  $\alpha$ , que permite assim definir uma relação de ordem entre os elementos de  $Cn(\alpha) \perp A$  que não esta-

<sup>8</sup>No caso das bases de crenças estes conjuntos não são fechados dedutivamente.

<sup>9</sup>Na operação de contracção de intersecção parcial, o resultado é dado pela intersecção dos elementos de  $Cn(\alpha) \perp A$  que forem escolhidos pela função de selecção.

vam ainda relacionados. Com base nesta nova ordem (total), é definida uma nova função de selecção, e a contracção de intersecção parcial definida com esta função satisfaz todos os postulados definidos na AGM para a contracção.

### 2.1.2.3 Problemas com a teoria de Nebel

A abordagem seguida por Nebel vem resolver alguns dos problemas existentes com a AGM (descritos na secção 2.1.1.3), principalmente por considerar que as crenças do agente são representadas em bases de crenças finitas.

No entanto, muitos dos problemas ficaram por resolver e outros foram introduzidos. Todos os problemas apontados relacionados com a utilização da lógica proposicional como mecanismo de representação das crenças do agente mantêm-se, bem como a não representação do conhecimento relativo ao valor epistémico das crenças. Para além disso, a operação de revisão continua a não ser adequada para modelar o raciocínio acerca de mundos em mudança.

Os problemas introduzidos por esta abordagem têm a ver com o facto de se considerar bases finitas de crenças, introduzindo assim uma dependência das operações em relação à forma como as crenças são representadas. Por exemplo, apesar de logicamente equivalentes, a contracção da proposição  $A$  em  $\{A, B\}$  vai ter um resultado diferente da contracção de  $A$  em  $\{A \wedge B\}$ . Neste último caso, o resultado obtido,  $\{\}$ , não é o esperado para uma operação de contracção racional. Por esta razão, uma boa escolha da forma da base de crenças utilizada é essencial para o bom comportamento das operações.

Nebel refere ainda o facto de existirem problemas relacionados com a aplicação sucessiva da contracção a uma base de crenças. Este assunto, no entanto, não foi discutido em [Nebel 1989].

### 2.1.3 A teoria de Fuhrmann

A teoria de revisão de crenças de [Fuhrmann 1991] é a que mais se assemelha à que vai ser utilizada neste trabalho, pois considera que as crenças do agente são representadas por bases finitas de crenças e não obriga à existência de uma ordem total entre as crenças. No entanto, esta teoria continua a definir a contracção de uma base de crenças de forma única, e a admitir a existência de apenas uma ordem parcial entre as crenças, quando na prática pode ser útil admitir a existência de várias ordens parciais, possivelmente contraditórias entre si e de uma preferência entre essas ordens (cf. [Doyle 1991]).

### 2.1.4 Teorias de coerência e teorias de fundamentos

Uma das críticas apontadas por Nebel à teoria AGM é que quando se faz a contracção de uma crença a um conjunto de crenças, as crenças que dependem dela continuam no conjunto, apesar de já não existirem razões para serem acreditadas. O resultado de contrair uma crença de uma base de crenças segundo Nebel, faz com que as suas conclusões deixem de ser acreditadas. No entanto, a abordagem seguida por Nebel não é consensual, e existem defensores para cada uma destas abordagens. A abordagem seguida pelas teorias de revisão de crenças separa-as em *teorias de coerência* e *teorias de fundamentos*.

Os defensores das teorias de coerência, das quais a AGM é um exemplo, defendem que um agente racional deve manter uma crença enquanto ela for consistente com o resto das suas crenças, isto é, as revisões de crenças respeitam o princípio da mudança mínima e mantêm o máximo possível de crenças à medida que outras vão sendo adicionadas ou removidas.

Por outro lado, os defensores das teorias de fundamentos, entre as quais se enquadra a apresentada por Nebel, argumentam que um agente racional apenas deve manter uma crença se tiver alguma razão para acreditar nela. Neste caso, as crenças do agente vão mudando à medida que outras vão sendo adicionadas ou removidas. Normalmente, estas abordagens utilizam um registo de dependências entre as várias crenças do agente, isto é, que crenças é que foram usadas para derivar outras. Este registo de dependências é usado para fazer a actualização no resto das crenças do agente de cada vez que se efectua uma mudança. No caso da teoria de Nebel, não é mantida a dependência entre as crenças, mas existe a distinção entre as crenças da base e as derivadas.

Os defensores de cada uma das abordagens têm motivações diferentes: enquanto uns têm motivações psicológicas e filosóficas (os defensores das teorias de coerência), outros pretendem modelar agentes racionais em sistemas computacionais (os defensores das teorias de fundamentos).

#### 2.1.4.1 Críticas às teorias dos fundamentos

Uma síntese dos problemas apontados às teorias de fundamentos e a resposta a esses problemas são dadas por Doyle em [Doyle 1992]:

**Crítica:** As teorias de fundamentos não são psicologicamente realistas, uma vez que os humanos raramente se lembram das razões para as suas crenças e muitas vezes mantêm crenças para as quais já não têm explicação.

**Resposta:** Segundo Doyle, o que se pretende pode não ser que o sistema funcione como os humanos, mas sim que se comporte duma forma racional, de preferência melhor que os humanos. Para além disso, se quisermos podemos usar uma teoria de coerência mesmo quando mantemos um registo das dependências entre crenças.

**Crítica:** As teorias de fundamentos não são conservativas: Doyle cita [Harman 1986] ao apresentar esta crítica: “*One is justified in continuing fully to accept something in the absence of a special reason not to.*” Isto corresponde a dizer que se deve manter o máximo possível de crenças de cada vez que é feita uma alteração.

**Resposta:** Considerando esta abordagem como correcta, Doyle mostra que se pode simular este comportamento nas teorias de fundamentos, fornecendo para cada fórmula uma justificação do género: “acreditar em qualquer fórmula  $A$  a não ser que  $A$  tenha sido directamente posta em causa.”

Considerando que o termo *conservatismo* significa mudança mínima (em vez de persistência), também as abordagens de coerência apresentam problemas, uma vez que o conceito de mudança mínima não está bem definido. Para além disso, as teorias de fundamentos satisfazem também um critério de mudança mínima, uma vez que o seu processo de revisão minimiza o conjunto das crenças alteradas, mas tendo em conta as razões existentes para as crenças.

**Crítica:** As teorias de fundamentos não são económicas. Esta é uma crítica fundamentalmente apontada aos sistemas que mantêm o registo de dependências entre as crenças. Segundo Gärdenfors, um dos críticos das teorias dos fundamentos, os custos de manter um registo de dependências entre crenças e de o usar são muito superiores aos seus benefícios.

**Resposta:** Doyle aborda esta crítica sob três perspectivas diferentes: os custos lógicos, os custos computacionais e a conveniência prática da sua utilização.

Do ponto de vista dos custos lógicos envolvidos, o peso subjacente à manutenção do registo de dependências não é significativo, quando comparado com as dificuldades de lidar com conjuntos infinitos de crenças, ou, mesmo considerando bases finitas de crenças, com a necessidade de efectuar verificações de consistência.<sup>10</sup> Para além disso, as teorias de coerência, como por exemplo a AGM, utilizam informação sobre o valor epistémico das crenças, e essa informação, expressa através de uma ordem, deve satisfazer certos critérios lógicos cuja complexidade é idêntica à da determinação da consistência.

---

<sup>10</sup>A verificação de consistência, para lógicas razoavelmente expressivas, é indecidível.

Relativamente aos custos computacionais, Doyle conclui que os custos associados à revisão usando as ordens de valores epistémicos da AGM são, pelo menos, tão elevados como os associados com a revisão usando o registo de dependências.

Para além destes dois aspectos, Doyle sugere que a utilização de justificações entre as crenças constitui uma representação mais adequada do que as ordens de valores epistémicos.

**Crítica:** As teorias de fundamentos são supérfluas, dada a noção de valor epistémico de uma crença: o valor epistémico das crenças pode ser usado para determinar que crenças fornecem razões para que outras. Desta forma, a principal vantagem das teorias de fundamentos, manter o registo de dependências para saber que crenças fornecem razões para que outras, desaparece.

**Resposta:** Doyle mostra que nem todos os tipos de dependências podem ser adequadamente descritos usando valores epistémicos. Para além disso, ainda está por provar que os valores epistémicos das crenças não têm uma tradução possível em termos de dependências, o que seria um argumento a favor da crítica.

**Conclusão:** Apesar de não haver uma resposta definitiva para qual das abordagens é melhor, as abordagens fundacionais conseguem incorporar muitos dos aspectos das abordagens de coerência, e ao mesmo tempo são um meio de mecanizar estas abordagens. Vistas de perto, as duas abordagens têm muitas semelhanças, e a melhor solução para a IA parece ser a preocupação com a eficiência computacional de cada uma das abordagens, e possivelmente de abordagens que incorporem aspectos de ambas.

#### 2.1.4.2 As teorias de fundamentos em Inteligência Artificial

Apesar de se reconhecer que as teorias dos fundamentos não são as mais adequadas do ponto de vista psicológico, estas têm a vantagem de permitir construir agentes racionais, que são menos sujeitos a erros do que os construídos usando teorias de coerência. Para além disso, são mais fáceis de mecanizar: por exemplo, com base nas justificações guardadas para cada crença, a manutenção da consistência da base de conhecimento quando é encontrada uma contradição está bastante facilitada, quando comparada com a mesma tarefa sem o registo de dependências (a única forma de o fazer é por tentativa e erro, tirando fórmulas do conjunto de fórmulas até que a contradição deixe de ser derivável).

Por estas razões, as teorias mais utilizadas em IA são as teorias dos fundamentos.

## 2.2 Lógicas Não Monótonas e Raciocínio Não Monótono

Quando se tenta modelar um agente racional para raciocinar acerca do mundo, é necessário dotá-lo de conhecimento. Um dos formalismos mais usados para representar o conhecimento de agentes é a lógica de primeira ordem. No entanto, esta lógica tem várias limitações, nomeadamente no que diz respeito à representação de conhecimento que pode não ser certo, como se pode verificar no seguinte exemplo:

*O Piupiu é uma ave*

*Todas as aves voam*

*O Piupiu não voa*

Neste caso, temos informação que *o Piupiu não voa*, mas também podemos derivar que *o Piupiu voa*; por isso, temos uma contradição. Se decidíssemos abandonar uma destas crenças para deixarmos de ter a contradição, mas quiséssemos manter toda a informação específica acerca do Piupiu, teríamos que abandonar a regra que diz que *todas as aves voam*.

Só que, assim, perdemos a informação que temos acerca da capacidade de voar de todas as outras aves que pudessem existir na base de conhecimento do agente. Tendo isto em consideração, poderíamos propor uma alternativa à abordagem anterior: em vez de abandonarmos a regra que diz que todas as aves voam, podemos substituí-la por outra, onde introduzimos as exceções encontradas.

No exemplo anterior isso poderia ser feito, por exemplo, substituindo a regra

*Todas as aves voam*

por

*Todas as aves que não sejam o Piupiu voam*

Com isto conseguimos resolver a contradição. No entanto, se durante a construção da base de conhecimento aparecerem mais aves que não voam e se verificar que todas elas são pinguins, incluindo o Piupiu, faz mais sentido dizer que as aves voam, a não ser que sejam pinguins, do que de considerar individualmente todas as exceções. Para isso, em vez de

*Todas as aves que não sejam o Piupiu voam*

passaremos a ter

*Todas as aves que não sejam pinguins voam*

*O Piupiu é um pinguim*

Desta forma, continuamos a resolver a contradição que referimos acima, e temos uma forma geral de tratar todos os pinguins. Só que, agora, surgem dois problemas: por um lado, não conseguimos enumerar explicitamente todas as exceções que a regra pode ter (pinguins, avestruzes, aves mortas, etc); por outro, não conseguimos aplicar esta regra a aves que não sabemos se são ou não pinguins (ou avestruzes, ou aves mortas, etc). Por isso, continuamos a perder muitas das conclusões da regra inicial: para todas as aves acerca das quais temos informação incompleta acerca do facto de serem ou não pinguins. Por esta razão, esta alternativa parece muito insuficiente.

Para modelar domínios acerca dos quais temos informação incompleta existem as lógicas não-monótonas. As lógicas não-monótonas têm duas propriedades interessantes: por um lado, permitem a representação de regras gerais, que podem ter exceções, como por exemplo, que *geralmente as aves voam*; por outro, permitem que crenças mantidas anteriormente possam ser abandonadas face a nova informação.

Nesta secção vamos descrever sucintamente a lógica da omissão de Reiter, por ser a primeira abordagem ao raciocínio não-monótono que não se baseia exclusivamente na lógica de primeira ordem. Outros formalismos não-monótonos são, por exemplo, a circunscrição [McCarthy 1980, McCarthy 1977], a lógica de Poole [Poole 1988], a lógica auto-epistémica de Moore [Moore 1983, Moore 1985] e os modelos preferidos de Shoam [Shoam 1987, Shoam 1988].

### 2.2.1 A lógica da omissão de Reiter

A lógica da omissão foi apresentada em [Reiter 1980]. A lógica da omissão é uma lógica não-monótona que permite a representação de tudo o que é possível representar na lógica de primeira ordem e ainda a representação de regras de omissão. As regras de omissão correspondem a sugestões plausíveis de como o agente pode expandir o seu conhecimento, mas cujas conclusões se pode vir a provar que estavam erradas.

Uma *regra de omissão* é da forma

$$\frac{A(x) : B_1(x), \dots, B_n(x)}{C(x)}$$

em que  $A(x), B_1(x), \dots, B_n(x), C(x)$  são fbfs da lógica de primeira ordem, e significa que, “se soubermos  $A(x)$  para um determinado  $x$  e for consistente assumir  $B_1(x), \dots, B_n(x)$  para esse  $x$ , então podemos concluir  $C(x)$ ”.

Na lógica da omissão a informação é representada em *teorias de omissão*. Uma teoria

de omissão é um par  $(\psi, \delta)$ , em que  $\delta$  é um conjunto de fbfs da lógica de primeira ordem e  $\psi$  é um conjunto de regras de omissão.

Os conjuntos de conclusões que podem ser obtidos duma teoria de omissão através da utilização das fbfs da lógica de primeira ordem e das regras de omissão chamam-se extensões. Uma teoria de omissão pode ter zero, uma ou mais extensões, uma vez que podem existir conflitos entre as várias regras de omissão.

A informação do exemplo acerca do Piupiu poderia ser representada pelas seguintes fbfs e regras de omissão:

informação	fbf ou regra
Os pinguins são aves	$\forall(x)[pinguim(x) \rightarrow ave(x)]$
O Piupiu é um pinguim	$pinguim(Piupiu)$
Tipicamente as aves voam	$\frac{ave(x):voa(x)}{voa(x)}$
Tipicamente os pinguins não voam	$\frac{pinguim(x):\neg voa(x)}{\neg voa(x)}$

Tabela 2.1: Fbfs e regras de omissão da lógica da omissão para o exemplo do Piupiu.

Em termos da lógica da omissão, ficaríamos com a teoria  $(\psi, \delta)$ , em que:

$$\psi = \{\forall(x)[pinguim(x) \rightarrow ave(x)], pinguim(Piupiu)\}$$

$$\delta = \left\{ \frac{ave(x) : voa(x)}{voa(x)}, \frac{pinguim(x) : \neg voa(x)}{\neg voa(x)} \right\}$$

Esta teoria tem duas extensões: uma em que o Piupiu é um pinguim, uma ave e voa; e outra em que o Piupiu é um pinguim, uma ave e não voa, dependendo da ordem pela qual forem aplicadas as regras de omissão. Destas duas extensões, será escolhida uma, mas a lógica da omissão de Reiter não especifica como.

Em [Lukaszewicz 1985] e [Etherington 1988] é apresentada uma semântica formal para a lógica da omissão. Para a sua semântica, Etherington prova também a coerência e a completude da lógica da omissão.

### 2.3 Sistemas de revisão de crenças

Se, de cada vez que fosse feita uma mudança nas crenças de um agente racional, todas as suas conclusões tivessem que ser recalculadas, os sistemas que implementassem agentes com capacidade de efectuar raciocínio seriam extremamente ineficientes.

Os sistemas de revisão de crenças preocupam-se com o modo como se podem propagar eficientemente as alterações efectuadas nas crenças de agentes racionais. Basicamente, os sistemas de revisão de crenças mantêm um registo das dependências entre as várias crenças, e com base nesse registo conseguem propagar as alterações de uma forma eficiente.

Para manter o registo das dependências entre as crenças, os sistemas de revisão de crenças usam redes de dependências [Doyle 1983, Reinfrank 1989], onde as crenças do agente são representadas através de nós, as dependências entre as várias crenças são representadas como justificações entre os vários nós e cada nó pode ter zero ou mais justificações. Aos sistemas de revisão de crenças em geral não interessa a crença que cada nó representa, mas apenas as relações que tem com os outros nós da rede. Por esta razão, quando falamos nas estruturas manipuladas pelos sistemas de revisão de crenças falamos geralmente de nós e não de crenças. As crenças em que o agente acredita são representadas pelos nós da rede de dependências que são acreditados.

Existem basicamente duas formas de fazer o registo das dependências entre os nós: ou associando com cada nó os nós de que ele depende directamente, no caso dos sistemas baseados em justificações (JTMSs); ou associando com cada nó os nós que correspondem às crenças básicas de que esse nó depende, no caso dos sistemas baseados em suposições (ATMSs).

Ortogonalmente ao tipo de registo de dependências que os sistemas fazem, existem mais duas perspectivas segundo as quais podem ser classificados:

- Quanto à monotonicidade, podem ser monótonos, se as justificações para os nós forem baseadas apenas no facto de se acreditar noutros nós; ou não-monótonos, no caso de as justificações para os nós poderem também considerar o facto de não se acreditar noutros nós.
- Quanto à capacidade de efectuar inferência, podem ou não ter capacidades dedutivas.

Em todos os sistemas de revisão de crenças, cada nó tem um rótulo associado, e é com base nesse rótulo que se sabe se o nó é ou não acreditado.

Estes sistemas têm também a capacidade de identificar que crenças é que deram origem a outras crenças, nomeadamente a contradições, o que os torna particularmente

adequados para resolver, ou ajudar a resolver contradições.

### 2.3.1 Sistemas baseados em justificações

Nos sistemas baseados em justificações (JTMSs), o rótulo de cada nó pode ser IN ou OUT, correspondendo ao facto de o nó ser ou não ser acreditado, respectivamente. Este rótulo é calculado com base nos rótulos dos nós de que ele depende directamente e no tipo de justificação que existe entre eles. Por esta razão, sempre que existe uma mudança nas crenças do agente, que se reflecte na mudança de algum rótulo de algum nó da rede, essa mudança tem que ser propagada a todos os nós que dependiam dele. Assim, é fácil de compreender porque é que estes sistemas apenas têm capacidade de manter um espaço de crenças para o agente, que corresponde a associar com cada nó da rede o rótulo IN ou OUT. Em compensação, podem tratar com facilidade crenças baseadas na ausência de informação.

O TMS de Doyle [Doyle 1979], que se considera ter iniciado o estudo da revisão de crenças em IA, é o JTMS mais conhecido. Este sistema é não-monótono e não tem capacidade de inferência.

### 2.3.2 Sistemas baseados em suposições

Nos sistemas baseados em suposições (ATMSs), o rótulo de cada nó corresponde ao conjunto de nós que representam as crenças básicas de que esse nó depende. No caso de existir mais do que uma justificação para um nó, vão existir vários conjuntos de nós no seu rótulo. Paralelamente à rede de dependências, existe uma estrutura, a que se dá geralmente o nome de *contexto*, que representa as crenças básicas que são acreditadas. A maneira de ver se um determinado nó é acreditado é verificar se algum dos conjuntos de nós do seu rótulo está contido no contexto, ou seja, se as crenças básicas de que esse nó depende são acreditadas. No caso dos sistemas não-monótonos, não é tão simples verificar se um determinado nó é acreditado e os rótulos são estruturas mais complexas, mas também existem sistemas com estas capacidades.

A principal vantagem dos ATMSs em relação aos JTMSs é o facto de ser muito simples considerar múltiplas alternativas para as crenças do sistema, bastando para isso modificar o contexto. Adicionalmente, quando existem alterações no contexto não é necessário efectuar nenhuma alteração à rede, continuando a ser suficiente consultar o rótulo de um nó (previamente calculado) para saber se ele é acreditado ou não dado o contexto.

O MBR de Martins [Martins 1983, Martins 1991] e o ATMS de de Kleer [de Kleer 1986] são dois sistemas do tipo ATMS monótonos, tendo o MBR capacidade de efectuar inferência.

O NMATMS de [Dressler 1989] estende o ATMS de de Kleer, adicionando-lhe a possibilidade de representar justificações não-monótonas.

O sistema de revisão de crenças usado nesta tese, o SRC de [Cravo 1992, Cravo 1995] é também um ATMS. Este sistema é não-monótono e tem capacidade de efectuar inferência.

## Capítulo 3

# Formalismos usados neste trabalho

Neste capítulo descrevemos a lógica não-monótona, a teoria de revisão de crenças e o sistema de revisão de crenças em que esta tese se baseia.

### 3.1 A lógica SWMC

A SWMC é uma lógica não-monótona que foi desenvolvida para apoiar um sistema de revisão de crenças do tipo ATMS.

Ao longo desta secção vamos fazer um breve resumo dos conceitos mais importantes desta lógica. Uma descrição completa pode ser encontrada em [Cravo 1992], e uma versão mais actualizada em [Cravo 1993b].

#### 3.1.1 A linguagem da SWMC — fórmulas bem formadas

Nesta secção vamos descrever a linguagem da lógica SWMC. Ao conjunto de todas as *fórmulas bem formadas* (fbfs) da SWMC, damos o nome de  $\mathcal{L}$ . Este conjunto de fbfs está particionado em vários sub-conjuntos disjuntos:

- O conjunto das fórmulas da lógica de primeira ordem (LPO), representado por  $\mathcal{L}_{FOL}$ .<sup>1</sup> Assumimos as regras de formação da LPO para as conectivas  $\neg$ ,  $\wedge$ ,  $\vee$  e  $\rightarrow$ , e para os quantificadores  $\forall$  e  $\exists$ . Estas fbfs vão representar as crenças do agente, como por exemplo, a bfb *Pinguim*(*Piupiu*), para representar o facto de que o Piupiu é um pinguim.

---

<sup>1</sup>Do inglês *First Order Logic*.

- O conjunto das regras de omissão, representado por  $\mathcal{L}_D$ .<sup>2</sup> Para estas regras, é preciso um novo quantificador, o quantificador de omissão, representado pelo símbolo “ $\nabla$ ”, e a regra de formação é a seguinte: se  $A(x)^3 \in \mathcal{L}_{FOL}$ , então  $\nabla(x)[A(x)]$  é uma fbf. Fbfs deste tipo são chamadas *regras de omissão*, e servem para representar regras gerais ou com exceções. As regras de omissão são sugestões de como o agente pode estender as suas crenças, mas não são crenças do agente, logo não se pode dizer que uma regra de omissão é acreditada ou não. Por exemplo,  $\nabla(x)[Ave(x) \rightarrow Voa(x)]$  é uma regra de omissão que representa o conhecimento de que normalmente as aves voam.
- O conjunto das suposições,<sup>4</sup> representado por  $\mathcal{L}_A$ .<sup>5</sup> Se  $D \in \mathcal{L}_D$ , e  $c$  é um símbolo individual, então  $Applicable(D, c)$  é uma fbf que representa a suposição de que a regra de omissão  $D$  é aplicável ao símbolo individual  $c$ . Por exemplo, se tivermos a regra  $D = \nabla(x)[Ave(x) \rightarrow Voa(x)]$  e *Piupiu* for um símbolo da linguagem, então  $Applicable(D, Piupiu)$  é uma suposição. As suposições são necessárias para se saber de que fórmulas depende uma conclusão obtida através da aplicação de uma regra de omissão: depende não só da regra de omissão utilizada, mas também da suposição de que a regra é aplicável àquela instância em particular.
- O conjunto das exceções, representado por  $\mathcal{L}_E$ .<sup>6</sup> Se  $D \in \mathcal{L}_D$ ,  $E(x) \in \mathcal{L}_{FOL}$ , então  $\forall(x)[E(x) \rightarrow \neg Applicable(D, x)]$  é uma exceção. Estas regras são usadas para exprimir exceções às regras de omissão, bloqueando a sua aplicação. Seguindo o exemplo anterior, temos que  $\forall(x)[Pinguim(x) \rightarrow \neg Applicable(\nabla(x)[Ave(x) \rightarrow Voa(x)], x)]$  é uma exceção, que significa que os pinguins são uma exceção à regra de que “normalmente as aves voam”.

