# LÓGICAS NÃO-MONOTÔNICAS NA FORMALIZAÇÃO DO SENSO-COMUM

João Carlos Pereira da Silva

TESE SUBMENTIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Aprovada por :

Sheila Regina Murgel Veloso

Profa. Sheila Regina Murgel Veloso, D.Sc.
(Presidente)

Prof. Paulo Augusto Silva Veloso, Ph.D.

Marco Antonio Casanova

Prof. Marco Antonio Casanova, Ph.D.

RIO DE JANEIRO, RJ - BRASIL

NOVEMBRO DE 1990

SILVA, JOÃO CARLOS PEREIRA DA

Lógicas Não-Monotônicas na Formalização do Senso-Comum [Rio de Janeiro] 1990

*viii*, 223 p., 29,7 cm (COPPE/UFRJ, M.Sc., Engenharia de Sistemas e Computação, 1990)

Tese-Universidade Federal do Rio de Janeiro, COPPE

1.Raciocínio Não-monotônico 2.Lógica de Default 3.Circunscrição 4.Manutenção da Verdade I.COPPE/UFRJ II.TÍTULO (série)

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para obtenção do grau de Mestre em Ciências (M.Sc.).

LÓGICAS NÃO-MONOTÔNICAS NA FORMALIZAÇÃO DO SENSO-COMUM

João Carlos Pereira da Silva

Novembro de 1990

Orientadora : Profa. Sheila Regina Murgel Veloso

Programa : Engenharia de Sistemas e Computação

*A proposta é examinar problemas que surgem na formalização do raciocínio do senso-comum através de uma linguagem que possua o rigor formal da lógica de predicados mas que tenha propriedades não-monotônicas.*

*Inicialmente, duas formas de formalização do raciocínio não-monotônico são estudadas : a lógica de default e a circunscrição. São apresentados dois programas de apoio a este tipo de raciocínio : o TMS e o ATMS.*

*A seguir são estudadas as dificuldades da utilização efetiva de tais mecanismos de formalização, em especial nas situações que envolvam tempo.*

*Finalizando, são apresentadas as várias propostas de solução, novamente com enfase nas questões temporais.*

Abstract of Thesis presented to COPPE/UFRJ as partial fulfillment of the requirements for the degree of Master of Science (M.Sc.).

NONMONOTONIC LOGICS ON THE FORMALIZATION OF COMMOM-SENSE

João Carlos Pereira da Silva

November, 1990

Thesis Supervisor : Profa. Sheila Regina Murgel Veloso

Department : Systems Engineering and Computing

*The proposal is to investigate problems that arise in the formalization of the common-sense reasoning through a language that has the formal rigor of the predicate logic, but that has nonmonotonic properties.*

*First, two kinds of formalization of the nonmonotonic reasoning are studied : the default logic and the circumscription. In this part , two programs supporting for this kind of reasoning are seen : TMS and ATMS.*

*Then, the difficulties in the effective utilization of such a mechanism of formalization are studied, especially in situations where time is involved.*

*Finally, various solutions are presented, again with emphasis on temporal question.*

*Para Antonio e Jacinta*

## AGRADECIMENTOS

*Gostaria de agradecer à minha orientadora, Profa. Sheila R. M. Veloso, pelo seu trabalho de orientação desta tese; à Leonor , pelos fins de semana perdidos para que eu pudesse usar um computador aos sábados e domingos; e a Anamaria , pelos artigos que me enviou da Inglaterra.*

*Também agradeço à* CAPES *e à* SOCIEDADE CULTURAL E BENEFICENTE GUILHERME GUINLE *pelo apoio financeiro dado a este trabalho.*

CAPÍTULO I : INTRODUÇÃO.................. 1

CAPÍTULO II : LÓGICA DE DEFAULTS......... 6

*i) Definição*.............................. 7

*ii) Teoria Default Normal (TDN)*............ 13

*iii) Teoria Default Semi-Normal (TDSN)*..... 16

*iv) Conclusão*............................. 23

CAPÍTULO III : TMS E ATMS................. 25

*i) TMS*................................... 25

*i.1) Justificativas*...................... 27

*i.2) DDB*................................. 33

*ii) ATMS*................................. 35

*ii.1) Estrutura de Dados*................. 36

*ii.2) Reticulado de Ambiente*............. 40

*ii.3) Operações Básicas*.................. 42

*iii) Conclusão*............................ 44

CAPÍTULO IV : CIRCUNSCRIÇÃO.............. 46

*i) Definição*............................. 47

*ii) Satisfatibilidade*..................... 56

*iii) Circunscrição Priorizada*............. 63

*iv) Circunscrição Ponto a Ponto*........... 69

*iv.1) Prioridade na Circunscrição Ponto a Ponto*..................... 78

*v) Circunscrição Suportada*............... 80

*vi) Conclusão*............................. 85

CAPÍTULO V : TRABALHANDO COM A CIRCUNSCRIÇÃO................. 86

*i) Fórmulas Separáveis* ..................... 86

*i.1) Paralela* ..................... 86

*i.2) Priorizada* ..................... 100

*ii) MILO-Resolução* ..................... 103

*ii.1) Paralela* ..................... 103

*ii.2) Priorizada* ..................... 118

*iii) Provador para Teorias Circunscritas* ..................... 122

*iii.1) Paralela* ..................... 122

*iii.2) Priorizada* ..................... 129

*iv) Conclusão* ..................... 138

CAPÍTULO VI : PROBLEMA DO TIRO ..................... 141

*i) O Problema do Tiro* ..................... 141

*ii) TMS e Defaults Não-Normais* ..................... 158

*iii) Minimização Cronológica* ..................... 170

*iv) Minimização Causal* ..................... 185

*v) Sistema Temporal Não-monotônico* ..................... 190

*vi) Nova Política de Circunscrição* ..................... 199

*vii) Conclusão* ..................... 206

CAPÍTULO VII : CONCLUSÃO ..................... 208

APENDICE A ..................... 211

APENDICE B ..................... 215

REFERÊNCIAS BIBLIOGRÁFICAS ..................... 217

# CAPITULO I

# INTRODUÇÃO

A representação do raciocínio humano através de uma linguagem que apresente um rigor formal tal qual a lógica é uma idéia atraente . Isto porque o formalismo lógico sugere de imediato uma boa maneira de derivar novos conhecimentos a partir de conhecimentos antigos (dedução).

Uma característica importante do raciocínio humano é a capacidade que as pessoas possuem de tomar decisões e chegar a conclusões a partir de um número restrito de fatos. O que acontece é que nossas ações e conclusões não são baseadas somente naquilo que conhecemos ou sabemos, mas também no que desconhecemos.

Utilizar a lógica de primeira-ordem na formalização deste tipo de raciocínio não é possível. Um dos problemas é a necessidade que teríamos de enumerar todas as exceções possíveis quando fatos do tipo "Normalmente ..." ou "Tipicamente ..." tivessem que ser representados. Assim, para representar fatos tais como "Normalmente as aves voam", seria necessário que todas as exceções (as aves que não voam e os problemas que poderiam fazer com que uma ave não voe) ficassem explicitamente expressas nos axiomas.

Mesmo que isto fosse possível, a manipulação pela lógica de primeira-ordem também não é satisfatória. Por exemplo, no caso de representarmos o fato "Normalmente as aves voam" pelo axioma :

$$\forall x(\text{AVE}(x) \wedge \neg\text{PINGÜIM}(x) \wedge \neg\text{MORTO}(x) \wedge \text{TEM-ASA}(x) \wedge \ldots \supset \text{VOA}(x)),$$

onde estamos supondo que todas as exceções que impedem que uma ave voe estão explicitamente expressas no antecedente da implicação anterior.

Sabendo que Tweety é uma ave e não tendo nenhuma informação adicional, pelo senso-comum, concluimos que Tweety voa, uma vez que para nós as aves normalmente voam. O mesmo não ocorre na lógica de primeira-ordem. Para derivar VOA(Tweety) , não basta ter AVE(Tweety). Além disso, é necessário saber que Tweety não é uma exceção entre as aves.

Outro problema apresentado pela lógica clássica é o seu caráter monotônico. Isto significa que todas as conclusões obtidas, a partir de qualquer conjunto de sentenças desta lógica, é preservada quando novos fatos são acrescidos a este conjunto. Neste caso, se concluirmos de um certo conjunto de fatos que Tweety voa e posteriormente aprendermos que ele é um pingüim (e, portanto, não pode voar), a primeira conclusão é preservada.

Isto não corresponde a nossa maneira de raciocinar , pois, por não termos um conhecimento completo a respeito do mundo, sempre existirão conclusões que, por terem sido obtidas com base em especificações incompletas, deverão ser revistas.

Para tratar o raciocínio não-monotônico, uma série de métodos foram desenvolvidos a partir do início da década de 80. Dentre as várias propostas, vamos apresentar duas abordagens que se destacaram : a *lógica de default* e a *circunscrição* .

A *lógica de default* foi definida por REITER(1980). Esta nova lógica busca ampliar o conhecimento incompleto de uma dada teoria de primeira-ordem através de regras defaults. Podemos situar este tipo de abordagem dentro de um grupo que considera a lógica de primeira-ordem como uma linguagem inadequada para formalizar o raciocínio não-monotônico. Por isto, a necessidade de definir uma nova lógica capaz de fazê-lo.

A *circunscrição* foi proposta por McCARTHY(1980). Esta solução pode ser colocada num grupo que não considera a lógica de primeira-ordem como a causadora do problema, mas a maneira como ela é utilizada. Antes de tentar deduzir qualquer coisa de uma base de conhecimento (normalmente incompleta), é necessário que as hipóteses implicitas presentes na base de conhecimento sejam explicitadas. Isto é o que faz o método proposto, através do acréscimo do *axioma de circunscrição*.

A utilização de tais métodos, infelizmente, não é de todo satisfatória. Em 1987, HANKS e McDERMOTT(1987) apresentaram um problema que ficou conhecido como problema do tiro, onde uma situação temporal foi formalizada através desses métodos não-monotônicos, e resultados não-intuitivos foram obtidos.

A motivação para a apresentação de tal problema veio do fato de que até aquele instante HANKS e McDERMOTT(1987) consideravam que os trabalhos realizados na área das lógicas não-monotônicas ou tinham uma preocupação extremamente técnica ou eram feitos , ao contrário, sem uma preocupação com detalhes teóricos.

O objetivo portanto era, a partir de um problema prático (o da projeção temporal) , formalizá-lo nas lógicas não-monotônicas e analisar o comportamento da teoria resultante.

Neste trabalho o que se pretende é dar uma introdução da lógica de default e da circunscrição, apresentar o problema do tiro e o problema da utilização daquelas lógicas na sua formalização, e analisar algumas das tentativas de solução do problema do tiro.

O trabalho está organizado da seguinte forma :

- No capítulo II, é apresentada a lógica de default , e resultados importantes para dois tipos de teorias defaults são vistos ;

- No capítulo III , são apresentados dois tipos de sistemas de manutenção da verdade : o TMS e o ATMS. O primeiro será utilizado no estudo do problema do tiro e o segundo sugere uma implementação para um dos métodos utilizados para determinar quando um fato segue ou não de uma teoria circunscrita.

- No capítulo IV, serão definidas algumas formas de circunscrição , e veremos duas classes de teorias às quais a aplicação do método não acarreta inconsistência;

- No capítulo V, apresentaremos três métodos que nos permitirão determinar quando um fato segue ou não de uma teoria circunscrita ;

- No capítulo VI, o problema do tiro é apresentado e formalizado através da lógica de default e da circunscrição . Várias propostas para solucionar o problema decorrente desta formalização serão analisadas.

# CAPITULO II

## LÓGICA DE DEFAULT

Um dos motivos que dificulta a utilização da lógica tradicional na representação do senso-comum é o fato dela possuir a caracteristica da monotonicidade, isto é , a inclusão de novos axiomas numa dada teoria preserva as conclusões já obtidas .

Como , normalmente , os sistemas não possuem um conhecimento completo dos fatos , a impossibilidade de retratar conclusões antigas diante de novos fatos pode levar a um conjunto inconsistente.

Para tratar este problema, REITER (1980) definiu a lógica de default, que procura através de regras defaults ampliar o conhecimento (incompleto) de uma determinada teoria.

Na seção (i) , definiremos formalmente o que é um default e uma teoria default .

Na seção (ii) , veremos duas importantes propriedades das teorias defaults normais que nos permitirão utilizá-las dentro do espírito a que se propõe a lógica de default .

Na seção (iii) , mostraremos a necessidade da utilização de defaults semi-normais na representação de determinados tipos de conhecimento , e que apenas um certo tipo de teoria default semi-normal apresenta as características encontradas nas teorias defaults normais .

i) *Definição*

A lógica tradicional possui uma característica que torna inadequada a sua utilização na formalização do raciocínio do senso-comum , que é a monotonicidade .

O que ocorre normalmente é que o conhecimento das pessoas é incompleto . Nossas conclusões ou ações estão baseadas em descrições incompletas das situações . Dessa maneira , suposições ( implícitas e explícitas ) são feitas sobre o estado dos fatos que desconhecemos . Tais suposições podem ser revistas mais tarde , pelo aparecimento de novas evidências que as tornam incorretas. Conseqüentemente , as conclusões obtidas que envolvam tais suposições não serão mais válidas .

Dessa forma , no raciocínio do senso-comum , as hipóteses estão baseadas nas evidências e na falta de evidências em contrário , fato este que impede o uso da lógica tradicional, por faltar a esta um modo adequado de tratar a ausência de conhecimento .

A *lógica de default* , definida por REITER (1980) , foi desenvolvida visando tratar este tipo de problema da seguinte maneira : ela amplia a lógica tradicional de primeira-ordem com um mecanismo que assume alguns fatos na falta de um conhecimento específico .

Utilizaremos neste capítulo e nos seguintes , a notação e nomenclatura usuais dos símbolos e conceitos básicos da lógica de primeira-ordem, salientando que :

i) o conectivo $\equiv$ será usado para representar a equivalência. Assim, a fórmula $A \equiv B$ é considerada logicamente válida ;

ii) o símbolo $\Leftrightarrow$ será usado para representar a igualdade sintática. Por exemplo, A $\Leftrightarrow \forall x\ (P(x) \supset Q(x))$ significa que A é da forma $\forall x\ (P(x) \supset Q(x))$.

Uma <u>linguagem de primeira-ordem</u> *L* sobre um alfabeto de primeira-ordem *A* é o conjunto de termos e fórmulas bem formadas (f.b.f.) de primeira-ordem sobre *A* .

Uma f.b.f. é dita <u>*fechada*</u> se, e somente se , ela não possui variáveis livres , ou seja , todas as variáveis da fórmula estão quantificadas .Para qualquer conjunto de f.b.f. $S$ e qualquer f.b.f. $w$ , $S \vdash w$ significa que $w$ é <u>*derivável*</u> num sistema formal de primeira-ordem a partir das premissas $S$ ; e $S \models w$ significa que w é <u>*consequência lógica*</u> de S. Para qualquer conjunto de f.b.f. $S \subseteq L$ , definiremos $Th(S) = \{ w \,/\, w \in L$, $w$ é fechada e $S \models w \}$ que é o fecho sob a relação $\models$ .

Formalmente , os defaults são definidos por :

**DEFINIÇÃO 2.1** *(Default)*

Um <u>*default*</u> é uma expressão composta pelas fórmulas bem formadas (f.b.f.) $\alpha(x)$ , $\beta_1(x)$ , ... , $\beta_m(x)$ e $\gamma(x)$ , da forma :

$$( \alpha(x) : \beta_1(x), \ldots, \beta_m(x) ) \,/\, \gamma(x)$$

onde $\alpha(x)$ é chamado de <u>*pré-requisito*</u> , os $\beta_i$'s ($i$ = 1,...,m) são as <u>*justificativas*</u> e $\gamma(x)$ é o <u>*consequente*</u>.

Quando $\alpha$ , $\beta_i$ e $\gamma$ não possuem variáveis livres , o default é dito *fechado*.

Os defaults são utilizados como regras de inferência e podem ser interpretados da seguinte maneira : se $\alpha$ é um fato conhecido de um dado conjunto W e os $\beta i$'s são fatos consistentes com aqueles de W , então podemos considerar $\gamma$ como um fato conhecido.

Quando num default ( $\alpha(x)$ : $\beta(x)$ ) / $\gamma(x)$ , tivermos $\beta(x) \Leftrightarrow \gamma(x)$ , vamos chamar tal default de *normal* . Caso $\beta(x) \Leftrightarrow \gamma(x) \wedge \psi(x)$ , para algum $\psi(x)$ , dizemos que este default é *semi-normal* . Caso contrário, o default é chamado de *não-normal* .

EXEMPLO 2.2

Os defaults :

$$D_1 \Leftrightarrow ( \text{AVE}(x) : \text{VOA}(x) ) / \text{VOA}(x) ;$$

$$D_2 \Leftrightarrow ( \text{AVE}(x) : (\neg\text{PINGÜIM}(x) \wedge \text{VOA}(x)) ) / \text{VOA}(x) ;$$

$$D_3 \Leftrightarrow ( \text{AVE}(x) : \neg\text{VOA}(x) ) / \text{PINGÜIM}(x) ;$$

são , respectivamente, defaults normal , semi-normal e não-normal ■

DEFINIÇÃO 2.3 *(Teoria Default)*

Uma *teoria default* é um par $\Delta = (D,W)$ onde D é um conjunto de defaults e W é um conjunto de f.b.f. . Quando todos os defaults de D são fechados , dizemos que a teoria $\Delta = (D,W)$ é fechada.

**DEFINIÇÃO 2.4** *(Teorias Normal, Semi-normal e Não-normal)*

Uma teoria default $\Delta = (D,W)$ é dita *normal* (*semi-normal* ; *não-normal*) se , e somente se , todos os defaults de D são normais (semi-normais ; não-normais).

**EXEMPLO 2.5**

Considerando $D_1$ , $D_2$ e $D_3$ do exemplo 2.2 e $W = \{\ \}$ , temos as seguintes teorias defaults , respectivamente , normal , semi-normal e não-normal : $\Delta_1 = (D_1,W)$ , $\Delta_2 = (D_2,W)$ , $\Delta_3 = (D_3,W)$ ∎

A função dos defaults de uma dada teoria $\Delta = (D,W)$ é estender o conhecimento que se tem a cerca de um dado mundo W, uma vez que,em geral, tal conjunto é incompleto . Assim, cada uma dessas extensões pode ser vista como um mundo no qual mais fatos são conhecidos .

Formalmente , as extensões são definidas da seguinte forma :

**DEFINIÇÃO 2.6** *(Extensão de uma Teoria Default)*

Seja $\Delta = (D,W)$ uma teoria default. Para qualquer conjunto de sentenças de primeira-ordem S , defina $\Gamma(S)$ como sendo o menor conjunto que satisfaz as seguintes condições :

(a) $W \subseteq \Gamma(S)$ ;

(b) $\Gamma(S) = Th(\Gamma(S))$ ;

(c) Para cada default $[(\ \alpha(x) : \beta_1(x),\dots,\beta_m(x)\ ) / \gamma(x)] \in D$ , se $\alpha(x) \in \Gamma(S)$ e $\neg\beta_1(x),\dots,\neg\beta_m(x) \notin S$

então $\gamma(x) \in \Gamma(S)$ .

Então , E é definido como sendo uma *extensão* da teoria default $\Delta = (D,W)$ se , e somente se, $\Gamma(E) = E$ .

**EXEMPLO 2.7**

A teoria default $\Delta = (D,W)$ , onde :

$$D = \{ (:\neg A) / \neg A \ , \ (:\neg B) / \neg B \} \text{ e } W = \{ A \vee B \}$$

possui duas extensões , que são :

$$E_1 = Th(W \cup \{\neg A\}) \text{ e } E_2 = Th(W \cup \{\neg B\}) \quad \blacksquare$$

As extensões podem ser caracterizadas pelo seguinte teorema :

**TEOREMA 2.8** *(REITER (1980))*

Sejam E um conjunto de f.b.f. e $\Delta = (D,W)$ uma teoria default . Defina :

$$E_0 = W ;$$

e para $i \geq 0$ , $E_{i+1} = Th(E_i) \cup \{ \gamma(x) \ / \ [( \alpha(x) : \beta_1(x), \ldots, \beta_m(x) ) / \gamma(x)] \in D$ onde $\alpha(x) \in E_i$ e $\neg\beta_1, \ldots, \neg\beta_m \notin E \}$ .

Então , E é uma extensão para $\Delta$ se, e somente se, $E = \bigcup_{i=0}^{\infty} E_i$ .

EXEMPLO 2.9

A teoria default $\Delta = ( \{(:A) / \neg A\} , \{ \} )$ não possui nenhuma extensão.

Uma possível extensão E desta teoria teria $\neg A \in E$. Pelo teorema 2.8, $E = \bigcup_{i=0}^{\infty} E_i$ tal que :

$$E_0 = W ;$$

$E_1 = Th(E_0) \cup \{ \gamma / [( \alpha : \beta) / \gamma] \in D$ onde $\alpha \in E_0$ e $\neg\beta \notin E \}$.

Como $\neg A \in E$, o segundo conjunto da união acima é vazio . Logo, a teoria $\Delta$ não possui nenhuma extensão com $\neg A$, nem sem $\neg A$.

O mesmo ocorre na teoria $\Delta' = (D', W)$ onde :

$$D' = \{ :(A \wedge \neg B) / A ; :(B \wedge \neg C) / B ; :(C \wedge \neg A) / C \} .$$

Note que $Th(\{A\})$ não pode ser uma extensão desta teoria pois , pelo teorema 2.8 temos :

$$E_0 = W ;$$

$$E_1 = Th(E_0) \cup \{A\}, \text{ pois } \neg A \vee B \notin Th(\{A\}) ;$$

$$E_2 = Th(E_1) \cup \{B\}, \text{ pois } \neg B \vee C \notin Th(\{A\}) ;$$

Assim, obrigatoriamente teríamos que aplicar o default $:(B \wedge \neg C) / B$ , impedindo que $Th(\{A\})$ seja uma extensão . O mesmo fato ocorre para $Th(\{B\})$ e $Th(\{C\})$.

Também $Th(\{A,B\})$ não poderia ser um extensão de $\Delta'$ pois neste caso $B \in Th(\{A,B\})$, o que acarreta que $B \vee \neg A \in Th(\{A,B\})$. Isto inibe a aplicação do default $:(A \wedge \neg B) / A$ uma vez que para aplicá-lo $\neg A \vee B \notin Th(\{A,B\})$. Pelo mesmo motivo Th({B,C}) e Th({A,C}) não podem ser uma extensão de $\Delta'$.

Logo, a teoria $\Delta'$ também não possui nenhuma extensão. As teorias defaults que não têm extensões são ditas <u>incoerentes</u> ■

As teorias defaults que têm uma extensão inconsistente são aquelas que possuem o conjunto de todas as fórmula bem formadas da linguagem L como uma de suas extensões . Dessa forma , temos que :

**TEOREMA 2.10** *(REITER (1980))*

Uma teoria default $\Delta = (D,W)$ possui uma extensão inconsistente se, e somente se, W for inconsistente .

*ii) <u>Teoria Default Normal (T.D.N.)</u>*

Como foi visto no exemplo 2.9 , existem teorias defaults que não possuem extensões . Obviamente , tais teorias não são atraentes para o objetivo ao qual se propõe a lógica de default . Assim , seria interessante caracterizar quais teorias defaults asseguram a existência de pelo menos uma extensão.

As teorias defaults normais (T.D.N.) fazem parte da classe de teorias bem-comportadas , no sentido da garantia de que existe , pelo menos ,uma extensão . Isto se deve ao fato de que os defaults normais não introduzem inconsistências, não refutam as suas próprias justificativas e nem a de outros já aplicados. Assim, temos :

**TEOREMA 2.11** *(REITER (1980))*

Toda T.D.N. possui uma extensão.

Outra importante propriedade das T.D.N. é a <u>semi-monotonicidade</u>, estabelecida pelo seguinte teorema :

**TEOREMA 2.12** *(REITER (1980))*

Suponha que D e D' são conjuntos de defaults normais com $D' \subseteq D$. Seja E' uma extensão para a teoria default normal $\Delta' = (D', W)$ e seja $\Delta = (D, W)$. Então $\Delta$ possui uma extensão E tal que $E' \subseteq E$ e $GD(E', \Delta') \subseteq GD(E, \Delta)$, onde $GD(E, \Delta) = \{ [( \alpha : \beta_1, \ldots, \beta_m ) / \gamma] \in D \ / \ \alpha \in E$ e $\neg\beta_1, \ldots, \neg\beta_m \notin E \}$.

A importância deste resultado reside no fato de que ele torna possível uma teoria de prova na qual as provas podem ser construídas levando-se em conta apenas alguns dos defaults do conjunto D. REITER (1980) desenvolveu tal teoria de prova para a classe dos defaults normais. Formalmente, a sua tarefa é determinar para uma dada teoria default normal $\Delta$ e uma f.b.f. $\psi$, se $\Delta$ possui uma extensão E tal que $\psi \in E$.

**DEFINIÇÃO 2.13** *(Prova Default)*

Sejam $\Delta = (D,W)$ uma T.D.N. e $\psi$ uma f.b.f. . Uma sequência finita $D_0,\ldots,D_k$ de subconjuntos finitos de D é uma *prova default* de $\psi$ com respeito a $\Delta$ se, e somente se:

(1) $W \cup CONSEQ(D_0) \vdash \psi$ , onde $CONSEQ(D_0)$ é o conjunto dos consequentes dos defaults de $D_0$ ;

(2) Para $1 \leq i \leq k$ , $W \cup CONSEQ(D_i) \vdash PREC(D_{i-1})$ , onde $PREC(D_{i-1})$ é a conjunção de todos os pré-requisitos dos defaults de $D_{i-1}$ ;

(3) $D_k = \{\ \}$ ;

(4) $W \cup \bigcup_{i=0}^{k} CONSEQ(D_i)$ é satisfatível .

**TEOREMA 2.14** *(Completude e Correção - REITER (1980))*

Seja $\psi$ uma f.b.f. . Uma T.D.N. $\Delta = (D,W)$ possui uma extensão E tal que $\beta \in E$ se, e somente se, $\beta$ tem uma prova default com respeito a $\Delta$ .

Por exemplo, para a teoria default normal $\Delta = (D,W)$, onde $D = \{ :A/A \ ;\ :\neg A/\neg A \ ;\ A:\neg B/\neg B \ ;\ C:F/F \}$ e $W = \{A, B \vee C\}$, uma prova default de $F \in E = Th(W \cup \{A,\neg B,F\})$ é dada pela seguinte sequência de defaults :

$D_0$ : $\{C:F/F\}$ ,

$D_1$ : $\{A:\neg B/\neg B\}$ ,

$D_2$ : $\{:A/A\}$ ,

$D_3$ : $\{\ \}$.

Note que $W \cup \{F,\neg B,A\}$ é satisfatível e as condições (1) e (2) são satisfeitas :

$$W \cup \{F\} \vdash F$$

$$W \cup \{\neg B\} \vdash C$$

$$W \cup \{A\} \vdash A$$

$$W \vdash true$$

iii) *Teoria Default Semi-Normal* (*T. D. SN.*)

No seu estudo inicial sobre lógica de default , REITER (1980) observou que desconhecia qualquer ocorrência natural de default que não pudesse ser representada através de defaults normais .

Posteriormente , ele e Criscuolo ( REITER e CRISCUOLO (1981) ) observaram que certos tipos de conhecimento não podiam ser completamente caracterizados por defaults normais. Isto porque os defaults interagem uns com os outros , ou com outras f.b.f. , de tal maneira que são obtidos resultados absolutamente indesejados.

Um exemplo dessa interação pode ser vista a seguir :

EXEMPLO 2.15

Considere as seguintes sentenças :

i) Todo pingüim é uma ave ;
ii) Normalmente as aves voam ;
iii) Normalmente os pingüins não voam ;

que podem ser formalizadas pela seguinte teoria default normal $\Delta = (D,W)$ :

$W = \{ \forall x ( PINGÜIM(x) \supset AVE(x) ) \}$ e

$D = \{ (AVE(x) : VOA(x)) / VOA(x) ; (PINGÜIM(x) : \neg VOA(x)) / \neg VOA(x) \}$

Tomando Tweety como uma instância de pingüim , PINGÜIM(Tweety) , obtemos duas extensões para a teoria acima : uma, intuitiva, na qual Tweety não voa; outra, onde Tweety voa.

Com vistas a eliminar a extensão não intuitiva , ou seja , aquela onde Tweety voa , Reiter e Criscuolo propuseram substituir o default :

$$(AVE(x) : VOA(x)) / VOA(x)$$

por :

$$(AVE(x) : (\neg PINGÜIM(x) \wedge VOA(x))) / VOA(x).$$

Assim , na nova teoria , onde :

$$D = \{ (AVE(x) : (\neg PINGÜIM(x) \wedge VOA(x))) / VOA(x) ; (PINGÜIM(x) : \neg VOA(x)) / \neg VOA(x) \} ,$$

obtemos somente uma extensão , justamente a que desejamos.

Neste exemplo , pudemos observar a interação ocorrida entre os defaults normais e uma f.b.f. . Exemplos de interações entre defaults podem ser encontrados em REITER e CRISCUOLO (1981) ■

Em todos os exemplos estudados , observou-se que os defaults que solucionavam os problemas surgidos com a interação entre defaults tinham a forma semi-normal. Assim, nesses casos , passando a trabalhar com teorias semi-normais , estaríamos nos restringindo à teorias bem comportadas , no sentido de termos as interações sob controle. Foi isso que fizemos no exemplo 2.15 , substituindo o default :

$$(AVE(x) : VOA(x)) \,/\, VOA(x)$$

pelo default :

$$(AVE(x) : (\neg PINGÜIM(x) \wedge VOA(x))) \,/\, VOA(x)$$

Apesar dos defaults normais serem um caso especial dos defaults semi-normais ( $\psi(x) \Leftrightarrow$ true , na definição de default semi-normal), a classe das teorias defaults semi-normais infelizmente não apresenta duas propriedades importantes garantidas às teorias normais , que são : (1) a de que toda teoria possui pelo menos uma extensão ; (2) a semi-monotonicidade .

**EXEMPLO 2.16**

A teoria default semi-normal $\Delta = (D',W)$ , onde :

$D' = \{ :(A \wedge \neg B) / A \,;\, :(B \wedge \neg C) / B \,;\, :(C \wedge \neg A) / C \}$
e $W = \{ \}$ , não possui nenhuma extensão (exemplo 2.9).

Já a teoria semi-normal $\Delta = (D,W)$ , onde :

$$D = \{ :(A \wedge B) / B \} ,$$

possui uma extensão $E_1 = Th(B)$ .

Quando acrescentamos o default $:\neg A / \neg A$ à D, obtemos uma nova teoria $\Delta'' = (D'',W)$ que possui uma única extensão $E'_1 = Th(\neg A)$ e não temos $Th(B) \subseteq Th(\neg A)$ , ou seja, a teoria semi-normal acima não é semi-monotônica ■

Ainda em seu trabalho, onde discutiram a necessidade das teorias semi-normais , Reiter e Criscuolo observaram que , em alguns casos, era possível retornar à uma teoria normal onde as interações entre defaults continuariam sob controle, devolvendo a estas teorias as propriedades (1) e (2) acima mencionadas .

Por outro lado , ETHERINGTON (1987) observou que as teorias semi-normais que são incoerentes ( que não possuem extensões ) apresentam uma característica especial : " a necessidade de que as extensões sejam fechadas sob as regras defaults força a aplicação de defaults cujos consequentes levam a contradição das justificativas de outros defaults já aplicados " .

Este é o caso da teoria $\Delta = (D,W)$ do exemplo 2.9, onde A é inconsistente com E , e portanto E não pode conter $\neg A$ .