As regras de formação das fbfs dos conjuntos  $\mathcal{L}_D$ ,  $\mathcal{L}_A$  e  $\mathcal{L}_E$  podem ser estendidas para uma sequência de variáveis  $\bar{x}$ . Apresentamos aqui a versão mais simples, sem perda de generalidade, para facilitar a compreensão dos aspectos fundamentais da lógica SWMC. A versão estendida pode ser encontrada em [Cravo 1993b].

### 3.1.2 As fbfs suportadas — registo de dependências

A SWMC mantém um registo de dependências entre as fórmulas, o que a torna adequada para apoiar um sistema de revisão de crenças do tipo ATMS. Para associar a cada fórmula as fórmulas de que ela depende, a SWMC utiliza *fbfs suportadas*. Uma fbf suportada é da

<sup>2</sup>Do inglês *Default*.

<sup>3</sup> $A(x)$  significa qualquer fbf da LPO cuja variável livre seja  $x$ .

<sup>4</sup>O termo “suposições” tem um significado diferente daquele utilizado na nomenclatura de sistemas de revisão de crenças. A SWMC tem mecanismos de suporte a um ATMS, mas o que corresponde às suposições do ATMS são as fórmulas não derivadas ou hipóteses da SWMC.

<sup>5</sup>Do inglês *Assumption*.

<sup>6</sup>Do inglês *Exception*.

forma  $\langle A, \tau, \alpha \rangle$ , onde  $A \in \mathcal{L}$  é uma fbf,  $\tau \in \{hyp, asp, der\}$  é um rótulo de origem, e  $\alpha \subset \mathcal{L}$  é um conjunto de origem. Consideramos definidas as funções *wff*, *ot* e *os*, cujo domínio é o conjunto das fbfs suportadas:<sup>7</sup>  $wff(\langle A, \tau, \alpha \rangle) = A$ ;  $ot(\langle A, \tau, \alpha \rangle) = \tau$  e  $os(\langle A, \tau, \alpha \rangle) = \alpha$ . A fbf suportada  $\langle A, \tau, \alpha \rangle$  corresponde a uma derivação particular da fbf  $A$ . Nesta fbf suportada:

- $\tau$ , o rótulo de origem, indica como é que a fbf foi gerada: *hyp* identifica hipóteses, *asp* identifica suposições, e *der* identifica fbfs derivadas.<sup>8</sup>
- $\alpha$ , o conjunto de origem, contém as hipóteses e/ou suposições que foram usadas numa derivação particular de  $A$ .

O registo de dependências pode servir para distinguir entre os vários tipos de crenças que o agente pode ter: por um lado, pode ter crenças certas, porque foram derivadas sem usar regras de omissão; por outro lado, pode ter crenças plausíveis, em que acredita “condicionalmente”, isto é, apenas se não tiver informação em contrário. Estas crenças plausíveis são representadas por fórmulas em cuja derivação foram usadas uma ou mais regras de omissão, o que pode ser observado a partir do conjunto de origem das fbfs suportadas que lhes correspondem. Vamos voltar a este assunto na secção 3.1.4, onde explicaremos em que condições é que se podem classificar as fbfs que representam as crenças como sólidas ou não.

### 3.1.3 As regras de inferência

As regras de inferência da SWMC definem como se podem derivar fbfs suportadas a partir de outras fbfs suportadas: caracterizam a fbf derivada, o seu rótulo de origem e o seu conjunto de origem a partir dos elementos correspondentes das fbfs suportadas que são as premissas da regra.<sup>9</sup> Há dois tipos de regras de inferência em SWMC:

- As regras de inferência clássicas, que são semelhantes às de um sistema de dedução natural para a Lógica de Primeira Ordem (LPO), tendo sido modificadas apenas para poderem lidar com fbfs suportadas.
- A regra de inferência estendida, que serve para permitir o raciocínio por omissão. O raciocínio por omissão caracteriza-se por utilizar regras (as regras de omissão) que não são universalmente verdadeiras porque têm excepções. A regra de inferência estendida é:

<sup>7</sup>Do inglês *well-formed formula (wff)*; *origin tag (ot)*; *origin set (os)*.

<sup>8</sup>Do inglês *hypothesis (hyp)*; *assumption (asp)*; *derived (der)*.

<sup>9</sup>Designamos por *premissas* as fbfs suportadas utilizadas na aplicação de uma regra de inferência para deduzir outra fbf suportada.

**Suposição e eliminação do quantificador de omissão (Sup-E $\nabla$ )**

A partir de  $\langle \nabla(x)[A(x)], hyp, \{\nabla(x)[A(x)]\} \rangle$ , podemos inferir, para qualquer símbolo individual  $c$ ,  $\langle Applicable(\nabla(x)[A(x)], c), asp, \alpha \rangle$  e  $\langle A(c), der, \alpha \rangle$ , em que  $\alpha = \{\nabla(x)[A(x)], Applicable(\nabla(x)[A(x)], c)\}$ .

Esta regra de inferência é semelhante à regra de eliminação do quantificador universal mas, uma vez que uma regra de omissão pode ter excepções, precisamos de assumir que o símbolo  $c$  utilizado não é uma excepção à regra, isto é, que a regra é aplicável a  $c$ .

A totalidade das regras de inferência da lógica SWMC, conforme são apresentadas em [Cravo 1993b], pode ser consultada no apêndice C.

**3.1.4 A noção de consequência — contextos e espaços de crenças**

Uma noção fundamental em todas as lógicas é a noção de consequência, definida entre um conjunto de fbfs e uma fbf. Esta noção, dada a informação que um agente dispõe, determina em que é que ele deve acreditar.

A informação de que um agente que utilize a SWMC dispõe é representada por um conjunto de hipóteses<sup>10</sup> designado por *contexto*. Assim, a noção de consequência em SWMC é definida entre um contexto e uma fbf.

O objectivo é, a partir de um contexto, chegar a todos os conjuntos de crenças em que é razoável um agente acreditar, designados por *espaços de crenças*. O conceito de espaço de crenças é fundamental para a definição de consequência. Por isso, primeiro vamos descrever como é que a partir de um contexto se chega aos espaços de crenças, e só depois damos a noção de consequência.

Um determinado contexto pode dar origem a mais do que um espaço de crenças (conjunto de fbfs da LPO), por regras de omissão diferentes “sugerirem” maneiras diferentes e contraditórias entre si de o agente estender as suas crenças; este facto corresponde à existência de *múltiplas extensões*<sup>11</sup> nas lógicas não-monótonas em geral. Em SWMC, para encontrar os espaços de crenças definidos por um contexto, existe um processo construtivo em quatro passos que parte do contexto e vai construindo estruturas intermédias até chegar aos espaços de crenças. No caso de o contexto ser inconsistente, não existe nenhum espaço de crenças definido por ele.

Uma noção auxiliar utilizada na construção dos espaços de crenças é a noção de derivabilidade:

<sup>10</sup>Hipóteses são todas as fbfs introduzidas pela regra de Introdução da Hipótese: fbfs da LPO, regras de omissão ou excepções, ou seja, todas as fbfs da SWMC menos as suposições.

<sup>11</sup>Em SWMC, as extensões designam-se por espaços de crenças.

**Definição 1** Derivabilidade em SWMC

Dado um conjunto de fbfs  $\alpha$  e uma fbf  $A$ , diz-se que  $A$  é derivável a partir de  $\alpha$  ( $\alpha \vdash_{SWMC} A$ ), sse existe uma fbf suportada  $\langle A, \tau, \gamma \rangle$ , com  $\gamma \subseteq \alpha$ .

Esta noção corresponde à noção de consequência da lógica clássica, mas, sendo monótona, não serve para capturar a noção de consequência em SWMC.

De seguida, vamos fazer a descrição intuitiva do processo de construção dos espaços de crenças a partir dos contextos; para uma descrição formal ver [Cravo 1993b].

O processo é constituído por quatro passos:

1. Estender o contexto com todas as suposições que individualmente são consistentes com o contexto inicial, obtendo-se assim o *contexto estendido*.
2. Dividir o contexto estendido em conjuntos máximos que sejam consistentes, os *núcleos primitivos*, que podem servir de base para a construção de um estado de crenças aceitável a partir das hipóteses do contexto. Este passo é necessário porque pode haver suposições ou pressupostos para a aplicação de regras de omissão que sejam contraditórios entre si.
3. Cada um dos núcleos primitivos é testado para ver se satisfaz determinados critérios de racionalidade, e os que não os satisfazem são eliminados. Os restantes são designados por *núcleos*. Os critérios de racionalidade utilizados podem ser encontrados em [Cravo 1993b].
4. Cada um dos núcleos define um *espaço de crenças*, que contém todas as fbfs da lógica de primeira ordem deriváveis em SWMC a partir do núcleo, e que corresponde a um estado de crença aceitável do agente.

Depois de apresentada a noção de espaço de crenças, podemos partir para o que nos propusemos no início desta secção: apresentar a noção de consequência em SWMC.

Em SWMC, dado um contexto  $\beta$ , existem três tipos distintos de consequências de  $\beta$ :

- *Consequência sólida*— dizemos que uma fbf  $A$  é uma consequência sólida de  $\beta$ ,  $\beta \vdash A$ , se na derivação de  $A$  apenas foram usadas fbfs da LPO que estão no contexto  $\beta$ . Neste caso, a fbf  $A$  pertence a todos os espaços de crenças definidos por  $\beta$ . Esta noção de consequência é monótona, isto é,  $A$  é acreditada dado qualquer contexto consistente que contenha  $\beta$ .
- *Consequência plausível*— dizemos que uma fbf  $A$  é uma consequência plausível de  $\beta$ ,  $\beta \vdash_P A$ , se na derivação de  $A$  foram eventualmente usadas uma ou mais suposições, e  $A$  pertence a todos os espaços de crenças definidos pelo contexto  $\beta$ . Intuitivamente, esta situação corresponde a existirem razões para acreditar em  $A$  a partir de  $\beta$

e não existirem razões em contrário. Esta noção de consequência, ao contrário da anterior, é não-monótona, uma vez que, ao adicionarmos outras hipóteses ao contexto,  $A$  pode deixar de ser acreditada em alguns ou mesmo em todos os espaços de crenças.

- *Consequência concebível*— dizemos que uma fbf  $A$  é uma consequência concebível de  $\beta$ ,  $\beta \vdash_C A$ , se na derivação de  $A$  foram eventualmente usadas uma ou mais suposições e  $A$  pertence a pelo menos um espaço de crenças definido pelo contexto  $\beta$ . Intuitivamente, esta situação corresponde a existirem razões para acreditar em  $A$  mas também existirem razões para não acreditar.

A partir destas definições, é fácil ver que uma consequência sólida também é plausível e que por sua vez uma consequência plausível também é concebível.

A distinção entre os vários tipos de consequência é importante, pois embora o agente continue a conseguir estender as suas crenças com consequências que não são sólidas, por vezes é importante saber se está a raciocinar com base em informação que é sólida ou que é apenas plausível ou concebível.

Este aspecto também é importante ao nível do sistema de revisão de crenças (ver secção 3.3.3, onde se descreve o revisor de crenças do SRC, o sistema de revisão de crenças utilizado nesta tese) quando nos encontramos perante uma contradição, pois o facto de termos uma contradição que é uma consequência sólida do contexto é grave, o que não acontece se tivermos uma contradição que é apenas uma conclusão plausível ou concebível desse contexto. No primeiro caso, vamos ter que alterar o contexto se quisermos que ele volte a ser consistente (de acordo com a teoria de revisão de crenças e o sistema de revisão de crenças descritos nas próximas secções), enquanto que o segundo caso é tratado automaticamente durante a construção dos núcleos primitivos definidos pelo contexto e não requer nenhuma acção específica para ser resolvido.

Em [Cravo 1993b] é apresentada uma semântica baseada em conjuntos de modelos para a SWMC, e são provados os resultados de correcção e completude para a lógica com essa semântica.

## 3.2 A teoria de revisão de crenças baseada na SWMC

Na secção anterior apresentámos uma lógica que permite efectuar raciocínio não-monótono, isto é, saltar para conclusões baseadas em conhecimento que não é certo, e que mais tarde podem ter que ser abandonadas. Nesta secção vamos fazer um breve resumo da Teoria de Revisão de Crenças (TRC), apresentada em [Cravo 1992, Cravo 1993a], baseada na lógica SWMC.

Qualquer teoria de revisão de crenças tem por objectivo caracterizar as mudanças a fazer nas crenças de um agente, de modo a incluir ou excluir determinada crença. As mudanças a efectuar nas crenças são guiadas por alguns critérios de racionalidade, um dos quais é o de que as mudanças a efectuar devem ser mínimas.

Como vimos na secção 2.1.1, em [Gärdenfors 1988] são apresentadas três operações sobre as crenças de um agente (vistas como conjuntos de crenças fechados em relação à noção de consequência lógica): expansão, contracção e revisão. A TRC de [Cravo 1993a] define também três operações — adição, remoção e revisão — correspondendo às operações de Gärdenfors, mas que apresentam algumas diferenças pelo facto de considerar bases de crenças finitas (no sentido de [Nebel 1989]) e não conjuntos de crenças fechados em relação à noção de consequência lógica. Para além disso, a lógica subjacente não é a Lógica Proposicional, mas sim uma lógica não-monótona, o que traz algumas alterações na definição das operações que serão descritas nas secções seguintes, e faz com que alguns dos postulados de Gärdenfors deixem de ser satisfeitos. Para uma descrição mais completa e formal, bem como para a demonstração de algumas propriedades destas operações, ver [Cravo 1993a].

As entidades sujeitas às operações de mudança são conjuntos finitos de crenças, que constituem a base para todas as crenças do agente. Na terminologia da SWMC estes conjuntos correspondem a contextos. Para que possamos caracterizar as operações é necessário definir primeiro quais são as consequências de um contexto. Consideramos como consequências de um contexto (para a definição das operações) a reunião de todos os espaços de crenças definidos por esse contexto em SWMC, ou seja todas as consequências concebíveis do contexto.

### 3.2.1 Adição de uma fbf a um contexto

A adição de uma fbf  $A$  a um contexto  $\beta$ , representada por  $(\beta + A)$ , corresponde a acrescentar  $A$  a  $\beta$ , e é trivialmente definida da seguinte forma:

$$(\beta + A) = \beta \cup \{A\}$$

Quando se acrescenta uma fbf a um contexto, é possível que passem a existir novas consequências desse contexto, mas por esta teoria ter subjacente uma lógica não-monótona, também é possível que algumas das consequências que existiam anteriormente deixem de ser acreditadas. No entanto, e como se pode ver pela sua definição, esta operação é sempre definida de forma única.

### 3.2.2 Remoção de uma fbf de um contexto

A remoção de uma fbf  $A$  de um contexto  $\beta$  corresponde a alterar  $\beta$  de modo a que  $A$  deixe de ser uma consequência desse contexto. Para isso, têm que ser invalidadas todas as derivações de  $A$  a partir de  $\beta$ .

As derivações de  $A$  a partir de  $\beta$  são definidas como sendo os conjuntos mínimos de fbfs  $\alpha$  tais que  $\alpha \vdash_{SWMC} A$ , considerando que  $\alpha$  está contido em algum núcleo definido por  $\beta$ .

A operação de remoção de uma fbf  $A$  de um contexto  $\beta$  é representada por  $(\beta - A)$  e pode ser feita de várias formas, quer retirando quer acrescentando fbfs a  $\beta$ , dependendo das derivações existentes para  $A$  a partir de  $\beta$ . Por este motivo, o resultado da operação de remoção de uma fbf de um contexto é um conjunto de contextos.

A forma de invalidar uma derivação  $\alpha$  de uma fbf  $A$  depende do conteúdo de  $\alpha$ :

- Se  $\alpha = \{\}$ , então  $A$  é um teorema, isto é,  $A$  é consequência de qualquer contexto consistente, e não existe nenhuma maneira de fazer com que  $A$  deixe de ser consequência de  $\beta$ . Neste caso, por definição,  $(\beta - A) = \{\beta\}$ .
- Se  $\alpha \cap \mathcal{L}_A = \{\}$ , isto é, se em  $\alpha$  só foram utilizadas hipóteses, então a única forma de invalidar  $\alpha$  é retirar de  $\beta$  uma das hipóteses que tenham sido usadas nesta derivação. Se quisermos que a mudança a fazer em  $\beta$  seja mínima, então devemos retirar exactamente uma dessas hipóteses.
- Se  $\alpha \cap \mathcal{L}_A \neq \{\}$ , o que significa que foram usadas uma ou mais regras de omissão na derivação  $\alpha$ , então  $\alpha$  pode ser invalidada, quer retirando uma das hipóteses em  $\alpha$ , quer acrescentando a  $\beta$  informação que bloqueie a aplicação de alguma regra de omissão que tenha sido usada em  $\alpha$ . Para bloquear a aplicação de uma regra de omissão, a informação a acrescentar consiste em assumir que estamos na presença do que corresponde ou a uma das excepções da regra ou então à negação da conclusão da regra.

No caso de  $A$  não ser um teorema, podem existir várias maneiras possíveis de invalidar todas as derivações de  $A$  em  $\beta$ . De todas essas maneiras, estamos interessados em escolher uma que seja mínima, isto é, que faça o mínimo de alterações possível em  $\beta$ . Para isso, de entre as várias alternativas para invalidar uma derivação de  $A$  que retiram fbfs de  $\beta$ , escolhemos a que retirar menos; se ainda existirem várias alternativas “mínimas” que retiram as mesmas fbfs de  $\beta$ , de entre elas escolhemos a que acrescentar menos.

Mas este critério de mudança mínima não é suficiente para escolher a mudança de forma única, como vimos pelas definições dadas acima. Tal como foi dito em 2.1.1, em [Gärdenfors 1988] é introduzido o conceito de *valor epistémico* de uma crença, de forma a

que se possam ordenar as crenças do agente de acordo com esse valor, que intuitivamente representa a importância dessa crença. Assim, quando existem várias alternativas para a remoção, pode-se escolher as que retiram as crenças de menor valor. Na abordagem de Gärdenfors, a ordem entre as crenças é total. No entanto, como [Doyle 1991] argumenta (ver secção 2.1.4.1), a informação sobre a importância relativa das várias crenças de um agente raramente é tão completa. Normalmente apenas existem várias ordens parciais que podem, inclusive, ser contraditórias entre si. Por este motivo, a TRC considera a possibilidade de existirem zero ou mais ordens parciais de preferência entre as crenças do agente, e ainda uma ordem parcial entre essas ordens, para resolver os casos em que elas se contradizem. Na TRC é criada uma ordem parcial entre as crenças que resulta da combinação destas várias ordens entre crenças e da ordem entre as ordens. Esta ordem é usada para guiar as alterações a efectuar nas crenças do agente.

Apesar da utilização destas ordens de preferência para guiar as alterações a efectuar nas crenças do agente, pode não ser possível determinar essas alterações de forma única. Neste caso dizemos que existem várias alternativas igualmente aceitáveis para remover  $A$  de  $\beta$ .

### 3.2.3 Revisão de um contexto com uma fbf

A operação de revisão de um contexto  $\beta$  com uma fbf  $A$ , representada por  $(\beta * A)$ , corresponde a alterar  $\beta$  de modo a que contenha  $A$  e seja consistente, e é definida à custa da remoção e da adição, usando a identidade de Levi [Levi 1977].

Tal como na operação de remoção, o resultado desta operação é um conjunto de contextos. Para definir a operação de revisão, consideramos  $\beta_{FOL} = \beta \cap \mathcal{L}_{FOL}$ ,  $\beta_D = \beta \cap \mathcal{L}_D$ , e  $\beta_E = \beta \cap \mathcal{L}_E$ :

$$(\beta * A) = \begin{cases} \{\beta' \cup \beta_D \cup \beta_E \cup \{A\} : \beta' \in (\beta_{FOL} - \neg A)\} & \text{se } A \in \mathcal{L}_{FOL} \\ \{\beta \cup \{A\}\} & \text{caso contrário} \end{cases}$$

Repare-se que:

- Se  $A$  for consistente com  $\beta$ , basta adicioná-la ao contexto e não fazer mais alterações.
- Se  $A$  for incompatível com uma consequência sólida de  $\beta$ , então  $\beta \cup \{A\}$  é inconsistente, e é necessário remover pelo menos uma hipótese que não seja  $A$  para o tornar consistente.
- Se  $A$  for incompatível com uma consequência plausível ou concebível (mas não sólida) de  $\beta$ , não é necessário fazer nenhuma alteração a  $\beta$ , porque a lógica subja-

cente (a SWMC) garante que a consequência que era incompatível com  $A$  já não é consequência de  $\beta \cup \{A\}$ .

### 3.3 O sistema de revisão de crenças SRC

Em [Cravo 1992, Cravo 1995] é apresentado um Sistema de Revisão de Crenças (daqui em diante designado por SRC) que se baseia na lógica não-monótona SWMC e na teoria de revisão de crenças TRC, descritas nas secções 3.1 e 3.2, respectivamente. Este sistema é não-monótono, uma vez que se baseia numa lógica não-monótona; é baseado em suposições, devido à forma como regista as dependências entre os vários nós; e tem ainda capacidade de fazer inferência.

Os sistemas de revisão de crenças são, fundamentalmente, sistemas computacionais. Como tal, apesar de o SRC ser apresentado de uma forma abstracta, não se comprometendo com uma implementação particular, a sua definição tem em consideração as limitações inerentes aos sistemas computacionais.

Uma das limitações evidentes destes sistemas é a sua incapacidade de lidar com espaços de crenças, tal como eles são definidos na lógica, uma vez que estes, sendo fechados em relação à operação de consequência lógica, são infinitos.

Por outro lado, a implementação de um sistema de dedução em Lógica de Primeira Ordem enfrenta sempre um problema: a semi-decidibilidade do processo de dedução. As lógicas não-monótonas, geralmente, permitem saltar para conclusões se isso não tornar as crenças do agente inconsistentes, o que significa que necessitam de efectuar testes de consistência numa lógica de primeira ordem, um problema semi-decidível. Este é um dos motivos que torna particularmente difícil a utilização computacional de formalismos não-monótonos.

Estes são alguns dos problemas que o SRC tem de considerar. Para além disso, tem que considerar as tarefas básicas de um sistema de revisão de crenças: manter dependências entre as crenças para que as alterações possam ser propagadas sem muito esforço computacional; e resolver (ou ajudar o utilizador a resolver) as contradições que surjam nas crenças do agente.

Tendo em conta o registo de dependências, o SRC é um sistema do tipo ATMS, associando a cada crença as crenças que estão na base da sua derivação. Repare-se que a SWMC mantém já esse registo nas fbfs suportadas, no conjunto de origem. No SRC utilizam-se *fbfs-justificadas* para fazer esse registo de dependências, para manter informação que permita a propagação de alterações nas crenças quando tal é necessário, e para permitir que o sistema explique as razões pelas quais acredita numa determinada fbf. Uma fbf-justificada é representada por  $\langle A, \tau, \alpha, \gamma \rangle$ , correspondendo à fbf suportada

da SWMC  $\langle A, \tau, \alpha \rangle$  mas em que se regista, em  $\gamma$ , quais foram as crenças imediatamente utilizadas na derivação de  $A$ . Este registo serve também para que o sistema consiga actualizar as crenças com todas as suas possíveis derivações. Para mais detalhes consultar [Cravo 1995].

### 3.3.1 Estruturas do SRC

A informação que representa as crenças do agente é mantida em quatro estruturas:

- *A base de conhecimento* onde estão todas as fbfs justificadas conhecidas do sistema (quer por introdução do exterior, quer por inferência). Repare-se que este conjunto é necessariamente finito.
- *O conjunto de contextos* onde estão todos os contextos considerados até ao momento pelo sistema. Associadas a cada contexto são armazenadas as estruturas que ele define de acordo com a teoria de revisão de crenças: contexto estendido, núcleos primitivos e núcleos. Estão também associadas as hipóteses, as ordens entre as hipóteses e as suposições, e a ordem entre as ordens. A interacção com o SRC é sempre associada com um contexto.
- *O conjunto de conjuntos inconsistentes* onde estão os conjuntos mínimos de fórmulas já detectados pelo sistema como sendo inconsistentes. De notar que podem existir outros conjuntos de fórmulas inconsistentes, sem o sistema ter “conhecimento” disso.
- *O conjunto de suposições* onde estão todas as suposições levantadas pelo sistema até ao momento.

Estas estruturas são actualizadas pelos dois componentes fundamentais do SRC: o *motor de inferência* e o *revisor de crenças*. É através da consulta e actualização destas estruturas que o motor de inferência e o revisor de crenças comunicam e cooperam na modelação das crenças de um agente.

### 3.3.2 O motor de inferência

O motor de inferência serve para adicionar nova informação à base de conhecimento, quer vinda do exterior, quer resultante da aplicação de regras de inferência para gerar novas fbfs.

As regras de inferência utilizadas pelo SRC são as da lógica SWMC, modificadas para trabalhar com as fbfs-justificadas, e podem ser utilizadas pelo SRC, quer para efectuar inferência para a frente, quer para efectuar inferência para trás.

São também tarefas do motor de inferência:

- A determinação das consequências de um contexto— calcular os contextos estendidos, os núcleos primitivos e os núcleos definidos por um contexto. Para isso baseia-se nas definições destes conceitos na SWMC.
- A explicação— dizer por que razões acredita numa determinada fbf. Utiliza as justificações das fbfs justificadas para saber quais os passos seguidos na sua derivação e dar uma explicação que depende das regras de inferência utilizadas.
- A propagação de dependências— para além da propagação de dependências que é feita automaticamente através das regras de inferência, e que corresponde a dizer de que fbfs depende uma fbf, o motor de inferência também se serve das justificações para actualizar os conjuntos de origem de todas as fbfs que dependem de uma fbf para a qual foi encontrada uma nova derivação.
- A actualização das estruturas do SRC— é através da actualização destas estruturas que se representam as alterações feitas na informação existente no sistema. A sua actualização é importante, não só para posteriores derivações do motor de inferência, mas também para o revisor de crenças. Por exemplo, o motor de inferência actualiza o conjunto de conjuntos inconsistentes sempre que é detectada uma contradição.

### 3.3.3 O revisor de crenças

O revisor de crenças é o componente responsável pela tarefa de revisão de crenças: sempre que existe uma contradição deve detectá-la e determinar quais os possíveis “culpados”; é também o revisor de crenças o responsável pela remoção de fbfs de contextos. Estas tarefas são efectuadas com base na TRC descrita na secção 3.2. Para além disso, o revisor também é o responsável pela escolha dos espaços de crenças preferidos pelo sistema.

Tal como já foi dito na secção 3.3.1, todas as interacções com o sistema estão associadas com um contexto da lógica SWMC — o *contexto corrente*. É através deste contexto que o agente sabe em que fbfs é que acredita. Por uma questão de simplicidade, e uma vez que quando se acrescenta uma fbf ao contexto corrente ela vai automaticamente passar a pertencer à base de conhecimento do agente, falamos em “acrescentar uma fbf à base de conhecimento do agente” em vez de “acrescentar uma fbf ao contexto corrente, cujo espaço de crenças define as fbfs acreditadas da base de conhecimento do agente”. Esta simplificação não se pode fazer quando se retiram fbfs do contexto, porque as fbfs nunca podem ser retiradas da base de conhecimento.

A determinação dos “culpados” de uma contradição pode ser feita de uma forma

relativamente simples, uma vez que a SWMC, na qual o SRC se baseia, associa a cada fbf as fórmulas utilizadas numa derivação (no conjunto de origem das fbfs suportadas): quando é detectada uma contradição, os “culpados” encontram-se entre as fbfs do conjunto de origem da fbf suportada que corresponde à contradição. Em SWMC existem dois tipos de contradição: *contradições reais* e *contradições aparentes*. As primeiras ocorrem entre conclusões sólidas, derivadas apenas a partir de hipóteses, e são tratadas pelo revisor de crenças. As contradições aparentes são aquelas em cuja derivação foram usadas suposições, e não constituem na realidade contradições, não dando lugar a uma revisão de crenças. Em qualquer dos casos, o conjunto de conjuntos inconsistentes é actualizado pelo motor de inferência. No caso de estarmos perante uma contradição real, vai ter que ser removida do contexto pelo menos uma fbf do seu conjunto origem.

A remoção de fbfs de contextos pode ser efectuada por duas razões distintas: ou porque foi feito, a partir do exterior, um pedido ao SRC para remover uma fbf; ou porque foi detectada uma contradição e é necessário abandonar uma fbf para a resolver.

Quando é feito, a partir do exterior, um pedido ao SRC para remover uma fbf, o revisor de crenças executa as seguintes tarefas:

1. Calcula as derivações válidas para a fbf, o que corresponde, aproximadamente, a determinar os conjuntos de fbfs do conjunto suporte da fbf a remover que estão contidos num dos núcleos definidos pelo contexto;
2. Escolhe, de entre as alternativas possíveis para remover a fbf do contexto, a (ou as) que corresponde(m) a uma mudança mínima;
3. Retira e acrescenta ao contexto corrente as fbfs que forem indicadas pela alternativa escolhida no ponto anterior.

Quando é detectada uma contradição real (o único tipo de contradição que é tratado ao nível do revisor de crenças), o revisor de crenças não só identifica os “culpados” da contradição (como faz a generalidade dos sistemas de revisão de crenças), mas utiliza a TRC para resolver a contradição, ou pelo menos auxiliar o utilizador do sistema nessa tarefa. Neste caso, vai ser necessário escolher pelo menos uma das fbfs que esteja entre os “culpados” da contradição para ser abandonada. Para essa escolha, o SRC utiliza as ordens de preferência que o utilizador tenha eventualmente fornecido. Para além dessas ordens, o sistema determina uma ordem de forma automática — a *ordem de especificidade* — segundo a qual são preferidas as regras de omissão mais específicas às regras de omissão menos específicas. A forma de determinar se uma regra de omissão é mais específica do que outra é a seguinte: dizemos que a regra de omissão  $\nabla(x)[A_1(x) \rightarrow B_1(x)]$  é mais específica que a regra de omissão  $\nabla(x)[A_2(x) \rightarrow B_2(x)]$ , ambas pertencentes a um contexto  $\beta$ , se e só se

$$\beta \vdash \forall(x)[A_1(x) \rightarrow A_2(x)] \text{ e } \beta \not\vdash \forall(x)[A_2(x) \rightarrow A_1(x)].$$

Esta ordem parcial, apesar de definida automaticamente pelo sistema, comporta-se como uma qualquer ordem introduzida pelo utilizador, na medida em que lhe pode ser dada uma importância relativamente às restantes, através da ordem entre as ordens introduzida pelo utilizador. A ordem entre as ordens é única, e serve para desambiguar os casos em que existe conflito entre as várias ordens de preferência parciais. A combinação de todas as ordens parciais (incluindo a ordem de especificidade), usando a ordem entre as ordens sempre que necessário, dá origem à *ordem de preferência combinada*.

Esta ordem é utilizada pelo revisor de crenças no processo de resolução de uma contradição real, bem como para determinar quais os espaços de crenças preferidos de entre os que são definidos por um determinado contexto.

Uma vez que cada espaço de crenças é determinado univocamente por um dos núcleos definidos pelo contexto, preferir um espaço de crenças vai corresponder, na realidade, a preferir um dos núcleos. A escolha do núcleo preferido vai ser feita apenas com base nas preferências existentes entre as suposições, pois todas as hipóteses do contexto estão em todos os núcleos definidos por ele, e por isso as preferências entre as hipóteses do contexto não vão servir para se saber qual é preferido. No entanto, como o sistema apenas admite a especificação de preferências entre hipóteses, para determinar as preferências entre as suposições vão ser usadas as preferências entre as regras de omissão que lhes correspondem.<sup>12</sup> Em [Cachopo & Cardoso 1994] foi feita uma extensão a este sistema, de modo a que ele passasse também a permitir a especificação de preferências directamente entre suposições.<sup>13</sup> Para além disso, nem todas as suposições dos núcleos vão servir para escolher o preferido. As únicas suposições importantes para escolher, de entre dois núcleos, qual o núcleo preferido são as suposições que estão em conflito nesses dois núcleos, isto é, as suposições que estão num dos núcleos e não estão no outro porque se estivessem o tornariam inconsistente. Assim, na escolha de qual é o núcleo preferido de entre dois núcleos, apenas são consideradas as preferências entre as suposições que estão em conflito dos dois núcleos.

---

<sup>12</sup>Como seria de esperar, uma suposição é preferida a outra se a regra de omissão que lhe corresponde for preferida em relação à da outra.

<sup>13</sup>Por exemplo, para exprimir o facto que se prefere a aplicação de uma regra de omissão a uma instância relativamente à aplicação da mesma regra de omissão a outra instância.

## Capítulo 4

# Revisão Adaptativa de Crenças

Neste capítulo, explicamos o que significa a revisão adaptativa de crenças, e motivamos a necessidade desta nova operação de revisão de crenças, em que se tenta manter o máximo de informação possível quando é feita uma revisão, em vez de simplesmente abandonar uma das proposições que pertencia ao conjunto em conflito.

### 4.1 Motivação

Um dos componentes principais de um agente inteligente é a sua base de conhecimento, na qual vai registando as suas crenças, à medida que vai percepcionando o mundo exterior e raciocinando acerca das suas crenças. Durante a construção da base de conhecimento, o agente raramente (ou nunca) tem todo o conhecimento relevante para o domínio, ou consegue antecipar todos os casos em que pode ter necessidade de utilizar a informação contida na sua base de conhecimento.

Assim, o agente vai construindo modelos aproximados do mundo, os quais correspondem a simplificações da realidade, e serão melhorados à medida que o agente vai recebendo mais informação do mundo exterior.

Para exemplificar a necessidade da existência da revisão adaptativa de crenças vamos voltar ao exemplo do Piupiu. Podemos supor que temos na base de conhecimento uma regra que representa o facto de que *todas as aves voam* e que temos também informação sobre a existência de várias aves, para as quais foi deduzido que voam. Neste caso, quando surgisse informação acerca de aves que não voam, como por exemplo que *o Piupiu é uma ave* e que *o Piupiu não voa*, passaríamos a ter uma base de conhecimento inconsistente. A causa deste problema está precisamente no facto de que o conhecimento do agente correspondia apenas a uma simplificação do mundo, que agora se veio a provar que estava errada.

Na realidade, existe no mundo muito pouco conhecimento que possa ser representado por regras universais.<sup>1</sup> Este facto foi devidamente reconhecido no livro “*Readings in Nonmonotonic Reasoning*” [Ginsberg 1987], onde logo na primeira página é citada uma frase de Benjamin Franklin, que data de 1798: “Nothing is certain, but death and taxes”. Se por um lado, a utilização de regras universais simplifica o raciocínio, na medida em que não é necessário efectuar testes de consistência para verificar se se pode deduzir o consequente de uma regra universal; por outro, causa problemas sempre que é detectada uma contradição na base de conhecimento, pois pode provocar o abandono de informação que poderia ser mantida na base de conhecimento.

A razão pela qual se pode abandonar mais informação do que a que é necessário é porque, quando é detectada uma contradição, as teorias de revisão de crenças sugerem o abandono de uma (ou mais) das crenças que esteja na base da contradição. No exemplo anterior, isso iria corresponder a abandonar uma das seguintes crenças: *todas as aves voam*, *o Piupiu é uma ave* ou *o Piupiu não voa*. Intuitivamente, parece-nos que uma regra universal (*todas as aves voam*) tem maior probabilidade de estar errada do que informação específica acerca do Piupiu (*o Piupiu é uma ave* ou *o Piupiu não voa*). Por esta razão, de acordo com as teorias de revisão de crenças, deve ser a regra universal a escolhida para ser abandonada. No entanto, com esta escolha abandona-se também a informação acerca da capacidade de voar das outras aves, o que claramente não é desejável. O abandono de todas as conclusões da regra universal deve-se ao facto de estarmos a usar uma teoria dos fundamentos<sup>2</sup> para fazer a revisão das crenças do agente.

O facto de se abandonar mais informação do que o necessário não foi devido a uma escolha errada de qual a regra a abandonar, mas sim pelo facto de as teorias de revisão de crenças não serem suficientemente flexíveis para tratar os casos em que se encontra uma excepção para uma regra universal, mas em que geralmente as consequências da regra até estão de acordo com o resto do conhecimento do agente. Neste caso, seria mais adequada a utilização de uma teoria da coerência; no entanto, estas teorias têm o problema de permitir que o agente tenha crenças para as quais não tem razões para acreditar.

Uma melhor solução para o problema anterior, em que tentamos aproveitar o máximo de conhecimento possível, seria, em vez de abandonar a regra universal, *enfraquece-la*, transformando-a numa regra de omissão, e passar a ter a informação de que *geralmente as aves voam*. Desta forma, passamos a ter uma regra de omissão com excepções, e as possíveis contradições que possam surgir com informação específica acerca de alguma ave podem ser automaticamente resolvidas pelos mecanismos da lógica não-monótona

---

<sup>1</sup>Para além das regras universais que são usadas para representar alguns domínios formais (como por exemplo a matemática), as poucas regras universais que representam conhecimento acerca do mundo correspondem geralmente a definições (como por exemplo que todas as mães são do sexo feminino) e/ou a representações de hierarquias (como por exemplo, na hierarquia das aves, que todos os pinguins são aves).

<sup>2</sup>Na secção 2.1.4 apresentamos uma discussão das diferenças entre as teorias dos fundamentos e as teorias da coerência.

utilizada. Todas as conclusões da regra universal anterior (*todas as aves voam*) que não entram em contradição com outras fórmulas da base de conhecimento continuam a poder ser derivadas, mas agora a partir da regra de omissão. Basicamente, estamos a criar uma nova operação para a teoria de revisão de crenças, que vai permitir, em certos casos, manter toda a informação que seria mantida por uma teoria da coerência, mas que continua a ser uma teoria dos fundamentos, pois a regra de omissão passa a ser a razão para acreditar nas conclusões da regra universal que não são abandonadas.

O enfraquecimento de regras universais quando é detectada uma contradição vai corresponder a mais uma operação da teoria de revisão de crenças utilizada. Esta operação apenas faz sentido se for aplicada a uma regra universal, pois corresponde ao facto de ter sido detectada uma contradição, identificada como uma excepção para a regra universal. Para a base de conhecimento do agente voltar a ser consistente, sem ser necessário abandonar conhecimento que ainda é consistente com o resto da base de conhecimento do agente, a regra universal tem que ser enfraquecida, isto é, tem que ser transformada numa regra de omissão que represente o mesmo conhecimento, mas que permita a existência de excepções.

O modo mais simples de efectuar este enfraquecimento é retirar a regra universal da base de conhecimento, “deitando fora” todas as suas consequências, e depois introduzir a regra de omissão, voltando a inferir a maior parte das consequências, mas agora dependendo da regra de omissão.

Embora ao nível das teorias de revisão de crenças não existam preocupações com a eficiência das operações, quando se pretende criar um sistema computacional que tenha a capacidade de efectuar revisão de crenças, é necessário utilizar uma outra estratégia para fazer revisão de crenças do agente. Vamos, então utilizar um sistema de revisão de crenças, que se preocupa em propagar as alterações nas crenças do agente numa forma eficiente, usando para isso o registo de dependências entre as várias crenças. Ao nível do sistema de revisão de crenças, talvez os resultados do enfraquecimento pudessem ser propagados com um novo algoritmo de revisão de crenças, que simplesmente não fizesse revisão enquanto não fosse necessário, isto é, enquanto não surgisse um novo conflito; só quando existisse outra contradição é que se iria verificar se as crenças em conflito tinham passado a depender de uma regra mais fraca, que poderia ter excepções. Outra alternativa é, em vez de abandonar as crenças que tivessem sido derivadas a partir da regra universal, marcá-las como passando a ser consequências de uma regra de omissão, o que faria com que algumas das contradições em que estivessem envolvidas pudessem ser resolvidas automaticamente pelos mecanismos da lógica não-monótona utilizada.

## 4.2 Critérios para a selecção de uma regra para enfraquecer

Um problema que é necessário considerar ao nível da teoria de revisão de crenças é o facto de que as consequências das regras universais enfraquecidas (que passam a ser consequências de uma regra de omissão) podem agora entrar em conflito com as conclusões de outras regras de omissão, e só com a informação que existe na base de conhecimento do agente estes conflitos não conseguem ser resolvidos, pois o agente não tem maneira de saber que informação preferir.

Adicionalmente, pode acontecer que existam várias regras universais em conflito. Nestes casos, é necessário escolher pelo menos uma delas para ser enfraquecida. No entanto, mais uma vez a base de conhecimento não contém informação que permita escolher a regra a ser enfraquecida.

Nas próximas secções introduzimos a utilização de preferências entre as crenças do agente, que em alguns casos pode ajudar a escolher o que manter na base de conhecimento e o que enfraquecer.

### 4.2.1 Utilização de preferências

Em teorias científicas, por exemplo, é comum a existência de um conjunto de crenças que constitui o núcleo da teoria e um outro conjunto de crenças auxiliares que providenciam explicações para partes da teoria, não sendo centrais. Este aspecto é descrito por Gärdenfors, ao explicar as origens do valor epistémico das crenças:

*“A common ingredient of many modern metatheories of scientific research is that certain statements of a scientific theory are never called into question. [Kuhn 1970] speaks of ‘symbolic generalizations’ that form a part of the paradigm of a science. [Lakatos 1970] introduces the ‘core’ of a research program that is surrounded by a ‘protective belt’ of auxiliary hypothesis. If the research program encounters problems, that is, when it needs to be revised, the sentences in the core are never rejected but perhaps some parts of the protective belt are changed. [Sneed 1971] and [Stegmüller 1976] also speak of the core and the extended core of a theory. [ . . . ] All these meta-scientific theories thus present a rudimentary hierarchy among the sentences of a scientific theory. This hierarchy can be interpreted as a rough sorting of these sentences into degrees of epistemic entrenchment, the most entrenched being the sentences that are included in the paradigm or the core of the theory.” [Gärdenfors 1988, pp. 92–93]*

O abandono de uma crença do núcleo da teoria obriga, normalmente, ao abandono de toda a teoria,<sup>3</sup> enquanto que o abandono de uma crença auxiliar pode causar apenas um pequeno ajustamento. É natural, nesta situação, considerar que as crenças do núcleo da teoria têm um valor epistémico superior ao das outras crenças. No entanto, se

---

<sup>3</sup>Aquilo que se costuma designar por uma “revolução científica”.

se verificar de facto a necessidade de abandonar uma das crenças do núcleo, obrigando eventualmente ao abandono de toda a teoria, o valor epistémico dessas crenças será concerteza alterado.

O domínio das teorias científicas é um exemplo onde é fácil identificar a existência de dois (ou mais) conjuntos de crenças que têm valores epistémicos claramente diferentes. Por exemplo, as crenças do núcleo deverão ter um valor epistémico mais elevado do que as crenças auxiliares. O que se torna difícil, neste caso, é a determinação do valor epistémico das crenças dentro de cada um desses grupos, a um nível de detalhe mais refinado. O mesmo acontece noutros domínios, não existindo nenhuma teoria que explique convenientemente a natureza do valor epistémico das várias crenças.

Apesar destas dificuldades, concordamos com o facto de o valor epistémico das crenças ser essencial para a definição das operações que efectuam alterações nas crenças do agente. Do nosso ponto de vista, o valor epistémico das crenças constitui conhecimento adicional que o agente deve manter a par das outras crenças e que só não está representado no conjunto de crenças do agente por inadequação da representação.

Se considerarmos, por exemplo, que as regras de omissão que resultam do enfraquecimento de uma regra universal devem ter menos excepções do que as outras regras de omissão, deveríamos poder exprimir este facto de alguma forma. Podemos justificar melhor esta afirmação com um exemplo: a regra que diz que *geralmente as aves voam* deve ter menos excepções do que a regra que diz que *geralmente as aves são amarelas*, uma vez que existem menos aves que não voam do que aves que não são amarelas, e as aves que não voam também não são amarelas.

Assim, de cada vez que transformarmos uma regra universal numa regra de omissão, podemos criar uma *preferência*, que diz que as conclusões desta regra são preferidas em relação às conclusões de outras regras de omissão que não tenham resultado do enfraquecimento de uma regra universal. Assim, os conflitos com conclusões de outras regras de omissão que não resultem do enfraquecimento de uma regra universal podem ser resolvidos automaticamente numa teoria de revisão de crenças que permita a especificação de preferências entre as regras.

As preferências entre regras de omissão podem ser expressas na base de conhecimento do agente ou fora dela, o importante é que sejam usadas quando existe um conflito, para escolher as consequências das regras de omissão que são mais importantes para o agente.

No entanto, podem existir várias regras universais em conflito, que eventualmente podem ser transformadas em regras de omissão, como se ilustra no seguinte exemplo:

*As aves voam*

*Os pinguins são aves*

*Os pinguins não voam*

*Os pinguins australianos são pinguins*

*Os pinguins australianos voam*

*O Piupiu é um pinguim*

*O Tweety é um pinguim australiano*

No caso de as consequências das regras universais estarem em conflito, as consequências das regras de omissão correspondentes também estão em conflito.

Com os dados deste exemplo, podemos derivar que o Piupiu e o Tweety voam e não voam. Depois de transformar a regra que diz que as aves voam numa regra de omissão, por o Piupiu ser um pinguim ficámos a saber que ele não voa, uma vez que a regra que diz que os pinguins não voam é uma regra universal. No entanto, continuamos a poder derivar que o Tweety voa e não voa. Para resolver o problema do Tweety, podemos transformar a regra que diz que os pinguins não voam em regra de omissão. Neste caso, passamos a ter que o Tweety voa, por ser um pinguim australiano. Mas agora, as consequências das regras de omissão resultantes do enfraquecimento das regras universais são ambas igualmente fracas (ou fortes). Neste caso, as possíveis contradições não podem ser resolvidas pela teoria, mesmo utilizando a preferência que diz que as regras de omissão que resultaram do enfraquecimento duma regra universal são preferidas em relação às outras.

Para além de nos basearmos no facto de uma regra de omissão resultar do enfraquecimento de uma regra universal para preferirmos as suas consequências relativamente às consequências de outras regras de omissão, podemos usar outros critérios. Um dos que nos parece mais razoável é dizer que preferimos as consequências de regras de omissão mais específicas relativamente às consequências de regras de omissão menos específicas, em que uma regra é mais específica que outra se estiver mais abaixo na hierarquia. Este critério foi sugerido em [Cravo 1992, pp. 205], e no caso do exemplo anterior seria suficiente para escolher as consequências a preferir: o Tweety voa, por ser um pinguim australiano e o Piupiu não voa por ser um pinguim, o que corresponde à nossa intuição.

Uma outra situação em que a utilização de preferências pode ser vantajosa é quando o agente tem várias fontes de informação, possivelmente contraditórias entre si, mas nas quais o agente pode ter diferentes graus de confiança. O agente pode, através da especificação de preferências, dizer que prefere as crenças provenientes de uma dada fonte às provenientes de outra fonte.

Quando, mesmo usando preferências entre as crenças não é possível resolver a contradição ao nível da teoria, a escolha de qual a regra a preferir em caso de conflito tem que ser feita exteriormente ao agente.

### 4.2.2 Preferências como número de exceções

A necessidade de intervenção exterior em caso de conflito faz com que o agente não possa ser autónomo, o que não é desejável. Assim, gostaríamos de conseguir automatizar o processo de determinação de preferências. Uma possibilidade seria considerar a acumulação de exceções às regras de omissão, e utilizar este facto para fazer com que as regras com menos exceções sejam preferidas em relação às outras.

Mas, antes de podermos considerar esta abordagem, é necessário definir como deve ser contabilizado o número de exceções. Algumas alternativas são:

1. O número de constantes da base de conhecimento que são excepção;
2. A percentagem de constantes da base de conhecimento que são excepção, em relação às constantes a que a regra pode ser aplicada;
3. O número de predicados da linguagem que são simultaneamente satisfeitos pelas constantes que são excepção.

Podemos voltar ao exemplo do Piupiu, para ver em que é que se traduz cada uma destas alternativas: *Piupiu* é uma constante da base de conhecimento, enquanto que *Ave* é um predicado. Se soubermos que o Piupiu é uma ave, isto é, que a constante *Piupiu* satisfaz o predicado *Ave*, gostaríamos de poder utilizar o conhecimento que temos acerca das aves para raciocinar acerca do Piupiu. Mesmo sabendo que o Piupiu é uma excepção no que diz respeito à sua capacidade de voar, podemos utilizar o resto do conhecimento acerca das aves, por exemplo, para deduzir que o Piupiu tem duas patas e o corpo coberto de penas. Este tipo de raciocínio é muito utilizado em sistemas de herança, em que existem classes (por exemplo, a classe das aves) e instâncias<sup>4</sup> das classes (por exemplo, o Piupiu). Na área da Representação do Conhecimento existem vários sistemas de herança, baseados em formalismos específicos, como as redes semânticas [Quillian 1968, Fahlman 1979, Pinto 1992], os sistemas de enquadramentos [Minsky 1975, Brachman & Schmolze 1985], ou que tentam formalizar a herança usando uma lógica [Etherington & Reiter 1983, Touretzky 1984]. Embora não esteja no âmbito desta tese discutir problemas relacionados com os tipos de raciocínio associados com a herança, os conceitos de classe e instância voltarão a ser utilizados em alguns dos exemplos apresentados. Para já, podemos reformular a terceira alternativa para contabilização de exceções de uma forma mais simples: as exceções correspondem ao número de classes que são excepção. Voltando ao exemplo, no caso da regra que diz que as aves voam, podemos agrupar as suas exceções em classes, como os pinguins, as avestruzes, etc.