Também no caso da teoria default $\Delta' = (D',W)$ do exemplo 2.9 , ocorre o mesmo . Aplicando qualquer um dos defaults do conjunto D , ainda é possível aplicar um outro. Entretanto , quaisquer dois defaults que sejam aplicados resulta na rejeição da justificativa do outro default . Exemplificando , podemos considerar o default :

$$:(A \wedge \neg B) \;/\; A \qquad (*)$$

que ainda nos permite a aplicação do default $:(B \wedge \neg C) / B$, uma vez que $:(C \wedge \neg A) / C$ não pode ser aplicado pois tem a sua justificativa inconsistente com (*).

Note que o default $:(B \wedge \neg C) / B$ não pode ser aplicado pois, como vimos no exemplo 2.9, torna a justificativa de (*) inconsistente com o seu consequente.

Como meio de identificar tal tipo de circularidade, que impede as teorias defaults semi-normais de possuirem extensões , ETHERINGTON (1987) definiu um método que determina a existência da mesma.

Observe que a justificativa e o consequente dos defaults normais são identicos , e nenhum default que seja aplicado pode refutar as justificativas dos que já tenham sido aplicados . Logo , qualquer default normal capaz de refutar aquelas justificativas é inaplicável . Isto acarreta que a parte da justificativa que diferencia os defaults normais e semi-normais está totalmente envolvida no processo de tornar uma teoria incoerente , uma vez que ela pode ser refutada.

O conflito entre a obrigatoriedade da aplicação de um default e a consistência das justificativas nas teorias semi-normais pode ocorrer somente se alguma fórmula depende da falta de outra e , ao mesmo tempo , serve como suporte a inferência dessa fórmula . Na teoria do exemplo 2.16 , A depende da falta de B , B da falta de C e C da falta de A . Donde, se inferirmos A , bloquearíamos a inferência C, o que possibilitaria inferir B . Isso invalidaria a inferência A (o mesmo se passa com B e C ).

As teorias defaults semi-normais que através do método proposto não apresentam este tipo de circularidade são ditas <u>ordenadas</u> . Como consequência , se uma T.D.SN. é ordenada , então ela possui pelo menos uma extensão .

OBSERVAÇÃO 1 : A definição do método considera que as fórmulas estão na forma de cláusulas. Porém, o método se aplica a fórmulas de primeira-ordem arbitrárias , embora a determinação das dependências seja mais complicada .

DEFINIÇÃO 2.17 *(Relações $\ll$ e $\underline{\ll}$)*

Considere a teoria default semi-normal fechada $\Delta$ = $(D,W)$ . Sem perda de generalidade , vamos considerar que todas as fórmulas estão na forma de cláusulas . As relações binárias $\ll$ e $\underline{\ll}$ sobre *literais* , são definidas da seguinte maneira :

(1) Seja $\alpha \Leftrightarrow (\alpha_1 \vee \ldots \vee \alpha_n) \in W$, para algum $n \geq 1$. Para todo $\alpha_i, \alpha_j$ que ocorre em $\alpha$, se $\alpha_i \neq \alpha_j$ , então $\neg\alpha_i \underline{\ll} \alpha_j$ ;

(2) Seja $\delta \in D$, com $\delta \Leftrightarrow (\alpha : (\beta \wedge \gamma)) / \beta$. Sejam $\alpha_1, \ldots, \alpha_r$ , $\beta_1, \ldots, \beta_s$ , $\gamma_1, \ldots, \gamma_t$ , os literais que formam, respectivamente , as cláusulas $\alpha$ , $\beta$ e $\gamma$. Então :

i) $\alpha_i \underline{\ll} \beta_j$ , para todo $\alpha_i \in \{\alpha_1,\ldots,\alpha_r\}$ e $\beta_j \in \{\beta_1,\ldots,\beta_s\}$ ;

ii) $\neg\gamma_i \ll \beta_j$ , para todo $\gamma_i \in \{\gamma_1,\ldots,\gamma_t\}$ , $\beta_j \in \{\beta_1,\ldots,\beta_s\}$ e $\gamma_i \notin \{\beta_1,\ldots,\beta_s\}$ ;

iii) Seja $\beta \Leftrightarrow \beta_1 \wedge \ldots \wedge \beta_m$ , para algum $m \geq 1$ ,onde para cada $i \leq m$, $\beta_i \Leftrightarrow (\beta_{i,1} \vee \ldots \vee \beta_{i,mi})$ , e $mi \geq 1$. $\neg\beta_{i,j} \underline{\ll} \beta_{i,k}$, para cada $\beta_{i,j}$ , $\beta_{i,k}$ ocorrendo em $\beta_m$ com $\beta_{i,j} \neq \beta_{i,k}$.

(3) As relações $\underline{\ll}$ e $\ll$ são transitivas :

i) Se $\theta \underline{\ll} \omega$ e $\omega \underline{\ll} \varepsilon$ então $\theta \underline{\ll} \varepsilon$

ii) Se $\theta \ll \omega$ e $\omega \ll \varepsilon$ então $\theta \ll \varepsilon$

iii) Se ( $\theta \ll \omega$ e $\omega \underline{\ll} \varepsilon$ ) ou ( $\theta \underline{\ll} \omega$ e $\omega \ll \varepsilon$ ) então $\theta \ll \varepsilon$ .

DEFINIÇÃO 2.18 *(Ordenação)*

Uma teoria default semi-normal é dita <u>ordenada</u> se, e somente se , não existe um literal $\xi$ tal que $\xi \ll \xi$ .

Para a classe de teorias defaults semi-normais ordenadas temos então a garantia da existência de pelo menos uma extensão :

TEOREMA 2.19 *(ETHERINGTON (1987))*

Se uma teoria default semi-normal é ordenada , então ela possui pelo menos uma extensão .

EXEMPLO 2.20

No caso da teoria default semi-normal $\Delta = (D,W)$ do exemplo 2.16 , ela não é ordenada pois :

$$B \ll A \text{ , } C \ll B \text{ e } A \ll C \text{ ,}$$

donde, por transitividade : $A \ll A$ .

Para a teoria semi-normal $\Delta' = (D',W')$, onde :

$D' = \{$AVE(Tweety) : (¬PINGÜIM(Tweety) ∧ VOA(Tweety)) / VOA(Tweety) ;
PINGÜIM(Tweety) : (¬VOA(Tweety) ∧ true) / ¬VOA(Tweety) $\}$ , e

$W = \{ \forall x$ (PINGÜIM(x) ⊃ AVE(x)) $\}$ ,

não existe nenhum $\xi$ tal que $\xi \ll \xi$ , o que garante a existência de pelo menos uma extensão, como vimos no exemplo 2.15 ■

*iv) Conclusão*

Do ponto de vista da análise do problema da formalização do raciocínio do senso-comum através de métodos não-monotônicos, principalmente no escopo da lógica de defaults, os trabalho de REITER e CRISCUOLO(1981) e de ETHERINGTON(1987) se destacam.

Principalmente o primeiro, que pode ser considerado o precursor do trabalho desenvolvido em HANKS e McDERMOTT(1987) , no sentido da percepção de que a lógica de default (em especial as teorias defaults normais) não trata de maneira adequada certos tipos de conhecimentos.

Como veremos no capítulo VI , a sugestão de trabalhar com teorias defaults semi-normais feitas por REITER e CRISCUOLO(1981) não resolve definitivamente o problema do tiro. Na verdade, mesmo defaults não-normais não o solucionam . Ambos os tipos de teorias defaults apresentam problemas quando novos fatos são acrescidos ao conjunto W.

# CAPÍTULO III

## SISTEMAS DE MANUTENÇÃO DA VERDADE - TMS e ATMS

Neste capítulo , vamos apresentar dois programas que servem como apoio à outro na manipulação do raciocínio não-monotônico.

Basicamente , estes dois programas trabalham na manutenção da consistência entre as várias declarações de um sistema, embora o façam de perspectivas distintas.

Na seção (i) , apresentaremos o TMS , que a partir de um conjunto de fatos e justificativas determina um subconjunto dito de crenças correntes. Tal conjunto é revisto quando novos fatos são apresentados ao TMS.

Na seção (ii) , apresentamos o ATMS, que para cada fato que lhe é transmitido, determina os conjuntos mínimos a partir dos quais ele pode ser derivado.

### *i) TMS (Truth Maintenance System)*

O Sistema de Manutenção da Verdade (TMS - Truth Maintenance System) de DOYLE (1979) é um programa de apoio ao raciocínio não-monotônico . Ele trabalha na manutenção da consistência entre declarações geradas por outros provadores de teorema , sem gerar novos fatos . Quando uma inconsistência é encontrada , o TMS aciona um mecanismo próprio que resolve tal inconsistência através da alteração de um conjunto mínimo de crenças .

No TMS , cada declaração ou regra é representada por um nó . Cada nó estará sempre em um dos dois estados :

(a) IN : quando se acreditar que o nó é verdadeiro. Os nós nesse estado são chamados de *válidos* ;

(b) OUT : quando não se acredita que o nó é verdadeiro. Isto ocorre quando não existirem razões para acreditá-lo ou quando algumas das razões que nos levariam a acreditar no nó não são verdadeiras naquele instante .

Um *rótulo* TMS é uma atribuição de estados IN/OUT à cada um dos nós de um dado conjunto.

Os nós que estão no estado IN , formam um conjunto que é chamado de *crenças correntes* . Tal conjunto é determinado a partir de um conjunto de justificativas , que representam as razões (ou seja, outros nós) para se acreditar em um determinado nó. Portanto , o TMS acredita em um nó se este possuir um argumento que faça com que o TMS o aceite.

No TMS , os argumentos para um nó não se baseiam somente na crença corrente em outros nós , mas também na falta de crença . Este tipo de justificativa é que dá ao TMS a característica da não-monotonicidade .

EXEMPLO 3.1

Considere que as declarações "Tweety é uma ave" e "Tweety não é uma ave" sejam , respectivamente , representados pelos nós $n_1$ e $n_2$. Uma forma de justificar $n_1$

dentro de um conjunto de crenças, poderia estar baseada na falta de crença no nó $n_2$, isto é, acredita-se que Tweety é uma ave porque nada nos leva a crer que ele não o seja. Assim , o nó $n_1$ pertenceria ao conjunto de crenças correntes (estado IN) enquanto o nó $n_2$ não pertencesse (estado OUT) ■

Para construir o conjunto de crenças correntes é preciso um tipo fundamental de crença que permita estabelecer todas as outras . Existem dois tipos fundamentais de crença que são : (1) as *premissas* , que não dependem da crença corrente, ou da sua falta, em qualquer outro nó ; e (2) as *suposições* , que são crenças correntes cujas razões dependem da falta de crença em outros nós. A partir dessas crenças fundamentais e das justificativas dos outros nós , é construído o conjunto de crenças correntes.

A seguir , vamos examinar os dois principais tipos de justificativas do TMS, que nos permitirão construir o conjunto de crenças correntes.

### *i.1) Justificativas*

As justificativas de cada nó são fornecidas ao TMS pelo programa principal , e os seus dois principais tipos são : SL (Support List) e CP (Conditional Proof) .

Uma justificativa SL é da forma :

(SL (*lista in*) (*lista out*))

e é <u>*válida*</u> ( o nó que ela justifica está no estado IN ) se, e somente se , cada um dos nós de sua *lista* in estiver no estado IN e cada nó da sua *lista* out estiver no estado OUT .

As premissas são definidas por justificativas SL onde as *listas* in e out são vazias, indicando que os nós que possuem este tipo de justificativa não depende de crenças atuais , nem da sua ausência . As premissas são sempre válidas e portanto , os nós que elas justificam estarão sempre no estado IN .

As suposições são representadas por justificativas SL onde a *lista* out é não-vazia. Elas podem ser interpretadas da seguinte maneira : os nós da *lista* in abrangem as razões que levam o TMS a assumir o nó justificado ; os nós da *lista* out representam o critério específico que autoriza tal suposição .

Um rótulo TMS de um conjunto de nós deve ser <u>bem-fundamentado</u>, ou seja, todos os nós no estado IN devem ser <u>bem-justificados</u>, o que significa ter todos os seus nós da *lista* out no estado OUT e ter os da sua *lista* in bem-justificados. Logo, um conjunto de nós pode ter vários rótulos mas o TMS só considera aqueles bem-fundamentados.

**EXEMPLO 3.2**

Considere os seguintes nós com as respectivas justificativas SL :

| | | |
|---|---|---|
| (1) | Tweety é uma ave | (SL ( ) ( )) |
| (2) | Tweety voa | (SL (1) ( )) |

Por definição , o nó (1) é uma premissa pois não depende da crença , nem da falta de crença , em qualquer outro nó . Na justificativa SL do nó (2) , a *lista* in contém o nó (1) . Isso indica que a cadeia de raciocínio que leva a conclusão que o nó (2) deve ser acreditado depende da crença em (1). Se , em algum momento mais adiante, o TMS precisar remover a premissa representada pelo nó (1) , deverá remover também o nó (2) do conjunto de crenças correntes .

Considere que a justificativa SL do nó (2) fosse :

(SL (1) (3))                (*)

Agora, o nó (2) passa a ser uma suposição , onde o nó (3) representa a declaração "Tweety é um pingüim" . Além disso, vamos considerar que o nó (3) não possui qualquer justificativa. Nesse caso, o estado do nó (2) seria IN se o nó (1) estivesse no estado IN e o nó (3) estivesse no estado OUT .

Podemos interpretar a justificativa (*) da seguinte maneira: "Se Tweety é uma ave e não existe evidência de que ele seja um pingüim , então conclua que Tweety voa" . Assim, caso a *lista* out de uma justificativa SL seja não-vazia , o TMS pode manipular o raciocínio default .

Caso apareça mais tarde uma evidência de que Tweety é um pingüim ( e forneça uma justificativa para o nó (3) ), então o TMS passará o estado do nó (2) para OUT, já que a sua justificativa não será mais válida ■

O *conjunto de suporte de um nó* é formado pelos nós que o TMS usa para determinar o status (estado IN ou OUT) do nó. Quando o nó está IN , o conjunto de suporte é formado pelos nós das *listas* in e out ; quando o nó está OUT , o conjunto suporte é composto ou por um nó OUT da *lista* in ou por um nó IN da *lista* out.

O conjunto de *consequências de um nó n* é formado pelos nós onde n aparece ou na *lista* in ou na *lista* out. O conjunto de *consequências-afetadas por um nó n* é formado pelos nós n' do conjunto de consequências de n tais que n pertence ao conjunto de suporte do nó n' . O conjunto de *repercussões de um nó* é o fecho transitivo das consequências afetadas do nó.

Vamos chamar de *antecedentes* de um nó no estado IN , todos os nós que formam as listas in e out , e o seu *fundamento* será o fecho transitivo dos antecedentes , ou seja , os antecedentes, os antecedentes dos antecedentes , e assim por diante .

**EXEMPLO 3.3**

Para fixar as definições acima, considere o seguinte conjunto de nós com suas respectivas justificativas SL (a), estados (b), conjunto de suporte (c), consequências (d), consequências afetadas (e) e repercussões (f):

| | (a) | (b) | (c) | (d) | (e) | (f) |
|---|---|---|---|---|---|---|
| 1 | (3) ( ) | out | 3 | 2,3 | 2,3 | 1,2,3,4,6 |
| 2 | ( ) (1) | in | 1 | 4 | 4 | 4 |
| 3 | (1) ( ) | out | 1 | 1,4,6 | 1,6 | 1,2,3,4,6 |
| 4 | (2) ( ) | in | 2 | - | - | - |
| 5 | ( ) ( ) | in | - | 6 | - | - |
| 6 | (3,5) ( ) | out | 3 | - | - | - |

■

OBSERVAÇÃO 1 : As justificativas não são criadas pelo TMS. Elas devem ser fornecidas pelo programa que usa o TMS. Por sua vez, o TMS utiliza as justificativas para manter um banco de dados consistente de crenças.

Uma justificativa CP é da forma :

(CP <*consequente*> (*hipótese* in) (*hipótese* out) )

e representa um argumento hipotético , sendo *válido* se o nó *consequente* está no estado IN sempre que : (a) cada nó da *hipótese* in está no estado IN ; e (2) cada nó da *hipótese* out está no estado OUT .

Ao invés de trabalhar com este tipo de justificativa, o TMS trabalha com justificativas SL equivalentes a estas. Isto porque para determinar a validade de uma justificativa CP (exceto quando o consequente e as *hipóteses* in estão no estado IN e as *hipóteses* out estão no estado OUT) é necessário mudar os estados dos nós das *hipóteses* in e out, de modo a "construir" a situação hipotética na qual a validade de CP possa ser avaliada. Isto pode requerer do processo de manutenção da verdade , a verificação de mais justificativas CP, tornando o processo muito complexo.

O procedimento que o TMS usa para transformar uma justificativa CP da forma (CP C (IH) (OH)) para uma justificativa SL equivalente encontra todos os nós N no fundamento do consequente C de uma justificativa CP tal que:

(1) N não é uma das hipóteses nem uma de suas <u>repercussões</u>;

(2) N é ou um antecedente do consequente ou um antecedente de algum outro nó que pertence as repercussões das hipóteses. Quando N está no estado IN ele formará a *lista* in da justificativa SL equivalente . Caso N esteja no estado OUT , ele formará na *lista* out .

A seguir temos um exemplo de como o TMS determina Justificativas SL equivalentes à CP :

**EXEMPLO 3.4**

Vamos considerar os seguintes nós e suas respectivas justificativas :

| NÓS | JUSTIFICATIVAS | ESTADO INICIAL |
|---|---|---|
| 1 | SL ( ) (2) | in |
| 2 | SL ( ) (8) | out |
| 3 | SL ( ) (4) | out |
| 4 | SL (6,1) ( ) | in |
| 5 | SL (1,3) ( ) | out |
| 6 | SL ( ) ( ) | in |
| 7 | SL (4) ( ) | in |
| 8 | CP 7 (1) ( ) | in |

O TMS inicialmente determina o fundamento do nó consequente da justificativa CP, que é o nó 7. O fundamento de 7 é o conjunto formado pelos nós :

{1,2,4,6}

A condição (1) acima elimina desse conjunto os nós 1 (*hipótese in* de CP) e 4 (pertence a repercussão do nó 1)

Dos nós que restaram ( 2 e 6 ) , somente 6 satisfaz a condição (2), pois ele é um antecedente do nó 4, que como vimos acima pertence à repercussão do nó 1 (*hipótese in*). Como 6 está no estado IN, a justificativas SL equivalente é da forma :

SL (6) ( ) ■

No apêndice A se encontra o algorítmo correspondente ao procedimento acima.

*i.2) <u>DEPENDENCY-DIRECTED BACKTRACKING (DDB)</u>*

O TMS possui um mecanismo próprio , chamado DDB (Dependency-Directed Backtracking), que permite retirar algumas das crenças correntes (nós que estão no estado IN) quando é encontrada uma inconsistência dentro desse conjunto. Assim, o DDB transforma um nó contradição que está no estado IN para o estado OUT , pela retirada de pelo menos uma das suposições correntes. Vamos ver através de um exemplo como o TMS faz este trabalho :

EXEMPLO 3.5 (DOYLE (1979))

Considere a seguinte situação apresentada pelo programa principal ao TMS :

(1) HORA = 10 (SL ( ) (2))

(2) HORA ≠ 10

(3) SALA = 813 (SL ( ) (4))

(4) SALA = 801

O programa principal procura marcar uma reunião , de preferência às 10 horas , ou na sala 813 ou na sala 801. A partir das justificativas apresentadas , o TMS faz com que os nós (1) e (3) fiquem no estado IN e os nós (2) e (4) fiquem no estado OUT.

Suponha que posteriormente o programa principal informe ao TMS que existe uma tabela anterior de horários que elimina a combinação de hora (10 horas) e local (sala 813) para a reunião , criando para isso um novo nó que vai indicar ao TMS essa contradição :

(5) CONTRADIÇÃO (SL (1,3) ( )) ,

que fica no estado IN .

Neste instante , o procedimento DDB é acionado. Inicialmente, é criado um conjunto de suposições máximas S. Uma suposição $A \in$ S se, e somente se, $A$ é um fundamento do nó contradição e não existir nenhuma outra suposição $B$ nos fundamentos do nó contradição tal que $A$ pertença ao fundamento de $B$. Neste exemplo, o conjunto S é composto pelos nós (1) e (3).

Como não houve nenhuma tentativa anterior de retrocesso a partir do nó (5), ou seja, é a primeira vez que a contradição aparece, um novo nó deve ser criado. Este novo nó e a sua respectiva justificativa são da forma :

(6) $NG_1$ (CP 5 (1,3) ( )) , onde a *hipótese* in é formada pelos elementos de S; a *hipótese* out é vazia e o consequente é o nó contradição.

Pelo *algoritmo* 1 (apêndice A), a justificativa do nó (6) equivale a (SL ( ) ( )).

A seguir, o TMS seleciona arbitrariamente um dos dois nós pertencentes à S. Considerando que o TMS escolha o nó (3), ele justifica o nó (4) (único elemento da *lista* out do nó (3)) com :

(4) SALA = 801 (SL (6,1) ( )) , onde a *lista* in é formada pelos nós contradição e pelos outros elementos de S e a *lista* out são os nós OUT da *lista* out do nó escolhido.

Dessa forma , o TMS faz com que os nós (1),(4) e (6) estejam no estado IN e (2),(3) e (5) estejam no estado OUT■

O *algoritmo DDB* está descrito no apêndice A (*ALGORITMO 2*).

*ii)* <u>*ATMS (Assumption-Based TMS)*</u>

Na seção *(i)* apresentamos o TMS que é um sistema de apoio ao raciocínio não-monotônico que grava as inferências feitas por um programa principal, permite que este faça inferências não-monotonicas e garante que a base de dados seja livre de contradições. Todo o processo no TMS está baseado na manipulação das justificativas.

O Sistema de Manutenção da Verdade Baseado em Suposições (ATMS - Assumption-Based TMS) de DE KLEER(1986-a) procura aperfeiçoar o TMS convencional, simplificando o processo de manutenção da verdade e evitando os casos em que é necessário a utilização do DDB.

O sistema do ATMS é formado por dois componentes : um resolvedor de problemas (RP) e um ATMS. O RP é normalmente de primeira-ordem e contém todo o conhecimento do domínio, além dos procedimentos de inferência. O ATMS rotula cada um dos dados apresentados pelo RP com os conjuntos de suposições sob os quais ele é válido . Estes conjuntos representam os contextos nos quais os dados são válidos , e são computados pelo ATMS a partir das justificativas fornecidas pelo RP. Este tipo de representação facilita a referência aos contextos e permite que os conjuntos sejam manipulados de uma maneira mais conveniente que os dados que eles representam.

*ii.1) Estrutura de Dados*

Cada dado do RP é representado por um *nó ATMS* , que é a estrutura de dados básica do sistema. Os nós ATMS são tratados como símbolos proposicionais e são da forma :

$$\gamma_{dado} : \langle dado, rótulo, justificativa\rangle .$$

onde <u>dado</u> é um elemento do domínio de conhecimento do RP que é transmitido ao ATMS; uma <u>justificativa ATMS</u> descreve como um nó é derivado de outros, e é composta por tuplas $(\gamma_1, \ldots, \gamma_n), \ldots, (\gamma'_1, \ldots, \gamma'_m)$ que representam :

$$(\gamma_1 \wedge \ldots \wedge \gamma_n) \vee \ldots \vee (\gamma'_1 \wedge \ldots \wedge \gamma'_m) \supset \gamma_{dado}$$

sendo que os nós que formam as tuplas $(\gamma_1, \ldots, \gamma_n), \ldots, (\gamma'_1, \ldots, \gamma'_m)$ , são chamados <u>antecedentes</u> e $\gamma_{dado}$ , que é o nó que está sendo justificado, é chamado de <u>consequente</u>. As justificativas são cláusulas de Horn (possuem no máximo um literal negativo) e são fornecidas pelo RP ao ATMS.

Por exemplo, a inferência de Q(a), feita pelo RP, a partir de P(a) e $\forall x(P(x) \supset Q(x))$ é fornecida ao ATMS através da justificativa :

$$\gamma_{P(a)} , \gamma_{\forall x(P(x) \supset Q(x))} \supset \gamma_{Q(a)}$$

que trata $\gamma_{P(a)}$ , $\gamma_{\forall x(P(x) \supset Q(x))}$ e $\gamma_{Q(a)}$ como proposições.

Um <u>rótulo ATMS</u> é um conjunto de ambientes, onde um <u>ambiente ATMS</u> é um conjunto de suposições , que são nós presumivelmente verdadeiros, ou seja, nós considerados verdadeiros a menos que exista alguma evidência em contrário.

**EXEMPLO 3.6**

Considere que um RP possui o dado AVE(Tweety) que denota que "Tweety é uma ave". Este dado é representado no ATMS por :

$$\gamma_{AVE(Tweety)} \quad : \langle AVE(Tweety), \{\{\ \}\}, \{(\ )\}\rangle \ ,$$

Neste caso , dizemos que o nó $\gamma_{AVE(Tweety)}$ é uma <u>*premissa*</u> pois possui uma justificativa sem antecedentes. As premissas são válidas universalmente.

Uma <u>*suposição*</u> é um nó cujo rótulo possui um ambiente unitário que contém o próprio nó. Se representarmos :

"Tweety voa" por V ;

"Tweety é uma ave" por P ;

"Tweety é um animal" por A ;

"Tweety é normal" por a ;

uma suposição $\Gamma_V$ pode ser da forma :

$$\langle V, \{\{V\}, \{P, A, a\}\}, \{(\Gamma_P, \gamma_a)\}\rangle \ ,$$

Por convenção, as letras maiúsculas e $\Gamma_x$ denotam suposições enquanto as letras minúsculas e $\gamma_x$ denotam os outros tipos de nós. Logo, são suposições <u>V</u>, <u>P</u> e <u>A</u> ; e <u>a</u> é um outro tipo qualquer de nó.

O rótulo ATMS de $\Gamma_V$ é formado por $\{V\}$ e $\{P, A, a\}$ , e ele é justificado por $(\Gamma_P, \gamma_a)$. Eles representam , respectivamente, as implicações :

$$V \vee (P \wedge A \wedge a) \supset V$$

$$\Gamma_P \wedge \gamma_a \supset V$$

O nó *false* ($\perp$) é representado por :

$$\gamma_{\perp} : \langle \perp , \{ \} , \{ \ldots \} \rangle .$$

onde $\{\ldots\}$ indica que podemos ter qualquer conjunto de justificativas ■

Dizemos que um nó ATMS n pertence a um ambiente E se n pode ser derivado de E e do conjunto de justificativas J :

$$E, J \vdash n .$$

Para cada nó n, o rótulo (ou seja, conjunto de ambientes $\{E_1, \ldots, E_k\}$) possui as seguintes propriedades :

(1) *CONSISTENCIA* : Para todo $i \in [1,k]$ , $E_i$ é consistente ;

(2) *CORREÇÃO* : Para todo $i \in [1,k]$ , n é derivável de $E_i$ ;

(3) *COMPLETUDE* : Todo ambiente E no qual n é derivável possui algum subconjunto E' tal que $E' \in \{E_1, \ldots E_k\}$;

(4) *MINIMALIDADE* : Nenhum ambiente $E_i$ é um subconjunto de qualquer outro.

Um ambiente E é inconsistente (dito *nogood*) quando *false* ($\perp$) é derivável de E e do conjunto de justificativas J :

$$E, J \vdash \perp$$

Os ambientes inconsistentes (nogood) representam conjunções inconsistentes de suposições . Tais ambientes formariam o rótulo de $\gamma_\perp$ caso os ambientes inconsistentes não fossem excluídos dos rótulos dos nós.

Um contexto ATMS é o conjunto formado pelas suposições de um ambiente consistente combinadas com todos os nós deriváveis dessas suposições .

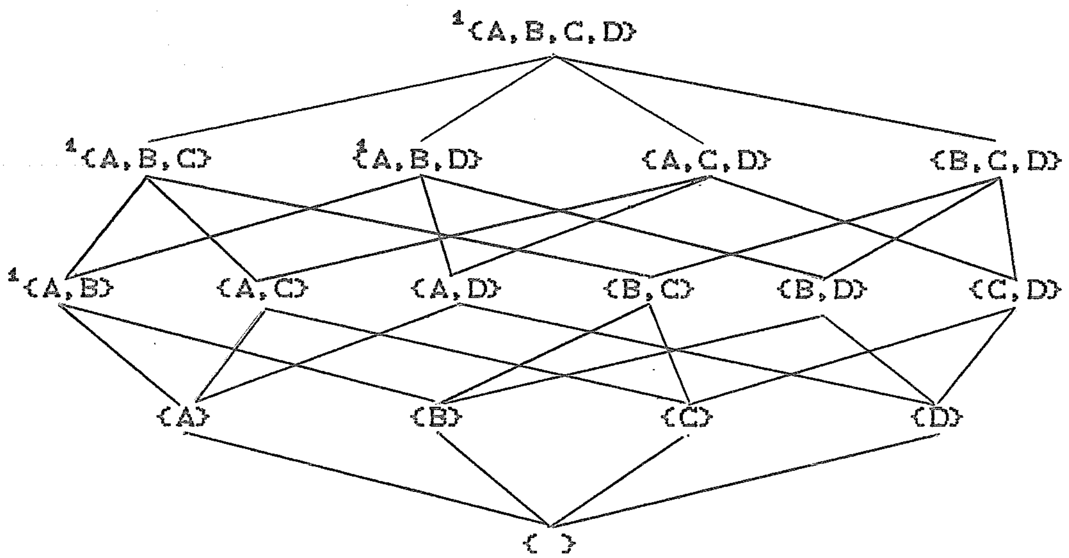
*ii.2) Reticulado do Ambiente*

A tarefa do ATMS é determinar , de uma maneira eficiente, todos os possíveis contextos e seus conteúdos a partir das inferências feitas até aquele momento pelo RP. A eficiência é obtida observando-se que se um dado é derivável de um conjunto particular de suposições , ele também será derivável de qualquer outro conjunto de suposições que tenha aquele como subconjunto . Inicialmente, veremos como o ATMS verifica se um certo dado pertence ou não à um contexto e a seguir , como os rótulo são atualizados a partir de novas justificativas.

O ATMS computa os conteúdos de cada contexto de forma indireta através do rótulo de cada dado. Logo, determinar se um dado d pertence à um certo contexto significa saber se pelo menos um dos ambientes do rótulo de d é um subconjunto das suposições do contexto em questão. A verificação de tal condição é feita de maneira simples , como mostra o seguinte exemplo :

**EXEMPLO 3.7**

Para um conjunto de n suposições , temos $2^n$ ambientes possíveis que são representados através de um reticulado de ambiente. Assim, no caso de termos A, B, C e D como suposições, os ambientes possíveis são :




Como qualquer ambiente E que permita a derivação de $\perp$ é inconsistente, ele e todos os ambientes que tenham E como um de seus subconjuntos, deverão ser retirados do reticulado. Por exemplo, se considerarmos o ambiente {A,B} como sendo inconsistente, todos os ambientes marcados com (1) devem ser retirados do reticulado acima. Logo, todos os ambientes <u>nogood</u> são gerados por {A,B}.

Para determinar se um dado <u>d</u> pertence à um contexto, cujo conjunto de suposições é {A,C,D}, o ATMS busca dentre os ambientes de <u>d</u> algum que seja subconjunto de {A,C,D}.

Caso $\gamma_d$ seja da forma :

$$\gamma_d : \langle d , \{\{C,D\}\} , \{\ldots\}\rangle$$

sendo $\{C,D\} \subset \{A,C,D\}$, d pertencerá a tal contexto. Porém, se :

$$\gamma_d : \langle d \;,\; \{\{B,D\}\} \;,\; \{\ldots\}\rangle$$

d não pertencerá ao contexto de suposições $\{A,C,D\}$ ■

*ii.3) Operações Básicas*

O ATMS executa três operações básicas : (1) cria um nó ATMS para um dado do RP ; (2) cria uma suposição ; e (3) acrescenta uma justificativa para um nó.