Provavelmente não existe uma alternativa que seja a melhor para a contabilização de exceções, pois isso vai depender do domínio que estivermos a considerar, e também

---

<sup>4</sup>Por vezes, às instâncias também se dá o nome de objectos.

do conhecimento que estiver expresso na regra: se for uma regra acerca de classes, faz sentido usar a terceira alternativa; se for uma regra acerca de instâncias, seria melhor uma das duas primeiras. O importante é que, depois de escolhida uma alternativa para a contabilização de excepções (possivelmente uma para cada regra de omissão), essa contabilização pode ser feita e actualizada automaticamente, à medida que o agente vai adquirindo mais conhecimento e efectuando mais raciocínio, sendo a informação acerca de excepções também representada na sua base de conhecimento.

Com a contabilização do número de excepções na base de conhecimento do agente, podemos representar explicitamente as preferências que o agente tem entre as suas crenças. Assim, passamos a ter a possibilidade de fazer revisão de crenças, não só ao nível do conhecimento do agente, mas também ao nível das suas preferências. No entanto, uma questão que se pode colocar é: será que as crenças e as preferências entre as crenças do agente devem ser representadas ao mesmo nível? Enquanto que as crenças reflectem o modelo que o agente tem do mundo, as preferências entre as crenças reflectem o conhecimento que o agente tem acerca do seu próprio conhecimento, chamado *meta-conhecimento*. Podemos reformular a questão inicial: onde contabilizar o número de excepções, ao nível da base de conhecimento do agente ou numa base de conhecimento à parte, onde está apenas informação acerca de preferências? Se por um lado as excepções estão intimamente ligadas com as regras de omissão que representam o conhecimento do agente, por outro serão usadas para calcular preferências entre esse conhecimento. Tendo em consideração o facto de as excepções serem também informação acerca do mundo e não acerca do conhecimento do agente, a nossa intuição é para as colocar na base de conhecimento, juntamente com a outra informação de que o agente dispõe.

Como meta-conhecimento deveria existir uma regra que dissesse que se preferem regras com menos excepções a regras com mais excepções (independentemente da forma como as excepções são calculadas), tal como uma regra que dissesse que se preferem regras de omissão mais específicas em relação às menos específicas.<sup>5</sup> Neste caso, também é necessário definir como deve ser a interacção entre as várias “meta-regras”.

Se considerarmos que o agente utiliza uma lógica para representar o seu conhecimento, temos ainda um outro problema para resolver: como tratar inteiros e como representar em lógica o facto de que uma regra tem mais excepções do que outra. Temos várias soluções possíveis para este problema:

1. Utilizar procedimentos, a um nível extra-lógico (por exemplo, na teoria de revisão de crenças), que contabilizam as excepções de cada regra e que fazem as comparações necessárias. Esta abordagem é relativamente simples, mas é também pouco formal, e os procedimentos utilizados não terão uma semântica bem definida, pelo que se perde uma grande vantagem da utilização da lógica para representar conhe-

---

<sup>5</sup>As regras de omissão mais específicas são as que estão mais perto das instâncias, ou nas classes inferiores da hierarquia.

cimento. Para uma discussão acerca das vantagens e desvantagens da utilização de procedimentos relativamente à utilização da lógica na representação de conhecimento, ver [Winograd 1975].

2. Utilizar unicamente a lógica, representar os inteiros usando constantes da lógica e explicitar um predicado “maior” para todos os pares de inteiros que possam ser relevantes para este domínio. Um dos problemas desta abordagem é determinar “os pares de inteiros relevantes para o domínio”; para além disso, também não é claro como é que poderia ser representado o conhecimento de que se preferem as regras com menos exceções em relação às regras com mais exceções.
3. Representar o domínio dos inteiros na base de conhecimento do agente usando a lógica. Esta abordagem teria uma semântica bem definida e ao mesmo tempo daria ao agente a possibilidade de raciocinar acerca de inteiros (por exemplo, para fazer cálculo numérico). No entanto, de acordo com o teorema da incompletude de Gödel [Gödel 1965], ao estender a linguagem para permitir a representação do princípio da indução, passarão a existir verdades lógicas que não poderão ser provadas, o que não é desejável.

Avaliámos as vantagens e desvantagens de cada uma destas soluções para este trabalho, e considerámos que as vantagens das duas últimas não superavam a simplicidade da primeira. Por isso, o tratamento das exceções será feito ao nível da teoria de revisão de crenças, em que as regras de omissão com menos exceções têm um valor epistémico mais elevado do que as regras de omissão com mais exceções. Assim, em caso de conflito, as consequências das primeiras são preferidas em relação às consequências das segundas.

## Capítulo 5

# Alterações aos formalismos usados neste trabalho

Neste capítulo descrevemos as alterações que são necessárias efectuar aos formalismos usados neste trabalho para podermos construir um sistema de revisão de crenças com a capacidade de efectuar a revisão adaptativa das suas crenças. Este sistema vai corresponder a uma extensão do sistema descrito na secção 3.3.

### 5.1 Alterações à teoria de revisão de crenças

Para um agente poder efectuar a revisão adaptativa das suas crenças, é necessário que ele tenha a capacidade de enfraquecer uma regra universal da sua base de conhecimento e de a transformar numa regra de omissão. Esta transformação corresponde simplesmente a passar de uma regra que não pode ter excepções (a regra universal) para uma regra que representa o mesmo conhecimento, mas que pode ter excepções (a regra de omissão).

Ao nível da base de conhecimento,<sup>1</sup> o agente deixa de acreditar na regra universal e passa a ter a regra de omissão como um meio de estender as suas crenças. Em termos dos contextos da lógica SWMC, a regra universal deixa de pertencer ao contexto e a regra de omissão passa a pertencer ao contexto.

Para efectuar a revisão adaptativa, a teoria de revisão de crenças passará a ter mais uma operação, a *revisão adaptativa*. Esta operação vai poder ser efectuada sobre qualquer contexto, sempre que se encontra uma contradição, mas apenas faz sentido enfraquecer regras universais.

---

<sup>1</sup>A base de conhecimento do agente corresponde ao espaço de crenças definido por um contexto da lógica SWMC. Por agora, vamos supor que o agente quando deixa de acreditar numa fbf a retira da sua base de conhecimento, embora isso não seja verdade.

A revisão adaptativa de um contexto com uma regra universal deverá ter como resultado um outro contexto, com todas as regras que continha o contexto inicial, mas em que a regra universal é substituída por uma regra de omissão que represente o mesmo conhecimento, mas que permita a existência de exceções. Ao enfraquecimento da regra universal  $\forall(x)[A(x)]$  vai corresponder a regra de omissão  $\nabla(x)[A(x)]$ . Desta forma, definimos uma função, *Enf*, que dada uma regra universal a transforma na regra de omissão correspondente:

$$Enf(\forall(x)[A(x)]) = \nabla(x)[A(x)]$$

### 5.1.1 Revisão adaptativa de um contexto com uma fbf

Tal como a revisão, a revisão adaptativa de crenças também pode ser definida à custa das outras operações da teoria de revisão de crenças. Neste caso é ainda necessário utilizar a função *Enf*, definida na secção anterior.

A revisão adaptativa do contexto  $\beta$  com uma fbf  $A$  é definida como:

$$(\beta \otimes A) = \begin{cases} (\beta - A) + Enf(A) & \text{se } A \text{ for uma regra universal} \\ \beta & \text{caso contrário} \end{cases}$$

Esta operação corresponde a retirar a regra universal do contexto que representa a base de conhecimento do agente e depois estendê-lo com a regra de omissão correspondente. A segunda operação pode ser uma extensão (em vez de uma revisão) porque uma regra de omissão nunca pode ser inconsistente com um contexto: as fórmulas que dependerem dela é que poderão estar envolvidas em contradições aparentes, mas a regra de omissão pode sempre ser adicionada a qualquer contexto, uma vez que só existem contradições entre fbfs da LPO.

Assim, todas as consequências da regra universal deixarão de ser deriváveis a partir do novo contexto. Gostaríamos que todas as que ainda forem consistentes com o resto do conhecimento do agente possam voltar a ser derivadas a partir da regra de omissão correspondente.

Se repararmos na definição da função *Enf*, é trivial verificar que, a não ser nos casos em que estamos perante uma exceção para a regra universal, tudo o que pode ser derivado a partir da regra universal também pode ser derivado a partir da regra de omissão e vice-versa.

Mas as consequências de uma regra de omissão não dependem apenas da regra de omissão: dependem também da suposição de que a regra de omissão é aplicável a uma instância particular. Numa primeira análise deste facto, poderia parecer que a operação

de revisão adaptativa descrita acima estava errada e que precisava de ser alterada, para passar também a adicionar a suposição correspondente. No entanto, todas as suposições de que a nova regra de omissão pode ser aplicada a cada uma das constantes existentes na base de conhecimento do agente são adicionadas ao contexto estendido definido pelo contexto. Logo, todas as consequências da regra de omissão que forem consistentes com o resto do conhecimento do agente poderão voltar a ser derivadas a partir do novo contexto.

Desta forma, todas as fórmulas que dependiam da regra universal que ainda são consistentes com o resto do conhecimento do agente continuam a poder pertencer ao espaço de crenças definido pelo contexto que corresponde à base de conhecimento do agente,<sup>2</sup> pois podem voltar a ser derivadas a partir da regra de omissão correspondente.<sup>3</sup>

Temos ainda o problema que as consequências da regra universal, que agora dependem da regra de omissão, podem entrar em conflito com outras crenças do agente. Este problema já foi abordado na secção 4.2.1, em que falámos das preferências entre as crenças do agente, e pode ser resolvido precisamente através da especificação de preferências que permitam escolher as crenças que vão ser mantidas pelo agente.

## 5.2 Alterações ao sistema de revisão de crenças

O SRC tem fundamentalmente duas tarefas: manter dependências entre as crenças do agente para que as alterações possam ser propagadas sem muito esforço computacional; e resolver, ou ajudar a resolver, as contradições que surjam nas crenças do agente. Estas (e outras) tarefas são efectuadas pelos dois componentes principais do SRC, o motor de inferência e o revisor de crenças. Para incorporar a revisão adaptativa de crenças, estes dois componentes vão ter que sofrer alterações.

### 5.2.1 Alterações ao motor de inferência

O motor de inferência é o responsável pela inferência efectuada pelo SRC, e uma das suas tarefas é a propagação das dependências entre as várias fbfs que estão na base de conhecimento do SRC. Para efectuar as inferências possíveis a partir dessas fbfs, e ao mesmo tempo ir registando as dependências entre elas, o motor de inferência utiliza as regras de inferência da lógica SWMC, alteradas para lidar com fbfs justificadas.<sup>4</sup>

<sup>2</sup>Embora possa existir mais do que um espaço de crenças definido por um contexto, vai ser escolhido apenas um para corresponder às crenças do agente.

<sup>3</sup>Esta formulação não traz problemas, pois na teoria de revisão de crenças não existem preocupações ao nível computacional, como por exemplo com a eficiência das operações.

<sup>4</sup>Como foi dito na secção 3.3, a fbf justificada  $\langle A, \tau, \alpha, \gamma \rangle$  corresponde à fbf suportada da SWMC  $\langle A, \tau, \alpha \rangle$  mas em que se regista, em  $\gamma$ , quais foram as crenças imediatamente utilizadas na derivação de  $A$ .

Intuitivamente, a transformação de uma regra universal na regra de omissão correspondente, conforme foi descrita na teoria de revisão de crenças, parece fazer sentido. No entanto, agora é necessário encontrar uma forma mais eficiente de fazer a propagação das dependências entre as várias fbfs da base de conhecimento do sistema: abandonar todas as consequências da regra universal e voltar a derivá-las como consequências da regra de omissão correspondente não é suficiente.

É ainda necessário provar que, em termos lógicos, as consequências da regra universal serão também consequências da regra de omissão correspondente.

### 5.2.1.1 Transformação de regras universais em regras de omissão

A transformação de uma regra universal na regra de omissão correspondente apenas faz sentido se o que se puder derivar a partir da primeira for exactamente igual ao que se pode derivar a partir da segunda, excluindo as excepções à regra universal. Isto corresponde a dizer que a transformação faz sentido se nos casos em que se puder aplicar uma das regras também se pode aplicar a outra (continuando a excluir as excepções à regra universal). As regras universais e as regras de omissão são aplicadas através da aplicação das regras de inferência do SRC, por isso é através das regras de inferência do sistema que vamos averiguar se as condições anteriores se verificam. Neste caso, isto corresponde à aplicação das regras de inferência de eliminação do quantificador universal e de eliminação do quantificador de omissão.<sup>5</sup>

As regras de inferência utilizadas pelo motor de inferência do sistema de revisão de crenças SRC são as mesmas que as da lógica SWMC. No entanto, o SRC utiliza fbfs justificadas em vez de fbfs suportadas, o que vai obrigar à modificação das regras de inferência para manipularem este tipo de fbfs. No entanto, essas modificações são suficientemente simples para não ser necessário introduzir cada uma das regras de inferência. Para facilitar a leitura, vamos reproduzir aqui as duas regras de inferência mencionadas anteriormente, mas agora com as alterações necessárias para a manipulação de fbfs justificadas:

#### Eliminação do quantificador universal ( $E\forall$ )

Se  $Der(\forall(x)[A(x)], BC) \neq \{\}$ , então, para cada  $\delta \in Der(\forall(x)[A(x)], BC)$ , para qualquer símbolo individual  $c$ , podemos adicionar à base de conhecimento  $BC$  a fbf  $\langle A(c), der, \delta, \{\forall(x)[A(x)]\} \rangle$ .

Para compreender esta regra, precisamos da definição de derivação de uma fbf numa base de conhecimento: dada uma fbf  $A$  e uma base de conhecimento  $BC$ , o conjunto das

---

<sup>5</sup>Como se pode verificar no apêndice C, estas são as regras de inferência da lógica SWMC que correspondem à aplicação de uma regra universal ou de uma regra de omissão.

derivações de  $A$  na base de conhecimento  $BC$  é

$$Der(A, BC) = \{\alpha : \langle A, \tau, \alpha, \gamma \rangle \in BC\}$$

isto é, o conjunto de todos os conjuntos de hipóteses e/ou suposições a partir dos quais a fbf  $A$  já foi derivada. Posto isto, a regra da eliminação do quantificador universal diz que, se existir alguma derivação para uma regra universal na base de conhecimento (o que inclui a possibilidade de a regra universal ser uma hipótese), então, para qualquer símbolo individual  $c$  podemos adicionar à base de conhecimento a instanciação da regra universal, tendo como conjunto origem cada uma das possíveis derivações da regra universal.

#### Suposição e eliminação do quantificador de omissão (Sup-E $\nabla$ )

Se  $Der(D, BC) \neq \{\}$ , onde  $D = \nabla(x)[A(x)]$ , então, para qualquer símbolo individual  $c$ , podemos adicionar à base de conhecimento  $BC$  as fbfs  $\langle Applicable(D, c), asp, \{D, Applicable(D, c)\}, \{D, Applicable(D, c)\} \rangle$ , e  $\langle A(c), der, \{D, Applicable(D, c)\}, \{D, Applicable(D, c)\} \rangle$ .

A regra da suposição e eliminação do quantificador de omissão diz que, se existir alguma derivação para uma regra de omissão na base de conhecimento (voltando a incluir a possibilidade de a regra de omissão ser uma hipótese), então, para qualquer símbolo individual  $c$  podemos adicionar à base de conhecimento a suposição que a regra de omissão é aplicável a  $c$  e também a instanciação da regra de omissão com  $c$ . Neste caso, e uma vez que as regras de omissão só podem ser hipóteses,<sup>6</sup> qualquer uma das fbfs adicionadas vai depender da regra de omissão e da suposição que a regra de omissão é aplicável a  $c$ .

Como podemos verificar, as condições de aplicabilidade das duas regras de inferência são muito semelhantes: existir pelo menos uma derivação para a fbf que estivermos a considerar (regra universal ou regra de omissão).

As regras universais podem ser introduzidas como hipóteses ou derivadas pelo motor de inferência, casos em que os conjuntos de origem serão diferentes. Para tentarmos saber qual o conjunto de origem da fbf suportada correspondente à regra universal que pretendemos enfraquecer, convém lembrar que a estamos a enfraquecer por termos chegado a uma contradição. Para resolver a contradição, temos que remover alguma fbf do conjunto de origem da fbf que corresponde à contradição. Como nos conjuntos de origem das fbfs suportadas só existem hipóteses, a regra universal que estamos a enfraquecer vai ser uma hipótese. Por isso, vai existir apenas um  $\delta$  em  $Der(\forall(x)[A(x)], BC)$ , que vai ser  $\delta = \{\forall(x)[A(x)]\}$ .

---

<sup>6</sup>Uma vez que não existe nenhuma regra de inferência que permita introduzir uma regra de omissão na base de conhecimento, para além da regra da introdução da hipótese.

As regras de omissão, como já dissemos, correspondem sempre a hipóteses. Por isso, sendo  $\nabla(x)[A(x)]$  uma regra de omissão, vai existir apenas um  $\delta$  em  $Der(\nabla(x)[A(x)], BC)$ , que vai ser  $\delta = \{\nabla(x)[A(x)]\}$ .

Tanto no caso da regra universal como no caso da regra de omissão, vai-se poder adicionar à base de conhecimento  $BC$  a instanciação da regra utilizada, o que, de acordo com a definição da função  $Enf$ , e utilizando a mesma instanciação para a variável, vai ser igual.

Assim, tudo o que poderia ser derivado a partir da regra universal vai poder ser derivado a partir da regra de omissão correspondente. No entanto, com a utilização da regra de omissão, vamos também poder (e dever) adicionar à base de conhecimento a suposição de que a regra de omissão é aplicável à instância que estamos a considerar. O problema que temos que resolver é garantir que usamos com a regra de omissão a mesma instanciação que foi usada com a regra universal, e que criamos a suposição com a instanciação correcta. Se ao nível da lógica este facto não traz problemas, pois todas as possíveis suposições são criadas na construção do contexto estendido, ao nível do sistema queremos efectuar esta operação de um modo mais eficiente. Para isso, precisamos de poder identificar rapidamente qual a suposição necessária e criá-la sem ser necessário voltar a construir as estruturas do sistema de revisão de crenças que dependem do contexto que está a ser utilizado.

### 5.2.1.2 Como determinar a suposição correcta?

Para garantir de uma forma eficiente a utilização da mesma instanciação com a regra universal e com a regra de omissão é necessário manter um registo, sempre que é feita a eliminação de um quantificador universal, de qual a instância que foi usada nessa eliminação. Se, na derivação de uma determinada fbf for aplicada a regra da eliminação do quantificador universal mais do que uma vez com a mesma regra universal, cada aplicação deve ser registada independentemente, e com a instanciação respectiva. A forma que escolhemos para fazer este registo é através da adição de um novo elemento à estrutura que representa as fbfs justificadas, onde será feito o registo das instanciações. Este novo elemento será um conjunto de instanciações.

As novas fbfs justificadas com mais este elemento, a que vamos chamar *instanciações*, passarão a ser representadas por  $\langle A, \tau, \alpha, \gamma, \epsilon \rangle$  e denominadas *fbfs super-justificadas*. A fbf super-justificada  $\langle A, \tau, \alpha, \gamma, \epsilon \rangle$  corresponde à fbf justificada  $\langle A, \tau, \alpha, \gamma \rangle$ , em que se regista, em  $\epsilon$ , as aplicações das regras universais utilizadas e respectivas instanciações. Uma instanciação é um par  $(R, c)$ , em que  $R$  é uma regra universal e  $c$  é uma constante, com a qual a regra universal foi instanciada.

Voltando mais uma vez ao exemplo do Piupiu, se tivermos a informação que o Piupiu

é uma ave e que as aves voam, podemos inferir que o Piupiu voa. Em termos de fbfs super-justificadas, esta informação pode ser representada pelas fbfs da figura 5.1.<sup>7</sup>

fbf super-justificada	abreviatura
$\langle ave(Piupiu), hyp, \{a(P)\}, \{\}, \{\} \rangle$	$a(P)$
$\langle \forall(x)[ave(x) \rightarrow voa(x)], hyp, \{a(x) \rightarrow v(x)\}, \{\}, \{\} \rangle$	$a(x) \rightarrow v(x)$
$\langle ave(Piupiu) \rightarrow voa(Piupiu), der, \{a(x) \rightarrow v(x)\}, -\{a(x) \rightarrow v(x)\}, \{(a(x) \rightarrow v(x), Piupiu)\} \rangle$	$a(P) \rightarrow v(P)$
$\langle voa(Piupiu), der, \{a(P), a(x) \rightarrow v(x)\}, -\{a(P), a(P) \rightarrow v(P)\}, \{(a(x) \rightarrow v(x), Piupiu)\} \rangle$	$v(P)$

Tabela 5.1: Fbfs super-justificadas da inferência de que o Piupiu voa.

Podemos agora enunciar a nova regra da eliminação do quantificador universal, que passará a usar fbfs super-justificadas:

#### Eliminação do quantificador universal ( $E\forall$ )

Se  $Der(\forall(x)[A(x)], BC) \neq \{\}$ , então, para cada  $\delta \in Der(\forall(x)[A(x)], BC)$ , para qualquer símbolo individual  $c$ , podemos adicionar à base de conhecimento  $BC$  a fbf  $\langle A(c), der, \delta, \{\forall(x)[A(x)]\}, \{(\forall(x)[A(x)], c)\} \rangle$ .

Todas as outras regras de inferência do motor de inferência do SRC serão alteradas para fazer a reunião das instanciações das fbfs super-justificadas utilizadas, e poderão ser consultadas no apêndice D. A regra da eliminação do quantificador universal apresentada no apêndice é mais geral do que esta, pois considera também o caso em que a regra universal pode ter sido derivada, e conseqüentemente as suas instanciações não serem o conjunto vazio.

Ao fazer o registo das instanciações que são usadas com a aplicação de cada regra universal, quando se pretender fazer o enfraquecimento de uma regra universal utilizada na derivação de uma fbf, sabe-se exactamente qual a suposição que deve ser levantada para a aplicação da regra da suposição e eliminação do quantificador de omissão.

#### 5.2.1.3 Enfraquecimento do quantificador universal

Para poder enfraquecer o quantificador universal, transformando-o num quantificador de omissão, o motor de inferência do SRC vai precisar de uma nova regra de inferência. Por um lado, pretendemos acrescentar a regra de omissão à base de conhecimento do agente; por outro, pretendemos retirar a regra universal da base de conhecimento do agente.

<sup>7</sup>Associámos com cada fbf uma abreviatura, para facilitar a sua referência ao longo do texto e nas estruturas associadas às fbfs.

Na realidade, na base de conhecimento do agente estão todas as fbfs que ele já considerou, tanto as inseridas a partir do exterior como as obtidas através de inferência. Isto significa que, uma vez introduzida na base de conhecimento do agente, uma fbf nunca vai ser retirada.

Assim, o motor de inferência apenas vai adicionar à base de conhecimento do agente a regra de omissão correspondente à regra universal a ser enfraquecida, deixando para o revisor de crenças a tarefa de fazer com que o agente deixe de ter a regra universal na sua base de conhecimento e de propagar esta alteração ao resto das crenças do agente.

Podemos agora enunciar a regra do enfraquecimento do quantificador universal:

#### **Enfraquecimento do quantificador universal (Enf $\forall$ )**

Se  $Der(\forall(x)[A(x)], BC) \neq \{\}$ , então podemos adicionar à base de conhecimento  $BC$  a fbf  $\langle \nabla(x)[A(x)], hyp, \{\nabla(x)[A(x)]\}, \{\nabla(x)[A(x)]\}, \{\} \rangle$ .

À parte da existência desta nova regra de inferência, todas as tarefas do motor de inferência continuam a ser feitas exactamente como eram feitas antes desta alteração, e conforme foram descritas na secção 3.3.2<sup>8</sup>, mas em que se ignora o último elemento das novas fbfs.

### **5.2.2 Alterações ao revisor de crenças**

Tal como foi dito na secção 3.3.3, o revisor de crenças é o componente do sistema de revisão de crenças responsável por todas as tarefas que envolvem uma mudança nas crenças do agente:

- detecção de contradições
- remoção de contradições
- remoção de fbfs de contextos
- escolha dos espaços de crenças preferidos pelo sistema

A detecção de contradições e a escolha dos espaços de crenças preferidos pelo sistema continuam a ser feitas como anteriormente, uma vez que as tarefas correspondentes da TRC em que o sistema se baseia não sofreram alterações.

A remoção de fbfs de contextos pode ser efectuada por duas razões distintas: ou porque foi feito, a partir do exterior, um pedido ao SRC para remover uma fbf; ou porque foi detectada uma contradição e é necessário abandonar uma fbf para a resolver.

<sup>8</sup>Para uma descrição completa ver [Cravo 1992, Cravo 1995].

Quando é feito, a partir do exterior, um pedido ao SRC para remover uma fbf, o SRC não vai entrar em linha de conta com a nova operação definida ao nível da teoria de revisão de crenças, e vai continuar a funcionar como anteriormente. Esta operação não sofreu alterações porque não nos parece fazer sentido remover uma fbf de um contexto eliminando uma das suas possíveis derivações e fornecendo em troca uma forma mais fraca de a derivar.

### 5.2.2.1 Remoção de uma fbf devido à detecção de uma contradição

Quando a remoção de uma fbf é feita devido à detecção de uma contradição nas crenças do agente, pode fazer sentido fazer uma revisão adaptativa das suas crenças, mas quem deve tomar esta decisão é o utilizador do sistema. Por isso, quando são calculadas as várias alternativas para remover uma fbf de um contexto, desde que alguma das alternativas incluía a remoção de uma regra universal, vai ser calculada uma outra alternativa, em tudo idêntica à anterior, mas em que se faz o enfraquecimento da regra universal em vez da sua remoção. Esta nova alternativa corresponde à nova operação da teoria de revisão de crenças. Se alguma das alternativas de revisão incluía a remoção de mais do que uma regra universal, o novo sistema calcula todos os seus subconjuntos, de modo a que o utilizador possa escolher entre remover e enfraquecer qualquer combinação de regras universais. No caso em que o utilizador escolhe uma alternativa que envolve o enfraquecimento de uma regra universal, isso vai corresponder à aplicação da nova regra do motor de inferência, a regra do enfraquecimento do quantificador universal.

Depois da aplicação desta regra de inferência, o motor de inferência deverá propagar os seus efeitos a todas as fbfs que dependiam da regra universal. Uma implementação deste sistema de revisão de crenças deverá ter um mecanismo de indexação que torne esta propagação trivial, por exemplo, através do registo, em cada fbf, das fbfs que a utilizaram na sua derivação.

Basicamente, o motor de inferência deverá procurar todas as fbfs justificadas que dependiam da regra universal e adicionar-lhe um novo conjunto de origem. Este conjunto, correspondendo a uma nova derivação para a fbf, será semelhante ao conjunto de origem que continha a regra universal, mas em que se substitui a regra universal pela regra de omissão correspondente e pela suposição de que se pode aplicar a regra de omissão à instância a que a regra universal tinha sido aplicada. Esta instância é determinada através do último elemento das fbfs super-justificadas (as instanciações), onde deverá existir um par com a regra universal enfraquecida e a instância que foi utilizada na sua aplicação.

Mais uma vez, convém referir que, se a conclusão da regra universal entrava em conflito com alguma outra crença do agente, a conclusão da regra de omissão também vai entrar em conflito com essa(s) crença(s). Se a outra crença for sólida (isto é, dependia apenas de fbfs da LPO — ver definição na página 35), então a conclusão da regra de omissão

não vai pertencer ao espaço de crenças que representa as crenças do agente; caso contrário, podem existir vários espaços de crenças que podem representar as crenças do agente, e um deles vai ser escolhido, ou com base nas preferências entre as crenças do agente ou então através duma intervenção exterior ao sistema.

Podemos ilustrar o novo método de remoção de contradições pelo sistema com um exemplo concreto, em que se chega a uma contradição e em que se mostram as alternativas para resolver a contradição, calculadas pelo sistema antes e depois da criação da revisão adaptativa de crenças.

Consideremos a seguinte informação:

*As aves voam*

*Quem tem asas voa*

*Quem tem as asas partidas não voa*

*O Piupiu é uma ave que tem asas, mas que estão partidas*

Esta informação pode ser esquematizada conforme a figura 5.1.

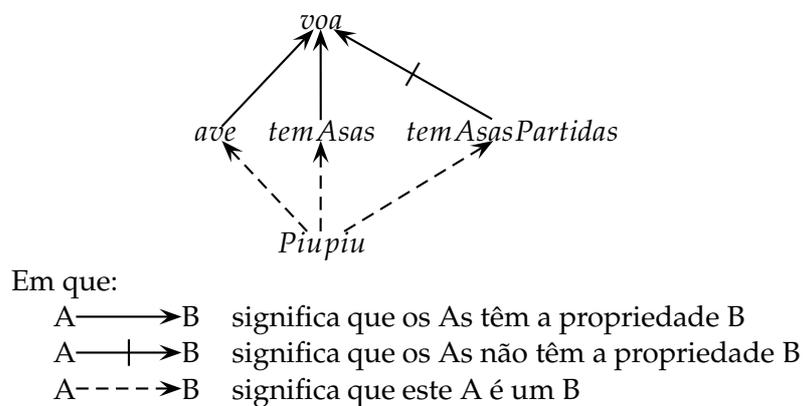


Figura 5.1: Informação esquemática acerca das aves.

Em termos de fbfs super-justificadas, esta informação seria representada pelas fbfs da tabela 5.2.

Ao derivar todas as consequências da informação da tabela 5.2, obteríamos as fbfs da tabela 5.3.

Como se pode verificar, chegámos a uma contradição, pois derivámos que o Piupiu voa e que ele não voa. Podemos derivar a contradição de duas maneiras distintas (usando a fbf  $v(P)1$  ou a fbf  $v(P)2$ ), por isso, se quisermos removê-la, temos que invalidar ambas as derivações.

fbf super-justificada	abreviatura
$\langle \forall(x)[ave(x) \rightarrow voa(x)], hyp, \{a(x) \rightarrow v(x)\}, \{\}, \{\} \rangle$	$a(x) \rightarrow v(x)$
$\langle \forall(x)[temAsas(x) \rightarrow voa(x)], hyp, \{tA(x) \rightarrow v(x)\}, \{\}, \{\} \rangle$	$tA(x) \rightarrow v(x)$
$\langle \forall(x)[temAsasPartidas(x) \rightarrow \neg voa(x)], hyp, -\{tAP(x) \rightarrow nV(x)\}, \{\}, \{\} \rangle$	$tAP(x) \rightarrow nV(x)$
$\langle ave(Piupiu), hyp, \{a(P)\}, \{\}, \{\} \rangle$	$a(P)$
$\langle temAsas(Piupiu), hyp, \{tA(P)\}, \{\}, \{\} \rangle$	$tA(P)$
$\langle temAsasPartidas(Piupiu), hyp, \{tAP(P)\}, \{\}, \{\} \rangle$	$tAP(P)$

Tabela 5.2: Fbfs que representam a informação da figura 5.1.

fbf super-justificada	abreviatura
$\langle ave(Piupiu) \rightarrow voa(Piupiu), der, \{a(x) \rightarrow v(x)\}, \{a(x) \rightarrow v(x)\}, -\{(a(x) \rightarrow v(x), Piupiu)\} \rangle$	$a(P) \rightarrow v(P)$
$\langle temAsas(Piupiu) \rightarrow voa(Piupiu), der, \{tA(x) \rightarrow v(x)\}, -\{tA(x) \rightarrow v(x)\}, \{(tA(x) \rightarrow v(x), Piupiu)\} \rangle$	$tA(P) \rightarrow v(P)$
$\langle temAsasPartidas(Piupiu) \rightarrow \neg voa(Piupiu), der, \{tAP(x) \rightarrow nV(x)\}, -\{tAP(x) \rightarrow nV(x)\}, \{(tAP(x) \rightarrow nV(x), Piupiu)\} \rangle$	$tAP(P) \rightarrow nV(P)$
$\langle voa(Piupiu), der, \{a(x) \rightarrow v(x), a(P)\}, \{a(P) \rightarrow v(P), a(P)\}, -\{(a(x) \rightarrow v(x), Piupiu)\} \rangle$	$v(P)1$
$\langle voa(Piupiu), der, \{tA(x) \rightarrow v(x), tA(P)\}, \{tA(P) \rightarrow v(P), tA(P)\}, -\{(tA(x) \rightarrow v(x), Piupiu)\} \rangle$	$v(P)2$
$\langle \neg voa(Piupiu), der, \{tAP(x) \rightarrow nV(x), tAP(P)\}, -\{tAP(x) \rightarrow nV(x), tAP(P)\}, \{(tAP(x) \rightarrow nV(x), Piupiu)\} \rangle$	$nV(P)$

Tabela 5.3: Consequências das fbfs da tabela 5.2.

Antes de ter sido introduzida no sistema a revisão adaptativa de crenças, o sistema iria calcular as seguintes alternativas para remover a contradição, que correspondem a invalidar simultaneamente as derivações de  $v(P)1$  e  $v(P)2$  ou então a invalidar a derivação de  $nV(P)$ :

1. Remover  $a(x) \rightarrow v(x)$  e  $tA(x) \rightarrow v(x)$
2. Remover  $a(x) \rightarrow v(x)$  e  $tA(P)$
3. Remover  $tA(x) \rightarrow v(x)$  e  $a(P)$
4. Remover  $tAP(x) \rightarrow nV(x)$
5. Remover  $a(P)$  e  $tA(P)$

6. Remover  $tAP(P)$ 

Com a revisão adaptativa de crenças, passamos a ter as alternativas anteriores e mais algumas, que são aquelas em que se pode enfraquecer uma ou mais regras universais:

7. Enfraquecer  $a(x) \rightarrow v(x)$  e  $tA(x) \rightarrow v(x)$
8. Enfraquecer  $a(x) \rightarrow v(x)$  e remover  $tA(x) \rightarrow v(x)$
9. Remover  $a(x) \rightarrow v(x)$  e enfraquecer  $tA(x) \rightarrow v(x)$
10. Enfraquecer  $a(x) \rightarrow v(x)$  e remover  $tA(P)$
11. Enfraquecer  $tA(x) \rightarrow v(x)$  e remover  $a(P)$
12. Enfraquecer  $tAP(x) \rightarrow nV(x)$

As alternativas 7, 8 e 9 foram calculadas a partir da alternativa 1, em que se enfraquecem as duas regras universais ou cada uma delas individualmente. As alternativas 10, 11 e 12 foram calculadas a partir das alternativas 2, 3 e 4, respectivamente, e correspondem ao enfraquecimento da regra universal em vez da sua remoção. No caso da alternativa 7, em que se enfraquecem as duas regras universais, pode ainda ser necessário estabelecer uma preferência entre as regras de omissão correspondentes, para o sistema poder escolher o espaço de crenças preferido, no caso de eventualmente algumas das suas conclusões entrarem em conflito.

Como podemos ver, com a revisão adaptativa de crenças podemos optar por fazer a revisão como anteriormente ou então por manter o máximo de informação possível, fazendo uma revisão mais permissiva, mas em que se resolve à mesma a contradição, tendo eventualmente que especificar preferências entre as crenças. Esta abordagem tem a vantagem de nos permitir manter tanto conhecimento como com uma teoria da coerência, mas continuando a usar uma teoria dos fundamentos, o que nos permite ter sempre uma justificação para as fbfs em que o sistema acredita.

## Capítulo 6

# Um sistema de revisão de crenças capaz de rever adaptativamente as suas crenças

Neste capítulo descrevemos a implementação de um sistema de revisão de crenças capaz de rever adaptativamente as suas crenças. Este sistema não foi criado de raiz, mas corresponde a uma extensão da implementação do sistema de revisão de crenças descrito ao nível abstracto na secção 3.3.

Neste capítulo, vamos ainda descrever os aspectos de algumas versões do sistema que são relevantes para este trabalho. Em cada secção serão dadas referências bibliográficas para descrições mais completas de cada versão.

### 6.1 O SNePS

O SNePS<sup>1</sup> [Shapiro 1979, Shapiro & Rapaport 1987, Shapiro & Martins 1990] é um sistema de representação do conhecimento e raciocínio baseado em redes semânticas, que foi desenhado para representar as crenças de um agente racional. Os vários antecessores e as várias versões do sistema são descritos em [Shapiro & Rapaport 1992]. Nesta secção vamos descrever muito sucintamente a versão de 1990 deste sistema, o SNePS 2.1 [Shapiro & Martins 1990].

Sendo o SNePS uma rede semântica proposicional, os nós representam proposições ou conceitos e os arcos servem para estruturar as proposições e conceitos. Por exemplo, para representar em SNePS que *o Piupiu é uma ave*, usaríamos os seguintes nós e arcos:

---

<sup>1</sup>Do inglês Semantic Network Processing System.

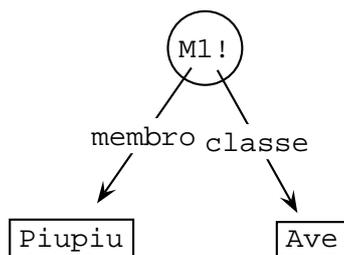


Figura 6.1: Representação da crença de que *o Piupiu é uma ave* em SNePS.

Em que os nós *Piupiu* e *Ave* representam os conceitos *Piupiu* e *Ave*, respectivamente e o nó *M1* representa a proposição que *o Piupiu é uma ave*. Os arcos *membro* e *classe* são definidos pelo utilizador do sistema, e quando usados em conjunto significam que o que estiver no destino do arco *membro* é membro da classe que estiver no destino do arco *classe*. O ponto de exclamação no nó *M1* significa que a proposição que ele representa é acreditada pelo sistema.

O SNePS também permite a representação de regras lógicas, com base na lógica que estiver subjacente a cada versão. No caso do SNePS 2.1, a lógica utilizada é a SWM [Shapiro & Martins 1990], uma lógica relevante<sup>2</sup> e monótona, que permite o registo de dependências entre as várias fbfs da linguagem. A SWM possui as conectivas convencionais da LPO e acrescenta outras novas, sobre as quais não vamos falar, por não serem directamente relevantes para o nosso trabalho. Para representar as várias conectivas, o SNePS usa arcos pré-definidos, aos quais atribui uma semântica meramente operacional, e nos quais se baseia para fazer inferência.

Um princípio importante em SNePS é o princípio da unicidade dos nós: cada nó representa um conceito e cada conceito é representado por um único nó (isto é, não há nós repetidos em SNePS). Isto faz com que todos os nós que representam conceitos relacionados com um dado conceito estejam ligados ao nó que o representa por uma sequência de arcos.

A interacção com o SNePS pode ser feita através de três linguagens distintas:

- O SNePSUL, que permite a estruturação de conceitos através dos nós e arcos que os constituem. Esta é a linguagem que está mais próxima da representação de uma rede de SNePS.
- O SNePSLOG [Shapiro, McKay, Martins & Morgado 1981, Matos & Martins 1989], que corresponde a uma linguagem semelhante à da lógica, com mais alguns comandos específicos para o sistema. Quando a interacção com o sistema é feita através desta linguagem, a criação de nós e arcos deixa de ser uma preocupação do utilizador.

<sup>2</sup>Baseada na lógica relevante de [Anderson & Belnap 1975].

- Uma língua natural (por exemplo o Português ou o Inglês), usando uma gramática ATN existente no sistema, e criando as regras gramaticais para a língua a utilizar.

Neste trabalho, vamos usar o SNePSLOG, por ser uma linguagem bastante semelhante à da lógica, que tem uma tradução directa para as estruturas de nós e arcos usadas pelo SNePS.

O SNePS possui um motor de inferência, o SNIP<sup>3</sup> [Mamede, Martins & Cravo 1988, Pinto-Ferreira, Mamede & Martins 1989]. Este motor de inferência pode efectuar três tipos de inferência [Shapiro, Martins & McKay 1982, Cravo & Martins 1989, Cravo & Martins 1990]:

- Inferência progressiva — quando é adicionada informação ao sistema e é explicitamente pedido para fazer inferência com ela, ou quando se pede ao sistema para fazer inferência com informação que ele já tinha.
- Inferência regressiva — quando é feita uma pergunta ao sistema e a resposta não está explicitamente na base de conhecimento, o sistema vai tentar inferir a resposta, usando inferência regressiva.
- Inferência baseada em caminhos — o sistema pode inferir uma sequência de arcos entre dois nós a partir de uma outra sequência de arcos entre esses nós existente na rede.

O SNePS também dispõe de um revisor de crenças, cujo objectivo é manter a consistência da base de conhecimento. Este revisor de crenças é conceptualmente diferente do que foi descrito neste trabalho, em particular por se basear numa lógica monótona.

## 6.2 O SNePSwD

Em [Cravo 1992] (e resumidamente em [Cravo & Martins 1993]) é feita a descrição de uma implementação do sistema de revisão de crenças SNePSwD.

O SNePSwD<sup>4</sup> foi implementado a partir do SNePS 2.1 e adicionou-lhe a capacidade de efectuar raciocínio de senso comum e de fazer revisão de crenças. O SNePSwD corresponde a uma implementação do sistema de revisão de crenças SRC, descrito na secção 3.3, que se baseia na lógica SWMC e na teoria de revisão de crenças TRC, descritas nas secções 3.1 e 3.2, respectivamente.

Relativamente ao SNePS, o SNePSwD acrescentou a possibilidade de representar o quantificador de omissão, passando a ter a possibilidade de efectuar raciocínio de senso

---

<sup>3</sup>Do Inglês SNePS Inference Package.

<sup>4</sup>Do inglês SNePS with Defaults.

comum numa lógica não-monótona. Assim, passámos a ter mais dois tipos de estruturas que não existiam em SNePS, as regras de omissão e as suposições.<sup>5</sup>

Para comportar esta alteração, os vários componentes do sistema tiveram que ser alterados:

- Foram criados novos arcos para representar as regras de omissão e as suposições da SWMC.
- A linguagem SNePSLOG foi alterada para permitir a representação das regras de omissão, das suposições e das preferências entre as regras de omissão.
- O motor de inferência foi alterado para passar a ter também a regra de inferência que lida com regras de omissão.
- O revisor de crenças foi alterado, de forma a funcionar de acordo com o sistema de revisão de crenças descrito na secção 3.3.

O SNePSwD tem dois modos de funcionamento, o modo manual e o modo automático.

No modo manual, quando é detectada uma contradição, o SNePSwD pergunta ao utilizador se quer ver alternativas para a remover. Estas alternativas são calculadas pelo sistema, e correspondem às mudanças mínimas a fazer na base de conhecimento do agente para que ela volte a ser consistente, geralmente através da remoção de uma ou mais fbfs. Se o utilizador não quiser seguir nenhuma das alternativas, pode fazer a alteração manualmente, e o SRC apenas indica os conjuntos que sabe serem inconsistentes, e dos quais tem que ser removida pelo menos uma fbf de modo a que o contexto volte a ser consistente.

No modo automático, e sempre que o sistema puder, usando as preferências entre as crenças, determinar de forma única qual (ou quais) a(s) crença(s) a remover, o sistema apenas informa o utilizador de que foi detectada uma contradição e de como é que ela foi removida.

### 6.2.1 Exemplo de utilização do SNePSwD

De seguida apresentamos um exemplo de interacção com o SNePSwD em modo manual. O exemplo foi ligeiramente editado para facilitar a sua leitura. O texto a **carregado** corresponde ao que é introduzido pelo utilizador do sistema, o texto entre linhas a *itálico* corresponde a comentários introduzidos posteriormente e o resto do texto ao que é produzido pelo sistema.

---

<sup>5</sup>As preferências entre as regras de omissão também têm que ser representadas, mas não aparecem ao nível da rede semântica.

A informação utilizada é a do exemplo da figura 5.1, da página 63, repetida aqui para facilitar a leitura. No final da interacção com o sistema é introduzida uma nova ave na base de conhecimento, e é perguntado ao sistema se ela voa.

*As aves voam*

*Quem tem asas voa*

*Quem tem as asas partidas não voa*

*O Piupiu é uma ave que tem asas, mas que estão partidas*

*O Tweety é uma ave*

A figura 5.1 é aqui apresentada, agora com a nova ave, e em cada seta é indicado o identificador numérico que lhe vai ser atribuído na interacção com o sistema. A consulta desta figura será particularmente útil para verificar o significado das alternativas sugeridas para remover a contradição detectada no exemplo.

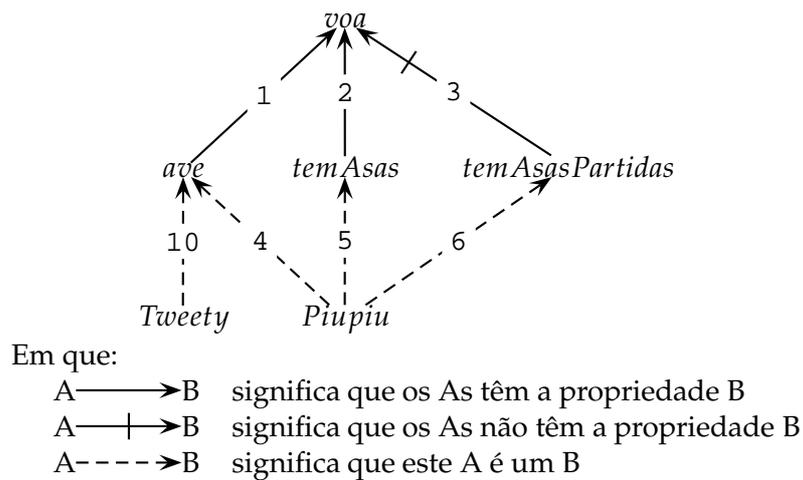


Figura 6.2: Informação esquemática acerca das aves.

USER: (snepslog)

Welcome to SNePSLOG (A logic interface to SNePS)

---

*Este comando serve para entrar na interface baseada em lógica para o SNePSwD.*

---



---

*Segue-se a introdução da informação no sistema. De reparar que o cálculo das estruturas associadas com as fbfs suportadas é feito*

*pelo sistema, o utilizador só tem que introduzir as fbfs em lógica. A cada fbf que é introduzida (ou derivada), o SNePSwD associa um identificador da forma 'WFFx', em que 'x' é um número inteiro único. Este identificador serve para o sistema ou o utilizador se poderem referir posteriormente à fbf.*

---

**: all(x) (ave(x) => voa(x))**

WFF1: all(X)(AVE(X) => VOA(X)) {<HYP,{WFF1}>}

**: all(x) (temAsas(x) => voa(x))**

WFF2: all(X)(TEMASAS(X) => VOA(X)) {<HYP,{WFF2}>}

**: all(x) (temAsasPartidas(x) => ~voa(x))**

WFF3: all(X)(TEMASASPARTIDAS(X) => (~VOA(X))) {<HYP,{WFF3}>}

**: ave(Piupiu)**

WFF4: AVE(PIUPIU) {<HYP,{WFF4}>}

**: temAsas(Piupiu)**

WFF5: TEMASAS(PIUPIU) {<HYP,{WFF5}>}

**: temAsasPartidas(Piupiu)**

WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP,{WFF6}>}

**: voa(Piupiu)?**

---

*É perguntado ao sistema se o Piupiu voa. Como essa informação não está explicitada na base de conhecimento, o sistema vai ter que fazer inferência para poder responder à pergunta.*

---

I wonder if WFF7: VOA(PIUPIU)

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF8: ~VOA(PIUPIU)

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF5: TEMASAS(PIUPIU) {<HYP,{WFF5}>}

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF4: AVE(PIUPIU) {<HYP,{WFF4}>}

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP,{WFF6}>}

holds within the BS defined by context DEFAULT-DEFAULTCT

I know WFF5: TEMASAS(PIUPIU) {<HYP, {WFF5}>}

Since WFF2: all(X)(TEMASAS(X) => VOA(X)) {<HYP, {WFF2}>}

and WFF5: TEMASAS(PIUPIU) {<HYP, {WFF5}>}

I infer WFF7: VOA(PIUPIU)

I know WFF4: AVE(PIUPIU) {<HYP, {WFF4}>}

Since WFF1: all(X)(AVE(X) => VOA(X)) {<HYP, {WFF1}>}

and WFF4: AVE(PIUPIU) {<HYP, {WFF4}>}

I infer WFF7: VOA(PIUPIU) {<DER, {WFF5, WFF2}>}

I know WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP, {WFF6}>}

Since WFF3: all(X)(TEMASASPARTIDAS(X) => (~VOA(X))) {<HYP, {WFF3}>}

and WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP, {WFF6}>}

I infer WFF8: ~VOA(PIUPIU)

A contradiction was detected within context DEFAULT-DEFAULTCT.