Criar um nó ATMS para cada dado do RP é a primeira atitude a ser feita . Aqui, é importante que dois dados "equivalentes" sejam associados ao mesmo nó ATMS, o que só pode ser feito se existir dentro do RP alguma maneira de indexação que permita que todas as formas de um mesmo dado sejam associados ao mesmo nó.

Criar uma suposição tal como criar um nó requer pouco do ATMS. Determinar quais suposições devem ser consideradas é uma tarefa que cabe ao RP.

A operação fundamental desempenhada pelo ATMS é o de acrescentar uma nova justificativa à um nó.

EXEMPLO 3.8

Considere a seguinte situação no ATMS :

(1) $\gamma_{P(a)}$ : $\langle P(a),\ \{\{A,B\},\{B,C,D\}\},\{\ldots\}\rangle$

(2) $\gamma_{\forall x(P(x) \supset Q(x))}$ : $\langle \forall x(P(x) \supset Q(x)),\ \{\{A,C\},\{D,E\}\},\{\ldots\}\rangle$

(3) $\gamma_{Q(a)}$ : <Q(a), { },{...}>

(4) nogood {A,B,E}

sendo A,B,C,D e E suposições.

Por (4), todos os ambientes que têm como um de seus subconjuntos {A,B,E} são retirados do reticulado de ambiente.

Quando o RP deduzir Q(a) de P(a) e $\forall x(P(x) \supset Q(x))$, ele informará tal fato ao ATMS através da seguinte justificativa :

$$\gamma_{P(a)} \wedge \gamma_{\forall x(P(x) \supset Q(x))} \supset \gamma_{Q(a)} \qquad (*)$$

Tal justificativa é acrescida ao nó $\gamma_{Q(a)}$, o que vai obrigar que o rótulo de $\gamma_{Q(a)}$ seja atualizado. O novo rótulo deve ser tal que preserve as propriedades de correção, completude, consistência e minimalidade.

Um rótulo completo e correto para $\gamma_{Q(a)}$ (que é o consequente de (*)) é formado pelos conjuntos obtidos da união de cada um dos ambientes de $\gamma_{P(a)}$ com cada ambiente de $\gamma_{\forall x(P(x) \supset Q(x))}$. Logo, os ambientes possíveis são :

{{A,B,C} , {A,B,C,D} , {A,B,D,E} , {B,C,D,E}} (**)

Retirando do conjunto acima o ambiente {A,B,C,D} que tem um subconjunto ({A,B,C}) também pertencente à (**), e o ambiente {A,B,D,E} que é inconsistente (contém o nó nogood {A,B,E}), obtemos :

{{A,B,C} , {B,C,D,E}}

que é mínimo e consistente, o que nos dá um rótulo que satisfaz todas as condições necessárias. Portanto :

$$\gamma_{Q(a)} \quad : \langle Q(a), \{\{A,B,C\} , \{B,C,D,E\}\}, \{(\gamma_{P(a)} , \gamma_{\forall x(P(x) \supset Q(x))})\}\rangle$$

Caso $\gamma_{Q(a)}$ pertencesse a justificativa de algum outro nó $\underline{n}$ , os rótulos dele também deveriam ser revistos ∎

O algorítmo correspondente a este procedimento está no apêndice A . *(ALGORITMO 3)*

*iii)* *Conclusão*

Apresentamos neste capítulo os sistemas de manutenção da verdade TMS e ATMS. A diferença entre tais sistemas está em como eles tratam as informações de que dispõem.

O TMS a partir de um conjuto de informações e de justificativas determina um conjunto de creças correntes. Tal conjunto é na verdade um subconjunto consistente do conjunto total de suas informações.

Já o ATMS determina para cada uma das informações de que dispõe, quais os vários conjunto de suposições mínimas a partir dos quais ele pode ser derivado. Neste caso a base de conhecimento pode ser inconsistente sem que isto afete o processo geral do ATMS.

Tais sistemas serão utilizados de maneira distinta neste trabalho. O TMS vai ser utilizado na formalização do problema do tiro (capítulo VI), onde tenta-se transportar para a lógica de default e a circunscrição , a característica que o leva a ter um bom comportamento (no sentido de nos fornecer uma resposta intuitiva) em relação ao problema do tiro.

O ATMS sugere uma maneira de implementação de um dos métodos que avaliam quando um fato segue ou não de uma teoria circunscrita (capítulo V - seção (iii)) .

# CAPITULO IV

## CIRCUNSCRIÇÃO

A circunscrição é um outro método utilizado na tentativa de formalizar o raciocínio não-monotônico, e tem sido considerada a mais poderosa ferramenta para esta tarefa, basicamente por estar fundamentada na lógica de predicados.

Definida inicialmente por McCARTHY (1980), a circunscrição vem sendo bastante estudada , o que acarretou a definição de uma série de outros tipos de circunscrição, que se inter-relacionam.

A idéia básica que está por trás dessa variedade de circunscrições é a de minimizar um determinado conjunto de predicados de modo que os objetos que possuem tais propriedades sejam os únicos conhecidos a possuí-las.

Para alcançar tal objetivo , a proposta é transformar uma dada teoria numa outra mais forte , através de um processo eminentemente sintático.

Na seção inicial , definiremos formalmente a idéia da circunscrição e apresentaremos uma caracterização de seus modelos.

Na seção (ii) caracterizaremos duas classes de teorias para as quais o processo de circunscrição não as torna inconsistentes.

Na seção (iii), veremos uma forma de circunscrição onde são atribuidas prioridades sobre os predicados que são circunscritos.

A seção (iv) apresenta uma forma de circunscrição obtida a partir da utilização de uma outra maneira de expressar a interpretação de um símbolo de predicado.

Na seção (v) é visto um tipo de circunscrição que utiliza um conceito especial de modelo usado na programação em lógica.

## *i ) Definição*

A circunscrição (McCARTHY (1980,1986)) é um método de formalização do raciocínio não-monotônico cuja idéia central é a de restringir os objetos que possuem determinadas propriedades , de forma que eles sejam os únicos objetos conhecidos a possuí-las.

O processo de circunscrição transforma uma teoria de primeira-ordem em outra mais forte , geralmente de segunda-ordem , através da representação sintática da idéia de minimização da extensão de um ou vários predicados da teoria original.

Assim, circunscrever um ou mais predicados numa dada teoria significa que pretendemos minimizar as suas extensões (estrutura de um símbolo de predicado n-ário, representada por um subconjunto de $|U|^n$, onde U é o universo do modelo) , de modo que os objetos que possuem as propriedades estabelecidas pelos predicados são os únicos conhecidos a possuí-las . Ao longo desse processo, alguns predicados (e/ou funções), diferentes dos que estão sendo circunscritos , podem variar , ou seja , é permitido que as suas extensões variem de modo a permitir que os predicados que estão sendo circunscritos se tornem mínimos.

Como estamos trabalhando com os predicados tal qual conjuntos , a condição de minimalidade deve ser vista com respeito a relação de inclusão de conjuntos ( $\subset$ ).

NOTAÇÃO 1 : Para dois predicados n-ários U e V , a fórmula $\forall x\ (U(x) \supset V(x))$ será representada por $U \leq V$ . Quando U e V forem , respectivamente, as tuplas de predicado $\langle U_1,\ldots,U_m\rangle$ e $\langle V_1,\ldots,V_m\rangle$, $U \leq V$ representa $U_1 \leq V_1 \wedge U_2 \leq V_2 \wedge \ldots \wedge U_m \leq V_m$. $U \leq V$ expressa que a extensão do predicado U é um subconjunto da extensão do predicado V . Além disso , temos : $U = V$ que representa $(U \leq V \wedge V \leq U)$ e $U < V$ que representa $(U \leq V \wedge \neg(V \leq U))$.

OBSERVAÇÃO 1 : As tuplas $U = \langle U_1,\ldots,U_n\rangle$ e $V = \langle V_1,\ldots,V_m\rangle$ são ditas <u>semelhantes</u> se $m = n$ e para todo i, $U_i$ e $V_i$ são símbolos do mesmo tipo (função ou predicado) e de mesma aridade.

Dessa maneira , a circunscrição fica formalmente definida pela :

DEFINIÇÃO 4.1 *(Circunscrição Paralela)*

Seja $T(P,Z)$ uma teoria tal que $P = \langle P_1,\ldots,P_n\rangle$ é uma tupla de constantes de predicado e $Z = \langle Z_1, \ldots, Z_m \rangle$ é uma tupla de constantes de função e/ou predicado, sendo que P e Z são disjuntos. A *circunscrição de P em T(P,Z) com Z variando* é representada pela seguinte expressão de segunda-ordem :

$$T(P,Z) \wedge \neg \exists P',Z' \; ( T(P',Z') \wedge P' < P ) \quad (1)$$

denotada por CIRC(T(P,Z);P;Z) ou CIRC(T;P;Z), onde P' e Z' são tuplas de variáveis de predicado (no caso de Z, e/ou função) semelhantes à P e Z. Este tipo de circunscrição é chamada de *paralela* e a expressão (1) é conhecida como *axioma da circunscrição* .

**OBSERVAÇÃO 2 :** A expressão (1) pode ser escrita da seguinte forma :

$$T(P,Z) \wedge \forall P',Z' \; [ \; (T(P',Z') \wedge P' \supset P) \supset P' \equiv P \; ] \quad (2)$$

onde $P' \supset P$ representa que " para toda tupla de variaveis x e todo $i \leq n$ , $(P_i'(x) \supset P_i(x))$ " .

**EXEMPLO 4.2**

Considere a teoria composta pelos seguintes axiomas :

(i) $\forall x$ ( AVE(x) $\wedge \neg$ AB(x) $\supset$ VOA(x) )

(ii) $\forall x$ ( PINGÜIM(x) $\supset$ AVE(x) )

(iii) $\forall x$ ( PINGÜIM(x) $\supset \neg$ VOA(x) ) onde : AVE(x) representa " x é uma ave" ;

PINGÜIM(x) representa " x é um pingüim" ;

VOA(x) representa " x voa" ;

AB(x) representa " x é anormal" ;

(i) representa que normalmente as aves voam ;

(ii) representa que os pingüins são aves ; e

(iii) representa que os pingüins não voam .

Vamos denotar tal teoria por T(P,Z) , sendo que P = <AB> e Z = <VOA> ,ou seja, T(AB,VOA). Note que as tuplas P e Z não contém todos os predicados pertencentes a teoria acima . Na verdade , a notação T(P,Z) objetiva apenas indicar quais os predicados que se quer circunscrever, que formam a tupla P, e quais serão variáveis ao longo do processo , que formam a tupla Z.

Dessa maneira , vamos circunscrever AB na teoria T(AB,VOA) com o predicado VOA variando, isto é, vamos fazer com que a extensão do predicado AB seja minima sob a condição T(AB,VOA),permitindo que extensão do predicado VOA varie.

Aplicando a expressão (2) da observação 2 , obtemos :

$$T(AB,VOA) \wedge \forall\, AB',VOA' \; [\; T(AB',VOA') \wedge (\; \forall x \;(\; AB'(x) \supset AB(x) \;)\;) \supset (\; \forall x \;(\; AB(x) \equiv AB'(x) \;)\;)\;] \equiv$$

$$\equiv T(AB,VOA) \wedge$$

$$\forall\, AB',VOA' \; [\; [\forall x \;(AVE(x) \wedge \neg\, AB'(x) \supset VOA'(x))\;] \wedge$$

$$[\; \forall x \;(PINGÜIM(x) \supset AVE(x))\;] \wedge$$

$$[\; \forall x \;(PINGÜIM(x) \supset \neg\, VOA'(x))\;] \wedge$$

$$[\; \forall x \;(AB'(x) \supset AB(x)\;] \supset [\; \forall x \;(\; AB(x) \equiv AB'(x) \;]\;] \quad (a)$$

Como a teoria T(AB,VOA) é considerada verdadeira, basta levar em conta somente a segunda parte da conjunção (a). Nela , o objetivo é encontrar substituições para as variáveis de predicado AB' e VOA', de modo a tornar o antecedente da implicação verdadeiro e assim determinar o que é anormal na teoria T(AB,VOA), através do seu consequente.

Considerando que os axiomas que formam a teoria T(AB,VOA) são as únicas informações disponíveis, podemos concluir, pelo senso comum, que as únicas aves que não voam são os pingüins e são , portanto, as únicas aves anormais. Assim, a substituição adequada seria :

$$VOA'(x) \Leftrightarrow AVE(x) \wedge \neg PINGÜIM(x), \text{ e}$$
$$AB'(x) \Leftrightarrow PINGÜIM(x) \wedge AVE(x),$$

ficando a fórmula entre colchetes da expressão (a) equivalente a :

$$\begin{aligned}
&\forall x\, (AVE(x) \wedge (\neg PINGÜIM(x) \vee \neg AVE(x)) \supset AVE(x) \wedge \\
&\quad \neg PINGÜIM(x)) \wedge \\
&\forall x\, (PINGÜIM(x) \supset AVE(x)) \wedge \\
&\forall x\, (PINGÜIM(x) \supset \neg AVE(x) \vee PINGÜIM(x)) \wedge \\
&\forall x\, (PINGÜIM(x) \wedge AVE(x) \supset AB(x)) \supset \\
&\quad \forall x\, ( AB(x) \equiv PINGÜIM(x) \wedge AVE(x)),
\end{aligned}$$

que tem o antecedente da implicação verdadeiro, já que T(AB,VOA) é considerada verdadeira. Donde obtemos que :

$$\forall x\, (AB(x) \equiv PINGÜIM(x) \wedge AVE(x)) \qquad (b)$$

Logo , a teoria circunscrita será formada por T(AB,VOA) mais a expressão (b). De tal teoria podemos concluir que $\forall x\ (AVE(x) \wedge \neg PINGÜIM(x) \supset VOA(x))$ (obtida de (i) e (b)), fato este que não é obtido se considerarmos apenas a teoria T(AB,VOA) ■

**DEFINIÇÃO 4.3** *(Circunscrição Básica)*

Quando na circunscrição paralela CIRC(T(P,Z);P;Z), a tupla Z for vazia, a circunscrição é chamada de <u>*básica*</u> , e as expressões (1) da definição 4.1 e (2) da observação 2 são escritas , respectivamente , na forma :

$$CIRC(T(P);P) \Leftrightarrow T(P) \wedge \neg \exists P' \ (\ T(P') \wedge P' < P) \quad (1'),\ e$$

$$CIRC(T(P);P) \Leftrightarrow T(P) \wedge \forall P' \ [(T(P') \wedge P' \supset P) \supset P' \equiv P] \quad (2').$$

**EXEMPLO 4.4**

Considere a teoria T(P) , onde P = <AVE> , composta por:

(i) AVE(Tweety)

(ii) AVE(Piu-Piu)

Aplicando a expressão (2') de definição 4.3 , de forma a circunscrever AVE na teoria T(AVE), obtemos :

$$T(AVE) \wedge \forall\ AVE'\ [\ AVE'(Tweety) \wedge AVE'(Piu\text{-}Piu) \wedge$$
$$\forall x\ (AVE'(x) \supset AVE(x)) \supset$$

$$[\forall x \ (AVE(x) \equiv AVE'(x))] \quad (c)$$

A teoria T(AVE) tem como informação que Tweety e Piu-Piu são as únicas aves conhecidas. Portanto, a substituição adequada é :

$$AVE'(x) \Leftrightarrow ( x = Tweety \vee x = Piu\text{-}Piu )$$

de forma que :

$$\begin{gathered}( Tweety = Tweety \vee Tweety = Piu\text{-}Piu ) \wedge \\ (Piu\text{-}Piu = Tweety \vee Piu\text{-}Piu = Piu\text{-}Piu ) \wedge \\ (\forall x \ (( x= Tweety \vee x = Piu\text{-}Piu ) \supset AVE(x))) \supset \\ (\forall x \ (AVE(x) \equiv ( x = Tweety \vee x = Piu\text{-}Piu ) ) ) ,\end{gathered}$$

onde o antecedente da implicação é verdadeiro , e segue daí que :

$$\forall x \ ( AVE(x) \equiv (x = Tweety \vee x = Piu\text{-}Piu) ) \blacksquare$$

As definições 4.1 e 4.3 podem ser entendidas como sendo a representação de que as extensões dos predicados pertencentes a tupla P são mínimas sob a condição T(P,Z), quando é permitido que as extensões dos predicados da tupla Z variem no processo de minimização, ou seja, a extensão de P não pode ser menor, mesmo que a extensão de Z varie.

Essa noção de minimalidade presente na definição da circunscrição pode ser expressa em termos de teoria de modelos. Para isto, vamos definir a relação $\leq$ mod (P,Z) que determina uma ordem parcial sobre a classe de todas as estruturas. Um conjunto S é dito parcialmente ordenado por uma relação binária $R$ se, para qualquer a, b e c $\in$ S , temos :

i) $R$ é reflexivo , i.e., a $R$ a ;

ii) $R$ é anti-simétrico , i.e., (a $R$ b) e (b $R$ a) se, e somente se, (a = b) ;

iii) $R$ é transitiva , i.e., (a $R$ b) e (b $R$ c) implicam que (a $R$ c) .

Por exemplo, a relação de divisibilidade determina uma ordem parcial entre os inteiros positivos.

**NOTAÇÃO 2 :** Para uma estrutura M, M[k] denota o objeto representado por k, sendo k uma constante ( de predicado ou função ) qualquer, de uma linguagem de primeira-ordem .

**DEFINIÇÃO 4.5** *( Relação $\leq$ mod (P,Z) )*

Sejam P uma tupla de constantes de predicado e Z uma tupla de constantes de função e/ou predicado, sendo que P e Z são disjuntos. Para quaisquer duas estruturas M e N , escrevemos M $\leq$ N mod (P,Z) se :

(a) $|M| = |N|$ (M e N têm o mesmo domínio) ;

(b) M[k] = N[k], para toda constante k não pertencente à $P \cup Z$ ;

(c) $M[P_i] \subset N[P_i]$ , para todo $P_i$ em P.

A definição 4.5 estabelece que os modelos M e N diferem somente na maneira de interpretar os predicados pertencentes à P e Z, e as extensões de cada predicado $P_i$ em M são subconjuntos de suas extensões em N.

OBSERVAÇÃO 3 : Se $M \leq N$ mod (P,Z), e não tivermos $N \leq M$ mod(P,Z), então $M < N$ mod (P,Z).

DEFINIÇÃO 4.6 *(Estrutura Mínima)*

Uma *estrutura* M *é mínima com respeito a relação ≤ mod (P,Z)* numa classe de estruturas S , se :

(a) $M \in S$ ;

(b) não existe $N \in S$ tal que $N < M$ mod (P,Z) .

OBSERVAÇÃO 4 : Note que dentro de uma classe de estruturas S, podemos ter mais de uma estrutura mínima com respeito a relação ≤ mod (P,Z).

Assim , a partir das definições 4.5 e 4.6 , os modelos da circunscrição podem ser caracterizados pela seguinte proposição :

PROPOSIÇÃO 4.7 *(LIFSCHITZ (1985-a))*

Uma estrutura M é um modelo de CIRC(T(P,Z);P;Z) se, e somente se, M é mínimo na classe de modelos de T(P,Z) com respeito a relação ≤ mod (P,Z).

A equivalência expressa pela proposição anterior segue do fato de que especificando valores de $P'$ e $Z'$ para os quais $T(P',Z') \wedge P' < P$ é verdadeiro em $M$, é equivalente a especificar um modelo $M'$ de $T(P,Z)$ tal que $M' < M$ (mod $P,Z$).

**OBSERVAÇÃO 5 :** Quando $Z$ for vazio , a relação $\leq$ mod $(P,Z)$ fica representada por $\leq$ mod $(P)$ .

*ii ) Satisfatibilidade*

O processo de fortalecimento de teorias efetuado pela circunscrição pode eventualmente resultar em teorias que apresentam problemas de satisfatibilidade . Nesta seção examinaremos exatamente este tipo de problema : dada uma teoria T consistente , o axioma de circunscrição obtido pode ser inconsistente. Tal fato foi observado por ETHERINGTON (1985), através do seguinte exemplo :

**EXEMPLO 4.8**

Considere a teoria $T(P)$ composta pelos axiomas :

(1) $\exists x\ (\ P(x) \wedge \forall y\ (\ P(y) \supset x \neq s(y)\ )\ )$

(2) $\forall x\ (\ P(x) \supset P(s(x))\ )$

(3) $\forall\ x,y\ (\ s(x) = s(y) \supset x = y\ )$

que pode ser vista como a descrição da função sucessor s no domínio dos números naturais, ou seja, $s : \mathbb{N} \rightarrow \mathbb{N}$ definida por $s(n) = n + 1$, onde (1) significa que existe um número natural ($P(x)$ representa "x é um número natural") que não é sucessor de nenhum outro número natural, (2) significa que todo número natural tem um sucessor e (3) que cada número natural possui um único sucessor.

Observe que em todos os modelos M da teoria T(P), M[P] contém uma sequência de elementos $S$ = ($x_1$, s'($x_1$), s'(s'($x_1$)),...) , sendo $x_1 \in |M|$ e s' = M[s], isomorfos aos números naturais. Eliminando-se um número finito de elementos iniciais de $S$ , é possível obter um modelo N tal que $|N| = |M|$ e $N[P] \subset M[P]$, isto é , $N < M$ mod(P). Deste modo, qualquer que seja o modelo de T(P), existirá um submodelo dele, o que acarretará a inexistência de um modelo mínimo para T(P) e,como consequência da proposição 4.7, tornará insatisfatível a CIRC(T(P);P) ■

Para enfrentar tal problema, **ETHERINGTON** (1985) buscou determinar uma classe de teorias na qual a circunscrição não acarretasse inconsistência. Tal classe é formada pelas teorias ditas <u>bem-fundamentadas</u>, que são aquelas onde para todo modelo M de uma dada teoria T(P,Z) existe um modelo mínimo N com respeito a relação $\leq$ mod (P,Z) ( $N \leq M$ mod (P,Z) ). Neste caso , diz-se que a teoria T(P,Z) é <u>bem-fundamentada com respeito a</u> (P,Z). Caso Z seja vazio, dizemos bem-fundamentada com respeito a P.

**EXEMPLO 4.9**

Considere a teoria T(AVE) composta por um único axioma :

AVE(Tweety) v AVE(Piu-Piu)

Tal teoria possui três modelos :

(i) $M_1[AVE] = \{ (Tweety) \}$ ;

(ii) $M_2[AVE] = \{ (Piu\text{-}piu) \}$ ;

(iii) $M_3[AVE] = \{ (Tweety) , (Piu\text{-}Piu) \}$ , sendo que os modelos $M_1$ e $M_2$ são mínimos.

Observe que para o modelo $M_3$ , existe um modelo , que pode ser $M_1$ ou $M_2$, mínimo com respeito a $\leq$ mod (P) tal que $M_1 \leq M_3$ (ou $M_2 \leq M_3$ ) mod(P). Logo , T(AVE) é uma teoria bem-fundamentada ■

A importância das teorias bem-fundamentadas fica estabelecida pelo :

**TEOREMA 4.10** *(LIFSCHITZ (1986))*

Se uma teoria T(P,Z) é consistente e bem-fundamentada com respeito a (P,Z), então CIRC(T(P,Z);P;Z) é satisfatível.

Observe que se a teoria T(P,Z) é bem-fundamentada , por definição, para todo modelo M de T(P,Z), existe um modelo mínimo N com respeito a relação $\leq$ mod(P,Z) (N $\leq$ M mod (P,Z)). Pela proposição 4.7, tal modelo mínimo N é um modelo da CIRC(T(P,Z);P;Z). Logo, como existe uma estrutura que satisfaz CIRC(T(P,Z);P;Z), ela é satisfatível.

Como veremos mais adiante (proposição 4.11), a circunscrição de P na teoria T(P,Z) com Z variando e a circunscrição de P na teoria $\exists z$ T(P,z) se inter-relacionam . Por isto, os problemas da satisfatibilidade e da boa-fundamentação para teorias $T(P,Z)$ podem ser reduzidos a questões semelhantes para $\exists z\, T(P,z)$. Para tal , Lifschitz estabeleceu o seguinte resultado :

**PROPOSIÇÃO 4.11** *(LIFSCHITZ (1985))*

Seja $T(P,Z)$ uma teoria onde P é uma tupla de constantes de predicado e Z é uma tupla de constantes de função e/ou predicado, sendo P e Z disjuntos. Então :

$$CIRC(T(P,Z);P;Z) \equiv T(P,Z) \wedge CIRC(\exists z T(P,z);P).$$

A partir da proposição 4.11, podemos mostrar que :

**TEOREMA 4.12** *(LIFSCHITZ (1986))*

A teoria $T(P,Z)$ é bem-fundamentada com respeito a $(P,Z)$ se, e somente se, $\exists z\, T(P,z)$ é bem-fundamentada com respeito a P.

**TEOREMA 4.13** *(LIFSCHITZ (1986))*

A $CIRC(T(P,Z);P;Z)$ é satisfatível se, e somente se, $CIRC(\exists z\, T(P,z);P)$ é satisfatível.

Dentro da classe das teorias bem-fundamentadas encontram-se as <u>teorias universais</u> ,que são aquelas teorias onde cada uma das fórmulas que as compõem não possui nenhum quantificador existencial quando escritas na forma prenex, ou seja, cada um dos axiomas é da forma $\forall x\ A$, onde $x$ é uma tupla de variáveis e $A$ é uma fórmula livre de quantificadores. Assim,

**TEOREMA 4.14** *(ETHERINGTON (1985))*

Qualquer teoria universal é bem-fundamentada com respeito a qualquer conjunto de constantes de predicado $P$.

Como consequência dos teoremas 4.10 e 4.14, temos o seguinte resultado :

**TEOREMA 4.15** *(ETHERINGTON (1985))*

Se $T(P)$ é uma teoria universal consistente, então $CIRC(T(P);P)$ é satisfatível.

Observe que o teorema 4.15 considera o conjunto $Z$ vazio. Isto porque mesmo para teorias universais, a circunscrição pode não preservar a satisfatibilidade quando $Z$ inclui constantes, como veremos no próximo exemplo.

**EXEMPLO 4.16**

Considere a teoria $T'(P,c)$ :

(1) $P(c) \wedge \forall y\ (P(y) \supset c \neq s(y))$ ;

(2) $\forall x\ (\ P(x) \supset P(s(x))\ )$ ;

(3) $\forall x,y\ (\ s(x) = s(y) \supset x = y)$ ;

onde c é uma constante.

Pelo teorema 4.13, a CIRC(T'(P,c);P;c) é satisfatível se, e somente se, a CIRC(∃z T'(P,z);P) também o for. Mas a CIRC(∃z T'(P,z);P) é equivalente a CIRC(T(P);P) onde a teoria T (P) é a mesma do exemplo 4.8.

Como já vimos, CIRC(T(P);P) é insatisfatível e, portanto, obtivemos que a circunscrição da teoria universal consistente T'(P) é insatisfatível ■

O exemplo anterior apresenta uma teoria satisfatível cuja circunscrição é insatisfatível. Os próximos resultados eliminam tal anomalia, quando o conjunto Z não contém símbolo de função ( note que no exemplo 4.16, Z possui símbolos de função ).

**TEOREMA 4.17** *(LIFSCHITZ (1986))*

Para quaisquer conjuntos disjuntos P e Z de constantes de predicado, toda teoria universal T(P,Z) é bem-fundamentada com respeito a (P,Z).

**TEOREMA 4.18** *(LIFSCHITZ (1986))*

Para quaisquer conjuntos disjuntos P e Z de constantes de predicado e qualquer teoria universal satisfatível T(P,Z), a CIRC(T(P,Z);P;Z) é satisfatível.

Posteriormente , LIFSCHITZ (1986) estendeu os resultados obtidos para teorias universais (teoremas 4.14 e

4.15) a uma nova classe de teorias , chamada quase-universais. Dizemos que uma teoria é quase-universal com respeito a uma tupla de constantes de predicado $P_1, P_2, \ldots$, se cada um dos axiomas for da forma $\forall x\ A$ , onde $x$ é uma tupla de variáveis e $A$ não contém ocorrências positivas de $P_1, P_2, \ldots$ no escopo dos quantificadores.

**EXEMPLO 4.19**

A teoria do exemplo 4.2 é quase-universal com respeito a AB :

(i) $\forall x\ (\ AVE(x) \wedge \neg\ AB(x) \supset VOA(x)\ )$

(ii) $\forall x\ (\ PINGÜIM(x) \supset AVE(x)\ )$

(iii) $\forall x\ (\ PINGÜIM(x) \supset \neg\ VOA(x)\ )$ .

Se escrevêssemos tal teoria na forma :

(i') $\forall x\ (\ AVE(x) \wedge \neg\ VOA(x) \supset AB(x)\ )$

(ii) $\forall x\ (\ PINGÜIM(x) \supset AVE(x)\ )$

(iii) $\forall x\ (\ PINGÜIM(x) \supset \neg\ VOA(x)\ )$

ela passa a ser quase-universal com respeito a VOA ■

Assim, os resultados 4.14 e 4.15 ficam estendidos para a classe de teorias quase-universais , respectivamente pelos seguintes teoremas :

**TEOREMA 4.20** *(LIFSCHITZ (1986))*

Se todo axioma da teoria T(P) é quase-universal com respeito a P , então T(P) é bem-fundamentada com respeito a P .

TEOREMA 4.21 *(LIFSCHITZ (1986))*

Se a teoria T(P) é satisfatível e quase-universal com respeito a P , então CIRC(T(P);P) é satisfatível .

*iii) Circunscrição Priorizada*

Consideremos a seguinte teoria $T(\langle AB_1,AB_2,AB_3\rangle;VOA)$ :

(1) $\neg AB_1(Tweety) \supset \neg VOA(Tweety)$

(2) $AVE(Tweety) \supset AB_1(Tweety)$

(3) $AVE(Tweety) \wedge \neg AB_2(Tweety) \supset VOA(Tweety)$

(4) $CANARIO(Tweety) \wedge \neg AB_3(Tweety) \supset AVE(Tweety)$

(5) $CANARIO(Tweety)$

onde (1) representa que a menos que Tweety seja anormal sob o aspecto 1 , ele não pode voar ; (2) representa que se Tweety é uma ave então é anormal sob o aspecto 1 ; (3) representa que a menos que Tweety seja uma ave anormal sob o aspecto 2 , ele pode voar ; (4) diz que a menos que Tweety seja canário anormal sob o aspecto 3, ele será uma ave ; e (5) diz que Tweety é um canário .

Na representação da teoria acima utilizamos três predicados de anormalidade ($AB_1,AB_2,AB_3$) ao invés de um único predicado AB . Isto porque todos os objetos são anormais em alguns aspectos , e este fato não implica na anormalidade de cada objeto em todos os aspectos. No caso da teoria acima , se as anormalidades fossem representadas por um único predicado AB , ao assumir que , por exemplo , Tweety é anormal sob o aspecto 1 , estaríamos assumindo que Tweety também o é sob o aspecto 2 e sob aspecto 3 . Isto contraria o nosso objetivo de representar a anormalidade de Tweety apenas sob o aspecto 1 .

Ao circunscrever $AB_1$, $AB_2$, e $AB_3$ na teoria acima , deveríamos obter que $AB_2$ e $AB_3$ são identicamente falsos , uma vez que não existe nenhuma informação na teoria que nos diga o contrário , e que Tweety é o único objeto que satisfaz $AB_1$. Dessa forma , poderíamos concluir que Tweety voa.

Entretanto, não é isto o que ocorre. O problema surge porque quando circunscrevemos os predicados $AB_1$, $AB_2$ e $AB_3$ , o fazemos em paralelo. No caso em questão , podemos ter Tweety anormal tanto no aspecto 1 como no aspecto 3 . A resposta "natural" para a consulta "Tweety voa" pode ser obtida se preferíssemos que Tweety fosse anormal no aspecto 1 ao invés de o ser no aspecto 3.