---

*É detectada uma contradição, o que é devidamente assinalado pelo sistema. São identificadas as fbfs que estão na base da contradição e é perguntado ao utilizador qual a acção a tomar.*

---

The contradiction involves the newly derived wff:

WFF8: ~VOA(PIUPIU) {<DER, {WFF6, WFF3}>}

and the previously existing wff:

WFF7: VOA(PIUPIU) {<DER, {WFF5, WFF2}>, <DER, {WFF4, WFF1}>}

You have the following options:

1. [C]ontinue anyway, knowing that a contradiction is derivable;
2. [R]e-start the exact same run in a different context which is not inconsistent;
3. [D]rop the run altogether.

(please type c, r or d)

=><= r

Do you want any suggestions ?

=><= y

---

*O utilizador escolhe resolver a contradição e de seguida pede sugestões para como o fazer.*

*As várias alternativas para remover a contradição são apresentadas*

pelo sistema, e correspondem às que foram descritas na página 64. Para facilitar a leitura, apresentamos, a seguir a cada alternativa, as fbfs que ela utiliza.

---

To remove the contradiction from the current context, you have the following alternatives:

Alternative 1:

Remove: (WFF1 WFF5)

---

WFF1:  $all(X)(AVE(X) \Rightarrow VOA(X)) \{<HYP, \{WFF1\}>\}$   
 WFF5:  $TEMASAS(PIUPIU) \{<HYP, \{WFF5\}>\}$

---

Alternative 2:

Remove: (WFF4 WFF5)

---

WFF4:  $AVE(PIUPIU) \{<HYP, \{WFF4\}>\}$   
 WFF5:  $TEMASAS(PIUPIU) \{<HYP, \{WFF5\}>\}$

---

Alternative 3:

Remove: (WFF1 WFF2)

---

WFF1:  $all(X)(AVE(X) \Rightarrow VOA(X)) \{<HYP, \{WFF1\}>\}$   
 WFF2:  $all(X)(TEMASAS(X) \Rightarrow VOA(X)) \{<HYP, \{WFF2\}>\}$

---

Alternative 4:

Remove: (WFF3)

---

WFF3:  $all(X)(TEMASASPARTIDAS(X) \Rightarrow (\sim VOA(X))) \{<HYP, \{WFF3\}>\}$

---

Alternative 5:

Remove: (WFF2 WFF4)

---

WFF2:  $all(X)(TEMASAS(X) \Rightarrow VOA(X)) \{<HYP, \{WFF2\}>\}$   
 WFF4:  $AVE(PIUPIU) \{<HYP, \{WFF4\}>\}$

---

Alternative 6:

Remove: (WFF6)

---

WFF6:  $TEMASASPARTIDAS(PIUPIU) \{<HYP, \{WFF6\}>\}$

---

Do you want to examine some hypotheses?

=><= n

Do you want to choose one of the alternatives?

=><= y

Which one ?

=><= 3

---

*Das várias alternativas, o utilizador escolhe a que lhe parece mais adequada, uma vez que o Piupiu, com as asas partidas, não deve voar, apesar de ser uma ave.*

---

Removed:

WFF1: all(X)(AVE(X) => VOA(X)) {<HYP,{WFF1}>}

WFF2: all(X)(TEMASAS(X) => VOA(X)) {<HYP,{WFF2}>}

New context:

WFF3: all(X)(TEMASASPARTIDAS(X) => (~VOA(X))) {<HYP,{WFF3}>}

WFF4: AVE(PIUPIU) {<HYP,{WFF4}>}

WFF5: TEMASAS(PIUPIU) {<HYP,{WFF5}>}

WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP,{WFF6}>}

---

*Depois de fazer as alterações no contexto, o sistema vai tentar derivar de novo a informação pedida anteriormente, para poder responder à pergunta que lhe foi feita.*

---

I wonder if WFF7: VOA(PIUPIU) {<DER,{WFF5,WFF2}>,<DER,{WFF4,WFF1}>}  
holds within the BS defined by context DEFAULT-DEFAULTCT

I know it is not the case that

WFF7: VOA(PIUPIU) {<DER,{WFF5,WFF2}>,<DER,{WFF4,WFF1}>}

I wonder if WFF8: ~VOA(PIUPIU) {<DER,{WFF6,WFF3}>}

holds within the BS defined by context DEFAULT-DEFAULTCT

I know WFF8: ~VOA(PIUPIU) {<DER,{WFF6,WFF3}>}

I wonder if WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP,{WFF6}>}

holds within the BS defined by context DEFAULT-DEFAULTCT

I know WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP, {WFF6}>}

Since WFF3: all(X)(TEMASASPARTIDAS(X) => (~VOA(X))) {<HYP, {WFF3}>}

and WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP, {WFF6}>}

I infer WFF8: ~VOA(PIUPIU) {<DER, {WFF6, WFF3}>}

WFF8: ~VOA(PIUPIU) {<DER, {WFF6, WFF3}>}

---

*No novo contexto, o Piupiu não voa, como seria de esperar.*

---

**: belief-spaces**

There is one belief space

Belief-space 1:

WFF8: ~VOA(PIUPIU) {<DER, {WFF6, WFF3}>}

---

*Existe apenas um espaço de crenças definido pelo contexto, em que o Piupiu não voa.*

---

**: ave(Tweety)**

WFF10: AVE(TWEETY) {<HYP, {WFF10}>}

---

*Mas, se o utilizador introduzir mais aves no sistema, elas voam ou não voam?*

---

**: voa(Tweety)?**

I wonder if WFF11: VOA(TWEETY)

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF12: ~VOA(TWEETY)

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF13: TEMASASPARTIDAS(TWEETY)

holds within the BS defined by context DEFAULT-DEFAULTCT

---

*Uma vez que a regra que diz que as aves voam foi abandonada, o sistema não consegue saber se o Tweety voa ou não.*

---

Este é um exemplo em que, face a uma contradição, é abandonada informação a mais. O facto de o utilizador ter representado que *as aves voam* e que *quem tem asas voa* com regras universais em vez de ser com regras de omissão fez com que, quando foi detectada

uma contradição, fosse abandonada mais informação do que aquela que seria desejável. Claro que nada impede o utilizador de, depois de resolvida a contradição, introduzir manualmente as regras de omissão correspondentes às regras universais abandonadas. No entanto, se num exemplo académico (e forçosamente reduzido) isso não é problemático, em situações reais, em que podem existir muitas contradições, isso não é prático.

### 6.3 O novo sistema de revisão de crenças — SNePSwDaAR

Ao incorporar no sistema a capacidade de fazer a revisão adaptativa das crenças, vão passar a existir mais alternativas para fazer a revisão das crenças do sistema: as alternativas calculadas pelo SNePSwD aumentadas com todas aquelas em que existe pelo menos uma regra universal que pode ser enfraquecida.

Em termos da representação em SNePS, o que tem que ser feito é criar um novo nó correspondente à regra de omissão, em tudo semelhante ao que representa a regra universal, mas em que se altera o arco que representa o quantificador; uma vez que o SNePS utiliza o princípio da unicidade dos nós, toda a restante estrutura do nó é partilhada pelas duas regras.

Se o utilizador escolher uma alternativa em que não se enfraquecem regras, a revisão é feita como anteriormente pelo SNePSwD. Se escolher uma revisão adaptativa, em que se enfraquece pelo menos uma regra, a BC vai ser tornada consistente através da remoção das regras universais escolhidas e da adição das regras de omissão correspondentes a cada uma delas. Neste caso, a regra de omissão vai passar a pertencer ao contexto e a regra universal vai deixar de pertencer ao contexto.

#### 6.3.1 O comando *weaken-wff*

Foi criado um novo comando de SNePSLOG, *weaken-wff*, que dada uma regra universal a enfraquece, isto é, a transforma em regra de omissão. Este comando só funciona para fbfs acreditadas, tal como as outras operações de modificação de contextos. Não é necessário fazer nada para considerar as excepções à regra de omissão, pois se existia uma regra universal acreditada e ainda não tinha sido detectada nenhuma contradição, é porque ainda não tinha sido encontrada nenhuma excepção a esta regra. Para além disso, se depois de fazer o enfraquecimento da regra universal for acrescentada alguma excepção à regra de omissão, isso vai ser tratado automaticamente pela lógica SWMC, que se vai encarregar de abandonar a crença menos importante, isto é, a conclusão da regra de omissão.

### 6.3.2 Exemplo de utilização do novo sistema SNePSwDaAR

A informação para este exemplo é a mesma do exemplo da figura 5.1, da página 63, à qual foi acrescentada na secção 6.2.1 uma nova ave, e que serviu como exemplo de utilização do SNePSwD. Mais uma vez, o exemplo foi ligeiramente editado para facilitar a sua leitura. O texto a **carregado** corresponde ao que é introduzido pelo utilizador do sistema, o texto entre linhas a *itálico* corresponde a comentários introduzidos posteriormente e o resto do texto ao que é produzido pelo sistema.

*As aves voam*

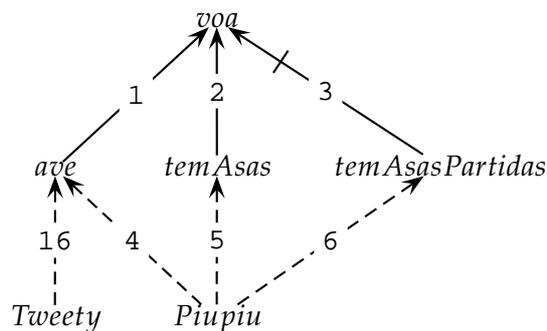
*Quem tem asas voa*

*Quem tem as asas partidas não voa*

*O Piupiu é uma ave que tem asas, mas que estão partidas*

*O Tweety é uma ave*

A figura 5.1 é aqui repetida, também com a nova ave e com a indicação do identificador numérico de cada seta.



Em que:

- A  $\longrightarrow$  B significa que os As têm a propriedade B
- A  $\longrightarrow|$  B significa que os As não têm a propriedade B
- A  $\dashrightarrow$  B significa que este A é um B

Figura 6.3: Informação esquemática acerca das aves.

---

*Toda a interacção com o sistema é igual à do exemplo utilizando o SNePSwD até ao próximo comentário. Está repetida aqui apenas para facilitar a leitura.*

---

USER: (snepslog)

Welcome to SNePSLOG (A logic interface to SNePS)

: **all(x) (ave(x) => voa(x))**

WFF1: all(X)(AVE(X) => VOA(X)) {<HYP, {WFF1}>}

: **all(x) (temAsas(x) => voa(x))**

WFF2: all(X)(TEMASAS(X) => VOA(X)) {<HYP, {WFF2}>}

: **all(x) (temAsasPartidas(x) => ~voa(x))**

WFF3: all(X)(TEMASASPARTIDAS(X) => (~VOA(X))) {<HYP, {WFF3}>}

: **ave(Piupiu)**

WFF4: AVE(PIUPIU) {<HYP, {WFF4}>}

: **temAsas(Piupiu)**

WFF5: TEMASAS(PIUPIU) {<HYP, {WFF5}>}

: **temAsasPartidas(Piupiu)**

WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP, {WFF6}>}

: **voa(Piupiu) ?**

I wonder if WFF7: VOA(PIUPIU)

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF8: ~VOA(PIUPIU)

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF5: TEMASAS(PIUPIU) {<HYP, {WFF5}>}

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF4: AVE(PIUPIU) {<HYP, {WFF4}>}

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP, {WFF6}>}

holds within the BS defined by context DEFAULT-DEFAULTCT

I know WFF5: TEMASAS(PIUPIU) {<HYP, {WFF5}>}

Since WFF2: all(X)(TEMASAS(X) => VOA(X)) {<HYP, {WFF2}>}

and WFF5: TEMASAS(PIUPIU) {<HYP, {WFF5}>}

I infer WFF7: VOA(PIUPIU)

I know WFF4: AVE(PIUPIU) {<HYP, {WFF4}>}

Since WFF1: all(X)(AVE(X) => VOA(X)) {<HYP, {WFF1}>}

and WFF4: AVE(PIUPIU) {<HYP, {WFF4}>}

I infer WFF7: VOA(PIUPIU) {<DER, {WFF5, WFF2}>}

I know WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP, {WFF6}>}

Since WFF3: all(X)(TEMASASPARTIDAS(X) => (~VOA(X))) {<HYP, {WFF3}>}

and WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP, {WFF6}>}

I infer WFF8: ~VOA(PIUPIU)

A contradiction was detected within context DEFAULT-DEFAULTCT.

The contradiction involves the newly derived wff:

WFF8: ~VOA(PIUPIU) {<DER, {WFF6, WFF3}>}

and the previously existing wff:

WFF7: VOA(PIUPIU) {<DER, {WFF5, WFF2}>, <DER, {WFF4, WFF1}>}

You have the following options:

1. [C]ontinue anyway, knowing that a contradiction is derivable;
2. [R]e-start the exact same run in a different context which is not inconsistent;
3. [D]rop the run altogether.

(please type c, r or d)

=><= r

Do you want any suggestions ?

=><= y

---

*Toda a interação com o sistema foi igual até ao ponto em que o utilizador pede ao sistema para mostrar as alternativas para remover a contradição. Com o novo sistema, existem mais alternativas, em que se enfraquecem regras universais em vez de as remover. Estas alternativas foram descritas na página 65.*

---

To remove the contradiction from the current context, you have the following alternatives:

Alternative 1:

Remove: (WFF5)

Weaken: (WFF1)

---

WFF1: all(X)(AVE(X) => VOA(X)) {<HYP, {WFF1}>}

WFF5: TEMASAS(PIUPIU) {<HYP, {WFF5}>}

---

Alternative 2:

Remove: (WFF1 WFF5)

Alternative 3:

Remove: (WFF4 WFF5)

---

WFF4:  $AVE(PIUPIU) \{<HYP, \{WFF4\}>\}$   
 WFF5:  $TEMASAS(PIUPIU) \{<HYP, \{WFF5\}>\}$

---

Alternative 4:

Weaken: (WFF1 WFF2)

---

WFF1:  $all(X)(AVE(X) \Rightarrow VOA(X)) \{<HYP, \{WFF1\}>\}$   
 WFF2:  $all(X)(TEMASAS(X) \Rightarrow VOA(X)) \{<HYP, \{WFF2\}>\}$

---

Alternative 5:

Remove: (WFF1)

Weaken: (WFF2)

Alternative 6:

Remove: (WFF2)

Weaken: (WFF1)

Alternative 7:

Remove: (WFF1 WFF2)

Alternative 8:

Weaken: (WFF3)

Alternative 9:

Remove: (WFF3)

---

WFF3:  $all(X)(TEMASASPARTIDAS(X) \Rightarrow (\sim VOA(X))) \{<HYP, \{WFF3\}>\}$

---

Alternative 10:

Remove: (WFF4)

Weaken: (WFF2)

---

WFF2:  $all(X)(TEMASAS(X) \Rightarrow VOA(X)) \{<HYP, \{WFF2\}>\}$   
 WFF4:  $AVE(PIUPIU) \{<HYP, \{WFF4\}>\}$

---

Alternative 11:

Remove: (WFF2 WFF4)

Alternative 12:

Remove: (WFF6)

---

WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP,{WFF6}>}

---

Do you want to examine some hypotheses?

=><= n

Do you want to choose one of the alternatives?

=><= y

Which one ?

=><= 4

---

*Mais uma vez, a informação menos importante para o utilizador é a que está representada em WFF1 e WFF2. No entanto, com o novo sistema, o utilizador pode optar por enfraquecer estas regras em vez de as abandonar.*

---

Weakened:

WFF1: all(X)(AVE(X) => VOA(X)) {<HYP,{WFF1}>}

WFF2: all(X)(TEMASAS(X) => VOA(X)) {<HYP,{WFF2}>}

New context:

WFF10: default(X)(TEMASAS(X) => VOA(X)) {<HYP,{WFF10}>}

WFF3: all(X)(TEMASASPARTIDAS(X) => (~VOA(X))) {<HYP,{WFF3}>}

WFF4: AVE(PIUPIU) {<HYP,{WFF4}>}

WFF5: TEMASAS(PIUPIU) {<HYP,{WFF5}>}

WFF6: TEMASASPARTIDAS(PIUPIU) {<HYP,{WFF6}>}

WFF9: default(X)(AVE(X) => VOA(X)) {<HYP,{WFF9}>}

I wonder if WFF7: VOA(PIUPIU) {<DER,{WFF5,WFF2}>,<DER,{WFF4,WFF1}>}  
holds within the BS defined by context DEFAULT-DEFAULTCT

I know it is not the case that

WFF7: VOA(PIUPIU) {<DER,{WFF5,WFF2}>,<DER,{WFF4,WFF1}>}

I wonder if WFF8: ~VOA(PIUPIU) {<DER,{WFF6,WFF3}>}

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF2:  $\text{all}(X)(\text{TEMASAS}(X) \Rightarrow \text{VOA}(X)) \{ \langle \text{HYP}, \{\text{WFF2}\} \rangle \}$   
 holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF1:  $\text{all}(X)(\text{AVE}(X) \Rightarrow \text{VOA}(X)) \{ \langle \text{HYP}, \{\text{WFF1}\} \rangle \}$   
 holds within the BS defined by context DEFAULT-DEFAULTCT

I know WFF8:  $\sim \text{VOA}(\text{PIUPIU}) \{ \langle \text{DER}, \{\text{WFF6}, \text{WFF3}\} \rangle \}$

I wonder if WFF5:  $\text{TEMASAS}(\text{PIUPIU}) \{ \langle \text{HYP}, \{\text{WFF5}\} \rangle \}$   
 holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF4:  $\text{AVE}(\text{PIUPIU}) \{ \langle \text{HYP}, \{\text{WFF4}\} \rangle \}$   
 holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF6:  $\text{TEMASASPARTIDAS}(\text{PIUPIU}) \{ \langle \text{HYP}, \{\text{WFF6}\} \rangle \}$   
 holds within the BS defined by context DEFAULT-DEFAULTCT

I know WFF5:  $\text{TEMASAS}(\text{PIUPIU}) \{ \langle \text{HYP}, \{\text{WFF5}\} \rangle \}$

Since WFF10:  $\text{default}(X)(\text{TEMASAS}(X) \Rightarrow \text{VOA}(X)) \{ \langle \text{HYP}, \{\text{WFF10}\} \rangle \}$   
 and WFF5:  $\text{TEMASAS}(\text{PIUPIU}) \{ \langle \text{HYP}, \{\text{WFF5}\} \rangle \}$   
 I infer WFF7:  $\text{VOA}(\text{PIUPIU}) \{ \langle \text{DER}, \{\text{WFF5}, \text{WFF2}\} \rangle, \langle \text{DER}, \{\text{WFF4}, \text{WFF1}\} \rangle \}$

I know WFF4:  $\text{AVE}(\text{PIUPIU}) \{ \langle \text{HYP}, \{\text{WFF4}\} \rangle \}$

Since WFF9:  $\text{default}(X)(\text{AVE}(X) \Rightarrow \text{VOA}(X)) \{ \langle \text{HYP}, \{\text{WFF9}\} \rangle \}$   
 and WFF4:  $\text{AVE}(\text{PIUPIU}) \{ \langle \text{HYP}, \{\text{WFF4}\} \rangle \}$   
 I infer WFF7:  $\text{VOA}(\text{PIUPIU}) \{ \langle \text{DER}, \{\text{WFF5}, \text{WFF2}\} \rangle, \langle \text{DER}, \{\text{WFF4}, \text{WFF1}\} \rangle, \langle \text{DER}, \{\text{WFF5}, \text{WFF11}, \text{WFF10}\} \rangle \}$

I know WFF6:  $\text{TEMASASPARTIDAS}(\text{PIUPIU}) \{ \langle \text{HYP}, \{\text{WFF6}\} \rangle \}$

Since WFF3:  $\text{all}(X)(\text{TEMASASPARTIDAS}(X) \Rightarrow (\sim \text{VOA}(X))) \{ \langle \text{HYP}, \{\text{WFF3}\} \rangle \}$   
 and WFF6:  $\text{TEMASASPARTIDAS}(\text{PIUPIU}) \{ \langle \text{HYP}, \{\text{WFF6}\} \rangle \}$   
 I infer WFF8:  $\sim \text{VOA}(\text{PIUPIU}) \{ \langle \text{DER}, \{\text{WFF6}, \text{WFF3}\} \rangle \}$   
 WFF8:  $\sim \text{VOA}(\text{PIUPIU}) \{ \langle \text{DER}, \{\text{WFF6}, \text{WFF3}\} \rangle \}$

**: belief-spaces**

There is one belief space

Belief-space 1:

WFF8:  $\sim \text{VOA}(\text{PIUPIU}) \{ \langle \text{DER}, \{\text{WFF6}, \text{WFF3}\} \rangle \}$

---

*Depois de enfraquecer as regras universais que permitem derivar que o Piupiu voa, e uma vez que a regra que diz que quem tem as asas partidas não voa continua a ser uma regra universal, o sistema apenas acredita que o Piupiu não voa.*

: **ave(Tweety)**

WFF16: AVE(TWEETY) {<HYP,{WFF16}>}

---

*Se acrescentarmos uma outra ave à base de conhecimento do sistema,  
será que ela voa?*

---

: **voa(Tweety)?**

I wonder if WFF17: VOA(TWEETY)

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF18: ~VOA(TWEETY)

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF2: all(X)(TEMASAS(X) => VOA(X)) {<HYP,{WFF2}>}

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF1: all(X)(AVE(X) => VOA(X)) {<HYP,{WFF1}>}

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF20: TEMASAS(TWEETY)

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF16: AVE(TWEETY) {<HYP,{WFF16}>}

holds within the BS defined by context DEFAULT-DEFAULTCT

I wonder if WFF24: TEMASASPARTIDAS(TWEETY)

holds within the BS defined by context DEFAULT-DEFAULTCT

I know WFF16: AVE(TWEETY) {<HYP,{WFF16}>}

Since WFF9: default(X)(AVE(X) => VOA(X)) {<HYP,{WFF9}>}

and WFF16: AVE(TWEETY) {<HYP,{WFF16}>}

I infer WFF17: VOA(TWEETY)

WFF17: VOA(TWEETY) {<DER,{WFF9,WFF22,WFF16}>}

: **belief-spaces**

There is one belief space

Belief-space 1:

WFF17: VOA(TWEETY) {<DER,{WFF9,WFF22,WFF16}>}

WFF21: TEMASAS(TWEETY) => VOA(TWEETY) {<DER,{WFF19,WFF10}>}

WFF23: AVE(TWEETY) => VOA(TWEETY) {<DER, {WFF9, WFF22}>}

WFF8: ~VOA(PIUPIU) {<DER, {WFF6, WFF3}>}

---

*No caso do Tweety, é razoável assumir que ele voa, uma vez que é uma ave e não há nada que diga que ele não deve voar.*

---

: **temasaspartidas(Tweety)!**

---

*Se mais tarde for dito ao sistema que o Tweety também tem as asas partidas, o que é que pode ser derivado? O '!' no fim da fbf que foi introduzida pede ao sistema para derivar tudo o que conseguir a partir dessa fbf.*

---

Since WFF3: all(X)(TEMASASPARTIDAS(X) => (~VOA(X))) {<HYP, {WFF3}>}

and WFF24: TEMASASPARTIDAS(TWEETY) {<HYP, {WFF24}>}

I infer WFF18: ~VOA(TWEETY)

I know it is not the case that

WFF17: VOA(TWEETY) {<DER, {WFF9, WFF22, WFF16}>}

WFF18: ~VOA(TWEETY) {<DER, {WFF3, WFF24}>}

WFF24: TEMASASPARTIDAS(TWEETY) {<HYP, {WFF24}>}

: **belief-spaces**

There is one belief space

Belief-space 1:

WFF18: ~VOA(TWEETY) {<DER, {WFF3, WFF24}>}

WFF8: ~VOA(PIUPIU) {<DER, {WFF6, WFF3}>}

---

*Uma vez que ambas as aves têm as asas partidas, nenhuma delas voa.*

---

Neste caso, quando foi detectada a contradição, o utilizador escolheu rever as suas crenças de modo a eliminar a contradição, mas mantendo o máximo de informação possível que ainda era consistente com o resto do conhecimento do sistema. Assim, quando foram introduzidas mais aves na base de conhecimento do sistema, o sistema foi capaz de derivar algumas das suas propriedades.

Embora este exemplo seja muito reduzido, julgamos que é suficiente para mostrar em que é que o novo sistema difere do anterior. Optámos por não mostrar mais exemplos porque os exemplos suficientemente curtos para poderem ser aqui apresentados não iriam mostrar nada de novo.

Realmente interessante seria mostrar como o sistema se comporta a “juntar” duas ou mais bases de conhecimento, em que o conhecimento representado seria forçosamente mais extenso, e possivelmente iriam existir mais interações entre crenças que dessem origem a contradições.

## 6.4 Aplicações do novo sistema de revisão de crenças

A revisão adaptativa de crenças tem a vantagem, relativamente à revisão de crenças tradicional, de manter mais informação na base de conhecimento do agente quando é detectada uma contradição que envolva uma ou mais regras universais. Em particular, quando é encontrada uma excepção para uma regra universal, é sugerido que a regra universal seja transformada numa regra de omissão. Assim, durante a construção de uma base de conhecimento, é possível usar regras universais mesmo quando não se tem a certeza de essa regra não ter excepções, pois temos a garantia que, se for encontrada alguma excepção isso será tratado sem problemas. Este tipo de comportamento do sistema torna-o particularmente adequado para algumas aplicações, que vamos descrever em seguida.

- Ao representar conhecimento nem sempre temos todo o conhecimento certo à partida, e é possível que existam erros de representação, por isso é natural que o conhecimento representado precise de ser corrigido. Para além disso, com grandes quantidades de conhecimento é natural que não se consigam prever todas as consequências de todas as crenças, que algumas delas entrem em conflito, e que tenham que ser corrigidas. Assim, um sistema com revisão adaptativa de crenças poria à disposição de quem representa conhecimento uma ferramenta que pode ajudar a verificar a consistência do conhecimento representado, fazendo inferência e tentando resolver contradições. O sistema deve resolver as contradições que conseguir e para as outras deve dar as várias alternativas e perguntar o que fazer.
- Em alguns casos, pode ser útil criar uma base de conhecimento que utilize conhecimento já representado noutras bases de conhecimento. Nestas situações, é natural que algumas das bases de conhecimento repitam a representação de algum conhecimento, o que pode não acontecer necessariamente de formas consistentes. Este sistema pode ajudar a detectar contradições existentes e sugerir formas de as resolver, o que facilita a tarefa de quem estiver a construir a nova base de conhecimento.
- Ao criar uma base de conhecimento a partir de várias fontes de informação, é natural que alguns dos aspectos transmitidos pelas várias fontes sejam contraditórios.

No entanto, podemos ter à partida diferentes graus de confiança em cada uma das fontes e saber que, em caso de conflito, devemos preferir a informação vinda de uma determinada fonte relativamente à informação de outras fontes. Neste caso, podemos usar preferências para expressar o nosso grau de confiança em cada fonte e usar o sistema de revisão de crenças para manter a base de conhecimento consistente.

- Uma vez que a funcionalidade deste sistema corresponde a uma extensão da funcionalidade do SNePSwD, este sistema continua a poder ser usado em todas as aplicações do SNePSwD, já descritas em [Cravo 1992, pp. 263–284].

# Capítulo 7

## Trabalho futuro

Neste capítulo descrevemos alguns pontos que podem ser objecto de estudo no futuro, e que correspondem a possíveis melhoramentos e extensões ao trabalho descrito nesta tese.

### 7.1 Evolução da lógica SWMC

Alguns dos aspectos merecedores de mais estudo ao nível da lógica SWMC são os seguintes:

- Em [Cravo 1992, sec. 9.2.1, pp. 291] é sugerido, como trabalho futuro, o estudo da derivação de regras de omissão. Neste trabalho criámos uma regra de inferência que permite inferir regras de omissão, mas apenas ao nível do sistema de revisão de crenças. Julgamos que seria interessante estudar como é que esta regra poderia ser incluída ao nível da lógica, considerando simultaneamente as propriedades da sua inclusão no sistema semântico.
- Na secção 4.2.2 vimos que a contabilização de excepções ao nível da lógica não é trivial, mas que a possibilidade de representar este tipo de conhecimento juntamente com o resto do conhecimento do agente traria várias vantagens. Assim, seria interessante fazer um estudo da possibilidade de contabilizar as excepções ao nível da lógica, e de representar este conhecimento juntamente com o resto do conhecimento do agente. Para formalizar a contabilização das excepções na base de conhecimento do agente seria provavelmente necessário criar uma meta-linguagem,<sup>1</sup> em que se pudessem representar meta-regras como: *em caso de conflito, preferir regras com menos excepções a regras com mais excepções; quando mais do que metade dos objectos a que a regra pode ser aplicada são excepções, fazer com que a regra desça na hierarquia, se isso*

---

<sup>1</sup>Uma meta-linguagem é uma linguagem que fala acerca de outra linguagem ou acerca dela própria.

*for possível; se todas as excepções a uma regra forem da mesma classe, condensar todas as excepções numa excepção referente à classe, etc.*

## 7.2 Tópicos em aberto na área da revisão de crenças

Nesta secção, vamos descrever alguns dos tópicos que consideramos importantes na área de revisão de crenças e que ainda não foram abordados:

- Como vimos na secção 3.2, em que descrevemos a teoria de revisão de crenças baseada na lógica SWMC, alguns dos postulados propostos por Gärdenfors não são satisfeitos pela TRC, por a lógica subjacente ser não-monótona. Seria interessante fazer um estudo de quais os postulados a sugerir para a TRC, e em particular determinar quais os postulados que a nova operação de revisão adaptativa de crenças deveria satisfazer.
- Na secção 3.2.2 vimos que a TRC considera uma ordem combinada entre as crenças para escolher uma alternativa para a operação de remoção, quando isso é possível. Pensamos que, à partida, essa mesma ordem poderia ser usada para guiar a escolha das crenças a enfraquecer, no caso de existir alguma preferência que possa ser aplicada. No entanto, ainda é necessário fazer um estudo formal para ver se esta afirmação é verdadeira. A(s) alternativa(s) seleccionada(s) pela TRC poderia(m) ser sugerida(s) pelo SRC ao utilizador do sistema, que se poderia basear nela(s) para efectuar a sua escolha. As preferências poderiam ser calculadas automaticamente através da contabilização de excepções (ver secção 7.3).
- Se considerarmos que um agente racional tem, para além das suas crenças, desejos de melhorar o seu estado actual e possivelmente intenções de chegar a um determinado estado do mundo, e que estes desejos e intenções são fortemente influenciados pelas crenças que o agente tem acerca do mundo, então é natural que a revisão de crenças esteja intimamente relacionada com os processos de raciocínio utilizados pelo agente para manipular os seus desejos e intenções. O trabalho de modelação de agentes na área de Inteligência Artificial Distribuída utiliza regularmente uma arquitectura de agente que contempla estes estados cognitivos. Por isto mesmo, para construir um agente racional, é importante caracterizar também a revisão de desejos e de intenções e relacioná-las com a revisão de crenças.
- Grande parte do trabalho feito actualmente em IA utiliza outras lógicas, para além das clássicas (lógica de primeira ordem ou lógica proposicional), para representar os vários tipos de informação que podemos ter acerca do mundo:
  - Informação típica: Normalmente as aves voam, mas nem todas...
  - Informação causal: Se estiver a chover, levo o chapéu de chuva

- Informação taxonómica: Os cães são mamíferos
- Informação tipificada: As pessoas que têm mais de 1,80m
- Etc...

Por exemplo, embora a informação causal e a informação taxonómica estejam intimamente relacionadas com determinados tipos de raciocínio, que são diferentes para cada uma, ambas são normalmente representadas em lógica de primeira ordem usando implicações, o que não permite distinguir qual o mecanismo de inferência mais adequado para cada um dos casos.

Seria interessante o estudo das operações de mudança nas crenças de um agente utilizando modelos para as crenças mais ricos (lógicas não-monótonas, lógicas descritivas,<sup>2</sup> lógicas tipificadas, etc.).

- O valor epistémico de uma crença, no qual nos devemos basear durante a especificação de preferências entre as crenças, é um conceito que não está perfeitamente definido. Intuitivamente, corresponde ao grau de confiança que o agente tem em cada uma das suas crenças, e pode estar relacionado com a sua fonte (enciclopédia, professor, vizinho do lado, etc.), com as crenças a partir das quais foi derivada, etc. Parece fazer sentido que ao longo do tempo, e/ou com a chegada de nova informação este valor possa ser alterado. Por exemplo, se o agente confirmar o que o vizinho lhe disse numa enciclopédia, faz sentido que o valor epistémico da crença confirmada seja reforçado. Por enquanto, ainda não existem teorias que estudem a natureza do valor epistémico das crenças e a sua evolução, mas pensamos que este é um aspecto importante que deveria ser considerado.

### 7.3 Criação de um módulo de aprendizagem para o SRC

Quando se pretende construir um agente racional, um dos principais objectivos é que ele seja o mais autónomo possível. Nesse sentido, incorporar no sistema a possibilidade de escolher automaticamente a(s) crença(s) a enfraquecer, baseado no que fosse ditado pela TRC, seria da maior importância.

Para aumentar ainda mais a autonomia do agente, o ideal seria que ele fosse aprendendo como fazer revisão de crenças à medida que o utilizador o vai guiando, e que no final da sua construção a revisão de crenças fosse totalmente automática, incluindo também a especificação de preferências entre crenças.

Um dos domínios em que nos parece particularmente interessante tentar automatizar o processo de revisão é quando representamos conhecimento acerca de hierarquias. Neste caso, em vez da simples passagem de regras universais para regras de omissão,

---

<sup>2</sup>Do Inglês *description logics*.

poderíamos fazer com que as regras fossem “subindo” ou “descendo” na hierarquia, conforme fossem encontradas confirmações ou exceções.

Entre outras, julgamos que as capacidades descritas de seguida seriam importantes num sistema com capacidades de aprendizagem.

### 7.3.1 Tratamento de exceções: indução de regras

Para ilustrar porque é que consideramos importante o sistema ter a capacidade de induzir novas regras, vamos voltar a considerar como exemplo a representação das aves e da sua forma de deslocação. Um possível começo seria considerar que *todas as aves voam*. Quando fosse encontrada a primeira exceção, passaríamos a considerar que *normalmente as aves voam*, e em princípio trataríamos as exceções subsequentes individualmente, dizendo que *o Piupiu não voa, o Oscar não voa*, etc.

Com o aparecimento de mais exceções, não deveria ser difícil agrupá-las em subclasses das aves: pinguins, avestruzes, aves mortas, etc. Seria interessante que o agente fosse capaz de determinar as características que estas aves têm em comum, por exemplo pertencerem a uma determinada classe, e que, quando tivesse um número significativo de exceções, as transformasse em regras como: *os pinguins não voam*.

Teríamos, assim, um agente com capacidades de indução. No entanto, com esta capacidade surge também um problema: como saber que induções é que são válidas. Por exemplo, a partir de cem corvos pretos, poderemos induzir que todos os corvos são pretos (apesar de poderem existir alguns corvos albinos, que seriam uma exceção a esta regra), enquanto que a mesma indução não é válida a partir de cem galinhas pretas, uma vez que existem demasiadas galinhas que seriam exceção a esta regra.

Seria também interessante estudar a possibilidade de o agente *descobrir* novas classes de objectos a partir dos predicados que são satisfeitos pelas constantes da linguagem. Poderia ser gerada uma hierarquia paralela à que tivesse sido introduzida directamente na base de conhecimento do agente, de modo a que a maior parte das regras tivesse menos exceções e possivelmente a inferência fosse mais eficiente.

### 7.3.2 Fortalecimento de regras: passagem do particular para o geral

Consideremos agora que fornecíamos ao agente informação acerca do tipo de cobertura de cada subclasse das aves: *os pinguins têm penas, os corvos têm penas, as galinhas têm penas*, etc.

Neste caso, seria desejável, por questões de economia de regras e até mesmo de exac-

tidão do conhecimento, que o agente inferisse que *as aves têm (cobertura de) penas*.<sup>3</sup>

Esta capacidade de generalização está intimamente ligada com a capacidade de determinar características comuns entre classes de objectos — a indução, e poderia ser muito útil em sistemas cujo funcionamento depende do reconhecimento de semelhanças, não só para descobrir novas propriedades, mas também para determinar possíveis aberrações.

### 7.3.3 Enfraquecimento de regras: passagem do geral para o particular

Numa possível expansão do exemplo anterior poderíamos representar conhecimento acerca da forma de deslocação dos animais. Uma vez que a grande maioria dos animais existentes até ao momento voavam, seria de esperar que o agente generalizasse a informação de que as aves voam e inferisse que *normalmente os animais voam*.

No entanto, à medida que fossem acrescentadas novas classes de animais, como por exemplo os peixes e os mamíferos, o número de excepções a esta nova regra deveria fazer com que o agente deixasse de ter confiança nas conclusões derivadas a partir dela. Neste ponto, gostaríamos que o agente tivesse a capacidade de decidir que a regra tinha demasiadas excepções e que a particularizasse, para cada uma das subclasses dos animais, e ficasse de novo com a regra acerca das aves, outra para os mamíferos, que andam, e outra para os peixes, que nadam.

Claro que estas novas regras também têm excepções:

Os pinguins, as avestruzes, etc. são aves e não voam.

As baleias, os morcegos, etc. são mamíferos e não andam.

Mas, intuitivamente, as excepções encontradas não parecem suficientes para fazer as regras acerca da forma de deslocação destas classes descer ainda mais um nível na hierarquia dos animais. Devemos reconhecer que esta intuição é particularmente difícil (ou completamente impossível) de implementar, até porque não somos capazes de saber exactamente do que se trata. No entanto, um compromisso razoável seria considerar a relação entre o número de novas regras necessárias para representar a forma de deslocação e o número de regras correspondentes a excepções a essas regras.

A determinação à partida de qual o nível de uma hierarquia mais adequado para colocar uma regra acerca de uma propriedade dessa hierarquia não é um problema simples. Numa abordagem simplista, poderíamos considerar qual o nível com a melhor relação entre o número de excepções e o número de confirmações abaixo da regra. Um bom ponto de partida para um estudo mais alargado é o algoritmo de classificação do KL-ONE [Schmolze & Lipkis 1983].

---

<sup>3</sup>A escolha de uma regra universal ou de omissão poderia ser feita pelo utilizador.

Muitas vezes somos forçados a rever as nossas representações iniciais, mas podemos criar mecanismos que nos permitam, até certos limites, testar a coerência das nossas representações. Neste caso, uma boa estratégia seria verificar que as exceções de uma regra não estão mais acima na hierarquia do que a própria regra.

#### **7.4 Aumento da funcionalidade do SNePSwDaAR**

Como seria de esperar, tudo o que foi dito na secção anterior que pode melhorar o funcionamento do SRC ao nível abstracto, pode também ser incluído numa outra implementação deste sistema que vise aumentar a sua funcionalidade.

Consideramos que incluir um modo de funcionamento automático no SNePSwDaAR, que retire do utilizador o peso de escolher a alternativa a seguir para efectuar as mudanças no contexto seria de particular importância. Uma vez que o SNePSwD já tem esta possibilidade, pensamos que seria simples inclui-la no novo sistema, desde que, ao nível da TRC, fosse especificado como é que deve ser feita a ordenação das alternativas que incluem o enfraquecimento de regras em relação às outras.

#### **7.5 Relação com a evolução de teorias científicas**

No início deste século começaram a ser feitas observações que contrariavam as leis da mecânica Newtoniana. Com a evolução dos equipamentos de observação e com a possibilidade de observar fenómenos que envolviam velocidades cada vez maiores, as previsões das leis de Newton começaram a falhar. Durante algumas décadas, o que foi feito foi acomodar essas observações à teoria existente, que “já tinha dado provas da sua correcção” ao prever correctamente os acontecimentos observáveis durante um longo período de tempo. As observações que não estavam de acordo com o que era previsto pela teoria eram consideradas como casos “anormais”, que não eram explicados correctamente pela teoria devido a erros na observação ou então eram considerados como exceções, mas a teoria não era alterada.

Podemos estabelecer um paralelo entre esta atitude e a teoria de revisão de crenças proposta neste trabalho, em que as exceções vão sendo acomodadas ao conhecimento existente, fazendo-se o mínimo possível de mudanças. Embora esta atitude não seja a mais correcta possível, tem a vantagem de permitir a manutenção de uma base de conhecimento consistente, em que todas as crenças têm (pelo menos) uma justificação.

Voltando às leis da mecânica, verificamos que foi necessário pôr em causa as noções de espaço e de tempo até então aceites como certas para se poder criar uma nova teoria que previsse correctamente os fenómenos observados. Estamos a falar da teoria da re-

latividade de Einstein, que tem como caso particular a mecânica Newtoniana, mas que provocou uma revolução científica ao nível da física.

Uma extensão a este trabalho poderia tentar modelar este tipo de comportamento, isto é, quando as regras da base de conhecimento começam a ter demasiadas excepções, pô-las em causa e sugerir outras que explicassem correctamente os factos conhecidos, incluindo os que correspondiam a excepções da teoria anterior.

# Apêndice A

## Notação utilizada

- os conjuntos de crenças, fechados dedutivamente, são representados por letras gregas maiúsculas:  $\Delta, \Sigma, \Lambda, \dots$
- os conjuntos finitos de crenças (ou bases de crenças, que não são fechadas dedutivamente) são representados por letras gregas minúsculas:  $\alpha, \beta, \gamma, \dots$
- as fórmulas (geralmente correspondentes a crenças), são representadas por letras romanas maiúsculas:  $A, B, C, \dots$ ;
- as operações de expansão, contracção e revisão de um conjunto de crenças  $\Delta$  com uma fórmula  $A$ , são representadas, respectivamente, por  $(\Delta + A)$ ,  $(\Delta - A)$ ,  $(\Delta * A)$  (ou  $(\alpha + A)$ ,  $(\alpha - A)$ ,  $(\alpha * A)$  para as bases finitas de crenças);
- o conjunto de todas as consequências de um conjunto finito de fórmulas  $\alpha$  é representado por  $Cn(\alpha)$ ;
- o conjunto  $\Delta_{\perp}$  representa o conjunto de crenças inconsistente, isto é, o conjunto de todas as fórmulas da linguagem.

## Apêndice B

### Abreviaturas utilizadas

ATMS	Assumption-based Truth Maintenance System Sistema de revisão de crenças baseado em suposições
fbf	fórmula bem formada
IA	Inteligência Artificial
JTMS	Justification-based Truth Maintenance System Sistema de revisão de crenças baseado em justificações
LPO	Lógica de Primeira Ordem
RC	Representação do Conhecimento
SNePS	Semantic Network Processing System
SNePSwD	SNePS with Defaults
SNePSwDaAR	SNePS with Defaults and Adaptative Revision
SRC	Sistema de Revisão de Crenças
SWM	Lógica desenvolvida por Shapiro, Wand e Martins
SWMC	Lógica não-monótona, desenvolvida por M. Cravo, a partir da SWM
TRC	Teoria de revisão de crenças

## Apêndice C

# Regras de inferência da lógica SWMC

### Regras clássicas

#### Hipótese (Hip)

Dada uma qualquer fbf  $A$ , em que  $A \in (\mathcal{L}_{FOL} \cup \mathcal{L}_D \cup \mathcal{L}_E)$ , podemos escrever a fbf suportada  $\langle A, hyp, \{A\} \rangle$ .

#### Introdução da implicação (I $\rightarrow$ )

A partir de  $\langle B, \tau, \alpha \rangle$ ,  $B \in \mathcal{L}_{FOL}$  e qualquer hipótese  $H \in (\alpha \cap \mathcal{L}_{FOL})$  podemos inferir  $\langle H \rightarrow B, der, \alpha - \{H\} \rangle$ .

#### Modus Ponens – Eliminação da implicação (MP)

A partir de  $\langle A, \tau_1, \alpha_1 \rangle$  e  $\langle A \rightarrow B, \tau_2, \alpha_2 \rangle$  podemos inferir  $\langle B, der, \alpha_1 \cup \alpha_2 \rangle$ .

#### Modus Tollens – Eliminação da implicação (MT)

A partir de  $\langle \neg B, \tau_1, \alpha_1 \rangle$  e  $\langle A \rightarrow B, \tau_2, \alpha_2 \rangle$  podemos inferir  $\langle \neg A, der, \alpha_1 \cup \alpha_2 \rangle$ .

#### Introdução da dupla negação (IDN)

A partir de  $\langle A, \tau, \alpha \rangle$ , se  $A \in \mathcal{L}_{FOL}$ , podemos inferir  $\langle \neg\neg A, der, \alpha \rangle$ .

#### Eliminação da dupla negação (EDN)

A partir de  $\langle \neg\neg A, \tau, \alpha \rangle$  podemos inferir  $\langle A, der, \alpha \rangle$ .

#### Introdução da negação (I $\neg$ )

A partir de  $\langle A \wedge \neg A, \tau, \alpha \rangle$ , e qualquer hipótese  $H \in \alpha \cap \mathcal{L}_{FOL}$ , podemos inferir  $\langle \neg H, der, \alpha - \{H\} \rangle$ .

**Introdução da conjunção (I $\wedge$ )**

A partir de  $\langle A, \tau_1, \alpha_1 \rangle, \langle B, \tau_2, \alpha_2 \rangle, A \in \mathcal{L}_{FOL}$  e  $B \in \mathcal{L}_{FOL}$  podemos inferir  $\langle A \wedge B, der, \alpha_1 \cup \alpha_2 \rangle$ .

**Eliminação da conjunção (E $\wedge$ )**

A partir de  $\langle A \wedge B, \tau, \alpha \rangle$  podemos inferir  $\langle A, der, \alpha \rangle$ , ou  $\langle B, der, \alpha \rangle$ , ou ambas.

**Introdução da disjunção (I $\vee$ )**

A partir de  $\langle A, \tau, \alpha \rangle, A \in \mathcal{L}_{FOL}$  e qualquer  $B \in \mathcal{L}_{FOL}$  podemos inferir  $\langle A \vee B, der, \alpha \rangle$ , ou  $\langle B \vee A, der, \alpha \rangle$ , ou ambas.

**Eliminação da disjunção (E $\vee$ )**

A partir de  $\langle A \vee B, \tau_1, \alpha_1 \rangle, \langle A \rightarrow C, \tau_2, \alpha_2 \rangle$  e  $\langle B \rightarrow C, \tau_3, \alpha_3 \rangle$  podemos inferir  $\langle C, der, \alpha_1 \cup \alpha_2 \cup \alpha_3 \rangle$ .

**Introdução do quantificador universal (I $\forall$ )**

A partir de  $\langle A(c), der, \alpha \rangle, A(c) \in \mathcal{L}_{FOL}$ , e  $A(c)$  usa um símbolo individual  $c$  que não ocorre em nenhuma fbf em  $\alpha$  podemos inferir  $\langle \forall(x)[A(x)], der, \alpha \rangle$ .

**Eliminação do quantificador universal (E $\forall$ )**

A partir de  $\langle \forall(x)[A(x)], \tau, \alpha \rangle$  e qualquer símbolo individual  $c$  podemos inferir  $\langle A(c), der, \alpha \rangle$ .

**Introdução do quantificador existencial (I $\exists$ )**

A partir de  $\langle A(c), \tau, \alpha \rangle$ , em que  $c$  é um símbolo individual, e  $A(c) \in \mathcal{L}_{FOL}$  podemos inferir  $\langle \exists(x)[A(x)], der, \alpha \rangle$ .

**Eliminação do quantificador existencial (E $\exists$ )**

A partir de  $\langle \exists(x)[A(x)], \tau, \alpha_1 \rangle, \langle B, der, \{A(c)\} \cup \alpha_2 \rangle$ , e  $A(c)$  usa um símbolo individual  $c$  que não ocorre em  $B$  nem em nenhuma fbf em  $\alpha_2$  podemos inferir  $\langle B, der, \alpha_1 \cup \alpha_2 \rangle$ .

**Regra estendida****Suposição e eliminação do quantificador de omissão (Sup-E $\nabla$ )**

A partir de  $\langle \nabla(x)[A(x)], hyp, \{\nabla(x)[A(x)]\} \rangle$ , podemos inferir, para qualquer símbolo individual  $c$ ,  $\langle Applicable(\nabla(x)[A(x)], c), asp, \alpha \rangle$ , e  $\langle A(c), der, \alpha \rangle$ , em que  $\alpha = \{\nabla(x)[A(x)], Applicable(\nabla(x)[A(x)], c)\}$ .

## Apêndice D

# Regras de inferência do novo sistema de revisão de crenças

Para podermos apresentar as regras de inferência do novo sistema de revisão de crenças, e uma vez que elas entram em linha de conta com todas as derivações que existem para cada fbf, é necessário usar as funções *Der* e *CombDer*, já definidas em [Cravo 1995, pp. 12]), mas ligeiramente adaptadas para lidarem com fbfs super-justificadas:

Dada uma fbf  $A$  e uma base de conhecimento  $BC$ , o conjunto das derivações de  $A$  na base de conhecimento  $BC$  é

$$Der(A, BC) = \{\alpha : \langle A, \tau, \alpha, \gamma, \epsilon \rangle \in BC\}$$

isto é, o conjunto de todos os conjuntos de hipóteses e/ou suposições a partir dos quais a fbf  $A$  já foi derivada.