Para tentar contornar esse problema, foi definida (McCARTHY (1986) e LIfSCHITZ(1985-a)) uma nova forma de circunscrição chamada de <u>circunscrição priorizada</u> . Para isto, tomamos a tupla de variáveis de predicado P da teoria T(P,Z) , e a dividimos em partes disjuntas P1,P2,...,PK . O objetivo é expressar que os predicados de P1 devem ser minimizados com prioridade maior que os de P2, os de P2 com prioridade maior que os de P3 , e assim por diante.

**OBSERVAÇÃO 6 :** A prioridade dos predicados de anormalidade (que representam exceções) deve ser maior quanto mais específicos forem os fatos à eles relacionados . Assim, se Tweety é um pingüim e uma ave, a anormalidade associada à pingüim ( que geralmente não voam ) deve ser minimizada com prioridade maior que a associada à pássaros (que geralmente voam).

Como iremos trabalhar com prioridades, ao invés de minimizar $(AB_1, AB_2, AB_3)$ com respeito a relação $\leq$ definida por :

$$(U_1, U_2, U_3) \leq (V_1, V_2, V_3) \Leftrightarrow U_1 \leq V_1 \wedge U_2 \leq V_2 \wedge U_3 \leq V_3$$

deve ser utilizada a relação $\leq'$ definida por :

$$(U_1, U_2, U_3) \leq' (V_1, V_2, V_3) \Leftrightarrow U_1 \leq V_1 \wedge (U_1 = V_1 \supset U_2 \leq V_2) \wedge (U_2 = V_2 \supset U_3 \leq V_3)$$

Assim, considerando que $p$ e $q$ representam, respectivamente, $p_1, \ldots, p_k$ e $q_1, \ldots q_k$ , onde $p_i$ e $q_i$ são tuplas de variáveis de predicado semelhantes à Pi , defina $p \leq' q$ como :

$$p \leq' q \Leftrightarrow \bigwedge_{i=1}^{k} \left( \bigwedge_{j=1}^{i-1} p_j = q_j \supset p_i \leq q_i \right) \qquad (*)$$

OBSERVAÇÃO 7 : $p <' q$ significa que $p \leq' q \wedge p \neq q$ .

**EXEMPLO 4.22**

Para P1 = $\{p_1\}$ , P2 =$\{p_2\}$ e P3 =$\{p_3\}$ , a expressão (*) fica da seguinte maneira :

$$p \leq' q \equiv \bigwedge_{i=1}^{3} ( \bigwedge_{j=1}^{i-1} p_j = q_j \supset p_i \leq q_i ) \equiv$$
$$\equiv p_1 \leq q_1 \wedge ( p_1 = q_1 \supset p_2 \leq q_2 ) \wedge$$
$$( p_1 = q_1 \wedge p_2 = q_2 \supset p_3 \leq q_3 )$$

Para k = 1 , a expressão (*) fica $p \leq' q \equiv p \leq q$ ■

Agora, podemos definir formalmente a circunscrição priorizada pela :

**DEFINIÇÃO 4.23** *(Circunscrição Priorizada)*

Sejam P uma tupla de constantes de predicado , Z uma tupla disjunta de P e T(P,Z) uma teoria. Além disso , seja $p \leq' q$ definida como (*) . A circunscrição de P em T(P,Z) com Z variando , em relação a $\leq'$ , denotada por $CIRC_{\leq'}(T(P,Z);P;Z)$ ou CIRC(T(P,Z);P1 > ... > Pk;Z) , é :

$$T(P,Z) \wedge \neg\exists\, p,z\ (T(p,z) \wedge p <' P)$$

Esta circunscrição é chamada de <u>*circunscrição priorizada*</u> .

Tal como no caso da circunscrição paralela , aqui também é possivel representar a noção de minimalidade da circunscrição priorizada em termos da teoria de modelos .

Para isto, vamos considerar a seguinte relação :

**DEFINIÇÃO 4.24** *(Relação $\leq$ mod(P;Z;$\leq$'))*

Sejam P uma tupla de constantes de predicado e Z uma tupla de constantes de função e/ou predicado , sendo P e Z disjuntos. Para quaisquer duas estruturas M e N , escrevemos M $\leq$ N mod(P;Z;$\leq$') se :

(i) $|M| = |N|$ ;

(ii) M[k] = N[k] , para toda constante k não pertencente à P U Z ;

(iii) Para todo j=1,...,k , se M[Pi] = N[Pi] para i < j, então M[Pj] $\subset$ N[Pj] .

**PROPOSIÇÃO 4.25** *(LIFSCHITZ(1985))*

Uma estrutura M é um modelo de CIRC(T(P,Z);P1 > ... > Pk;Z) se, e somente se, M é mínimo na classe dos modelos de T(P,Z) com relação a $\leq$ mod(P;Z;$\leq$') .

Quanto à satisfatibilidade da circunscrição priorizada, os seguintes resultados são generalizações, respectivamente, dos teoremas 4.10, 4.17 e 4.18 e da proposição 4.11, obtidos na seção *(ii)* :

**TEOREMA 4.26** *(LIFSCHITZ(1985))*

Se a teoria T(P,Z) é satisfatível e bem-fundamentada com respeito a (P1 > ... > Pk ;Z) , então CIRC(T(P,Z);(P1 > ... > Pk ;Z) é satisfatível .

TEOREMA 4.27 *(LIFSCHITZ(1985))*

Para quaisquer tuplas disjuntas P1,...Pk,Z de constantes de predicado distintos , toda teoria universal T(P,Z) é bem-fundamentada com respeito a (P1 > ... > Pk ;Z).

TEOREMA 4.28 *(LIFSCHITZ(1985))*

Para quaisquer tuplas disjuntas P1,...,Pk ,Z de constantes de predicado distintos , e qualquer teoria universal satisfatível T(P,Z) , a CIRC(T(P,Z);P1 > ... > Pk; Z) é satisfatível .

PROPOSIÇÃO 4.29 *(LIFSCHITZ (1985))*

Seja T(P,Z) uma teoria onde P é uma tupla de constantes de predicado e Z é uma tupla de constantes de função e/ou predicado, sendo P e Z disjuntos. Então :

$$CIRC_{\leq}(T(P,Z);P;Z) \equiv T(P,Z) \wedge CIRC_{\leq}(\exists z\ T(P,z);P) .$$

A relação entre as circunscrições paralela e a priorizada é tal que esta pode ser escrita como uma conjunção de circunscrições paralelas :

TEOREMA 4.30 *(LIFSCHITZ(1985))*

A circunscrição priorizada pode ser escrita como a conjunção de circunscrições paralelas da seguinte maneira :

$$CIRC(T(P,Z);P1 > \ldots > Pk;Z) \Leftrightarrow \bigwedge_{i=1}^{k} CIRC(T(P,Z);Pi;Pi+1 \cup \ldots \cup Pk \cup Z).$$

*iv) Circunscrição Ponto a Ponto*

A abordagem vista até aqui identifica um predicado com a sua extensão , ou seja, a interpretação de um símbolo de predicado n-ário num modelo é representada por um subconjunto de $U^n$ , onde U é o universo do modelo . Uma outra maneira de representar um predicado é através de uma função Booleana sobre $U^n$ , isto é , cada predicado é uma família dos elementos do conjunto {*false*, *true*}. A partir dessa nova perspectiva, LIFSCHITZ (1987-a) propôs uma outra noção de circunscrição , chamada ponto a ponto .

Segundo a semântica denotacional , um predicado n-ário é uma função da forma $\lambda x\ A(x)$ , onde x é um tupla de n variáveis e A(x) é uma fórmula . Caso $\mathbb{U}$ seja $\lambda x\ A(x)$ e t uma tupla de n termos , então o valor de $\mathbb{U}$ no ponto t (ou seja, $\mathbb{U}t$) representa A(t).

Para uma teoria T(P,Z), sendo P e Z uma tupla de constantes de predicado , vamos denotar $T(\lambda x(P(x) \wedge y \neq x), Z)$ como sendo uma teoria onde todas as ocorrências de P(x) em T(P,Z) são substituídas por $\lambda x(P(x) \wedge y \neq x)$. Por exemplo, para a teoria T(P) composta por $\forall x(P(x))$, $T(\lambda x(P(x) \wedge y \neq x))$ representará $\forall x(P(x) \wedge y \neq x)$. O significado desta mudança será discutido adiante.

**DEFINIÇÃO 4.31** *(Circunscrição Ponto a Ponto)*

Considere uma teoria T(P,Z) , onde P é uma constante de predicado e Z é uma tupla de constantes de predicado e/ou função. A *circunscrição de* P *em* T(P,Z) *com* Z *variando*, denotada por $C_P(T(P,Z);Z)$ , é dada pela expressão :

$$T(P,Z) \wedge \forall y,z \neg [\ P(y) \wedge T(\ \lambda x(P(x) \wedge y \neq x), z\ )] \qquad (*)$$

Quando Z for vazio , a circunscrição ponto a ponto de P em T(P), denotada por $C_P(T(P))$ , é dada pela expressão :

$$T(P) \wedge \forall y \neg [\ P(y) \wedge T(\ \lambda x(P(x) \wedge y \neq x)\ )] \qquad (**)$$

**OBSERVAÇÃO 8 :** O subscrito P em $C_P(T(P,Z);Z)$ e $C_P(T(P))$ indica o predicado que está sendo minimizado .

**OBSERVAÇÃO 9 :** A fórmula (**) é de primeira-ordem , enquanto (*) , pela presença da variável z, é geralmente uma fórmula de segunda-ordem .

**EXEMPLO 4.32**

Considere a teoria $T(P) \equiv \forall x\ (Q(x) \supset P(x))$ . Pela definição 4.31 , a circunscrição ponto a ponto de P em T(P) $C_P(T(P))$ é dada por :

$$T(P) \wedge \forall y \neg [P(y) \wedge T(\ \lambda x(P(x) \wedge y \neq x)\ )]$$

que é equivalente a :

$$T(P) \wedge \forall y \neg [P(y) \wedge \forall x( Q(x) \supset (P(x) \wedge y \neq x) )] \quad (\#)$$

A segunda parte da conjunção anterior equivale a :

$$\forall y [P(y) \supset \exists x( (Q(x) \wedge \neg P(x)) \vee (Q(x) \wedge x=y) ) \quad (\#\#)$$

Como $(Q(x) \wedge \neg P(x))$ é falso ( por $T(P)$ ) e $\exists x(Q(x) \wedge x=y)$ é equivalente a $Q(y)$, temos que (##) equivale a :

$$\forall y [P(y) \supset Q(y)]$$

Logo, $C_P(T(P))$ é equivalente a : $\forall x (Q(x) \equiv P(x))$ ■

**EXEMPLO 4.33**

Considere a teoria T(AB,VOA) composta por :

(1) $\forall x (AVE(x) \wedge \neg AB(x) \supset VOA(x))$

(2) $\forall x (PINGÜIM(x) \supset AVE(x))$

(3) $\forall x (PINGÜIM(x) \supset \neg VOA(x))$

A circunscrição ponto a ponto de AB em T(AB,VOA) com VOA variando, $C_{AB}(T(AB,VOA);VOA)$ é :

$$\begin{aligned}T(AB,VOA) \wedge \forall y,VOA' \neg [\ & AB(y) \wedge \\ & \forall x (AVE(x) \wedge \neg (AB(x) \wedge y\neq x) \supset VOA'(x)) \wedge \\ & \forall x (PINGÜIM(x) \supset AVE(x)) \wedge \\ & \forall x (PINGÜIM(x) \supset \neg VOA'(x)) \ ]\end{aligned}$$ ■

Um modelo M da expressão (*) (repectivamente (**)) é um modelo mínimo de T(P,Z) (resp. T(P)) que não pode ser transformado em outro modelo de T(P,Z) (resp. T(P)) pela mudança do valor de P de *true* para *false* em um ponto (o que representaria retirar um elemento da extensão do predicado circunscrito). Esta nova forma de circunscrição pode ser vista como uma conjunção infinita de condições de minimalidades locais , representada na definição 4.31 , tanto em (*) quanto em (**) , pelo quantificador $\forall x$ . A fórmula que segue este quantificador pode ser vista como a condição de minimalidade do valor de P no ponto x.

Um resultado importante é o que relaciona as circunscrições básica e paralela com a circunscrição ponto a ponto . Inicialmente , vamos estabelecer esse resultado considerando , como na definição 4.31 , que um único predicado está sendo circunscrito . Assim temos :

**PROPOSIÇÃO 4.34** *(LIFSCHITZ(1987-a))*

Considere uma teoria T(P,Z) , onde P é uma constante de predicado e Z é uma tupla de constantes de predicado e/ou função. Além disso, considere CIRC(T(P,Z);P;Z) e $C_P$(T(P,Z);Z) como sendo , respectivamente, as expressões (1) da definição 4.1 e (*) da definição 4.31 . Então :

$$CIRC(T(P,Z);P;Z) \text{ implica } C_P(T(P,Z);Z) .$$

Qunado Z for vazio , temos que :

$$CIRC(T(P);P) \text{ implica } C_P(T(P)) .$$

**PROVA** : Para demonstrar a proposição anterior, basta mostrar que :

$\forall p\ \neg(T(p) \wedge p < P)$ implica $\forall y\ \neg[P(y) \wedge T(\lambda x(P(x) \wedge y \neq x))]$

Para isto , vamos supor que :

$$P(yo) \wedge T(\lambda x(P(x) \wedge yo \neq x)) . \quad (*)$$

Fazendo $p = \lambda x(P(x) \wedge yo \neq x)$, segue que :

$\neg \{ T(\lambda x(P(x) \wedge yo \neq x)) \wedge \forall x((P(x) \wedge yo \neq x) \supset P(x)) \wedge \neg \forall x(P(x) \supset (P(x) \wedge yo \neq x)) \}$ que equivale a :

$\neg T(\lambda x(P(x) \wedge yo \neq x)) \vee \neg \forall x((P(x) \wedge yo \neq x) \supset P(x)) \vee \forall x(P(x) \supset (P(x) \wedge yo \neq x)) \}$

A segunda parte da disjunção acima é falsa pois é equivalente a $\exists x\ (P(x) \wedge yo \neq x \wedge \neg P(x))$.

A terceira parte equivale a : $\forall x\ (P(x) \supset yo \neq x)$.

Logo, temos : $\neg T(\lambda x(P(x) \wedge yo \neq x)) \vee \forall x\ (P(x) \supset yo \neq x)$ , o que contradiz a conjunção (*) ∎

A recíproca do teorema anterior nem sempre é veradeira. Por exemplo, considere a teoria composta por $P(a) \equiv P(b)$. Note que um modelo onde o predicado P seja identicamente falso satisfaz tanto CIRC(T;P) quanto $C_P(T)$. Entretanto, o modelo de T no qual P é verdadeiro em exatamente dois pontos (a e b) é um modelo de $C_P(T)$, mas não será um modelo da CIRC(T;P) , pois {a,b} não é mínimo ( { } ⊂ {a,b} ). O teorema seguinte estabelece quando a recíproca é verdadeira :

PROPOSIÇÃO 4.35 *(LIFSCHITZ(1987-a))*

Sejam T(P,Z), P , Z , CIRC(T(P,Z);P;Z) , $C_P$(T(P,Z);Z), CIRC(T(P);P) e $C_P$(T(P)) definidos como na proposição 4.34. Se todas as ocorrências de P em T(P,Z) (resp., T(P)) são positivas , então CIRC(T(P,Z);P;Z) é equivalente a $C_P$(T(P,Z);Z) (resp., CIRC(T(P);P) é equivalente a $C_P$(T(P))).

**PROVA :** Basta mostrar que :

$$\forall y\ \neg[P(y) \wedge T(\lambda x(P(x) \wedge y \neq x))] \text{ implica } \forall p\ \neg(T(p) \wedge p < P),$$

pois a outra parte foi provada na desmonstração anterior.

Suponha que $T(p) \wedge p < P$. Então existe $y_0$ tal que $P(y_0)$ e $\neg p(y_0)$.

Defina $p' = \lambda x(P(x) \wedge y_0 \neq x)$ . Temos então que $\forall y(p(y) \supset p'(y))$, ou seja, $\forall x(p(x) \supset (P(x) \wedge y_0 \neq x))$. Como todas as ocorrências de P em T(P) são positivas, segue que $T(p')$, o que contradiz $\forall y\ \neg[P(y) \wedge T(\lambda x(P(x) \wedge y \neq x))]$ ■

A partir da observação 9 e da proposição 4.35 (no caso de Z ser vazio) , temos um novo caso no qual a circunscrição básica se torna equivalente a uma fórmula de primeira-ordem :

**COROLARIO 4.36** *(LIFSCHITZ(1987-a))*

Sejam $T(P)$ uma teoria , $P$ uma constante de predicado e $CIRC(T(P);P)$ como na proposição 4.34. Se todas as ocorrências de $P$ em $T(P)$ são positivas então $CIRC(T(P);P)$ é equivalente a uma fórmula de primeira-ordem .

**EXEMPLO 4.37**

Considere a teoria $T(P)$ formada por $P(a)$. Temos que a $C_P(T(P))$ é equivalente a :

$$T(P) \wedge \forall y \neg[P(y) \wedge Pa \wedge y \neq a))] \equiv$$

$$T(P) \wedge \{ \forall y[P(y) \supset y=a] \vee \neg P(a) \} \equiv$$

$$T(P) \wedge \forall y[P(y) \supset y=a]$$

Como $P(a)$ é equivalente a $\forall y(y=a \supset P(y))$, a $C_P(T(P))$ equivale a : $\forall y[P(y) \equiv y=a]$, que por sua vez é equivalente a $CIRC(T(P);P)$ ■

Até agora, consideramos o caso em que um único predicado é circunscrito. Quando $P$ for uma tupla de predicados $P_1,\ldots,P_n$ , a circunscrição paralela $CIRC(T(P,Z);P;Z)$ pode ser escrita como a conjunção de circunscrições ponto a ponto , desde que a tupla de predicados $P$ só ocorra positivamente na teoria $T(P,Z)$. Para isto, vamos utilizar os seguintes resultados :

**PROPOSIÇÃO 4.38** *(LIFSCHITZ(1987-a))*

Sejam $T(P,Z)$ uma teoria , $P$ uma tupla de constantes de predicado e Z uma tupla de constantes de predicado e/ou função . Então :

$$CIRC(T(P,Z);P;Z) \text{ implica } CIRC(T(P,Z);P_i;Z) \quad (i=1,\ldots,n)$$

**PROPOSIÇÃO 4.39** *(LIFSCHITZ(1987-a))*

Sejam T(P,Z) , P e Z como na proposição 4.38. Se todas as ocorrências de $P = \langle P_1,\ldots,P_n\rangle$ em T(P,Z) são positivas , então :

$$CIRC(T(P,Z);P;Z) \text{ é equivalente a } \bigwedge_{i=1}^{n} CIRC(T(P,Z);P_i;Z)$$

Assim, a partir das proposições 4.35 e 4.39, temos :

**TEOREMA 4.40** *(LIFSCHITZ(1987-a))*

Sejam T(P,Z), P e Z definidos como na proposição 4.38. Se todas as ocorrências de $P = \langle P_1,\ldots,P_n\rangle$ em T(P,Z) são positivas , então :

$$CIRC(T(P,Z);P;Z) \text{ é equivalente a } \bigwedge_{i=1}^{n} C_{P_i}(T(P,Z);Z) .$$

**EXEMPLO 4.41**

Considere a teoria T(P) formada por $P_1(a) \wedge P_2(b)$ . Como $P_1$ e $P_2$ só ocorrem positivamente em T(P) , podemos utilizar a proposição 4.39, fazendo Z vazio, para obter que a CIRC(T(P);P) equivale a :

$$CIRC(T(P);P_1) \land CIRC(T(P);P_2) \Leftrightarrow \forall x\ (P_1(x) \equiv x=a) \land \forall x\ (P_2(x) \equiv x=b)\ (*).$$

Utilizando o teorema 4.40 , temos que a $CIRC(T(P);P)$ equivale a :

$$C_{P_1}(T(P)) \land C_{P_2}(T(P)) \Leftrightarrow$$

$$\Leftrightarrow T(P) \land \forall y\ \neg\ [P_1(y) \land P_1(a) \land y \neq a \land P_2(b)] \land \forall y\ \neg\ [P_2(y) \land P_2(b) \land y \neq b \land P_1(a)]\ ,$$

que é equivalente a :

$$T(P) \land \forall y\ (P_1(y) \supset y=a) \land \forall y\ (P_2(y) \supset y=b)$$

Como $P_1(a)$ é equivalente a $\forall y(y=a \supset P_1(y))$ e $P_2(b)$ é equivalente a $\forall y(y=b \supset P_2(y))$, temos que a $C_{P_1}(T(P)) \land C_{P_2}(T(P))$ equivale a $(*)$ ■

O corolário 4.36 pode ser estendido para o caso de P ser uma tupla de constantes de predicado. Dessa forma , tanto o corolário 4.36 , quanto o corolário 4.42 (abaixo) , determinam quais os tipos de teorias que quando são circunscritas apresentam como resultado fórmulas de primeira-ordem.

COROLARIO 4.42 *(LIFSCHITZ(1987-a))*

Sejam $T(P)$ uma teoria e P uma tupla de constantes de predicado. Considere $A(P)$ e $B(P)$ , tais que $T(P) \equiv A(P) \land B(P)$ , onde $P_1,\ldots,P_n$ só ocorrem positivamente em $A(P)$ e negativamente em $B(P)$. Então, a $CIRC(A(P) \wedge B(P);P)$ é equivalente a uma fórmula de primeira-ordem.

**EXEMPLO 4.43**

Considere a teoria $T(P)$, onde $P = \langle P_1, P_2 \rangle$, composto por :

(1) $\forall x\ (Q(x) \supset P_1(x)) \Leftrightarrow A(P)$ ;

(2) $\forall x\ (\neg P_2(x)) \Leftrightarrow B(P)$ .

Temos que:

$$CIRC(T(P);P) \equiv CIRC(A(P) \wedge B(P);P) \equiv CIRC(A(P);P) \wedge B(P).$$

Pela proposição 4.35, $CIRC(A(P);P) \equiv C_{P_1}(A(P)) \equiv A(P) \wedge \forall y\ \neg[P_1(y) \wedge \forall x(Q(x) \supset (P_1(x) \wedge y \neq x))]$, que é uma fórmula de primeira-ordem.

Como $B(P)$ também é de primeira-ordem, obtemos :

$$CIRC(T(P);P) \equiv A(P) \wedge B(P) \wedge \forall y\ \neg[P_1(y) \wedge \forall x(Q(x) \supset (P_1(x) \wedge y \neq x))]$$

é uma fórmula de primeira-ordem ■

*iv.1) Prioridade na Circunscirção Ponto a Ponto*

Para trabalhar com prioridades na circunscrição ponto a ponto, utiliza-se simplesmente dois resultados vistos anteriormente . O primeiro deles é o obtido no teorema 4.30, onde :

$$CIRC(T(P,Z);P1 > \ldots > Pk\ ;Z) \equiv \bigwedge_{i=1}^{k} CIRC(T(P,Z);Pi;Pi+1 \cup \ldots \cup Pk \cup Z)$$

$$\equiv CIRC(T(P,Z);P1;P2 \cup \ldots \cup Pk \cup Z) \wedge$$
$$CIRC(T(P,Z);P2;P3 \cup \ldots \cup Pk \cup Z) \wedge$$
$$\wedge \ldots \wedge$$
$$CIRC(T(P,Z);Pk;Z) \qquad (1)$$

Aplicando a cada termo de (1) a proposição 4.35 ( supondo obviamente que cada Pi além de só ocorrer positivamente em T(P,Z) , possui uma única constante de predicado) , obtemos :

$$CIRC(T(P,Z);P1 > \ldots > Pk\ ;Z) \equiv \bigwedge_{i=1}^{k} C_{Pi}\ (T(P,Z);Pi+1 \cup \ldots \cup Pk \cup Z)$$

$$\equiv C_{P1}\ (T(P,Z);P2 \cup \ldots \cup Pk \cup Z) \wedge$$
$$C_{P2}\ (T(P,Z);P3 \cup \ldots \cup Pk \cup Z) \wedge$$
$$\wedge \ldots \wedge$$
$$C_{Pk}\ (T(P,Z);Z) \qquad (2)$$

No caso de cada $Pi = \langle Pi_1, \ldots ,Pi_{ni} \rangle$ , ou seja, possuir ni constantes de predicado , devemos utilizar o teorema 4.40 na conjunção (1) , o que nos dará :

$$CIRC(T(P,Z);P1 > \ldots > Pk\ ;Z) \equiv \bigwedge_{i=1}^{k} [ \bigwedge_{j=1}^{ni} C_{Pi_j}(T(P,Z);Pi+1 \cup \ldots \cup Pk \cup Z)$$

$$\equiv [ \bigwedge_{j=1}^{n1} C_{P1_j}(T(P,Z);P2 \cup \ldots \cup Pk \cup Z) ] \wedge$$

$$[\bigwedge_{j=1}^{n2} C_{P2_j}(T(P,Z);P3 \cup \ldots \cup Pk \cup Z)] \wedge \ldots \wedge$$

$$[\bigwedge_{j=1}^{nk} C_{Pk_j}(T(P,Z);Z)] \qquad (3)$$

**EXEMPLO 4.44**

Considere a teoria T(P) , onde P = $\langle Q_1,Q_2 \rangle$ , composta por : $\forall x\ (Q_1(x) \vee Q_2(x))$ . Fazendo P1 = $\langle Q_1 \rangle$ e P2 = $\langle Q_2 \rangle$ , a expressão (2) fica sendo :

$$CIRC(T(P);P1 > P2) \equiv C_{P1}(T(P);P2) \wedge C_{P2}(T(P))$$
$$\equiv T(P) \wedge \forall y,q_2\ \neg[Q_1(y) \wedge \forall x((Q_1(x) \wedge y\neq x) \vee q_2(x))] \wedge$$
$$\forall y\ \neg[Q_2(y) \wedge \forall x(Q_1(x) \vee (Q_2(x) \wedge y\neq x))]$$

Se acrescentarmos a teoria T(P) o axioma $\forall x\ Q_3(x)$ , e fizéssemos P1 = $\langle Q_1,Q_2 \rangle$ e P2 = $\langle Q_3 \rangle$ , obteríamos pela expressão (3) :

$$CIRC(T(P);P1 > P2) \equiv [\bigwedge_{j=1}^{2} C_{P1_j}(T(P);P2)] \wedge$$
$$[\bigwedge_{j=1}^{1} C_{P2_j}(T(P))] \equiv$$
$$\equiv C_{P1_1}(T(P);P2) \wedge C_{P1_2}(T(P);P2) \wedge$$
$$C_{P2_1}(T(P)) \equiv$$
$$\equiv C_{Q1}(T(P);Q_3) \wedge C_{Q2}(T(P);Q_3) \wedge C_{Q3}(T(P)) \quad \blacksquare$$

v) *Circunscrição Suportada*

A circunscrição suportada foi definida por YOU(1989) com o objetivo de introduzir na circunscrição o conceito de modelos suportados, que é utilizado na programação em lógica na definição de uma semântica para uma classe especial de programas . Aqui, vamos nos restringir a definição deste novo tipo de circunscrição, não examinando a sua relação com a programação em lógica .(ver YOU(1989))

Vamos considerar que uma *cláusula orientada* é uma cláusula escrita na forma :

$$A_1 \wedge \ldots \wedge A_n \rightarrow B_1 \vee \ldots \vee B_m$$

onde $n,m \geq 0$ e para $i \in [1,n]$ e $j \in [1,m]$ , $A_i$ e $B_j$ são literais. A *cabeça* de uma cláusula orientada é a disjunção dos literais $B_j$ e o *corpo* é a conjunção dos $A_i$ . Quando $n = 0$ , a cláusula orientada é chamada de fato . Caso contrário , ela é uma regra .

Um modelo M é dito suportado se cada átomo pertencente a M ou é um fato ou é uma instância básica de um dos literais que formam a cabeça de uma regra, cujo corpo é *true* em M.

O que é pretendido ao se transportar a noção de suporte para a circunscrição é impor a idéia de que todo átomo de um modelo pode ser derivável a partir dos fatos conhecidos, seguindo a orientação das regras. Formalmente :

**DEFINIÇÃO 4.45**

Sejam T(P,Z) um conjunto de cláusulas orientadas e P e Z tuplas disjuntas de constantes de predicado. Uma estrutura de Herbrand M é um modelo suportado da CIRC(T(P,Z);P;Z) se, e somente se, M é um modelo de CIRC(T(P,Z);P;Z) e para qualquer átomo $q(t_1,\ldots,t_n)$ em M tal que $q \in P$, existe uma cláusula orientada :

$$A_1 \wedge \ldots \wedge A_n \rightarrow B_1 \vee \ldots \vee B_m$$

e uma substituição $\theta$ tal que $\theta B_i \equiv q(t_1,\ldots,t_n)$ para algum $i$ , e $\theta A_1 \wedge \ldots \wedge \theta A_n$ é true em M .

**OBSERVAÇÃO 10** : Como nem todos os modelos suportados são mínimos , é necessário que as estruturas da definição 4.45 sejam modelos de CIRC(T(P,Z) ;P;Z).

**OBSERVAÇÃO 11** : Para estender este ( e os outros ) resultado para modelos quaisquer , basta acrescentar as hipóteses do nome único, o fecho sob o domínio e o axioma da igualdade.

**EXEMPLO 4.46**

Considere a teoria T composta por :

(1) PINGÜIM(Tweety)
(2) PINGÜIM(Tweety) $\rightarrow$ AB(Tweety) $\vee$ ¬VOA(Tweety) ,

onde (1) representa que Tweety é um pingüim e (2) que se Tweety é um pingüim , ou é anormal ou não voa. A circunscrição de AB em T com VOA variando possui um modelo

(AB,VOA)-mínimo : $M_1$ = {PINGÜIM(Tweety)} . Este modelo é suportado pois para ¬VOA(Tweety) existe uma cláusula orientada , no caso (2) , tal que PINGÜIM(Tweety) é *true* em M , e existe um átomo na cabeça da cláusula equivalente a ¬ VOA(Tweety) ■

Observe que se M é um modelo de CIRC(T(P,Z);P;Z), onde T(P,Z) é uma teoria cujas cláusulas orientadas são disjunções de literais , e P e Z são quaisquer tuplas de constantes de predicado , M é um modelo (P,Z)-mínimo. Para algum $q(t_1,\ldots,t_n) \in M$ e $q \in P$ , existe uma cláusula orientada cuja satisfatibilidade depende de $q(t_1,\ldots,t_n)$ ser *true*. Logo, $q(t_1,\ldots,t_n)$ é suportada por esta cláusula, o que acarreta que a CIRC(T;P;Z) é suportada. Assim, para teorias clausais de primeira-ordem, a circunscrição suportada é pelo menos tão poderosa quanto a circunscrição padrão.

Existe um problema quando as cláusulas são escritas na forma de implicação : alguns modelos da teoria original podem se transformar em não-modelos , uma vez que alguns de seus átomos podem se tornar não-deriváveis.

**EXEMPLO 4.47**

Considere a teoria composta por uma única cláusula :

$$P(a) \vee P(b)$$

Se escrevermos $\neg P(a) \rightarrow P(b)$ , o único modelo suportado de CIRC(T;P) é {P(b)} , enquanto se escrevermos

$\neg P(b) \rightarrow P(a)$ , CIRC(T;P) só possuirá o modelo {P(a)} ■

O efeito geral de escrever cláusulas na forma de implicação é descrito pela seguinte proposição :

**PROPOSIÇÃO 4.48** *(YOU(1989))*

Seja T um conjunto de cláusulas orientadas com uma cláusula da forma:

$$(i) \quad A_1 \wedge \ldots \wedge A_n \rightarrow B_1 \vee \ldots \vee B_m \vee C$$

onde $A_i, B_j$ e C são literais. Seja T' o conjunto de cláusulas orientadas obtidas de T pela substituição de (i) por :

$$(ii) \quad A_1 \wedge \ldots \wedge A_n \wedge \neg C \rightarrow B_1 \vee \ldots \vee B_m .$$

Então , para qualquer P e Z , temos :

(1) Todo modelo suportado de CIRC(T';P;Z) é um modelo suportado de CIRC(T;P;Z) ;

(2) Se m=0 e C é um literal negativo , então M é um modelo suportado de CIRC(T;P;Z) se, e somente se, M é um modelo suportado de CIRC(T';P;Z).