Dado um conjunto de fbfs  $\beta$  e uma base de conhecimento  $BC$ , o conjunto das combinações das derivações das fbfs de  $\beta$  em  $BC$  é o conjunto de conjuntos mínimos,  $\delta$ , que satisfazem a seguinte condição:

$$CombDer(\beta, BC) = \{\delta : \forall_{(A \in \beta)} [\exists_{(\alpha \in Der(A, BC))} [\alpha \subseteq \delta]]\}$$

isto é, o conjunto de todos os conjuntos mínimos obtidos pela união de uma e uma só derivação de cada uma das fbfs em  $\beta$ .

Nas nossas regras de inferência, vamos querer combinar directamente derivações de fbfs super-justificadas, isto é, em vez de termos as fbfs e de precisarmos de calcular as suas derivações, temos as derivações que queremos considerar para cada uma das fbfs. Neste caso, a combinação das derivações corresponde à sua união, uma vez que a união de conjuntos já garante que não existem elementos repetidos, e portanto já é um conjunto

mínimo. Vamos usar o símbolo  $\cup$  para representar a união de vários conjuntos:

$$\bigcup(\alpha_1, \dots, \alpha_n) = \alpha_1 \cup \dots \cup \alpha_n$$

Usando o exemplo de [Cravo 1995, pp. 12], e supondo que  $BC$  contém as seguintes fbfs super-justificadas,

$$\langle A, \tau_{A1}, \alpha_{A1}, \gamma_{A1}, \epsilon_{A1} \rangle$$

$$\langle A, \tau_{A2}, \alpha_{A2}, \gamma_{A2}, \epsilon_{A2} \rangle$$

$$\langle B, \tau_{B1}, \alpha_{B1}, \gamma_{B1}, \epsilon_{B1} \rangle$$

$$\langle B, \tau_{B2}, \alpha_{B2}, \gamma_{B2}, \epsilon_{B2} \rangle$$

Temos:

$$Der(A, BC) = \{\alpha_{A1}, \alpha_{A2}\}$$

$$Der(B, BC) = \{\alpha_{B1}, \alpha_{B2}\}$$

$$CombDer(\{A, B\}, BC) = \{\alpha_{A1} \cup \alpha_{B1}, \alpha_{A1} \cup \alpha_{B2}, \alpha_{A2} \cup \alpha_{B1}, \alpha_{A2} \cup \alpha_{B2}\}$$

$$\bigcup(\alpha_{A1}, \alpha_{B1}) = \alpha_{A1} \cup \alpha_{B1}$$

supondo que nenhum destes conjuntos contém outro.

Precisamos ainda de nos poder referir às instanciações das fbfs super-justificadas. Para isso, vamos definir uma função  $inst^1$ , que dada uma fbf super-justificada dá o seu conjunto de instanciações:

$$inst(\langle A, \tau, \alpha, \gamma, \epsilon \rangle) = \epsilon$$

Convém notar que as instanciações  $\epsilon$ , que são criadas ao longo da derivação de uma fbf super-justificada, dependem apenas da derivação particular da fbf, que pode ser obtida a partir de  $\alpha$ . Assim, ao longo da apresentação das regras de inferência vamos usar um abuso de notação, e usar  $inst(A, \alpha)$  em vez de  $inst(\langle A, \tau, \alpha, \gamma, \epsilon \rangle)$ .

---

<sup>1</sup>Do inglês *instantiation*.

## Regras clássicas

### Hipótese (Hip)

Dada qualquer fbf  $A \in \cup(\mathcal{L}_{FOL}, \mathcal{L}_D, \mathcal{L}_E)$ , podemos adicionar à base de conhecimento  $BC$  a fbf super-justificada  $\langle A, hyp, \{A\}, \{\}, \{\} \rangle$ .

### Introdução da implicação (I $\rightarrow$ )

Se  $Der(B, BC) \neq \{\}$  e  $B \in \mathcal{L}_{FOL}$ , então, para cada  $\delta \in Der(B, BC)$  tal que  $H \in \delta$ , podemos adicionar à  $BC$  a fbf super-justificada  $\langle H \rightarrow B, der, \delta - \{H\}, (\{H\}, B), inst(B, \delta) \rangle^2$ .

### Modus Ponens – Eliminação da implicação (MP)

Se  $Der(A, BC) \neq \{\}$  e  $Der(A \rightarrow B, BC) \neq \{\}$ , então, para cada  $\delta_1 \in Der(A, BC)$ , e para cada  $\delta_2 \in Der(A \rightarrow B, BC)$ , podemos adicionar à base de conhecimento  $BC$  a fbf super-justificada  $\langle B, der, \cup(\delta_1, \delta_2), \{A, A \rightarrow B\}, \cup(inst(A, \delta_1), inst(A \rightarrow B, \delta_2)) \rangle$ .

### Modus Tollens – Eliminação da implicação (MT)

Se  $Der(\neg B, BC) \neq \{\}$  e  $Der(A \rightarrow B, BC) \neq \{\}$ , então, para cada  $\delta_1 \in Der(\neg B, BC)$  e para cada  $\delta_2 \in Der(A \rightarrow B, BC)$ , podemos adicionar à base de conhecimento  $BC$  a fbf super-justificada  $\langle \neg A, der, \cup(\delta_1, \delta_2), \{\neg B, A \rightarrow B\}, \cup(inst(\neg B, \delta_1), inst(A \rightarrow B, \delta_2)) \rangle$ .

### Introdução da dupla negação (IDN)

Se  $Der(A, BC) \neq \{\}$  e  $A \in \mathcal{L}_{FOL}$ , então, para cada  $\delta \in Der(A, BC)$ , podemos adicionar à base de conhecimento  $BC$  a fbf super-justificada  $\langle \neg\neg A, der, \delta, \{A\}, inst(A, \delta) \rangle$ .

### Eliminação da dupla negação (EDN)

Se  $Der(\neg\neg A, BC) \neq \{\}$ , então, para cada  $\delta \in Der(\neg\neg A, BC)$ , podemos adicionar à base de conhecimento  $BC$  a fbf super-justificada  $\langle A, der, \delta, \{\neg\neg A\}, inst(\neg\neg A, \delta) \rangle$ .

### Introdução da negação (I $\neg$ )

Se  $Der(A \wedge \neg A, BC) \neq \{\}$ , então, para cada  $\delta \in Der(A \wedge \neg A, BC)$ , tal que  $H \in \delta \cap \mathcal{L}_{FOL}$ , podemos adicionar à base de conhecimento  $BC$  a fbf super-justificada  $\langle \neg H, der, \delta - \{H\}, \{(\{H\}, A \wedge \neg A)\}, inst(A \wedge \neg A, \delta) \rangle$ .

### Introdução da conjunção (I $\wedge$ )

Se  $Der(A, BC) \neq \{\}$ ,  $Der(B, BC) \neq \{\}$ ,  $A \in \mathcal{L}_{FOL}$  e  $B \in \mathcal{L}_{FOL}$ , então, para cada  $\delta_1 \in Der(A, BC)$ , e para cada  $\delta_2 \in Der(B, BC)$ , podemos adicionar à base de conhecimento  $BC$  a fbf super-justificada  $\langle A \wedge B, der, \cup(\delta_1, \delta_2), \{A, B\}, \cup(inst(A, \delta_1), inst(B, \delta_2)) \rangle$ .

<sup>2</sup>Uma vez que  $H \in \delta$  e  $\delta$  só contém hipóteses, as intanciações de  $H$  são o conjunto vazio.

**Eliminação da conjunção (E $\wedge$ )**

Se  $Der(A \wedge B, BC) \neq \{\}$ , então, para cada  $\delta \in Der(A \wedge B, BC)$ , podemos adicionar à base de conhecimento  $BC$  as fbfs super-justificadas  $\langle A, der, \delta, \{A \wedge B\}, inst(A \wedge B, \delta) \rangle$ , ou  $\langle B, der, \delta, \{A \wedge B\}, inst(A \wedge B, \delta) \rangle$ , ou ambas.

**Introdução da disjunção (I $\vee$ )**

Se  $Der(A, BC) \neq \{\}$  e  $A \in \mathcal{L}_{FOL}$ , então, para qualquer  $B \in \mathcal{L}_{FOL}$ , e para qualquer  $\delta \in Der(A, BC)$ , podemos adicionar à base de conhecimento  $BC$  as fbfs super-justificadas  $\langle A \vee B, der, \delta, \{A\}, inst(A, \delta) \rangle$ , ou  $\langle B \vee A, der, \delta, \{A\}, inst(A, \delta) \rangle$ , ou ambas.

**Eliminação da disjunção (E $\vee$ )**

Se  $Der(A \vee B, BC) \neq \{\}$ ,  $Der(A \rightarrow C, BC) \neq \{\}$  e  $Der(B \rightarrow C, BC) \neq \{\}$ , então, para cada  $\delta_1 \in Der(A \vee B, BC)$ , para cada  $\delta_2 \in Der(A \rightarrow C, BC)$  e para cada  $\delta_3 \in Der(B \rightarrow C, BC)$ , podemos adicionar à  $BC$  a fbf super-justificada  $\langle C, der, \bigcup(\delta_1, \delta_2, \delta_3), \{A \vee B, A \rightarrow C, B \rightarrow C\}, \bigcup(inst(A \vee B, \delta_1), inst(A \rightarrow C, \delta_2), inst(B \rightarrow C, \delta_3)) \rangle$ .

**Introdução do quantificador universal (I $\forall$ )**

Se  $Der(A(c), BC) \neq \{\}$  e  $A(c) \in \mathcal{L}_{FOL}$ , então, para cada  $\delta \in Der(A(c), BC)$ , tal que  $A(c)$  usa um símbolo individual  $c$  que não ocorre em nenhuma fbf em  $\delta$ , podemos adicionar à base de conhecimento  $BC$  a fbf super-justificada  $\langle \forall(x)[A(x)], der, \delta, \{(\{c\}, A(c))\}, inst(A(c), \delta) \rangle$ .

**Eliminação do quantificador universal (E $\forall$ )**

Se  $Der(\forall(x)[A(x)], BC) \neq \{\}$ , então, para cada  $\delta \in Der(\forall(x)[A(x)], BC)$ , para qualquer símbolo individual  $c$ , podemos adicionar à base de conhecimento  $BC$  a fbf super-justificada  $\langle A(c), der, \delta, \{\forall(x)[A(x)]\}, inst(\forall(x)[A(x)], \delta) \cup \{(\forall(x)[A(x)], c)\} \rangle$ .

**Introdução do quantificador existencial (I $\exists$ )**

Se  $Der(A(c), BC) \neq \{\}$ , em que  $c$  é um símbolo individual, e  $A(c) \in \mathcal{L}_{FOL}$ , então, para cada  $\delta \in Der(A(c), BC)$ , podemos adicionar à base de conhecimento  $BC$  a fbf super-justificada  $\langle \exists(x)[A(x)], der, \delta, \{A(c)\}, inst(A(c), \delta) \rangle$ .

**Eliminação do quantificador existencial (E $\exists$ )**

Se  $Der(\exists(x)[A(x)], BC) \neq \{\}$  e  $Der(B, BC) \neq \{\}$ , então, para cada  $\delta_1 \in Der(\exists(x)[A(x)], BC)$  e para cada  $\delta_2 \in Der(B, BC)$ , tal que  $\delta_2$  contém  $A(c)$  e o símbolo individual  $c$  não ocorre em  $B$  nem em nenhuma fbf em  $(\delta_2 - \{A(c)\})$ , podemos adicionar à base de conhecimento  $BC$  a fbf super-justificada  $\langle B, der, \bigcup(\delta_1, (\delta_2 - \{A(c)\}), \{(\{A(c)\}, B), \exists(x)[A(x)]\}, \bigcup(inst(\exists(x)[A(x)], \delta_1), inst(B, \delta_2)) \rangle$ .

## Regras estendidas

### Suposição e eliminação do quantificador de omissão (Sup-E $\nabla$ )

Se  $Der(D, BC) \neq \{\}$ , onde  $D = \nabla(x)[A(x)]$ , então, para cada  $c$  que seja uma instância de  $x$ , podemos adicionar à base de conhecimento  $BC$  as fbfs super-justificadas  $\langle Applicable(D, c), asp, \{D, Applicable(D, c)\}, \{D, Applicable(D, c)\}, \{\} \rangle^3$  e  $\langle A(c), der, \{D, Applicable(D, c)\}, \{D, Applicable(D, c)\}, \{\} \rangle$ .

### Enfraquecimento do quantificador universal (Enf $\forall$ )

Se  $Der(\forall(x)[A(x)], BC) \neq \{\}$ , então, para cada  $\delta \in Der(\forall(x)[A(x)], BC)$ , podemos adicionar à base de conhecimento  $BC$  a fbf super-justificada  $\langle \nabla(x)[A(x)], hyp, \{\nabla(x)[A(x)]\}, \{\}, \{\} \rangle$ .

---

<sup>3</sup>Sabemos que as regras de omissão são sempre hipóteses, por isso as suas instanciações são sempre vazias.

# Bibliografia

- Alchourrón, C. A., Gärdenfors, P. & Makinson, D. [1985], 'On the logic of theory change: partial meet functions for contraction and revision', *The Journal of Symbolic Logic* 50(2), 510–530.
- Anderson, A. R. & Belnap, N. D. J. [1975], *Entailment, the Logic of Relevance and Necessity*, Princeton University Press, Princeton.
- Brachman, R. J. & Schmolze, J. G. [1985], 'An overview of the KL-ONE knowledge representation system', *Cognitive Science* 9, 171–216.
- Brachman, R. & Levesque, H., eds [1985], *Readings in Knowledge Representation*, Morgan Kaufmann Publishers, Inc., Los Altos, USA.
- Cachopo, J. & Cardoso, A. [1994], Lep: (meta-)lógica de especificação de preferências, Trabalho final de curso, Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisbon, Portugal.
- Cravo, M. R. [1992], Raciocínio por Omissão e Revisão de Crenças: Dois Aspectos do Senso Comum, PhD thesis, Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisbon, Portugal.
- Cravo, M. R. [1993a], A belief revision theory based on SWMC, Technical Report GIA 93/03, Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisbon, Portugal.
- Cravo, M. R. [1993b], SWMC: A logic for default reasoning and belief revision (a new version), Technical Report GIA 93/02, Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisbon, Portugal.
- Cravo, M. R. [1995], BRS: A belief revision system capable of non monotonic reasoning and belief revision, Technical Report GIA 95/01, Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisbon, Portugal.
- Cravo, M. R. & Martins, J. P. [1989], Path-based inference in snebr, in J. P. Martins & E. M. Morgado, eds, 'EPIA 89: Proc. of the 4th Portuguese Conference on Artificial Intelligence', Springer-Verlag, Heidelberg, Germany, pp. 97–106.

- Cravo, M. R. & Martins, J. P. [1990], Path-based inference revisited, in D. Kumar, ed., 'Current Trends in SNePS: Sematic Network Processing System', Springer-Verlag, Heidelberg, Germany, pp. 15–26.
- Cravo, M. R. & Martins, J. P. [1993], 'SNePSwD, a newcomer to the SNePS family', *Journal of Experimental and Theoretical Artificial Intelligence* 5, 135–148.
- de Kleer, J. [1986], 'An assumption-based truth maintenance system', *Artificial Intelligence* 28(2), 127–162.
- Doyle, J. [1979], 'A truth maintenance system', *Artificial Intelligence* 12(3), 231–272.
- Doyle, J. [1983], The ins and outs of reason maintenance, in A. Bundy, ed., 'Proceedings of IJCAI-83', Morgan Kaufmann Publishers, Inc., Karlsruhe, FRG, pp. 349–351.
- Doyle, J. [1991], Rational belief revision, in 'Proceedings of KR-91', Morgan Kaufmann Publishers, Inc., Cambridge, MA, pp. 163–174.
- Doyle, J. [1992], Reason maintenance and belief revision: Foundations vs. coherence theories, in Gärdenfors [1992], pp. 29–51.
- Dressler, O. [1989], An extended basic ATMS, in Reinfrank, de Kleer, Ginsberg & Sandewall, eds, 'Non-Monotonic Reasoning', Vol. 346 of *Lecture Notes in Artificial Intelligence*, Springer-Verlag, Heidelberg, Germany, pp. 143–163.
- Etherington, D. W. [1988], *Reasoning with incomplete information*, Pitman Publishing, London, UK.
- Etherington, D. W. & Reiter, R. [1983], On inheritance hierarchies with exceptions, in 'Proceedings of AAAI-83', Morgan Kaufmann Publishers, Inc., Washington, DC, pp. 104–108.
- Fahlman, S. E. [1979], *NETL: A System for Representing and Using Real-World Knowledge*, MIT Press, Cambridge, MA.
- Fuhrmann, A. [1991], 'Theory contraction through base contraction', *Journal of Philosophical Logic* 20(2), 175–203.
- Gärdenfors, P. [1988], *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, MIT Press, Cambridge, MA.
- Gödel, K. [1965], On formally undecidable propositions of the principia mathematica and related systems., in M. Davis, ed., 'The Undecidable: Basic Papers on Undecidable Propositions, Unsolvability Problems and Computable Functions', Raven Press, Hewlett, New York, USA, pp. 39–71.
- Ginsberg, M. L., ed. [1987], *Readings in Nonmonotonic Reasoning*, 1 edn, Morgan Kaufmann Publishers, Inc., Los Altos, USA.

- Gärdenfors, P., ed. [1992], *Belief Revision*, number 29 in 'Cambridge Tracts in Theoretical Computer Science', Cambridge University Press.
- Harman, G. H. [1986], *Change in View: Principles of Reasoning*, MIT Press, Cambridge, MA.
- Harper, W. L. [1987], Ramsey test conditionals and iterated belief change, in Harper & Hooker, eds, 'Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science', Vol. 1, D. Reidel Publishing Company, Dordrecht, The Netherlands, pp. 117–135.
- Katsuno, H. & Mendelzon, A. O. [1992], On the difference between updating a knowledge base and revising it, in Gärdenfors [1992], pp. 183–203.
- Kuhn, T. S. [1970], *The Structure of Scientific Revolutions*, 2, enlarged edn, University of Chicago Press, Chicago, Illinois.
- Lakatos, I. [1970], Falsification and the methodology of scientific research programmes, in I. Lakatos & A. Musgrave, eds, 'Criticism and the growth of knowledge', Cambridge University Press, Cambridge, MA, pp. 91–196.
- Levi, I. [1977], 'Subjunctives, dispositions, and chances', *Synthese* 34, 423–455.
- Lukaszewicz, W. [1985], Two results on default logic, in 'Proceedings of IJCAI-85', Morgan Kaufmann Publishers, Inc., Los Angeles, USA, pp. 459–461.
- Lukaszewicz, W. [1990], *Non-monotonic reasoning: formalization of commonsense reasoning*, Ellis Horwood, Chichester, England.
- Mamede, N., Martins, J. P. & Cravo, M. R. [1988], The SNeBR inference engine, Technical Report GIA 88/02, Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisbon, Portugal.
- Martins, J. P. [1983], Reasoning in multiple belief spaces, PhD thesis, State University of New York at Buffalo, Buffalo, NY.
- Martins, J. P. [1991], Computational issues in belief revision, in A. Fuhrmann & M. Morreau, eds, 'Proceedings of the Workshop on The Logic of Theory Change', Vol. 465 of *Lecture Notes in Artificial Intelligence*, Springer-Verlag, Heidelberg, Germany, pp. 51–71.
- Matos, P. & Martins, J. P. [1989], SNePSLOG – a logic interface to SNePS, Technical Report GIA 89/03, Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisbon, Portugal.
- McCarthy, J. [1977], Epistemological problems of artificial intelligence, in 'Proceedings of IJCAI-77', Morgan Kaufmann Publishers, Inc., Cambridge, MA, pp. 1038–1044.

- McCarthy, J. [1980], 'Circumscription: A form of nonmonotonic reasoning', *Artificial Intelligence* **13**(1-2), 27-39.
- Minsky, M. [1975], A framework for representing knowledge, in Winston, ed., 'The Psychology of Computer Vision', McGraw-Hill Book Co., New York, USA, pp. 211-277. Also in [Brachman & Levesque 1985, pp. 245-262].
- Moore, R. C. [1983], Semantical considerations on nonmonotonic logic, in 'Proceedings of IJCAI-83', Morgan Kaufmann Publishers, Inc., Karlsruhe, FRG, pp. 272-279.
- Moore, R. C. [1985], 'Semantical considerations on nonmonotonic logic', *Artificial Intelligence* **25**(1), 75-94.
- Nebel, B. [1989], A knowledge level analysis of belief revision, in 'Proceedings of KR-89', Morgan Kaufmann Publishers, Inc., Toronto, Canada, pp. 301-311.
- Nebel, B. [1990], *Reasoning and revision in hybrid representation systems*, Vol. 422 of *Lecture Notes in Artificial Intelligence*, Springer-Verlag, Heidelberg, Germany.
- Newell, A. [1982], 'The knowledge level', *Artificial Intelligence* **18**, 87-127.
- Pinto-Ferreira, C., Mamede, N. & Martins, J. P. [1989], SNIP 2.1 – the SNePS inference package, Technical Report GIA 89/05, Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisbon, Portugal.
- Pinto, H. S. [1992], Formalização da herança em SNePS, Master's thesis, Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisbon, Portugal.
- Poole, D. [1988], 'A logical framework for default reasoning', *Artificial Intelligence* **36**(1), 27-47.
- Quillian, M. R. [1968], Semantic memory, in M. Minsky, ed., 'Semantic Information Processing', MIT Press, Cambridge, MA, pp. 227-270.
- Quine, W. V. O. & Ullian, J. S. [1978], *The Web of Belief*, 2 edn, Random House.
- Reinfrank, M. [1989], Fundamentals and logical foundations of truth maintenance, Linköping studies in science and technology, 221, Linköping University, Linköping, Sweden.
- Reiter, R. [1980], 'A logic for default reasoning', *Artificial Intelligence* **13**(1-2), 81-132.
- Russell, S. J. & Norvig, P. [1995], *Artificial Intelligence: A Modern Approach*, Prentice-Hall, Inc.
- Schmolze, J. G. & Lipkis, T. A. [1983], Classification in the KL-ONE knowledge representation system, in 'Proceedings of IJCAI-83', Morgan Kaufmann Publishers, Inc., Karlsruhe, FRG, pp. 330-332.

- Shapiro, S. C. [1979], The SNePS Semantic Network Processing System, in N. V. Findler, ed., 'Associative Networks: Representation and Use of Knowledge by Computers', Academic Press, New York, USA, pp. 179–203.
- Shapiro, S. C. & Martins, J. P. [1990], Recent advances and developments: The SNePS 2.1 report, in Kumar, ed., 'Current Trends in SNePS - Semantic Network Processing System: Proceedings of the First Annual SNePS Workshop', number 437 in 'Lecture Notes in Artificial Intelligence', Springer-Verlag, Heidelberg, Germany, pp. 1–13.
- Shapiro, S. C., McKay, D. P., Martins, J. P. & Morgado, E. M. [1981], SNePSLOG: A "higher order" logic programming language, SNeRG Technical Note 8, Department of Computer Science, University at Buffalo. Presented at the Workshop on Logic Programming for Intelligent Systems, R.M.S. Queen Mary, Long Beach, CA.
- Shapiro, S. C. & Rapaport, W. J. [1987], SNePS considered as a fully intensional propositional semantic network, in N. Cercone & G. McCalla, eds, 'The Knowledge Frontier', Springer-Verlag, Heidelberg, Germany, pp. 263–315.
- Shapiro, S. C. & Rapaport, W. J. [1992], The SNePS family, in F. Lehman, ed., 'Semantic Networks in Artificial Intelligence', Pergamon, pp. 243–275.
- Shapiro, S., Martins, J. P. & McKay, D. [1982], Bi-directional inference, in 'Proceedings of the Fourth Annual Conference of the Cognitive Science Society', the Program in Cognitive Science of The University of Chicago and The University of Michigan, Ann Arbor, MI, pp. 90–93.
- Shoam, Y. [1987], A semantical approach to nonmonotonic logics, in Ginsberg [1987], pp. 227–250.
- Shoam, Y. [1988], *Reasoning about change: time and causation from the standpoint of Artificial Intelligence*, MIT Press, Cambridge, MA.
- Sneed, J. D. [1971], *The logical structure of mathematical physics*, Dordrecht, The Netherlands, Dordrecht, The Netherlands.
- Stegmüller, W. [1976], *The structure and dynamics of theories*, Springer-Verlag, Heidelberg, Germany.
- Touretzky, D. S. [1984], Implicit ordering of defaults in inheritance systems, in 'Proceedings of AAAI-84', Morgan Kaufmann Publishers, Inc., Austin, USA, pp. 322–325.
- Winograd, T. [1975], Frame representations and the declarative/procedural controversy, in Bobrow & Colins, eds, 'Representation and understanding: studies in Cognitive Science', Academic Press, New York, USA, pp. 185–210.