*vi) Conclusão*

Neste capítulo apresentamos uma série de tipos de circunscrição (paralela, priorizada, ponto a ponto e suportada) que se inter-relacionam : a priorizada pode ser escrita como uma conjunção de circunscrições paralelas; a ponto a ponto equivale a paralela quando os predicados circunscritos só ocorrem positivamente na teoria; os modelos da suportada são modelos da circunscrição paralela que são suportados pelos fatos.

Além disso, procurou-se determinar classes de teorias (universal e quase-universal) nas quais a aplicação do método não transformasse teorias consistentes em inconsistentes.

Note que em todas elas (exceto em casos especiais) obtemos uma teoria de segunda-ordem. No próximo capítulo apresentaremos métodos que nos permitirão determinar quando um fato segue ou não de uma teoria circunscrita .

# CAPITULO V

# TRABALHANDO COM A CIRCUNSCRIÇÃO

No capítulo anterior , apresentamos a definição formal da circunscrição e uma série de outras formas de circunscrição. Por normalmente a teoria circunscrita ser de segunda-ordem , é difícil utilizar um provador padrão de teoremas na derivação das consequências lógicas de tais teorias.

Neste capítulo examinaremos três abordagens que nos permitem determinar quando uma fórmula segue ou não de uma teoria circunscrita. Os métodos que veremos são utilizados no caso das circunscrições paralela e priorizada.

Na seção (i) , uma classe de teorias nas quais o axioma de circunscrição pode ser transformado para primeira-ordem é caracterizada. Nestes casos, podemos depois utilizar um provador padrão.

Nas seções (ii) e (iii) , apresentaremos métodos que determinam as consequências lógicas de teorias circunscritas em geral.

## *i) Fórmulas Separáveis e a Z-Resolução*

### *i.1) Paralela*

Esta primeira abordagem, proposta por LIFSCHITZ (1985-a), procura caracterizar uma classe de teorias cujos axiomas de circunscrição, correspondentes a cada uma delas, podem ser transformados em axiomas equivalentes de primeira-ordem.

Inicialmente, Lifschitz observou que dois tipos de teorias podiam ser circunscritas de uma maneira bem simples, sendo a resultante expressa em termos de primeira-ordem. São elas :

<u>CASO 1</u> : Teorias onde o predicado P não ocorre na forma positiva . Por exemplo , para a teoria $T(P) \Leftrightarrow \neg P(a)$ , obtemos que $CIRC(T;P) \Leftrightarrow \forall x \neg P(x)$, visto que, como não possuímos qualquer informação positiva a respeito de P, devemos considerar que nenhum elemento conhecido possui tal propriedade.

<u>CASO 2</u> : Teorias da forma $U \leq P$ , onde U não contém P. Por exemplo, a teoria $T(P) \Leftrightarrow Q \leq P \Leftrightarrow \forall x\, (Q(x) \supset P(x))$, onde obtemos que $CIRC(T;P) \Leftrightarrow \forall x\, (Q(x) \equiv P(x)) \Leftrightarrow (Q = P)$ , pois os elementos conhecidos que possuem a propriedade P são os mesmos, e só aqueles , que possuem a propriedade Q.

OBSERVAÇÃO 1 : Estes são caso particulares do corolário 4.42 ( capítulo IV - seção (iv) ) .

A partir dessas observações, Lifschitz definiu dois tipos de fórmulas :

**DEFINIÇÃO 5.1** *(Fórmulas Solitárias)*

Seja $P = \langle P_1, \ldots, P_m \rangle$ uma tupla de constantes de predicado. Dizemos que a fórmula $A(P)$ é *solitária com respeito a P*, caso ela seja da forma :

$$N(P) \wedge (U \leq P) \quad (a)$$

onde $N(P)$ é uma fórmula que não contém ocorrências positivas de $P$ , e $U$ é uma tupla de constantes de predicado que não contém $P_1, \ldots, P_m$ .

**DEFINIÇÃO 5.2** *(Fórmulas Separáveis)*

Uma fórmula $A(P)$ é chamada de *separável* quando ela for equivalente a disjunção de fórmulas solitárias , ou seja , pode ser escrita da forma :

$$\bigvee_i \left[ N_i(P) \wedge (U_i \leq P) \right] \quad (b)$$

onde $N_i(P)$ é uma fórmula que não contém ocorrências positivas de $P$ e cada $U_i$ é uma tupla de constantes de predicado que não contém $P_1, \ldots, P_m$ .

**EXEMPLO 5.3**

A fórmula

$$A(P) \Leftrightarrow [\, \neg P(a) \wedge \forall x\, (Q(x) \supset P(x)) \,] \vee$$
$$[\, \neg P(b) \wedge \forall x\, (R(x) \supset P(x)) \,]$$

é uma fórmula separável pois é composta pela disjunção das fórmulas solitárias :

$$\neg\ P(a) \wedge \forall x\ (Q(x) \supset P(x)) \quad \text{e}$$

$$\neg\ P(b) \wedge \forall x\ (R(x) \supset P(x)) \ \blacksquare$$

Considere a fórmula $A(P)$ como sendo a conjunção de todos os axiomas de uma teoria $T(P)$, onde $P = \langle P_1, \ldots, P_m \rangle$ é uma tupla de constantes de predicado. Quando $A(P)$ for uma fórmula solitária, ou seja, $A(P)$ equivale a $(a)$, a circunscrição da teoria $T(P)$ com relação a $P$ será dada por:

**TEOREMA 5.4** *(LIFSCHITZ(1985))*

$$CIRC(N(P) \wedge (U \leq P);P) \Leftrightarrow N(U) \wedge (U = P) \ .$$

No caso de $A(P)$ ser da forma $(b)$, ou seja, ser uma fórmula separável, vale o seguinte resultado :

**TEOREMA 5.5** *(LIFSCHITZ(1985))*

$$CIRC(\bigvee_i [N_i(P) \wedge (U_i \leq P)];P) \Leftrightarrow \bigvee_i [D_i \wedge (U_i = P)] \ ,$$

onde :

$$D_i \Leftrightarrow N_i(U_i) \wedge \bigwedge_{j \neq i} \neg\ [N_j(U_j) \wedge (U_j < U_i)].$$

**EXEMPLO 5.6**

Considere a teoria $T(P)$ tal que :

$$A(P) \Leftrightarrow [\ \neg P(a) \wedge \forall x\ (Q(x) \supset P(x))\ ] \vee$$
$$[\ \neg P(b) \wedge \forall x\ (R(x) \supset P(x))\ ].$$

Aplicando o teorema 5.5, temos que :

$$CIRC(A(P);P) \Leftrightarrow [D_1 \wedge (Q = P)] \vee [D_2 \wedge (R = P)], \text{ onde:}$$
$$D_1 \Leftrightarrow \neg\, Q(a) \wedge \neg\, [\ \neg\, R(b) \wedge (R < Q)]\ \text{e}$$
$$D_2 \Leftrightarrow \neg\, R(b) \wedge \neg\, [\ \neg\, Q(a) \wedge (Q < R)] \quad \blacksquare$$

Os teoremas 5.4 e 5.5 valem quando a circunscrição é básica. Uma maneira de estender estes resultados para a circunscrição paralela é utilizando a proposição 4.11 (capítulo IV - seção (ii)) . Infelizmente , tal proposição introduz um novo quantificador de segunda-ordem e com ele mais dificuldades.

Para solucionar este problema , YUAN (1988) definiu um processo, chamado Z-resolução que tem como objetivo caracterizar uma classe de teorias de primeira-ordem, na qual é possível transformar a circunscrição paralela de tais teorias em básicas, sem que se introduza um novo quantificador de segunda-ordem. Esta classe será composta pelas teorias ditas <u>livre de Z-conflito</u> .

O processo de Z-resolução inicialmente transforma todas as ocorrências negativas dos predicados da tupla Z em ocorrências positivas. Para isto , vamos considerar que os axiomas de uma dada teoria $T(P,Z)$ , onde P e Z são tuplas disjuntas de constantes de predicado, estão na forma de cláusulas, que $\Phi$ é uma das cláusulas de $T(P,Z)$, que $\neg\mathcal{Z}$ é um literal negativo de $\Phi$ e que $A'(\Phi,\mathcal{Z})$ é o conjunto de todos os Z-resolventes de $\Phi$ com cada uma das cláusulas de $T(P,Z)$ que contém ocorrências positivas dos predicados de Z. O conjunto de Z-resolução $\mathcal{R}(T(P,Z),\Phi,\mathcal{Z})$ é definido como sendo $\mathcal{A}'(\Phi,\mathcal{Z}) \cup ( T(P,Z) / \Phi )$, ou seja, é o conjunto de todos os Z-resolventes de $\Phi$ mais a teoria $T(P,Z)$ sem a cláusula $\Phi$.

**EXEMPLO 5.7**

Considere a teoria $T(AB,VOA)$ do exemplo 4.2 , que é composta pelos seguintes axiomas na forma clausal :

(1) $\neg$ AVE(x) v AB(x) v VOA(x)

(2) $\neg$ PINGÜIM(x) v AVE(x)

(3) $\neg$ PINGÜIM(x) v $\neg$ VOA(x)

Lembrando que P = <AB> e Z = <VOA> , vamos transformar a ocorrência negativa do predicado VOA em ocorrência positiva . Para isto , tome :

$\Phi$ = { $\neg$ PINGÜIM(x) v $\neg$ VOA(x) }, que é a cláusula que possui ocorrência negativa do predicado VOA ;

$\neg\, \mathcal{Z} \Leftrightarrow \neg$ VOA(x) , que é o literal negativo de $\Phi$ a ser eliminado ;

$\mathcal{A}'(\Phi,\mathcal{Z})$ = {$\neg$ AVE(x) v AB(x) v $\neg$ PINGÜIM(x)} , obtido pela resolução das cláusulas (1) e (3). Como esse conjunto só possui um elemento, vamos chamá-lo de (4).

O conjunto de Z-resolução obtido será composto pela cláusulas (1) , (2) e (4) , já que :

$\mathcal{R}(T(P,Z),\Phi,\mathcal{Z}) = \mathcal{A}'(\Phi,\mathcal{Z}) \cup (T(P,Z) / \Phi) =$

= { $\neg$ AVE(x) v AB(x) v $\neg$ PINGÜIM(x)}

$\cup$ { $\neg$ AVE(x) v AB(x) v VOA(x) ,

$\neg$ PINGÜIM(x) v AVE(x)}

Esta nova teoria formada pelas cláusulas (1) , (2) e (4) não possui ocorrências negativas de Z = <VOA> ∎

**EXEMPLO 5.8 (YUAN(1988))**

Considere a teoria T(P,Z) , onde P = $\langle P_1, P_2 \rangle$ e Z = $\langle Z_1, Z_2 \rangle$, formada pelas cláusulas :

(1) $\neg\ Q_1(x,y)$ v $Z_1(x,y)$ v $Z_2(x,y)$

(2) $\neg\ Z_2(x,y)$ v $Q_2(x,y)$ v $P_1(x,y)$

(3) $\neg\ Z_1(x,y)$ v $\neg\ Z_1(y,z)$ v $P_2(x,z)$

Note que os predicados $Z_1$ e $Z_2$ ocorrem negativamente na teoria T(P,Z) , e por isso vamos transformar tais ocorrências em positivas.

Inicialmente , tome :

$\Phi$ = { $\neg\ Z_1(x,y)$ v $\neg\ Z_1(y,z)$ v $P_2(x,z)$ }, que é a cláusula onde $Z_1$ ocorre negativamente ;

$\neg\ \mathcal{Z} \Leftrightarrow \neg\ Z_1(x,y)$ , que é uma das ocorrências negativas do predicado $Z_1$ em $\Phi$ .

Pela resolução das cláusulas (1) e (3) , obtemos :

$\mathcal{A}'(\Phi,\mathcal{Z})$ = { $\neg Q_1(x,y)$ v $Z_2(x,y)$ v $\neg\ Z_1(y,z)$ v $P_2(x,z)$ } , cuja única cláusula vamos chamar de (4).

O conjunto de Z-resolução será composto pelas cláusulas (1) ,(2) e (4), já que :

$$\mathcal{R}(T(P,Z),\Phi,\mathcal{Z}) = \mathcal{A}'(\Phi,\mathcal{Z}) \cup (T(P,Z) \,/\, \Phi) =$$
$$= \{ \neg Q_1(x,y) \vee Z_2(x,y) \vee \neg Z_1(y,z) \vee P_2(x,z) \} \cup$$
$$\{ \neg Q_1(x,y) \vee Z_1(x,y) \vee Z_2(x,y) \,,$$
$$\neg Z_2(x,y) \vee Q_2(x,y) \vee P_1(x,y) \} =$$
$$= T'(P,Z) \,.$$

Como a cláusula (4) ainda possui ocorrência negativa de $Z_1$ , vamos considerar agora :

$\Phi = \{ \neg Q_1(x,y) \vee Z_2(x,y) \vee \neg Z_1(y,z) \vee P_2(x,z) \}$, que é a cláusula (4) ;

$\neg \mathcal{Z} \Leftrightarrow \neg Z_1(y,z)$, que é a ocorrência negativa do predicado $Z_1$ em $\Phi$ .

Logo,

$$\mathcal{A}'(\Phi,\mathcal{Z}) = \{ \neg Q_1(x,y) \vee \neg Q_1(z,x) \vee P_2(z,y) \vee Z_2(z,x) \vee Z_2(x,y) \}$$

, que é o resolvente das cláusulas (1) e (4) , e vamos chamá-lo de (5).

O conjunto de Z-resolução obtido é composto pelas cláusulas (1) , (2) e (5) , pois :

$$\mathcal{R}(T'(P,Z),\Phi,\mathcal{Z}) = \mathcal{A}'(\Phi,\mathcal{Z}) \cup (T(P,Z) \,/\, \Phi) =$$
$$= \{ \neg Q_1(x,y) \vee \neg Q_1(z,x) \vee P_2(z,y) \vee Z_2(z,x) \vee Z_2(x,y) \} \cup$$
$$\{ \neg Q_1(x,y) \vee Z_1(x,y) \vee Z_2(x,y) \,,$$
$$\neg Z_2(x,y) \vee Q_2(x,y) \vee P_1(x,y) \} =$$
$$= T''(P,Z) \,.$$

A teoria $T''(P,.Z)$ obtida não possui mais nenhuma outra ocorrência negativa de $Z_1$ , mas possui ainda uma ocorrência negativa de $Z_2$ ( cláusula (2) ) que precisa ser transformada em ocorrência positiva.

Antes , observe que na teoria $T''(P,Z)$, $Z_1$ ocorre (positivamente) somente na cláusula (1) . Como na circunscrição os predicados da tupla Z podem variar , podemos considerar o predicado $Z_1$ verdadeiro, sem que isto afete o resultado da circunscrição . Assim, a cláusula (1) será verdadeira , independente dos outros predicados , e podemos , portanto , retirá-la da teoria $T''(P,Z)$, que agora passa a ser composta pelas cláusulas (2) e (5) , e denotada por $T''(P,Z')$ , onde $Z' = \langle Z_2 \rangle$ .

Agora , vamos transformar as ocorrência negativas de $Z_2$ , na teoria $T''(P,Z')$, em ocorrências positivas . Para isso, vamos fazer :

$\Phi = \{ \neg Z_2(x,y) \vee Q_2(x,y) \vee P_1(x,y) \}$ , que é a cláusula (2) da teoria $T''(P,Z')$ onde $Z_2$ ocorre negativamente ;

$\neg \mathcal{Z} = \neg Z_2(x,y)$ , que é a ocorrência negativa de $Z_2$ em $\Phi$ .

Daí, segue que :

$A'(\Phi,\mathcal{Z}) = \{ \neg Q_1(x,y) \vee \neg Q_1(z,x) \vee P_2(z,y) \vee Q_2(z,x) \vee P_1(z,x) \vee Q_2(x,y) \vee P_1(x,y) \}$ , que é o resolvente das cláusulas (2) e (5) , e que será chamado de cláusula (6) ; e

$$R(T''(P,Z'),\Phi,\mathcal{Z}) = A'(\Phi,\mathcal{Z}) \cup ( T''(P,Z') / \Phi ) =$$
$$= \{ \neg Q_1(x,y) \vee \neg Q_1(z,x) \vee P_2(z,y) \vee Q_2(z,x) \vee P_1(z,x) \vee Q_2(x,y) \vee P_1(x,y) \}$$
$$\cup \{ \neg Q_1(x,y) \vee \neg Q_1(z,x) \vee P_2(z,y) \vee Z_2(z,x) \vee Z_2(x,y) \}$$
, que é o conjunto de Z-resolução, formado pelas cláusulas (5) e (6).

Observe que agora $T''(P,Z')$ não possui mais nenhuma ocorrência negativa de $Z_2$ ∎

Nos exemplo 5.7 e 5.8 , o processo de Z-resolução foi bem sucedido , uma vez que em ambas as teorias conseguimos transformar as ocorrências negativas dos predicados de Z em ocorrências positivas. Com isso, a circunscrição paralela pode ser reduzida a circunscrição básica como consequência do seguinte resultado :

**TEOREMA 5.9** *(YUAN(1988))*

Sejam $T(P,Z)$ uma teoria de primeira-ordem , $z$ um subconjunto de constantes de predicado de $Z$ , tal que os elementos de $z$ só ocorrem positivamente em $T(P,Z)$ , e $T'(P,Z')$ a teoria resultante do processo de Z-resolução menos os axiomas que contenham algum dos predicados pertencentes a $z$ . Então :

$$CIRC(T(P,Z);P;Z) \Leftrightarrow T(P,Z) \wedge CIRC(T'(P,Z');P;Z') ,$$

onde $Z' = Z / z$ , ou seja, a tupla de predicados $Z$ menos aqueles que só ocorrem positivamente na teoria $T(P,Z)$.

Quando $z = Z$ na definição acima , a circunscrição paralela fica reduzida a básica , ou seja ,

$$CIRC(T(P,Z);P;Z) \Leftrightarrow T(P,Z) \wedge CIRC(T'(P);P)$$

**CONTINUAÇÃO DO EXEMPLO 5.7**

No caso do exemplo 5.7, teremos pelo teorema anterior que :

$$\mathrm{CIRC}(T(AB,VOA);AB;VOA) \Leftrightarrow T(AB,VOA) \wedge \mathrm{CIRC}(T'(AB);AB),$$

onde $T'(AB) = \{ (\neg\, \mathrm{AVE}(x) \vee \mathrm{AB}(x) \vee \neg\, \mathrm{PINGÜIM}(x)) \wedge$
$(\neg\, \mathrm{PINGÜIM}(x) \vee \mathrm{AVE}(x)) \} =$

Tomando $A(AB)$ como a conjunção que forma a teoria $T'(AB)$, obtemos que $A(AB)$ é equivalente a :

$$(\neg\, \mathrm{PINGÜIM}(x) \vee \mathrm{AVE}(x)) \wedge ( (\mathrm{AVE}(x) \wedge \mathrm{PINGÜIM}(x)) \supset \mathrm{AB}(x) ),$$

que é uma fórmula solitária (def. 5.1). Logo, pelo teorema 5.4, temos :

$$\mathrm{CIRC}(A(AB);AB) \Leftrightarrow (\neg\, \mathrm{PINGÜIM}(x) \vee \mathrm{AVE}(x)) \wedge ( \mathrm{AB}(x) \equiv \mathrm{AVE}(x) \wedge \mathrm{PINGÜIM}(x) ).$$

Como $(\neg\, \mathrm{PINGÜIM}(x) \vee \mathrm{AVE}(x))$ pertence a teoria $T(AB,VOA)$ (axioma (2)) , temos que :

$$\mathrm{CIRC}(T(AB,VOA);AB;VOA) \Leftrightarrow T(AB,VOA) \wedge (AB = \mathrm{PINGÜIM} \wedge \mathrm{AVE}) ;$$

resultado idêntico ao obtido no exemplo 4.2 ∎

CONTINUAÇÃO DO EXEMPLO 5.8

No exemplo 5.8, temos que :

$$CIRC(T(P,Z);P;Z) \Leftrightarrow T(P,Z) \wedge CIRC(T'(P);P) ,$$

onde T'(P) é formada somente pela cláusula (6) , pois a cláusula (5) possui uma ocorrência positiva de $Z_2$. Observe que as cláusulas onde $Z_1$ ocorre positivamente foram eliminadas no decorrer do processo de Z-resolução, o que é possível graças ao teorema 5.9 ■

A aplicação do teorema 5.9 só é possível quando o processo de Z-resolução for bem-sucedido. Isto só é garantido para uma classe de teorias , que são chamadas livres de Z-conflito.

Considere dois predicados $Z_i$ e $Z_j$ pertencentes a Z. Dizemos que $Z_i$ deriva $Z_j$ , representado por $Z_i \rightarrow Z_j$ , se existir uma cláusula C na teoria T(P,Z) tal que $Z_i$ ocorre negativamente em C e $Z_j$ ocorre positivamente em C. A relação binária => sobre Z é definida da seguinte forma : para $Z_i,Z_j \in Z$, $Z_i \Rightarrow Z_j$ se, ou $Z_i \rightarrow Z_j$, ou existe um predicado $Z_k \in Z$ e duas cláusulas $C_1$ e $C_2$ na teoria T(P,Z) tais que $Z_i$ e $Z_k$ ocorrem negativamente em $C_1$ e $Z_j$ e $Z_k$ ocorrem positivamente em $C_2$ . Definimos =>> como sendo o fecho transitivo da relação => .

DEFINIÇÃO 5.10 *(Extensão Z-recursiva)*

Dados $Z_i,Z_j \in Z$ , dizemos que eles são *estendidos Z-recursivamente* se $Z_i$ =>> $Z_j$ e $Z_j$ =>> $Z_i$ . Um predicado $Z_i$

é estendido Z-recursivamente se $Z_i \Rightarrow\!\!\!> Z_i$

**DEFINIÇÃO 5.11** *(Teoria Livre de Z-Conflito)*

Sejam $T(P,Z)$ uma teoria onde os axiomas estão representados na forma clausal e P e Z são tuplas de predicados. Dizemos que $T(P,Z)$ é *livre de Z-conflito* se sempre que existir uma cláusula C, e dois predicados $Z_i$ e $Z_j$ que ocorrem negativamente em C , então $Z_i$ e $Z_j$ não são Z-recursivamente estendidos.

**EXEMPLO 5.12** ( YUAN (1988) )

Considere a seguinte teoria $T(P,Z)$ , onde $P = \langle P_1 \rangle$ e $Z = \langle Z_1, Z_2 \rangle$ :

(1) $\neg Z_1(x,y) \vee \neg Z_2(y,x) \vee P_1(x,y)$

(2) $\neg Q_1(x,y) \vee Z_1(x,y)$

(3) $\neg Q_2(x,y) \vee Z_2(x,y)$

(4) $\neg Z_1(x,y) \vee Z_2(x,y) \vee Q_3(x,y)$

(5) $\neg Z_2(x,y) \vee Z_1(x,y) \vee Q_4(x,y)$

Observe que $Z_1 \rightarrow Z_2$ ( expressão (4) ) e $Z_2 \rightarrow Z_1$ (expressão (5)), e portanto $Z_1 \Rightarrow\!\!\!> Z_2$ e $Z_2 \Rightarrow\!\!\!> Z_1$. Logo , a teoria $T(P,Z)$ acima é estendida Z-recursivamente e, consequentemente , não é livre de Z-conflito . Tais teorias são ditas <u>Z-conflitantes</u> ■

**OBSERVAÇÃO 2 :** Obviamente o teorema 5.9 só se aplica a teorias livres de Z-conflito , como as dos exemplos 5.7 e 5.8.

Algumas teorias Z-conflitantes podem ser transformada em teorias livres de Z-conflito sem afetar o resultado da circunscrição . Para isto , considere uma teoria T(P,Z) onde P e Z são tuplas de predicados. Seja a teoria T'(P,Z) obtida pela negação de algum z ∈ Z em T(P,Z), isto é , tomando o complementar dos literais pertencentes ao conjunto z.

No caso do exemplo 5.12, teríamos , negando $Z_2$, a teoria T'(P,Z) composta por :

$(1') \quad \neg Z_1(x,y) \vee Z_2(y,x) \vee P_1(x,y)$

$(2') \quad \neg Q_1(x,y) \vee Z_1(x,y)$

$(3') \quad \neg Q_2(x,y) \vee \neg Z_2(x,y)$

$(4') \quad \neg Z_1(x,y) \vee \neg Z_2(x,y) \vee Q_3(x,y)$

$(5') \quad Z_2(x,y) \vee Z_1(x,y) \vee Q_4(x,y)$

A teoria T'(P,Z) é livre de Z-conflito e podemos utilizar o teorema 5.9 . Observe que os modelo da circunscrição de T(P,Z) e T'(P,Z) diferem apenas nos z que foram negados, e portanto vale :

**TEOREMA 5.13** *(YUAN(1989))*

Sejam uma teoria T(P,Z) e uma teoria T'(P,Z) obtida a partir de T(P,Z) pela substituição dos predicados de um subconjunto z de Z por sua negação. Então :

$$CIRC(T(P,Z);P;Z) \Leftrightarrow CIRC(T'(P,Z);P;Z') \wedge \bigwedge_{z_i \in z} \forall x \; ( \; Z_i(x) \equiv \neg \; Z_i'(x)) \; .$$

Um exemplo de teoria Z-conflitante que não pode ser transformada em uma livre de Z-conflito é :

$$(1)\ \neg Q(x) \vee Z(x,y) \vee Z(y,x)$$

$$(2)\ \neg Z(x,y) \vee \neg Z(y,x) \vee P(x,y)$$

Negando Z, obtemos :

$$(1)\ \neg Q(x) \vee \neg Z(x,y) \vee \neg Z(y,x)$$

$$(2)\ Z(x,y) \vee Z(y,x) \vee P(x,y)$$

Ambas as teorias são estendidas Z-recursivamente.

*i.2) Priorizada*

No caso da circunscrição ser priorizada, o resultado obtido para fórmulas separáveis no teorema 5.5, fica agora estabelecido pelo :

TEOREMA 5.14 *(LIFSCHITZ(1985))*

Seja A(P) a fórmula obtida pela conjunção de todos os axiomas de uma teoria T(P), onde $P = \langle P_1,\ldots,P_m\rangle$ é uma tupla de constantes de predicado . Se A(P) é uma fórmula separável , ou seja :

$$A(P) \Leftrightarrow \bigvee_i \left[ N_i(P) \wedge (U_i \leq P) \right]$$

então :

$$CIRC_{\leq'}(T(P);P) \Leftrightarrow \bigvee_i \left[D_i \wedge (U_i = P)\right]$$

onde :

$$D_i \Leftrightarrow N_i(U_i) \wedge \left( \bigwedge_{j\neq i} \neg\left[N_j(U_j) \wedge (U_j <' U_i)\right] \right) .$$

Observe que a utilização do teorema 5.14 e da proposição 4.29 para obter a teoria circunscrita priorizada traz consigo o problema de que agora a separação diz respeito as partições P1,...,Pk de P e não mais ao conjunto de todos os predicados que integram P , como no caso da circunscrição paralela . O teorema 4.30, no qual a circunscrição priorizada é escrita como uma conjunção de circunscrições paralelas, permite que o procedimento de YUAN (1988) seja utilizado em cada uma das partes de tal conjunção, como veremos a seguir .

**EXEMPLO 5.15**

Vamos empregar o processo de YUAN (1988) para o caso da circunscrição priorizada .

Para isso, considere a teoria T representada na forma clausal :

(1) $AB_1(Tweety) \vee \neg VOA(Tweety)$

(2) $\neg AVE(Tweety) \vee AB_1(Tweety)$

(3) $\neg AVE(Tweety) \vee AB_2(Tweety) \vee VOA(Tweety)$

(4) $\neg CANARIO(Tweety) \vee AB_3(Tweety) \vee AVE(Tweety)$

(5) $CANARIO(Tweety)$

O que vamos verificar é se :

$$CIRC(T; AB_3 > AB_2 > AB_1; VOA, AVE, CANARIO) \models VOA(Tweety)$$

Pelo teorema 4.70 , temos que :

$CIRC(T; AB_3 > AB_2 > AB_1; VOA, AVE, CANARIO) \Leftrightarrow$

$\Leftrightarrow CIRC(T; AB_3; AB_2, AB_1, VOA, AVE, CANARIO) \wedge$

$CIRC(T;AB_2;AB_1,VOA,AVE,CANARIO) \wedge$

$CIRC(T;AB_1;VOA,AVE,CANARIO)$ (A)

O que devemos fazer é eliminar da teoria T, em cada uma das circunscrições que formam a conjunção (A) , as ocorrências negativas dos predicados AVE,VOA,CANARIO que são variáveis no processo de circunscrição. Assim, após o processo de Z-resolução e aplicando o teorema 5.9 , teremos que :

$$CIRC(T;AB_3 > AB_2 > AB_1;VOA,AVE,CANARIO) \Leftrightarrow T \wedge$$
$$\wedge\ CIRC(T';AB_3;AB_2,AB_1) \wedge CIRC(T';AB_2;AB_1) \wedge$$
$$\wedge\ CIRC(T';AB_1) \qquad (B)$$

onde T' é a teoria composta por (teorema 5.9) :

*i)* $AB_1(Tweety) \vee AB_2(Tweety) \vee AB_3(Tweety)$

*ii)* $AB_1(Tweety) \vee AB_3(Tweety)$

Note que a $CIRC(T';AB_3;AB_2,AB_1)$ tem como modelos mínimos :

$\{AB_1(Tweety), \neg AB_2(Tweety), \neg AB_3(Tweety)\}$ e

$\{AB_1(Tweety), AB_2(Tweety), \neg AB_3(Tweety)\}$.

Somente o primeiro modelo também é modelo das outras duas circunscrições (o segundo não é modelo da $CIRC(T';AB_2;AB_1)$ ) que formam a conjunção (B) . Logo, a teoria T mais $AB_3(Tweety) \equiv false$ , $AB_1(Tweety) \equiv true$ e $AB_2(Tweety) \equiv false$ nos fornece a resposta desejada, ou seja ,

$$CIRC(T; AB_3 > AB_2 > AB_1; VOA, AVE, CANARIO) \models VOA(Tweety) \blacksquare$$

*ii) MILO- Resolução*

*ii.1) Paralela*

Na seção anterior , caracterizamos uma classe de teorias de primeira-ordem cuja circunscrição podia ser escrita também em linguagem de primeira-ordem. Desta forma, uma consulta a estas teorias circunscritas podia ser feita através de qualquer um dos métodos conhecidos para teorias de primeira-ordem. Nesta e na próxima seção , veremos métodos que nos permitam obter respostas à consultas feitas a teorias circunscritas em geral.

O primeiro desses métodos foi proposto por **PRZYMUSINSKI** (1989) e está baseado numa versão modificada da resolução linear ordenada ( OL - resolution , CHANG (1971) ), chamada *MILO-Resolução* (*MI*nimal model *L*inear *O*rdered resolution). De modo a simplificar a apresentação dos resultados , vamos considerar as teorias compostas apenas por cláusulas básicas, observando que os resultados podem ser estendidos para cláusulas em geral.

Vamos considerar aqui que as cláusulas são da forma $l_1 \vee \ldots \vee l_m$ , onde os literais que as compõem estão ordenados. Além disso, alguns desses literais podem estar marcados , caso em que serão denotados por $[l_i]$. Os literais marcados servem somente para identificar quais literais já foram resolvidos. Dizemos que uma cláusula é uma tautologia se existir entre seus literais não-marcados, um par de literais complementares . Uma cláusula C subjulga outra D se todo literal não-marcado de C ocorre , igualmente não-marcado em D.

Considere uma teoria T e uma cláusula $C_0$ de T, chamada de cláusula topo. Uma OL-dedução de $C_n$ a partir de T com a cláusula topo $C_0$ é uma sequência de cláusulas $C_i$ , i = 0,...,n , onde $C_0 \Leftrightarrow C$ e para cada i = 0,...,n-1, $C_{i+1}$ é o resolvente das cláusulas $C_i$ e $B_i$ , onde $B_i \in T$ e o literal a ser resolvido em $C_i$ é sempre o primeiro mais à esquerda . Além disso, quando no resolvente aparecerem um literal marcado e o seu complementar não-marcado, devemos retirar de $C_{i+1}$ este literal marcado (*redução*).

Caso o literal mais à esquerda esteja marcado ( não existindo portanto nenhum literal não-marcado no início da cláusula que possa ser resolvido) , devemos retirá-lo da cláusula (*contração*).

**OBSERVAÇÃO 3 :** Considerando uma cláusula como uma sequência de literais (ditas cadeias) , este método é chamado *eliminação de modelos* (CASANOVA(1987)).

**EXEMPLO 5.16**

Considere a seguinte teoria T :

$b_1 \Leftrightarrow S$ $\qquad\qquad$ $b_3 \Leftrightarrow \neg R \vee Q$

$b_2 \Leftrightarrow \neg S \vee Q \vee R$ $\qquad\qquad$ $b_4 \Leftrightarrow \neg O$

Uma OL-dedução da cláusula vazia $\square$ (sem nenhum literal) a partir de T e da cláusula topo $\neg Q \vee O$ , é :

$C_0 \Leftrightarrow \neg Q \vee O$ $\qquad\qquad$ $\neg S \vee Q \vee R \Leftrightarrow b_2$

$\downarrow$

$C_1 \Leftrightarrow \neg S \vee R \vee [\neg Q] \vee O$ $\qquad\qquad$ $S \Leftrightarrow b_1$

$\downarrow$

$C_2 \Leftrightarrow R \vee [\neg Q] \vee O \vee [\neg S]$ $\qquad\qquad$ $\neg R \vee Q \Leftrightarrow b_3$

$\downarrow$

$C_3 \Leftrightarrow [\neg Q] \vee O \vee [\neg S] \vee [R] \vee Q$

De $C_3$ devemos eliminar Q (possui o seu complementar marcado) , donde obtemos $[\neg Q] \vee O \vee [\neg S] \vee [R]$ . Como o literal mais à esquerda está marcado , devemos retirá-lo . Assim , a cláusula a ser resolvida agora é $C_3' \Leftrightarrow O \vee [\neg S] \vee [R]$ :

$C_3' \Leftrightarrow O \vee [\neg S] \vee [R]$ $\qquad\qquad$ $\neg O \Leftrightarrow b_4$

$\downarrow$

$[O] \vee [\neg S] \vee [R] \Leftrightarrow C_4$ ,ou seja,

ou seja, $C_4 \Leftrightarrow \square$ ■

A MILO-dedução diferencia-se da OL-dedução em dois pontos :

(1) na MILO, a resolução é aplicada somente a literais pertencentes à $\mathbb{Z} \cup \mathbb{P}$ , onde $\mathbb{P}$ é o conjunto de todos os literais *negativos* cujos símbolos de predicado estão em P e

$\mathbb{Z}$ é o conjunto dos literais (positivos e negativos) cujos símbolos de predicado estão em Z ; enquanto na OL , a resolução é aplicada a todo literal ;

(2) na MILO, todos os literais marcados que não tenham um literal não-marcado , pertencentes à $\mathbb{Z} \cup \mathbb{P}$ , à sua esquerda podem ser removidos , ao contrário da OL, onde somente os literais marcados , sem literais não-marcados à sua esquerda, podiam ser retirados.

Formalmente, a MILO-dedução é definida por :

**DEFINIÇÃO 5.17** *(MILO-dedução)*

Dadas uma teoria T(P,Z), onde P e Z são tuplas disjuntas de constantes de predicado, e uma cláusula C, uma *MILO-dedução de uma cláusula* $C_n$ *a partir de* T(P,Z) ∪ {C} , é uma sequência de cláusulas $C_0,\ldots,C_n$ tal que $C_0 \Leftrightarrow C$ e $C_{i+1}$ é gerada a partir de $C_i$ da seguinte maneira :

(A) uma cláusula $D_{i+1}$, chamada de <u>resolvente ordenado</u>, é obtida resolvendo-se o primeiro literal mais à esquerda $l_j$ de $C_i \Leftrightarrow l_1 \vee \ldots \vee l_m$ , que pertença à $\mathbb{Z} \cup \mathbb{P}$, e alguma cláusula de entrada $C \Leftrightarrow k_1 \vee \ldots \vee k_u \vee \ldots \vee k_s$ da teoria T(P,Z) , de modo que :

$$D_{i+1} \Leftrightarrow l_1 \vee \ldots \vee l_{j-1} \vee k_1 \vee \ldots \vee k_{u-1} \vee k_{u+1} \vee \ldots \vee k_s \vee [l_j] \vee l_{j+1} \vee \ldots l_m \text{ , onde } k_u = \neg l_j \text{ ;}$$

ou seja, marcando-se o literal $l_j$ com [ ], e inserindo a cláusula C, sem o complementar de $l_j$ ($[k_u]$), em $C_i$, na frente do literal marcado $[l_j]$;

(B) a cláusula $C_{i+1}$ é obtida de $D_{i+1}$ através das seguintes reduções, pela ordem :

(B1) eliminar qualquer literal não-marcado k em $D_{i+1}$ quando existir um literal marcado [¬k] em $D_{i+1}$ ;

(B2) fundir quaisquer literais idênticos em $D_{i+1}$ ;

(B3) remover quaisquer literais marcados em $D_{i+1}$ que não são precedidos por literais não-marcados de $\mathbb{Z} \cup \mathbb{P}$ .

(C) nenhuma das cláusulas $C_i$ pode ser uma tautologia e nem pode ser subjulgada por qualquer das cláusulas $C_j$ , $j<i$. ■

**EXEMPLO 5.18**

Considere a teoria T(P,Z) que é composta pelas cláusulas :

$$c_1 \Leftrightarrow S(a)$$

$$c_2 \Leftrightarrow \neg S(a) \vee Q(a) \vee R(a)$$

$$c_3 \Leftrightarrow \neg R(a) \vee Q(a)$$

$$c_4 \Leftrightarrow \neg O(a)$$

onde P = < Q >, Z = < R > e , consequentemente, $\mathbb{Z} \cup \mathbb{P}$ = { ¬R(a) , R(a) , ¬Q(a) }.

Vamos encontrar uma MILO-dedução a partir de T(P,Z) ∪ { ¬ Q(a) v O(a) }. Inicialmente, temos :

$$C_0 \Leftrightarrow \neg Q(a) \vee O(a).$$

Para gerar $C_1$, devemos construir $D_1$ a partir de $C_0$ e uma cláusula de entrada. O literal a ser resolvido em $C_0$ é

$\neg Q(a)$ , pois é o primeiro literal da cláusula que pertence à $\mathbb{Z} \cup \mathbb{P}$ . Tomando (2) como sendo a cláusula de entrada que deve ser resolvida com $C_0$ , temos , por (A) da definição 5.17, que :

$$C_0 \Leftrightarrow \neg Q(a) \vee O(a) \qquad\qquad \neg S(a) \vee Q(a) \vee R(a) \Leftrightarrow c_2$$

$$\downarrow$$

$$\neg S(a) \vee R(a) \vee [\neg Q(a)] \vee O(a) \Leftrightarrow D_1$$

Como $D_1$ não pode ser reduzido por nenhuma das maneiras presentes em (B) (definição 5.17) , temos que $C_1 \Leftrightarrow D_1$ , que não é uma tautologia nem é subjulgada por $C_0$.

Em $C_1 \Leftrightarrow \neg S(a) \vee R(a) \vee [\neg Q(a)] \vee O(a)$ , o primeiro literal pertencente à $\mathbb{Z} \cup \mathbb{P}$ é $R(a)$, que resolvido com (3) nos dá :

$$C_1 \Leftrightarrow \neg S(a) \vee R(a) \vee [\neg Q(a)] \vee O(a) \qquad \neg R(a) \vee Q(a) \Leftrightarrow (c3)$$

$$\downarrow$$

$$\neg S(a) \vee Q(a) \vee [R(a)] \vee [\neg Q(a)] \vee O(a) \Leftrightarrow D_2$$

Observe que $C_2$ é obtido de $D_2$, após reduzir esta aplicando (B1), passando a ser $\neg S(a) \vee [R(a)] \vee [\neg Q(a)] \vee O(a)$ , e aplicando (B2) , tornando $C_2 \Leftrightarrow \neg S(a) \vee O(a)$, que satisfaz (C).

Como $C_2$ não contém nenhum literal pertencente à $\mathbb{Z} \cup \mathbb{P}$, não podemos resolvê-lo com nenhuma cláusula da teoria $T(P,Z)$, terminando aqui o processo de MILO-dedução . Note que $C_0$ poderia ser resolvido inicialmente com (3) , o que nos daria a seguinte MILO-resolução :

$$C_0 \Leftrightarrow \neg Q(a) \vee O(a) \qquad\qquad \neg R(a) \vee Q(a) \Leftrightarrow (c3)$$

$$\downarrow$$

$$C_1 \Leftrightarrow \neg R(a) \vee [\neg Q(a)] \vee O(a) \qquad \neg S(a) \vee Q(a) \vee R(a) \Leftrightarrow (c2)$$

$$\downarrow$$

$$C_2 \Leftrightarrow \neg S(a) \vee O(a) \;,$$

obtida pela redução de $\neg S(a) \vee Q(a) \vee [\neg R(a)] \vee [\neg Q(a)] \vee O(a)$ , através da aplicação de (B1) e (B2) da definição 5.17.

Partindo de $C_0 \Leftrightarrow \neg Q(a) \vee O(a)$, nenhuma outra MILO-dedução é possível ■

Uma cláusula MILO-dedutível a partir de $T(P,Z) \cup \{C\}$ e que não contenha literais pertencentes à $\mathbb{Z} \cup \mathbb{P}$ é chamada de <u>MILO-folha</u> de $T(P,Z) \cup \{C\}$ . No exemplo anterior , $\neg S(a) \vee O(a)$ é uma MILO-folha de $T(P,Z) \cup \{\neg Q(a) \vee O(a)\}$, onde $T(P,Z)$ é composta pelos axiomas (c1),(c2),(c3),(c4) .

A <u>MILO-derivada</u> de $T(P,Z) \cup \{C\}$, denotada por $DERIV(T(P,Z),C)$ ou $DERIV(T,C)$ , é a conjunção de todas as MILO-folhas de $T(P,Z) \cup \{C\}$ . No exemplo 5.18, $T(P,Z) \cup \{\neg Q(a) \vee O(a)\}$ possui duas MILO-folhas , que são iguais a $\neg S(a) \vee O(a)$,e $DERIV(T,\neg Q(a) \vee O(a)) \Leftrightarrow \neg S(a) \vee O(a)$. Quando uma dada teoria $T(P,Z) \cup \{C\}$ não possuir nenhuma MILO-folha, $DERIV(T,C) \Leftrightarrow$ true.

EXEMPLO 5.19

Para $T(P,Z)$ igual a do exemplo 5.18 e a cláusula $R(a) \vee \neg Q(a)$, se aplicarmos a definição 5.17 a $T(P,Z) \cup \{R(a) \vee$

$\neg Q(a)\}$, obteremos :

$$\begin{array}{ccc} Co \Leftrightarrow R(a) \vee \neg Q(a) & & \neg R(a) \vee Q(a) \Leftrightarrow (c3) \\ & \downarrow & \\ & Q(a) \vee [R(a)] \vee \neg Q(a) & \end{array}$$

que é uma tautologia e , pelo ítem (c) da definição 5.17 , não pode ser uma das cláusulas presentes na sequência de MILO-dedução. Como $R(a)$ de $Co$ não pode ser resolvida com nenhuma outra cláusula , $T(P,Z) \cup \{Co\}$ não possui nenhuma MILO-folha e portanto, $DERIV(T,Co) \Leftrightarrow true$ ■

**DEFINIÇÃO 5.20** *(MILO-derivada de uma Conjunção de Cláusulas)*

Seja $H = C_1 \wedge \ldots \wedge C_m$ uma conjunção de cláusulas. A *MILO-derivada* $DERIV(T,H)$ de $T(P,Z) \cup \{H\}$ é definida por :

$$DERIV(T,H) = DERIV(T_1,C_1) \wedge \ldots \wedge DERIV(T_m,C_m) ,$$

onde $T_1 = T(P,Z)$ e $T_{i+1} = T_i \cup \{C_i\}$ , para $i = 1,\ldots,m-1$ .

O algorítmo baseado na MILO-resolução está fundamentado nos dois próximos teoremas . O primeiro deles transforma o problema de determinar se uma dada fórmula F segue de uma teoria circunscrita, ao de determinar se a negação da MILO-derivada $DERIV(T,\neg F)$ segue da teoria circunscrita.

Formalmente :

**TEOREMA 5.21** *(PRZYMUSINSKI (1989))*

Para qualquer fórmula F temos que :

$$CIRC(T(P,Z);P;Z) \models F \text{ se , e somente se, }$$
$$CIRC(T(P,Z);P;Z) \models \neg\, DERIV(T(P,Z),\neg F) .$$

Note que $\neg$ DERIV(T(P,Z),$\neg$F) não possui literais de Z, o que nos permite utilizar o seguinte resultado :

**TEOREMA 5.22** *(PRZYMUSINSKI (1989))*

Suponha que G é qualquer fórmula que não contém literais de $\mathbb{Z}$ e que possa ser representado por $D_1 \wedge \ldots \wedge D_m$ , onde cada $D_i$ , i = 1,...,m , são cláusulas . Então : CIRC(T(P,Z);P;Z) $\models$ G se, e somente se, T $\models$ DERIV(T,$D_i$) , para todo i $\leq$ m .

A partir desses resultados, obtemos o seguinte algorítmo que fornece respostas sobre consultas à teorias circunscritas :

<u>ALGORITMO</u> <u>1</u> : *Decidir se a teoria CIRC(T(P,Z);P;Z) implica numa dada fórmula F*

P1 - Caso F não contenha literais de $\mathbb{Z}$ , faça P2 . Caso contrário , faça P3.

P2 - (a) Represente F $\Leftrightarrow$ $G_1 \wedge \ldots \wedge G_m$ , onde $G_i$ são cláusulas ;

(b) Para j = 1,...,m , use o <u>ALGORITMO</u> <u>2</u> para encontrar DERIV(T,$G_j$) ;

(c) Para $j = 1,\ldots,m$ , verifique se

$T \models DERIV(T,G_j)$ ;

(d) Se existir $j$ tal que $T \not\models DERIV(T,G_j)$ então a resposta é *NÃO* . Caso contrário , *SIM* .

P3 - (a) Represente $F \Leftrightarrow K_1 \vee \ldots \vee K_m$ , onde $K_i$ são conjunções de literais ;

(b) Faça $T_1 = T(P,Z)$ e para $i = 1,\ldots,m-1$, use o ALGORITMO 2 para encontrar $DERIV(T_i,\neg K_i)$ , onde $T_{i+1} = T_i + \neg K_i$ ;

(c) Utilize P2 para verificar se :

$$CIRC(T(P,Z);P;Z) \models \neg [ \bigwedge_{i=1}^{m} DERIV(T_i,\neg K_i) ] ;$$

(d) Caso a resposta em (c) for afirmativa, a resposta da consulta será *SIM*. Caso contrário, será *NÃO* .

ALGORITMO 2 : *Dadas uma teoria T e uma cláusula C , computar DERIV(T,C)*

P1 - Encontrar todas as MILO-folhas de $T \cup \{C\}$ ;

P2 - Faça $DERIV(T,C)$ ser a conjunção de todas as MILO-folhas obtidas em P1.

**OBSERVAÇÃO 4 :** No algorítmo 1 , o passo P2 (c) pode ser feito por qualquer um dos métodos conhecidos para verificação da insatisfatibilidade de $T \cup \{ \neg DERIV(T,G_j) \}$.

**OBSERVAÇÃO 5 :** Note que ao final , o problema inicial será reduzido ao de determinar se uma certa fórmula é ou não consequência lógica da teoria $T(P,Z)$ ( passo P2 (c) ). Assim sendo, o algorítmo sempre terminará para teorias (finitas) compostas por cláusulas básicas. O algorítmo pode ser generalizado para teorias em geral , o que acarretará a impossibilidade de decidir se uma dada fórmula é derivável de uma certa teoria.

**EXEMPLO 5.23**

Novamente , vamos considerar a teoria $T(AB,VOA)$ vista no exemplo 4.2, representada aqui na forma clausal :

(a) $\neg$ AVE(x) v AB(x) v VOA(x)

(b) $\neg$ PINGÜIM(x) v AVE(x)

(c) $\neg$ PINGÜIM(x) v $\neg$ VOA(x)

A consulta que vamos fazer a teoria circunscrita CIRC(T(AB,VOA);AB;VOA) é se os pingüins , e só eles, são anormais. Ou seja , queremos saber se :

$$CIRC(T(AB,VOA);AB;VOA) \models \forall x\ (AB(x) \equiv PINGÜIM(x) \wedge AVE(x)).$$

Para a teoria circunscrita que estamos considerando, temos $\mathbb{Z}$ = { VOA , $\neg$ VOA } e $\mathbb{P}$ = { $\neg$ AB }. Fazendo F(x) $\Leftrightarrow$ (AB(x) $\equiv$ PINGÜIM(x) $\wedge$ AVE(x)),e aplicando o passo P1 do algorítmo 1, como F(x) não possui literais de $\mathbb{Z}$ , passamos ao passo P2 do algorítmo 1.

Representando F(x) como uma conjunção de cláusulas , temos :

$F(x) \Leftrightarrow ( (\neg ABC(x) \vee PINGÜIM(x)) \wedge (ABC(x) \vee \neg PINGÜIM(x) \vee \neg AVE(x)) \wedge (\neg ABC(x) \vee AVE(x)) )$,

onde fazemos :

$$G_1(x) \Leftrightarrow (\neg ABC(x) \vee PINGÜIM(x))$$
$$G_2(x) \Leftrightarrow (ABC(x) \vee \neg PINGÜIM(x) \vee \neg AVE(x))$$
$$G_3(x) \Leftrightarrow (\neg ABC(x) \vee AVE(x))$$

No passo seguinte (P2 (b)) , vamos utilizar o algorítmo 2 para encontrar DERIV(T,$G_1$(x)) e DERIV(T,$G_2$(x)). Assim, temos que :

i) DERIV(T,$G_1$(x)) $\Leftrightarrow$ true , pois :

$$\begin{array}{c} \neg ABC(x) \vee PINGÜIM(x) \\ \downarrow \quad (a) \\ \neg AVE(x) \vee VOA(x) \vee [\neg ABC(x)] \vee PINGÜIM(x) \quad (*) \\ \downarrow \quad (c) \end{array}$$

$\neg$ AVE(x) v $\neg$PINGÜIM(x) v [VOA(x)] v [$\neg$ABC(x)] v PINGÜIM(x) , que por ser uma tautologia não pode fazer parte de uma MILO-dedução . Como (*) possui um literal pertencente à $\mathbb{Z} \cup \mathbb{P}$ , que é VOA(x) , e não pode ser resolvido com outra cláusula além de (c) , (*) não pode ser considerada uma MILO-folha. Portanto , DERIV(T,$G_1$(x)) $\Leftrightarrow$ true.

ii) DERIV(T,$G_2$(x)) $\Leftrightarrow$ $G_2$(x) $\Leftrightarrow$ ABC(x) v $\neg$PINGÜIM(x) v $\neg$AVE(x), já que $G_2$(x) não contém nenhum literal pertencente ao conjunto $\mathbb{Z} \cup \mathbb{P} = \{ VOA , \neg VOA, \neg AB \}$.

iii) $DERIV(T,G_3(x)) \Leftrightarrow true$ , pois :

$$\begin{array}{c} \neg ABC(x) \vee AVE(x) \quad (**) \\ \Big\downarrow (a) \\ \neg AVE(x) \vee VOA(x) \vee [\neg AB(x)] \vee AVE(x), \end{array}$$

que por ser uma tautologia não pode fazer parte de uma MILO-dedução. Como (**) possui um literal pertencente à $\mathbb{Z} \cup \mathbb{P}$ , que é AB(x) , e não pode ser resolvido com outra cláusula além de (a), (**) não pode ser considerada uma MILO-folha. Logo, $DERIV(T,G_3(x)) \Leftrightarrow true$ .

Agora, pelo passo P2 (c), temos que verifica se $T \models DERIV(T,G_i(x))$, para $i=1,2,3$ . Obviamente , $T \models DERIV(T,G_1(x))$ e $T \models DERIV(T,G_3(x))$ , pois $DERIV(T,G_1(x)) \Leftrightarrow DERIV(T,G_3(x)) \Leftrightarrow true$ . Para verificar que $T \models G_2(x)$, vamos utilizar a resolução e mostrar que $T \cup \{ \neg G_2(x) \}$ é insatisfatível. Assim ,

| | | |
|---|---|---|
| (a) | $\neg$ AVE(x) v AB(x) v VOA(x) | axioma |
| (b) | $\neg$ PINGÜIM(x) v AVE(x) | axioma |
| (c) | $\neg$ PINGÜIM(x) v $\neg$ VOA(x) | axioma |
| (d) | $\neg$ AB(x) | negação de $G_2(x)$ |
| (e) | PINGÜIM(x) | negação de $G_2(x)$ |
| (f) | AVE(x) | negação de $G_2(x)$ |
| (g) | $\neg$ AVE(x) v VOA(x) | resolvente de ((a),(d)) |
| (h) | VOA(x) | resolvente de ((g),(f)) |
| (i) | $\neg$ PINGÜIM(x) | resolvente de ((h),(c)) |
| (j) | $\square$ | resolvente de ((e),(i)) |

Portanto, para i=1,2,3 , T $\models$ DERIV(T,$G_i$(x)) , e dessa maneira a resposta para a consulta feita é SIM . Logo ,

$$CIRC(T(AB,VOA);AB;VOA) \models \forall x \ (AB(x) \equiv PINGÜIM(x) \land AVE(x)).$$

Observe que tal resultado está de acordo com o obtido no exemplo 4.2, onde vimos que CIRC(T(AB,VOA);AB;VOA) $\Leftrightarrow$ T(P,Z) $\land$ (AB = PINGÜIM $\land$ AVE) .

Se , nesta mesma teoria , fizéssemos a consulta :

$$F(x) \Leftrightarrow (AVE(x) \land \neg VOA(x) \supset PINGÜIM(x)) ,$$

ou seja , se as aves que não voam são pingüins.

Lembrando que $\mathbb{Z} \cup \mathbb{P}$ = { VOA , ¬VOA , ¬AB }, observe que agora F(x) possui um literal pertencente à $\mathbb{Z} \cup \mathbb{P}$ . Logo, devemos utilizar o passo P3 do algorítimo 1 . Como :

$$F(x) \Leftrightarrow ( \neg AVE(x) \lor VOA(x) \lor PINGÜIM(x)) ,$$

fazemos :

$K_1 \Leftrightarrow \neg$ AVE(x) ; $T_1$ = T(AB,VOA) ;

$K_2 \Leftrightarrow$ VOA(x) ; $T_2 = T_1$ U {AVE(x)} ;

$K_3 \Leftrightarrow$ PINGÜIM(x) ; $T_3 = T_2$ U {¬ VOA(x)} .

Calculando as DERIV($T_i$,¬$K_i$) , para i=1,2,3, através do algoritmo 2, obtemos :

a) DERIV($T_1$,¬$K_1$) $\Leftrightarrow$ ¬$K_1$ e DERIV($T_3$,¬$K_3$) $\Leftrightarrow$ ¬$K_3$ , pois em ambos os casos não temos literais pertencentes à $\mathbb{Z} \cup \mathbb{P}$ para resolver ;

b) DERIV(T2,¬K2) $\Leftrightarrow$ ¬ AVE(x) v AB(x) , obtida de :

$$\begin{array}{c} \neg\ VOA(x) \\ \big\downarrow\ (a) \\ \neg\ AVE(x)\ \vee\ AB(x)\ , \end{array}$$

que é a única MILO-folha .

Logo :

$$\begin{array}{c} DERIV(T_1,\neg K_1)\ \wedge\ DERIV(T_2,\neg K_2)\ \wedge\ DERIV(T_3,\neg K_3)\ \Leftrightarrow \\ AVE(x)\ \wedge\ (\ \neg AVE(x)\ \vee\ AB(x))\ \wedge\ \neg PINGÜIM(x)\ \equiv \\ (\ AVE(x)\ \wedge\ AB(x)\ \wedge\ \neg PINGÜIM(x)\ )\ . \end{array}$$

Pelo passo P3 (c) do algorítmo 1 , devemos utilizar o passo P2 do mesmo algorítmo para verificar se :

$$CIRC(T(AB,VOA);AB;VOA) \models \neg(AVE(x) \wedge AB(x)\ \wedge \neg PINGÜIM(x))$$

Utilizando P2, temos :

$$\begin{array}{c} \neg AVE(x)\ \vee\ \neg AB(x)\ \vee\ PINGÜIM(x) \\ \big\downarrow\ (a) \\ \neg AVE(x)\ \vee\ [\neg AB(x)]\ \vee\ VOA(x)\ \vee\ PINGÜIM(x) \\ \big\downarrow\ (c) \\ \neg AVE(x)\ \vee\ [\neg AB(x)]\ \vee\ \neg PINGÜIM(x)\ \vee\ [VOA(x)]\ \vee\ PINGÜIM(x) \end{array}$$

que é uma tautologia e portanto :

$$DERIV(T, \neg AVE(x) \vee \neg AB(x) \vee PINGÜIM(x)) \Leftrightarrow true.$$

Obviamente , T $\models$ true (passo P2 (c)), e assim a resposta a consulta feita é *SIM* , ou seja ,

$$CIRC(T(AB, VOA); AB; VOA) \models (AVE(x) \wedge \neg VOA(x) \supset PINGÜIM(x))_\blacksquare$$

*ii.2) __Priorizada__*

O objetivo agora é saber se dada uma fórmula F , ela é consequência lógica da teoria circunscrita priorizada CIRC(T(P,Z);P1 > ... > Pk;Z) . Considerando o teorema 4.30, queremos saber se :

$$CIRC(T;P1;P2 \cup \ldots \cup Pk \cup Z) \wedge \ldots \wedge CIRC(T;Pk;Z) \models F \quad (*)$$

Para isso, vamos considerar a k-ésima circunscrição da expressão (*) . Calcula-se DERIV(T(P,Z),¬F) com relação a CIRC(T;Pk;Z) , ou seja , levando em conta o conjunto $\mathbb{Z} \cup \mathbb{P}k$ quando for utilizar a MILO-resolução . Pelo teorema 5.21 , CIRC(T;Pk;Z) $\models$ F se, e somente se, CIRC(T;Pk;Z) $\models$ ¬DERIV(T,¬F) .

Como ¬D ≡ DERIV(T,¬F) não possui literais pertencentes à $\mathbb{Z}$ , escrevendo-o na forma $C_1 \wedge \ldots \wedge C_m$ e calculando DERIV(T,Ci) para i=1,...,m , obtemos pelo teorema 5.22 :

(*) se, e somente se, CIRC(T;P1;P2 U ... U Pk U Z) ∧ ... ∧

$$CIRC(T;Pk-1;Pk \cup Z) \models \bigwedge_{i=1}^{m} DERIV(T,C_i) .$$

Repetindo este processo , tomando sempre a última circunscrição da conjunção , chegaremos ao problema de determinar se uma certa fórmula , obtida através do processo , é derivável da teoria T(P,Z) .

EXEMPLO 5.24

Vamos aplicar o procedimento acima para a mesma teoria do exemplo 5.15 :

(1) $AB_1(Tweety) \vee \neg VOA(Tweety)$

(2) $\neg AVE(Tweety) \vee AB_1(Tweety)$

(3) $\neg AVE(Tweety) \vee AB_2(Tweety) \vee VOA(Tweety)$

(4) $\neg CANARIO(Tweety) \vee AB_3(Tweety) \vee AVE(Tweety)$

(5) $CANARIO(Tweety)$

e vamos ver se VOA(Tweety) segue da teoria circunscrita :

$CIRC(T(AB,VOA);AB_3 > AB_2 > AB_1;\ VOA,AVE,CANARIO) \Leftrightarrow$

$CIRC(T(AB,VOA);\ AB_3;\ AB_2,AB_1,VOA,AVE,CANARIO) \wedge$

$CIRC(T(AB,VOA);\ AB_2;\ AB_1,VOA,AVE,CANARIO) \wedge$

$CIRC(T(AB,VOA);\ AB_1;\ VOA,AVE,CANARIO)$

(i) Calculando a $DERIV(T,\neg VOA(Tweety))$, sendo $\mathbb{Z}_1 \cup \mathbb{P}_1 = \{$ $\neg AB_1(Tweety)$, $VOA(Tweety)$, $\neg VOA(Tweety)$, $AVE(Tweety)$, $\neg AVE(Tweety)$, $CANARIO(Tweety)$, $\neg CANARIO(Tweety)$ $\}$,

temos :

$$\neg VOA(Tweety)$$
$$\downarrow 3$$
$$\neg AVE(Tweety) \vee AB_2(Tweety) \vee [\neg VOA(Tweety)]$$
$$\downarrow 4$$
$$\neg CANARIO(Tweety) \vee AB_3(Tweety) \vee AB_2(Tweety) \vee$$
$$[\neg VOA(Tweety)] \vee [\neg AVE(Tweety)]$$
$$\downarrow 5$$
$$AB_3(Tweety) \vee AB_2(Tweety)$$

Portanto:

$$DERIV(T,\neg VOA(Tweety)) \Leftrightarrow AB_3(Tweety) \vee AB_2(Tweety) \ .$$

Calculando a DERIV de cada uma das partes da conjunção $\neg DERIV(T,\neg VOA(Tweety))$ com relação a $\mathbb{Z}_1 \cup \mathbb{P}_1$, obtemos, respectivamente, $\neg AB_3(Tweety)$ e $\neg AB_2(Tweety)$. Logo, o nosso problema passa a ser :

$$CIRC(T(AB,VOA);\ AB_3;\ AB_2,AB_1,VOA,AVE,CANARIO) \wedge$$
$$CIRC(T(AB,VOA);AB_2;AB_1,VOA,AVE,CANARIO) \models \neg AB_3(Tweety) \wedge \neg AB_2(Tweety)$$

(ii) Fazendo $\mathbb{Z}_2 \cup \mathbb{P}_2 = \mathbb{Z}_1 \cup \mathbb{P}_1 \cup \{AB_1(Tweety), \neg AB_2(Tweety)\}$, obtemos que a $DERIV(T,AB_3(Tweety) \vee AB_2(Tweety)) \Leftrightarrow AB_3(Tweety) \vee AB_2(Tweety)$, pois nenhum elemento que pertença a $\mathbb{Z}_2 \cup \mathbb{P}_2$ que possa ser resolvido.

A seguir devemos calcular a DERIV de cada termo da conjunção $\neg DERIV(T,\ AB_3(Tweety) \vee AB_2(Tweety))$. Segue que :

$$DERIV(T,\neg AB_3(Tweety)) \Leftrightarrow \neg AB_3(Tweety) \text{ e}$$

$$DERIV(T,\neg AB_2(Tweety)) \Leftrightarrow true \text{ , pois :}$$

$$\neg AB_2(Tweety)$$
$$\downarrow 3$$
$$\neg AVE(Tweety) \vee VOA(Tweety) \vee [\neg AB_2(Tweety)]$$
$$\downarrow 4$$
$$\neg CANARIO(Tweety) \vee AB_3(Tweety) \vee VOA(Tweety) \vee$$
$$[\neg AVE(Tweety)] \vee [\neg AB_2(Tweety)]$$
$$\downarrow 5$$
$$AB_3(Tweety) \vee VOA(Tweety) \vee [\neg AVE(Tweety)]$$
$$[\neg AB_2(Tweety)] \vee [\neg CANARIO(Tweety)]$$
$$\downarrow 1$$
$$AB_3(Tweety) \vee AB_1(Tweety)$$

a sua única MILO-folha possui um elemento que pertence a $\mathbb{Z}_2 \cup \mathbb{P}_2$ e que não pode ser resolvido ($AB_1(Tweety)$).

Assim, o nosso problema agora é saber se :

$$CIRC(T(AB,VOA);AB_3;AB_2,AB_1,VOA,AVE,CANARIO) \models \neg AB_3(Tweety)$$

ou seja, pelo teorema 5.22 :

$$T \models DERIV(T,\neg AB_3(Tweety)) .$$

Nesse caso, $\mathbb{Z}_3 \cup \mathbb{P}_3 = \mathbb{Z}_2 \cup \mathbb{P}_2 \cup \{\neg AB_3(Tweety), AB_2(Tweety)\}$, donde $DERIV(T,\neg AB_3(Tweety)) \Leftrightarrow true$ pois :

$\neg AB_3(Tweety)$

$\downarrow 4$

$\neg CANARIO(Tweety) \vee AVE(Tweety) \vee [\neg AB_3(Tweety)]$

$\downarrow 5$

$AVE(Tweety) \vee [\neg AB_3(Tweety)]\ [\neg CANARIO(Tweety)]$

3 $\swarrow$ $\searrow$ 2

$AB_2(Tweety) \vee VOA(Tweety)$ $\qquad$ $AB_1(Tweety)$

$\vee [\neg AB_3(Tweety)]\ [\neg CANARIO(Tweety)]$

$\vee [AVE(Tweety)]$

1 $\downarrow$

$AB_1(Tweety) \vee AB_2(Tweety)$

que não pode ser resolvido com nenhuma cláusula de T.

Como T $\models$ true, temos que :

$CIRC(T(AB,VOA); AB_3 > AB_2 > AB_1;$

$VOA,AVE,CANARIO) \models VOA(Tweety)$ ■

*iii) Provador para Teorias Circunscritas*

*iii.1) Paralela*

Proposto por GINSBERG (1989) , este provador utiliza , para determinar se uma dada sentença segue ou não de uma certa teoria circunscrita , o seguinte processo : inicialmente é construída uma *prova d* para a consulta feita, sendo *d* uma fórmula obtida pela conjunção e disjunção das instâncias das negações dos predicados que estão sendo circunscritos . A seguir , tenta-se enfraquecer tal prova construindo-se uma *prova* para $\neg d$ .Caso não se consiga tal prova, a consulta feita será consequência lógica da teoria circunscrita em questão.

Nos resultados desta seção , vamos considerar somente interpretações de Herbrand tais que o universo de Herbrand em questão é finito,o que vai nos permitir estabelecer uma nova relação de ordem sobre os modelos. Para isto , vamos definir um conjunto D como sendo formado por todas as proposições $\neg P(t)$, onde P é um predicado a ser circunscrito numa dada teoria T e t é uma instância básica dos argumentos de P e o conjunto E como sendo formado por todas as proposições P(t). Se M é um modelo, M[D] e M[E] representam, respectivamente, o conjunto de proposições em D e E que são satisfeitas por M.

A nova relação ($\leq_D$) será definida sobre o conjunto D, o que permitirá caracterizar os modelos da circunscrição como sendo aqueles máximos sob a ordem $\leq_D$ . Formalmente :

DEFINIÇÃO 5.25 (*Relação* $\leq_D$)

Seja D um conjunto composto por todas as proposições da forma $\neg P(t)$ , onde P é um predicado a ser circunscrito e t é uma instância básica dos argumentos de P . Dados dois modelos M e N , escreve-se M $\leq_D$ N quando M[D] $\subset$ N[D], ou seja, o conjunto de proposições em D que são satisfeitas por M é um subconjunto das proposições em D satisfeitas por N.

OBSERVAÇÃO 6 : No artigo de GINSBERG (1989) , o conjunto D é chamado de conjunto de default , embora ele não seja composto por defaults tal como definidos no capítulo 2 .

OBSERVAÇÃO 7 : Os elementos D-máximos de uma coleção de modelos serão aqueles máximos sob a ordem parcial $\leq_D$ .

A partir desta nova relação de ordem definida acima , Ginsberg estabeleceu um resultado que permite determinar se uma dada sentença segue ou não de uma teoria circunscrita onde os predicados que não estão sendo circunscritos , têm necessariamente que variar .

Note que um modelo M da CIRC(T(P,Z);P;Z) é mínimo sob a relação ≤ mod(P,Z), ou seja, M[E] será mínimo sob a relação ≤ mod(P,Z). Como estamos supondo que o universo de Herbrand é finito, teremos que M[D] será <u>máximo</u> sob a relação $\leq_D$ . Logo :

TEOREMA 5.26 *(GINSBERG(1989))*

Considere uma teoria T sem símbolos de função , P um predicado e D o conjunto de todas as proposições da forma ¬ P(t) , onde t é uma instância básica dos argumentos de P. Para qualquer sentença Q , Q é uma consequência lógica de circunscrever P em T com <u>todos os outros predicados variando</u> se , e somente se, Q é satisfeita por todos os modelos D-máximos de T .

**OBSERVAÇÃO 8** : Para garantir que todas as instâncias de P são enumeráveis e não equivalentes, considera-se que a teoria T do teorema 5.26 é fechada sob o domínio e satisfaz as hipóteses do nome único , ou seja , existem um conjunto finito $\{t_1,\ldots,t_n\}$ e axiomas que nos permitem concluir que qualquer objeto no domínio é igual a um , e somente um , dos $t_i$ , para $i = 1,\ldots,n$ .

**EXEMPLO 5.27**

Considere a teoria T(AB,<AVE,PINGÜIM,VOA>) formada pelos axiomas :

(1) AVE(Tweety)

(2) PINGÜIM(Fred)

(3) $\forall x$ (AVE(x) $\wedge$ $\neg$ AB(x) $\supset$ VOA(x))

(4) $\forall x$ (PINGÜIM(x) $\supset$ AVE(x) $\wedge$ $\neg$ VOA(x))

onde : (1) significa que Tweety é uma ave ;

(2) significa que Fred é um pingüim ;

(3) significa que normalmente as aves voam ; e

(4) significa que os pingüins são aves que não voam.

Nesta teoria temos que :

$$D = \{ \neg \text{AB(Tweety)} , \neg\text{AB(Fred)} \} .$$

Neste caso, a circunscrição de AB na teoria acima, com os outros predicados variando , implica em VOA(Tweety) . Isto porque esta teoria possui um único modelo mínimo , cuja extensão do predicado AB é {AB(Fred)} e , portanto, o modelo D-máximo não pode incluir ¬ AB(Fred) .

Logo, os modelos D-máximos da teoria T(AB,<AVE,PINGÜIM,VOA>) são aqueles que incluem ¬ AB(Tweety) ∈ D e, como consequência, VOA(Tweety) ■

O motivo pelo qual Ginsberg preferiu trabalhar com a noção de modelos D-máximos ao invés dos modelos mínimos , se deve ao fato daqueles permitirem que o teorma 5.26 seja reestruturado de uma forma vantajosa. Para isto, é necessário que seja definida a noção de confirmação .

**DEFINIÇÃO 5.28** *(Forma Normal Disjuntiva (f.n.d.) com respeito à D )*

Considere D um conjunto de sentenças . Dizemos uma sentença R está na *forma normal disjuntiva (f.n.d.) com respeito a* D se R é da forma :

$$\bigvee_i \bigwedge_j d_{ij} \text{ , onde } d_{ij} \in D \text{ .}$$

**DEFINIÇÃO 5.29** *(Sentenças confirmadas, sem confirmação e não-confirmadas)*

Sejam os conjuntos de senteças T e D. Dizemos que uma sentença Q é *confirmada por* R *para* T e D se valem as seguintes condições :

(1) T U {R} é satisfatível ;

(2) T U {R} $\models$ Q ;

(3) R está na f.n.d. com respeito a D .

Quando não existir R que confirme Q , dizemos que Q está *sem confirmação*. Caso R confirme ¬Q , dizemos que R *não-confirma* Q .

**EXEMPLO 5.30**

Considere a teoria do exemplo 5.27, onde D = {¬AB(Tweety) , ¬AB(Fred)}. Observe que VOA(Tweety) é confirmado por ¬AB(Tweety), já que :

i) o modelo mínimo de T , cuja extensão de AB é {AB(Fred)}, satisfaz T ∪ {¬AB(Tweety)} ;

ii) dos axiomas (1), (3) e ¬AB(Tweety), deduzimos VOA(Tweety) , ou seja, T ∪ {¬AB(Tweety)} $\models$ VOA(Tweety) ;

iii) ¬AB(tweety) está na f.n.d. com respeito a D.

Note que não existe nenhum R que esteja na f.n.d. com respeito a D tal que T ∪ {R} seja satisfatível e T ∪ {R} $\models$ ¬AB(Fred), uma vez que AB(Fred) pertence ao modelo mínimo de T . Logo, ¬AB(Fred) fica sem confirmação. ■

**TEOREMA 5.31** *(GINSBERG (1989))*

Sejam T e D conjuntos de sentenças, e Q uma sentença . Então Q é satisfeita por todos os modelos D-máximos de T se, e somente se, existe alguma R que confirme Q tal que ¬R fique sem confirmação.

**EXEMPLO 5.32**

Observe que no exemplo 5.30, VOA(Tweety) é confirmado por ¬AB(Tweety). Além disso, AB(Tweety) é não-confirmado porque , para o conjunto D = {¬AB(Tweety) , ¬AB(Fred)} , o único R que poderia confirmar AB(Tweety) seria ¬AB(Fred) . Mas, pelo exemplo 5.30, T U { ¬AB(Fred) } é insatisfatível.

Logo, pelo teorema 5.31 , VOA(Tweety) é satisfeita por todos os modelos D-máximos de T , resultado que já havia sido obtido no exemplo 5.27 ∎

Como consequência dos teoremas 5.26 e 5.31 , temos :

**PROPOSIÇÃO 5.33**

Sejam uma teoria T sem símbolos de função , P um predicado e D o conjunto de todas as proposições da forma ¬P(t) , onde t é uma instância básica dos argumentos de P. Para qualquer sentença Q , Q é uma consequência de circunscrever P em T com todos os outros predicados variáveis se, e somente se, existe uma sentença R que confirme Q tal que ¬R seja sem confirmação .

A utilização da proposição 5.33 requer um meio através do qual seja possível determinar as várias sentenças que confirmam uma dada consulta . Observe que se considerarmos a coleção de todas as sentenças que confirmam uma certa consulta, a disjunção de todas essas sentenças também a confirmará. Além disso , a negação dessa disjunção será sem confirmação se,e somente se, a consulta for satisfeita por todos os modelos D-máximos da teoria T . Dessa forma , podemos determinar se uma consulta Q é ou não satisfeita por todos os modelos D-máximos de T considerando apenas a mais fraca das sentenças que a confirma.

Para determinar a sentença mais fraca que confirma uma dada consulta, GINSBERG (1989) utilizou o ATMS (Capítulo 3), que executa este tipo de cálculo como foi demonstrado por REITER (1987). Neste caso, a resposta à uma consulta Q será dada pelos ambientes do nó Q correspondente, que serão *as* mais fracas das sentenças a confirmar Q.

*iii.2)* *Priorizada*

No caso examinado até aqui, GINSBERG(1989) procurou construir um provador para teorias circunscritas , onde os elementos do conjunto D eram tratados igualmente , ou seja, o provador trabalhava com a circunscrição paralela de teorias . O que vamos ver nesta seção é como um provador , dentro da idéia exposta anteriormente , deve manipular a circunscrição priorizada .

Agora , os elementos do conjunto D devem ser tratados distintamente de acordo com a prioridade que lhes é atribuída . Para circunscrição priorizada , o procedimento pode ser visto como uma disputa entre um *crente* $\mathbb{B}$ e um *discrente* $\mathbb{U}$ . Primeiramente , $\mathbb{B}$ apresenta um argumento que leve a conclusão desejada . A seguir , $\mathbb{U}$ tenta replicar esse argumento , apresentando um contra-argumento cuja prioridade seja maior ou igual a do argumento que $\mathbb{B}$ obteve. Então , $\mathbb{B}$ procura refutar o contra-argumento , buscando um contra-contra-argumento de prioridade estritamente maior que o apresentado por $\mathbb{U}$ . O processo prossegue até que um dos dois não tenha mais resposta . A sentença original será uma consequência lógica da teoria circunscrita (com prioridade) quando a última palavra for dada por $\mathbb{B}$ . No caso de $\mathbb{U}$ dar a última palavra , a sentença não segue da teoria circunscrita .

Novamente , ao invés de trabalhar com o axioma da circunscrição, BAKER e GINSBERG (1989-a) trabalharam com conjuntos. A diferença em relação a circunscrição paralela é que aqui utiliza-se os conjuntos $D_1,\ldots,D_n$ , onde $D_i$ $(i=1,\ldots,n)$ são conjuntos cujos elementos têm prioridade i. Logo , $D_n$ possui os elementos de maior prioridade e $D_1$ os de menor prioridade .

**DEFINIÇÃO 5.34** *(Relação $>_D$)*

Sejam T um conjunto de sentenças e $D_1,\ldots,D_n$ conjuntos finitos de proposições. Definimos uma ordem parcial sob os modelos de T , escrevendo $M_1 >_D M_2$ , se vale a seguinte condição : para algum $i \in [1,n]$ , $M_2[D_i] \subset M_1[D_i]$ ; e para todo $j \in (i,n]$ , $M_1[D_j] = M_2[D_j]$ , sendo que $M_k[D_l]$ denota o conjunto das sentenças em $D_l$ que são satisfeitas por $M_k$ .

**EXEMPLO 5.35**

Por exemplo, se tivéssemos os seguintes conjuntos :

$D_1 = \{\neg AB_1(Tweety), \neg AB_4(Tweety)\}$

$D_1 = \{\neg AB_2(Tweety)\}$

$D_3 = \{\neg AB_3(Tweety)\}$

e considere os modelos $M_1$ e $M_2$ de um conjunto de sentenças T, que satisfazem, respectivamente, $\{\neg AB_2(Tweety), \neg AB_3(Tweety), \neg AB_4(Tweety)\}$ e $\{\neg AB_3\}$. Assim, $M_1 >_D M_2$ pois para $i = 2$, $M_2[D_2] \subset M_1[D_2]$ , e para $j = 3$ , $M_2[D_3] = M_1[D_3]$. Note que para $i = 1$, $M_2[D_1] \subset M_1[D_1]$ mas não temos $M_2[D_2] = M_1[D_2]$ ■

Um modelo máximo na ordem parcial $>_D$ será chamado de *D-maximal* .

A partir da definição 5.34 , podemos estabelecer o seguinte resultado , que equivale ao teorema 5.26 no caso da circunscrição paralela :

**TEOREMA 5.36** *(BAKER e GINSBERG (1989-a))*

Considere um teoria T sem símbolos de função e *P* uma coleção de predicados que estão priorizados em n níveis tais que os de nível n têm maior prioridade e os de nível 1 têm menor prioridade . Seja $D_i$ o conjunto de todas as proposições da forma $\neg P(t)$ onde P é um predicado de nível i e t é uma instância básica dos argumentos de P . Para qualquer sentença Q , Q segue da circunscrição priorizada de todos os predicados de P em T com todos os outros predicados variando se , e somente se , Q é satisfeita por todos os modelos D-máximos de T .

A demonstração deste teorema é obtida utilizando-se o teorema 5.26 (quando tinhamos um único conjunto D) e o teorema 4.30 (que relaciona a circuncrição priorizada com a paralela).

OBSERVAÇÃO 9 : Como no teorema 5.26 , a teoria T deve ser considerada fechada sob o domínio e satisfazendo as hipóteses do nome único.

DEFINIÇÃO 5.37 (*Argumento*)

Seja P uma sentença . Dizemos que P é um *argumento* se as seguintes condições valem :

1) P é da forma $\bigvee_i \bigwedge_j d_{ij}$ , para alguma coleção de $d_{ij} \in D_1 \cup \ldots \cup D_n$ ;

2) $T \cup \{P\}$ é satisfatível .

Se Q é uma sentença , dizemos que P é um argumento para Q se P é um argumento e $T \cup \{P\} \models Q$ .

EXEMPLO 5.38

Considerando a teoria T( <$AB_1$,$AB_2$,$AB_3$> ; <VOA,AVE, CANARIO> ) :

(1) $AB_1$(Tweety) ∨ ¬VOA(Tweety)

(2) ¬AVE(Tweety) ∨ $AB_1$(Tweety)

(3) ¬AVE(Tweety) ∨ $AB_2$(Tweety) ∨ VOA(Tweety)

(4) ¬CANARIO(Tweety) ∨ $AB_3$(Tweety) ∨ AVE(Tweety)

(5) CANARIO(Tweety)

e $D_1$ = {¬$AB_1$(Tweety)} , $D_2$ = {¬$AB_2$(Tweety)} e $D_3$ = {¬$AB_3$(Tweety)}.

Temos T $\models$ $AB_1(Tweety) \vee AB_2(Tweety) \vee AB_3(Tweety)$, uma vez que :

(6) $\neg AB_1(Tweety)$

(7) $\neg AB_2(Tweety)$

(8) $\neg AB_3(Tweety)$

(9) $AB_3(Tweety) \vee AVE(Tweety)$ resolvente de ((4),(5))

(10) $AB_3(Tweety) \vee AB_2(Tweety) \vee VOA(Tweety)$ resolvente de ((3),(9))

(11) $AB_3(Tweety) \vee AB_2(Tweety) \vee AB_1(Tweety)$ resolvente de ((1),(10))

(12) $\square$ resolvente de (6),(7),(8),(11)

Com isso, um argumento para esta teoria seria $\neg AB_3(Tweety) \vee \neg AB_2(Tweety)$, pois satisfaz as condições da definição 5.37 . Além disso, $\neg AB_3(Tweety) \vee \neg AB_2(Tweety)$ é um argumento para $AB_1(Tweety)$, pois $T \cup \{\neg AB_3(Tweety) \vee \neg AB_2(Tweety)\} \models AB_1(Tweety)$ .

Já $(\neg AB_1(Tweety) \wedge \neg AB_3(Tweety)) \vee (\neg AB_3(Tweety) \wedge \neg AB_2(Tweety))$ não é argumento da teoria , pois tal sentença é equivalente a :

$$(\neg AB_1(Tweety) \vee \neg AB_2(Tweety)) \wedge$$
$$(\neg AB_1(Tweety) \vee \neg AB_3(Tweety)) \wedge$$
$$(\neg AB_3(Tweety) \vee \neg AB_2(Tweety)) \wedge$$
$$\neg AB_3(Tweety)$$

que junto com a teoria T( $\langle AB_1, AB_2, AB_3\rangle$ ; $\langle VOA, AVE, CANARIO\rangle$ ) é insatisfatível . Aliás, qualquer P que inclua na teoria $\neg AB_1(Tweety)$, $\neg AB_2(Tweety)$ e $\neg AB_3(Tweety)$ não será argumento pois $T \cup \{\neg AB_1(Tweety), \neg AB_2(Tweety), \neg AB_3(Tweety)\}$ é insatisfatível ■

**OBSERVAÇÃO 10 :** A prioridade de um argumento é o menor K tal que o argumento contém algum $d_{ij} \in D_k$ . No exemplo acima , a prioridade do argumento $\neg AB_3(Tweety) \vee \neg AB_2(Tweety)$ é 2 .

Para utilizarmos a idéia descrita no início da seção de como o processo funciona , precisamos definir as noções de replicar e refutar .

**DEFINIÇÃO 5.39** *(Replicar e Refutar Argumento)*

Sejam P e Q argumentos . Dizemos que P *replica* Q se, e somente se, P é um argumento para $\neg Q$ , e a prioridade de P *é maior ou igual a prioridade de Q .*

Dizemos que P *refuta* Q se, e somente se , P é um argumento para $\neg Q$ , e a prioridade de P é estritamente maior que a prioridade de Q .

Lembre que o processo de replica e refutação de argumentos é na verdade uma disputa. Em um determinado momento um argumento ficará sem resposta, ou seja , poderá não ser replicado (ou refutado). No caso de um argumento que replique (resp. refute) outro e que não possa ser refutado (resp. replicado), teremos uma réplica (resp. refutação) definitiva. Formalmente, temos :

**DEFINÇÃO 5.40** *(Réplica e Refutação Definitivas)*

Sejam P e Q argumentos . Dizemos que P *replica definitivamente* Q se, e somente se, P replica Q e P não é refutado definitivamente .

Dizemos que P *refuta definitivamente* Q se, e somente se, P refuta Q e P não é replicado definitivamente .

Na verdade, na definição 5.40, uma réplica definitiva no nível K é definida em termos de refutações definitivas de nível (pelo menos) K+1, já que um argumento refutado deve ter sempre uma prioridade maior que o argumento que ele refuta.

**EXEMPLO 5.41**

Voltando a teoria do exemplo 5.38 :

(1) $AB_1(Tweety) \vee \neg VOA(Tweety)$

(2) $\neg AVE(Tweety) \vee AB_1(Tweety)$

(3) $\neg AVE(Tweety) \vee AB_2(Tweety) \vee VOA(Tweety)$

(4) $\neg CANARIO(Tweety) \vee AB_3(Tweety) \vee AVE(Tweety)$

(5) $CANARIO(Tweety)$

Se fizermos a consulta VOA(Tweety) à teoria acima , circunscrevendo $AB_1$, $AB_2$ e $AB_3$ de acordo com os conjuntos $D_1$, $D_2$ e $D_3$ do exemplo 5.38 , e permitindo que os outros predicados variem , devemos inicialmente encontrar um argumento para VOA(Tweety) .

Note que , pela definição 5.37 , $\neg AB_3(Tweety) \wedge \neg AB_2(Tweety)$ é um argumento para VOA(Tweety) , pois como vimos no exemplo 5.38, além de ser um argumento da teoria, temos T ∪ {¬$AB_2$(Tweety) ∨ ¬$AB_3$(Tweety)} ⊨ VOA(Tweety) . De fato, por resolução :

(7) ¬$AB_3$(Tweety)
(8) ¬$AB_2$(Tweety) } argumento
(9) ¬ VOA(Tweety)
(10) $AB_3$(Tweety) ∨ AVE(Tweety) resolvente de ((4),(5))
(11) $AB_3$(Tweety) ∨ $AB_2$(Tweety) ∨ VOA(Tweety) resolvente de ((3),(10))
(12) VOA(Tweety) resolvente de ((7),(8),(11))
(13) □ resolvente de ((9),(11))

O próximo passo seria procurar um argumento que replicasse ¬$AB_3$(Tweety) ∧ ¬$AB_2$(Tweety) , isto é , um argumento para $AB_3$(Tweety) ∨ $AB_2$(Tweety) que tenha prioridade maior ou igual a 2. O único argumento para $AB_3$(Tweety) ∨ $AB_2$(Tweety) é ¬$AB_1$(Tweety) , pois T ⊨ $AB_1$(Tweety) ∨ $AB_2$(Tweety) ∨ $AB_3$(Tweety) , e portanto T ∪ {¬$AB_1$(Tweety)} ⊨ $AB_2$(Tweety) ∨ $AB_3$(Tweety), mas este tem prioridade 1 e , portanto , não nos serve .

Logo , não é possível replicar ¬$AB_3$(Tweety) ∧ ¬$AB_2$(Tweety), ou seja, não é replicado definitivamente ■

A próxima proposição estabelece que quando existe um argumento para uma sentença Q, que não seja replicado definitivamente , Q será satisfeita por todos os modelos D-máximos. Logo, VOA(Tweety) pertencerá à todos os modelos

D-máximos , o que pelo teorema 5.36, corresponde a tê-lo como consequência lógica da teoria circunscrita. Novamente, este resultado corresponde ao que foi obtido anteriormente.

PROPOSIÇÃO 5.42 *(BAKER e GINSBERG (1989-a))*

Seja Q uma sentença . Então Q é satisfeita por todos os modelos de T se, e somente se, existe um argumento para Q que não é replicado definitivamente .

Tal como no caso da circunscrição paralela, para implementar esta proposição , é necessário encontrar os vários argumentos para uma dada sentença. Ao invés de considerarmos cada um deles individualmente , pelo grande número de argumentos e de contra-argumentos, vamos nos restringir ao argumento mais fraco , que é obtido pela disjunção de todas as conjunções de argumentos mínimos que possuem as devidas prioridades .

PROPOSIÇÃO 5.43 *(BAKER e GINSBERG (1989-a))*

Sejam Q uma sentença e i um nível de prioridade. Assuma que existe algum argumento para Q com, pelo menos , essa prioridade . Seja P a disjunção de todas as conjunções de argumentos mínimos para Q que tenha prioridade de , pelo menos , i . Então :

(1) P é um argumento para Q com prioridade de, pelo menos , i ;

(2) P é replicado definitivamente (refutado) se, e somente se, todo argumento para Q com prioridade de , pelo menos, i é replicado definitivamente (refutado)

Os argumentos mais fracos são obtidos através de um ATMS, por um processo semelhante ao caso da circunscrição paralela.

*iv) Conclusão*

Neste capítulo, apresentamos três procedimentos especiais que foram desenvolvidos com o objetivo de determinar quando um certo fato segue ou não de teorias circunscritas, uma vez que a circunscrição geralmente resulta em teorias de segunda-ordem.

Apesar de todo desenvolvimento apresentado pelo estudo da circunscrição ao longo dos seus dez anos de vida, um problema pouco abordado foi o de determinar qual política de circunscrição deve ser utilizada, ou seja , na aplicação do método, quais devem ser os predicados a serem circunscritos, quais devem variar e quais devem permanecer fixos.

Nos exemplos em geral, os predicados de anormalidade são os escolhidos para serem circunscritos (ou minimizados). A idéia aqui é considerar que a teoria nos diz, implicitamente, quais dos seus objetos são anormais. Circunscrevê-los significa explicitá-los na teoria resultante do processo.

Quanto a quais dos predicados devem variar no processo de circunscrição é uma pergunta que não possui uma resposta objetiva. Tal problema surgiu de forma mais clara quando aplicamos os procedimentos apresentados neste capítulo no caso da circunscrição priorizada.

No exemplo inicial da seção (iii) do capítulo IV, McCARTHY (1986) sugeriu que com uma simples alteração de política de circunscrição, na qual atribuiam-se prioridades sobre os predicados de anormalidade $AB_1$, $AB_2$ e $AB_3$, e tendo o predicado VOA como o único a variar, seria possível solucionar o problema surgido, ou seja, poderíamos concluir da teoria envolvida que Tweety voa (exemplos 5.15, 5.24 e 5.38).

A escolha do predicado VOA como variável foi baseada em uma intuição de McCARTHY (1986) : como a proposta do conjunto de axiomas é descrever o que voa, o predicado VOA deve variar.

Infelizmente, esta política intuitiva de McCarthy parece não ser adequada. Empregando a MILO-Resolução ou a Z-Resolução com tal política, não é possível obter que Tweety voa. Em ambos os métodos, a resposta adequada é obtida tomando os predicados VOA,AVE e CANARIO como variáveis, embora nenhuma restrição nesse sentido seja imposta pelos procedimentos.

Já no caso do provador de GINSBERG (1989) é imposta a restrição de que todos os predicados que não são circunscritos devem variar. Aqui, embora indicado quais predicados devem variar, não é explicado o por que.

Como veremos no próximo capítulo (seção (vi)), definir uma política de circunscrição adequada pode ser fundamental para a solução correta dos problemas que surgem quando da utilização de métodos não-monotônicos na formalização do senso-comum. É importante ressaltar que a análise para a definição de tal política deve ser abrangente , de maneira a estabelecer quem deve ser minimizado, quem deve variar, quem deve ser fixo e qual a prioridade deve ser usada.

# CAPÍTULO VI

## O Problema do Tiro

Neste capítulo vamos examinar o problema do tiro (YSP - Yale Shooting Problem), apresentado por HANKS e McDERMOTT(1987), onde é estudada a aplicação de lógicas não-monotômicas na formalização de situações temporais.

A seção (i) apresenta o YSP, e mostra que as anomalias apresentadas por este problema não são exclusivas dos casos temporais.

Na seção (ii), é apresentada uma tentativa de incorporação à lógica de default e à circunscrição, da característica que faz com que o TMS trate de forma adequada o YSP.

Nas seções (iii) e (iv) são apresentados novos critérios de seleção dos modelos mínimos, e que são chamados de minimização (resp.) cronológica e causal.

Na seção (v) é apresentada uma lógica temporal não-monotônica e na seção (vi) é examinada uma nova política de circunscrição.

### *i) O Problema do Tiro ( YSP - Yale Shooting Problem)*

Neste capítulo, vamos examinar o problema do tiro (YSP - Yale Shooting Problem) apresentado por HANKS e

McDERMOTT(1987) , onde são analisados os resultados obtidos quando da utilização de lógicas não-monotônicas na formalização de problemas temporais.

A motivação para esta análise surgiu do fato de que até aquele instante, HANKS e McDERMOTT(1987) consideravam que os estudos no campo das lógicas não-monotônicas podiam ser divididos em duas classes : na primeira, onde havia uma preocupação mais técnica e os exemplos apresentados produziam (propositalmente) resultados não-intuitivos. Na segunda classe, ao contrário, as lógicas não-monotônicas eram usadas sem uma preocupação maior com detalhes teóricos.

Assim, eles desenvolveram este trabalho buscando um meio termo: apresentando um problema (o da projeção temporal), formalizando-o através das lógicas não-monotônicas e então analisando a teoria resultante.

Dois motivos os levaram a escolher um problema temporal : o fato do problema de *frame* (tendência dos fatos persistirem ao longo do tempo) ser considerado como tendo solução através da utilização dessas lógicas; e a forte intuição de como tais problemas devem ser formalizados e que *inferências podem ser feitas*.

HANKS e McDERMOTT (1987) consideraram a seguinte situação para estudar o problema da projeção temporal : inicialmente uma pessoa, chamada Fred, está viva e uma arma está descarregada . Sabemos que se a arma estiver carregada e for disparada, Fred morrerá. O problema da projeção temporal é determinar quais fatos serão verdadeiros em cada situação, depois da ocorrência da seguinte sequência de eventos :

i) carregar a arma ;

ii) esperar alguns momentos , onde nada de relevante ocorre;

iii) disparar a arma;

O problema de se saber se é possível deduzir não-monotonicamente que a pessoa estará morta no final da sequência de eventos i , ii e iii é conhecido como o *problema do tiro* (YSP - Yale Shooting Problem) .

A representação temporal utilizada por HANKS e McDERMOTT(1987) no tratamento do YSP é o cálculo situacional (McCARTHY e HAYES(1969)). Nele, as situações representam instantâneos do mundo em um determinado instante de tempo . A transição de uma situação para outra se dá como resultado da ocorrência de um evento. Na representação do YSP são usadas as seguites variáveis : v para valores verdade, e para os eventos, s para situações e f para os fatos (ou fluentes) ; e as seguintes constantes : valores verdade *true* e *false* , os eventos *carregar*, *esperar* e *disparar* , as situações $s_0$, $s_1$, $s_2$, $s_3$, os fatos *carregada*, *vivo* e *morto*.

Um fato (ou fluente) f que é verdadeiro em uma situação s será representado por T(f,s) . A transição de uma situação para outra será representada pela função r(e,s) que mapeia um evento e e uma situação s em outra situação . O predicado AB(f,e,s) representa que o fato f é anormal (ou afetado) quando o evento e ocorre no estado s .

HANKS e McDERMOTT (1987) representaram o YSP através dos seguintes axiomas :

(1) $T(vivo, s_0)$ ;

(2) $\forall s\ [\ T(carregada, r(carregar, s))\ ]$ ;

(3) $\forall s\ [\ T(carregada, s) \supset AB(vivo, disparar, s) \wedge$
$T(morto, r(disparar, s))\ ]$ ;

(4) $\forall f, e, s\ [\ T(f, s) \wedge \neg AB(f, e, s) \supset T(f, r(e, s))\ ]$ ;

onde (1) representa que numa dada situação $s_0$, Fred está vivo ; (2) que a arma fica carregada quando o evento carregar ocorre ; (3) que em qualquer estado que a arma estiver carregada, o evento disparar afeta o fato de Fred estar vivo e, como resultado desse evento, Fred morrerá ; e (4) é o axioma *frame* que diz que fatos *normais* persistem ao longo da ocorrência de eventos *normais* .

As situações nas quais vamos estudar o problema da projeção temporal são :

$$s_0;$$
$$s_1 = r(carregar, s_0)\ ;$$
$$s_2 = r(esperar, s_1)\ ;$$
$$s_3 = r(disparar, s_2)\ .$$

**OBSERVAÇÃO 1 :** Os axiomas não representam nenhum fato anormal com respeito ao evento esperar , refletindo a idéia de que os fatos verdadeiros antes deste evento ocorrer, também o serão depois.

Se considerarmos apenas a teoria $\tau$ composta pelos axiomas (1)-(4) , os únicos fatos que poderemos deduzir são os seguintes :

(a) pelo axioma (1), Fred está vivo na situação inicial $s_0$ ( T(vivo,$s_0$) );

(b) pelo axioma (2), que a arma está carregada na situação $s_1$ ( T(carregada,$s_1$) ).

Observe que os fatos (a) e (b) não persistem ao longo do tempo, respectivamente, até a situação $s_1$ e $s_2$. Isto porque não é possível deduzir que tais fatos são normais com relação (resp.) aos eventos <u>carregar</u> e <u>esperar</u>, ou seja, não temos como obter da teoria $\tau$ que :

$$\neg \text{AB(vivo,carregar,}s_0\text{) e}$$
$$\neg \text{AB(carregada,esperar,}s_1\text{)}$$

Consequentemente , não é possível saber quais fatos são verdadeiros nas situações $s_2$ e $s_3$. Para tentar resolver este problema , vamos minimizar as anormalidades, circunscrevendo AB na teoria $\tau$ com T variando. Assim temos:

(a) Como nada na teoria $\tau$ nos leva a crer que o fato de Fred estar vivo seja anormal com relação ao evento <u>carregar</u> no estado $s_0$, pelo senso comum, podemos assumir a sua normalidade :

$$\neg \text{AB(vivo,carregar,}s_0\text{)} \qquad \text{(a.1)}$$

Dos axiomas (1) e (4) e mais (a.1) , deduzimos :

$$\text{T(vivo,}s_1\text{)} \qquad \text{(a.2)}$$

(b) Agora, temos no estado $s_1$ os seguinte fatos :

$$T(vivo,s_1) \text{ e } T(carregada,s_1)$$

Novamente, podemos assumir, pelo senso comum, que os fatos <u>vivo</u> e <u>carregada</u> são normais com relação ao evento <u>esperar</u> , já que a teoria $\tau$ não nos diz nada em contrário. Logo,

$$\neg\ AB(vivo,esperar,s_1) \quad (b.1)$$

$$\neg\ AB(carregada,esperar,s_1) \quad (b.2)$$

que juntos com o axioma (4) nos fornecem :

$$T(vivo,s_2) \quad (b.3)$$

$$T(carregada,s_2) \quad (b.4)$$

(c) Na situação $s_2$, o axioma (3) e o fato (b.4) nos dão que:

$$T(morto,s_3) \quad (c.1)$$

$$AB(vivo,disparar,s_2) \quad (c.2)$$

Portanto, é anormal que Fred permaneça vivo quando o evento <u>disparar</u> ocorre no estado $s_2$. Já o fato da arma estar carregada é normal com respeito a este mesmo evento, donde, pelo senso comum :

$$\neg\ AB(carregada,disparar,s_2) \quad (c.3)$$

De (c.3) e do axioma (4) , temos :

T(carregada,$s_3$) (c.4)

Resumindo , temos o seguinte modelo $M_1$:

i) Fred está vivo nos estados $s_0$, $s_1$ e $s_2$ , justificados respectivamente pelo axioma (1), por (a.2) e (b.3);

ii) a arma está carregada nos estados $s_1$, $s_2$ e $s_3$, justificados respectivamente pelo axioma (2), por (b.4) e (c.4);

iii) Fred está morto em $s_3$ ((c.1));

iv) Fred estar vivo é um fato anormal com relação ao evento disparar em $s_2$ ((c.2));

| $s_0$ | $s_1$ | $s_2$ | $s_3$ |
|---|---|---|---|
| vivo | vivo<br><br>carregada | vivo<br><br>carregada<br><br>AB(vivo,disparar,$s_2$) | $\neg$ vivo<br><br>carregada |

Fig. 6.1 : representação dos fluentes verdadeiros em $M_1$

Essa estrutura $M_1$ é um modelo de $\tau$ e , além disso , é mínima com relação ao predicado AB. Isto porque a extensão do predicado AB teria que ser vazia para que existisse um submodelo de $M_1$ , o que não acontece. Observe que :

T(vivo,$s_2$)

é verdadeiro no modelo $M_1$ . Caso este fato fosse falso, AB(vivo,disparar,$s_2$) não pertenceria a extensão de AB, mas esta não seria vazia já que uma das seguintes anormalidades deveria pertencer a extensão de AB :

$$AB(vivo, carregar, s_0)$$

$$AB(vivo, esperar, s_1) ,$$

pois no estado $s_2$ Fred não mais estaria vivo por causa da ocorrência ou do evento carregar ou esperar .

O modelo mínimo $M_1$ é intuitivo , uma vez que no final da sequência de eventos i, ii e iii, Fred estará morto como era esperado. O problema é que tal fato será consequência de circunscrever AB em $\tau$ com T variando se, e somente se, ele for satisfeito por todos os modelos mínimos da teoria $\tau$ ou caso o modelo $M_1$ seja o único mínimo (proposição 4.7), o que não acontece .

Um outro modelo mínimo $M_2$ pode ser contruído e, para isto considere o seguinte cenário :

| $s_0$ | $s_1$ | $s_2$ | $s_3$ |
|---|---|---|---|
| vivo | vivo<br><br>carregada | vivo | vivo |

Fig 6.2 : Representação dos fluentes verdadeiros em $M_2$

onde Fred está vivo em todos os estados e a arma está carregada só no estado $s_1$.

Observe que o resultado da ocorrência do evento esperar faz com que a arma fique descarregada (estado $s_2$). Desta forma, Fred agora não morrerá como resultado do disparo da arma em $s_2$, já que esta não se encontra carregada em tal situação.

Embora seja contraditório com o fato expresso pelos axiomas de que nenhuma anormalidade com relação ao evento <u>esperar</u> esteja explícita (observação 1) , e portanto, tudo deveria ser normal com respeito a tal evento, assumir a anormalidade :

$$AB(carregada, esperar, s_1)$$

gera um modelo mínimo com relação a AB.

Fazendo o mesmo tipo de construção do modelo anterior, vamos trabalhar de trás para frente, começando em $s_3$, onde Fred está vivo. Assim,

(a') Como nada na teoria $\tau$ nos faz crer que o fato de Fred estar vivo seja anormal com relação ao evento <u>atirar</u> em $s_2$, podemos assumir :

$$\neg\ AB(vivo, disparar, s_2) \qquad (a'.1)$$

Isto mais o axioma (3) escrito na forma :

$$(3')\ \forall s\ [\ \neg\ AB(vivo, disparar, s) \vee T(morto, r(disparar, s)) \supset \neg\ T(carregada, s)\ ]\ ,$$

nos permite deduzir que :

$$\neg\ T(carregada, s_2) \qquad (a'.2)\ .$$

(b') Como $T(carregada, s_1)$, temos pelo axioma (4) mais (a'.2) que :

AB(carregada, esperar, $s_1$) (b'.1) .

Além disso, pelo senso comum, podemos assumir que estar vivo é normal com relação aos eventos carregar e esperar , e portanto Fred estará vivo em $s_1$ e $s_2$.

Resumindo :

i) Fred está vivo nos estados $s_0$, $s_1$, $s_2$ e $s_3$;

ii) a arma está carregada no estado $s_1$ e descarregada nos demais ;

iii) a arma carregada é um fato anormal com respeito ao evento esperar no estado $s_1$ ((b'.1)) ;

que corresponde ao cenário descrito na figura 6.2.

Assim, temos dois modelos mínimos com relação ao predicado AB e portanto, os fatos que são consequências de circunscrever AB com T variando, deverão pertencer à ambos os modelos, o que só nos permite concluir que :

i) em $s_1$, Fred está vivo e a arma está carregada;

ii) em $s_2$, que Fred está vivo;

iii) em $s_3$, nada podemos deduzir além de fatos tais como: Fred está vivo ou morto. ( T(vivo, $s_3$) $\vee$ T(morto, $s_3$) )

| $s_0$ | $s_1$ | $s_2$ | $s_3$ |
|---|---|---|---|
| vivo | vivo<br><br>carregada | vivo | carregada |

Fig. 6.3 : fluentes comuns a $M_1$ e $M_2$

Observe que ambos os modelos foram construídos de maneira semelhante : uma "normalidade" é assumida, todas as deduções possíveis são feitas, outra "normalidade" é assumida, e assim por diante. Como resultado deste processo, em ambos os casos, obtemos uma anormalidade.

A única diferença na construção dos modelos está na normalidade que é assumida : no modelo intuitivo $M_1$, por estarmos trabalhando no sentido natural do tempo, as *normalidades* assumidas são aquelas que ocorrem o mais cedo possível; enquanto no modelo não-intuitivo $M_2$, onde trabalhamos de trás para frente no tempo, as *normalidades* assumidas são aquelas que ocorrem o mais tarde possível.

Assim , HANKS e McDERMOTT (1987) sugeriram que o problema surgido pela utilização da circunscrição na formalização do problema da projeção temporal seria o seu critério de minimalidade, que seleciona os modelos através da noção de inclusão de conjuntos da extensão do predicado que está sendo circunscrito. Na realidade, o critério que eles julgaram correto é o da *minimização cronológica* (SHOHAM (1988)), no qual as anormalidades ocorrem tão mais tarde quanto possível. Tal critério não pode ser expresso em termos do conceito de minimalidade da circunscrição.

O problema da geração de vários modelos no YSP não fica restrito a circunscrição. Outros métodos de formalização não-monotônoica apresentam esta mesma dificuldade (REITER e CRISCUOLO(1981) ; MORRIS(1988)). Consideremos que o YSP fosse representado na lógica de default através da teoria $\Delta' = (D', W')$ onde W' é composta pelos axiomas (1)-(4) e D' é o conjunto formado pelos defaults da forma $\langle\ :\neg AB(f,e,s)\ /\ \neg AB(f,e,s)\ \rangle$.

A teoria default $\Delta' = (D', W')$ possui duas extensões, que são :

$E_1 = Th(W' \cup D_1)$ onde $D_1 = \{ \neg ABC(vivo, carregar, s_0),$

$\neg ABC(vivo, esperar, s_1),$

$\neg ABC(carregada, esperar, s_1),$

$\neg ABC(carregada, disparar, s_2) \},$

que corresponde ao modelo intuitivo $M_1$; e

$E_2 = Th(W' \cup D_2)$ onde $D_2 = \{ \neg ABC(vivo, carregar, s_0),$

$\neg ABC(vivo, esperar, s_1),$

$\neg ABC(vivo, disparar, s_2),$

$\neg ABC(carregada, disparar, s_2) \},$

correspondente ao modelo não-intuitivo $M_2$.

Observe que os conjuntos $D_1$ e $D_2$ possuem os mesmos elementos com exceção de $\neg ABC(vivo, disparar, s_2)$ e $\neg ABC(carregada, esperar, s_1)$. Vamos examinar o que ocorre quando da aplicação dos defaults correspondentes à esses dois elementos , nos restringindo aos axiomas envolvidos :

(1') $T(carregada, s_1)$ (axioma (2))

(2') $T(carregada, s_1) \wedge \neg ABC(carregada, esperar, s_1) \supset T(carregada, s_2)$ (axioma (4))

(3') $T(vivo, s_2)$ (ax. (1), (4) e defaults $\neg ABC(vivo, carregar, s_0)$ e $\neg ABC(vivo, esperar, s_1)$)

(4') $T(vivo, s_2) \wedge \neg ABC(vivo, disparar, s_2) \supset T(vivo, s_3)$ (ax. (4))

(5') $T(carregada, s_2) \supset T(morto, s_3) \wedge ABC(vivo, disparar, s_2)$ (ax. (3))

e os respectivos defaults :

$D_1 \Leftrightarrow (:\neg AB(carregada,esperar,s_1) \ / \ \neg AB(carregada,esperar,s_1))$

$D_2 \Leftrightarrow (:\neg AB(vivo,disparar,s_2) \ / \ \neg AB(vivo,disparar,s_2))$ .

Aplicando o default $D_1$ ao conjunto formado por (1')-(5') , obtemos a partir de (1'), (2') e $\neg AB(carregada,esperar,s_1)$ que :

$$T(carregada,s_2) \qquad (6')$$

De (6') e do axioma (5') obtemos que :

$$T(morto,s_3) \wedge AB(vivo,disparar,s_2) ,$$

onde a segunda fórmula da conjunção acima inibe a aplicação do default $D_2$.

Caso aplicássemos o default $D_2$ às fórmulas (1')-(5') antes do default $D_1$, obteríamos de (5') e $\neg AB(vivo,disparar,s_2)$ que :

$$\neg \ T(carregada,s_2) \qquad (7')$$

De (7') , (1') e (2') , temos :

$$\neg AB(carregada,esperar,s_1) ,$$

que impede a aplicação do default $D_1$.

Logo, a aplicação de uma regra default inibe a outra (e vice-versa), e não existe nenhum meio próprio da lógica de default que trabalhe como mediador na aplicação das regras de forma cronológica, que segundo sugerido por HANKS e McDERMOTT(1987), nos possibilitaria obter apenas a extensão intuitiva $E_1$.

Esse tipo de problema apresentado pelo YSP não é uma característica própria deste problema. A seguir veremos uma versão simplificada do problema do mundo dos blocos (McCARTHY(1986)).

O problema do mundo dos blocos, como o próprio nome diz, consiste de um mundo composto por blocos que podem ser pintados e movimentados. Considere um mundo composto por apenas dois blocos <u>A</u> e <u>B</u>. Tais blocos se encontram sobre uma mesa e podem ser movidos para qualquer posição na mesa que esteja livre ( ou seja, posições não ocupadas pelo outro bloco ) e para cima do outro bloco, caso não haja nada que o impeça.

Vamos utilizar o seguinte conjunto de axiomas na formalização do problema :

(1) local(A,so) = $m_1$

(2) local(B,so) = $m_2$

(3) LIVRE(topo A,so)

(4) LIVRE(topo B,so)

Esses axiomas expressam as posições iniciais dos blocos A e B (respectivamente, $m_1$ e $m_2$) e indicam que não existe nada sobre nenhum dos blocos no estado inicial so. A situação acima pode ser representada pela seguinte figura :


situação so

Os outros axiomas necessários são :

$$(5)\ \forall x,l,s\ (\ \neg AB(x,l,s) \supset local(x,r(mover(x,l),s)) = l)$$

$$(6)\ \forall x,l,s\ (\ \neg LIVRE(topo\ x,s) \vee \neg LIVRE(l,s) \vee l = topo\ x \supset AB(x,mover(x,l),s) \wedge r(mover(x,l),s) = s)$$

$$(7)\ \forall l,s\ (\ LIVRE(l,s) \equiv (\forall x\ local(x,s) \neq l))$$

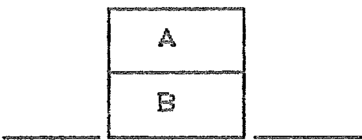
onde (5) representa que normalmente um bloco $\underline{x}$, quando movido para uma posição $\underline{l}$, num estado $\underline{s}$, estará localizado em $\underline{l}$; (6) representa que não se pode mover um bloco $\underline{x}$ para $\underline{l}$ no estado $\underline{s}$ quando ou o topo do bloco $\underline{x}$ está ocupado ou quando $\underline{l}$ está ocupado ou quando $\underline{l}$ for o topo do próprio bloco que está sendo movimentado; (7) uma posição está livre quando nenhum bloco a ocupa.

Vamos considerar além da situação inicial $s_0$ , os seguintes casos :

$$s_1 = r(mover(A,topo\ B),s_0)\ ,\ e$$

$$s_2 = r(mover(B,topo\ A),s_1)\ .$$

Observe que a primeira ação (mover o bloco A para o topo de B) é bem sucedida pois não há nada que impeça tal ação (pelo senso-comum, podemos assumir $\neg AB(A,topo\ B,s_0)$). Logo, a situação resultante $s_1$ seria :




situação $s_1$

Agora não é possível mover o bloco B para o topo de A, isto porque temos, do axioma (7), que $\neg LIVRE(topo\ B,s_1)$.

Com este fato mais o axioma (6) temos :

$$AB(B,mover(B,topo\ A),s_1)\ e$$

$$s_2 = r(mover(B,topo\ A),s_1) = s_1$$

ou seja, a ação acima não pode ser executada e portanto as situações $s_1$ e $s_2$ são idênticas.

O modelo que obtivemos acima é intuitivo. Entretanto, ao circunscrevermos o predicado AB na teoria composta por (1)-(7) , tal como no YSP, obtemos dois modelos mínimos : o intuitivo (acima, onde a extensão de AB é igual a { (B,mover(B,topo A),$s_1$)}) , e outro não-intuitivo onde a primeira ação é que não pode ser feita e cuja extensão de AB é igual a {(A,mover(A,topo B),$s_0$)}. Neste caso, as situações $s_0$ e $s_1$ são identicas.

Na verdade , como vimos no capítulo 2 (exemplo 2.15) , o surgimento de extensões anômalas não é uma característica dos problemas que envolvem tempo ( como no caso do YSP e do mundo dos blocos), e podem surgir mesmo para problemas não temporais. MORRIS (1988) salientou este fato apresentando o seguinte exemplo não-temporal que partilha com o YSP, a característica de que ao assumirmos uma normalidade, podemos obter uma anormalidade, e consequentemente, chegar a uma extensão não-intuitiva.

EXEMPLO 6.1 (MORRIS (1988))

Sejam :

A : "Tweety é um animal" ; F : "Tweety voa" ;

B : "Yweety é uma ave" ; S : "Tweety tem asas" ;

AB(A) : "Tweety é uma animal anormal com respeito a não voar" ;

AB(B) : "Tweety é uma ave anormal com respeito a ter asas"

Considere a seguinte teoria default Δ" = (D",W"), onde W" é composta por :

(1") $A \wedge \neg AB(A) \supset \neg F$

(2") $S \supset AB(A)$

(3") $B \supset A$

(4") $B \wedge \neg AB(B) \supset S$

(5") $S \supset F$

(6") $B$

onde (1") representa que se Tweety for um animal normal com respeito a não voar, então ele não voa ; (2") representa que se Tweety tem asas então ele é um animal anormal com respeito a não voar ; (3") representa que se Tweety é uma ave, ele é um animal ; (4") representa que se Tweety é uma ave normal com respeito a ter asas, então ele tem asas ; (5") representa que se Tweety tem asas, ele voa .

O conjunto D" é formado por :

$$\{ :(\neg AB(A))/\neg AB(A) \ , \ :(\neg AB(B))/\neg AB(B) \} .$$

Esta teoria possui duas extensões :

$E"_1 = Th(W" \cup \{\neg AB(B)\})$, que é a extensão intuitiva pois nela podemos concluir que Tweety voa , já que :

i) de (4"), (6") e $\neg AB(B)$, deduzimos S ;

ii) de (5") e S, deduzimos F.

Além disso, de (2") e S, obtemos AB(A) que inibe a aplicação do default :(¬AB(A))/¬AB(A) ; e

E"$_2$ = Th(W" ∪ {¬AB(A)}), que é a extensão não-intuitiva, pois nela podemos concluir que Tweety não voa, já que :

i) de (3") e (6"), deduzimos : A ;

ii) de (1"), A e ¬AB(A), deduzimos : ¬F.

Além disso, a aplicação de :(¬AB(B))/¬AB(B) é inibida pois :

i) de ¬F e (5"), deduzimos ¬S ;

ii) de ¬S e (4"), deduzimos ¬B ∨ AB(B) ;

iii) de B, ¬B ∨ AB(B), deduzimos AB(B).

Novamente, surge o problema de obtermos várias extensões, neste caso para um exemplo não-temporal ■

A seguir , vamos examinar uma série de propostas apresentadas para tentar solucionar o YSP.

*ii) <u>TMS e Defaults Não-Normais</u>*

No mesmo artigo onde aponta que a extensão anômala que aparece no YSP não se deve ao fato deste ser um problema que envolve questões temporais (exemplo 6.1), **MORRIS(1988)** mostrou que, reformulando ambos os problema (YSP e exemplo 6.1) em termos de um TMS, a extensão anômala desaparece.

A notação utilizada para as justificativas em **MORRIS(1988)** (e que será utilizada aqui para maior clareza na discussão dos resultados) é diferente da utilizada no capítulo 3. Considera-se que uma *justificativa* é da forma:

$$A \wedge out(B) \succ C \qquad (*)$$

onde A é chamado de *justificador* in (correspondendo a *lista* in da justificativa SL), out(B) é um *justificador* out (correspondendo a *lista* out da justificativa SL) e representa a ignorância a cerca do fato B. A conclusão da justificativa (*) é representada por C (corresponde ao nó que está sendo justificado) .

Assim, reformulando o exemplo 6.1 em termos do TMS, temos que:

(1") $A \wedge out(AB(A)) \succ \neg F$ *(suposição TMS)*

(2") $S \succ AB(A)$

(3") $B \succ A$

(4") $B \wedge out(AB(B)) \succ S$ *(suposição TMS)*

(5") $S \succ F$

(6") B *(premissa TMS)*

Lembre que uma justificativa SL é válida (o nó que ela justifica está no estado IN) se, e somente se, os nós de sua *lista* in (no nosso caso, *justificador* in) estão no estado IN e os da sua *lista* out (no nosso caso, *justificadores* out) estão OUT. Assim, para a teoria composta por (1")-(6"), teríamos dois *rótulos* (estados IN e OUT de cada uma das proposições) correspondentes as duas extensões vistas no exemplo 6.1 :

| | B | A | S | F | ¬F | AB(A) | AB(B) |
|---|---|---|---|---|---|---|---|
| $E_1$ | in | in | in | in | out | in | out |
| $E_2$ | in | in | out | out | in | out | in |

O mecanismo utilizado pelo TMS para rotular as proposições tem a propriedade de que somente rótulos bem-fundamentados são obtidos, ou seja, onde toda proposição rotulada IN é bem-justificada (é a conclusão de alguma justificativa cujos *justificadores*-out são todos rotulados OUT e cujos *justificadores*-in são eles próprios bem-justificados).

No caso do exemplo, como AB(B) não é conclusão de nenhuma justificativa, não existe rótulo bem-fundamentado no qual AB(B) esteja no estado IN. Logo, para o TMS, existe uma única extensão acessível no nosso exemplo (rótulo $E_1$) , que é justamente a que corresponde a nossa intuição, pois nela podemos concluir que Tweety voa (F está no estado IN).

No caso do YSP, se nos restringirmos a (1')-(5'), a representação natural em termos do TMS seria :

(1') $T(carregada, s_1)$

(2') $T(carregada, s_1) \wedge out(AB(carregada, esperar, s_1)) \succ$

$T(carregada, s_2)$

(3') $T(vivo, s_2)$

(4') $T(vivo, s_2) \wedge out(AB(vivo, disparar, s_2)) \succ T(vivo, s_3)$

(5'.1) $T(carregada, s_2) \succ AB(vivo, disparar, s_2)$

(5'.2) $T(carregada, s_2) \succ T(morto, s_3)$

onde as proposições têm os seguintes rótulos :

| | | $\overline{E}_1$ | $\overline{E}_2$ |
|---|---|---|---|
| *1* | T(carregada, $s_1$) | in | in |
| *2* | T(vivo, $s_2$) | in | in |
| *3* | T(carregada, $s_2$) | in | out |
| *4* | T(vivo, $s_3$) | out | in |
| *5* | T(morto, $s_3$) | in | out |
| *6* | AB(carregada, esperar, $s_1$) | out | in |
| *7* | AB(vivo, disparar, $s_2$) | in | out |

onde novamente um deles é eliminado (no caso, $\overline{E}_2$) pois AB(carregada,esperar,$s_1$) não é a conclusão de qualquer justificativa .

Observe que se mais tarde aprendermos que a arma não está carregada em $s_2$ (¬T(carregada,$s_2$) fica no estado IN), uma contradição é produzida, fazendo com que o TMS utilize o DDB (*algorítmo 2* - capítulo 3 e apêndice A). Escrevendo as justificativas dos nós (*1*)-(*7*) na forma utilizada no capítulo 3, temos :

| | | |
|---|---|---|
| *1* | ( ) ( ) | in |
| *2* | ( ) ( ) | in |
| *3* | (1) (6) | in |
| *4* | (2) (7) | out |
| *5* | (3) ( ) | in |
| *6* | | out |
| *7* | (3) ( ) | in |

Aprendendo ¬T(carregada,$s_2$), que será considerado o nó (8) e cuja justificativa será (SL ( ) ( )) , teremos uma contradição representada pelo nó (9) e justificada por :

(SL (3 8) ( ))

que passa a estar no estado IN.

Neste instante, o DDB é acionado e inicialmente, o conjunto de suposições máximas é determinado. No nosso caso tal conjunto é formado por um único nó:

$S = \{3\}$ .

A seguir, é criado o nó (nogood) (10), cuja justificativa é : (CP 9 (3) ( )).

Como o conjunto S só possui o nó (3), o TMS escolhe tal nó e justifica o seu único nó **OUT** da *lista* out (nó (6) - AB(carregada,esperar,$s_1$)), com (SL (10) ( )). Logo, os nós (1),(2),(4),(6) e (10) ficam no estado IN e os nós (3),(5),(7) e (9) ficam no estado **OUT** .

Assim, pelo procedimento DDB, o único rótulo a que o TMS teria acesso agora seria $\bar{E}_2$. O mesmo ocorre quando, no exemplo não-temporal, aprendemos que Tweety não tem asas ($\neg S$ passando para o estado IN) , fazendo com que o rótulo $E_2$ seja o único acessível.

Esse bom comportamento ( no que se refere ao aprendizado de novos fatos ) apresentado pelo TMS, em ambos os casos se deve, segundo **MORRIS(1988)**, a natureza unidirecional da inferência por ele efetuada (fato assegurado pela exigência de seus rótulos serem bem-fundamentados). Observando os exemplos de forma mais restrita :

(a) $\neg$ AB(carregada,esperar,$s_1$) $\supset$ AB(vivo,disparar,$s_2$)

(b) $\neg$ AB(carregada,esperar,$s_1$) $\supset$ T(morto,$s_3$)

(c) $\neg$ AB(vivo,disparar,$s_2$) $\supset$ T(vivo,$s_3$)

(a') $\neg$ AB(B) $\supset$ AB(A)

(b') $\neg$ AB(A) $\supset \neg$ F

(c') $\neg$ AB(B) $\supset$ F ,

note que temos dois defaults $D_1$ e $D_2$ tais que $D_1 \supset \neg D_2$ ((a) e (a')), que apesar de ser logicamente equivalente a $D_2 \supset \neg D_1$, não tem este fato explícito na extensão intuitiva, que inclui $D_1$ mas não $D_2$.

MORRIS(1988) procurou, dentro da lógica de default, encontrar uma maneira de incorporar essa característica do TMS. Para isso, ele observou que as justificativas da forma:

$$A \wedge \text{out}(B) \succ C$$

são satisfeitas quando B não se encontra no estado IN. Quando isto ocorre, significa que B não é derivável, o que equivale a dizer que $\neg$B é consistente com os outros fatos conhecidos. Isto sugere que as justificativas TMS sejam vistas como regras defaults da forma :

$$A : (\neg B) \;/\; C \;,$$

ou seja, como defaults não-normais.

Se considerarmos o exemplo não-temporal, utilizando a representação por defaults não-normais, teríamos $\Delta_\alpha$ =

$(D_a, W_a)$ onde $W_a$ é formada pelos axiomas :

(i) $S \supset AB(A)$

(ii) $B \supset A$

(iii) $S \supset F$

(iv) $B$

e $D_a$ é composta pelos defaults :

$$D_{a1} \Leftrightarrow A : (\neg AB(A)) \;/\; \neg F$$

$$D_{a2} \Leftrightarrow B : (\neg AB(B)) \;/\; S$$

que substituem , respectivamente, as justificativas :

$$A \wedge out(AB(A)) \succ \neg F$$

$$B \wedge out(AB(B)) \succ S \quad .$$

Com esta representação , a teoria default $\Delta_a$ possui uma única extensão, que é :

$E_1 = Th(W_a \cup \{S\})$, pois a partir de $D_{a2}$ e (i), obtemos :

$$AB(A) \quad ,$$

fazendo com que o default $D_{a1}$ não possa ser aplicado. Além disso, tal extensão está de acordo com o que esperávamos, já que de $D_{a2}$ e (iii) , deduzimos F (Tweety voa).

A outra possível extensão seria obtida se aplicássemos inicialmente o default $D_{a1}$ à teoria $W_a$, onde :

de $D_{a1}$ e (iii), deduzimos $\neg S$ .(*)

Agora, note que o default $D_{a2}$ poderia ser aplicado, o que colocaria em nossa extensão a proposição S, que junto com (*) , a tornaria inconsistente. Nesse caso, pela presença de uma inconsistência, tal extensão se auto-elimina, pois o teorema 2.10 nos garante que uma teoria default possui uma extensão inconsistente se, e somente se, o conjunto $W_a$ for inconsistente, o que não é verdade. Logo, uma única extensão ($E_1$) seria obtida.

O mesmo ocorre no YSP, onde teríamos a teoria default $\Delta_b$ = ($D_b$,$W_b$), sendo $W_b$ composta por :

(i') T(carregada,$s_1$)

(ii') T(vivo,$s_2$)

(iii') T(carregada,$s_2$) $\supset$ AB(vivo,disparar,$s_2$)

(iv') T(carregada,$s_2$) $\supset$ T(morto,$s_3$)

e $D_b$ tem os seguintes defaults :

$D_{b1}$ $\Leftrightarrow$ T(carregada,$s_1$) :

(¬AB(carregada,esperar,$s_1$)) / T(carregada,$s_2$) ,

e

$D_{b2}$ $\Leftrightarrow$ T(vivo,$s_2$) :(¬AB(vivo,disparar,$s_2$)) / T(vivo,$s_3$) .

A extensão intuitiva é obtida sem dificuldade, aplicando-se o default $D_{b1}$ a teoria $W_b$. Este default, junto com os axiomas (iii') e (iv') nos fornece :

T(morto,$s_3$) e AB(vivo,disparar,$s_2$) ,

que inibe a aplicação do outro default.

Caso o default $D_{b2}$ fosse aplicado primeiro, teríamos :

$$T(vivo,s_3) \text{ e } \neg T(carregada,s_2) . \qquad (\#)$$

Como nada impede que o primeiro default seja aplicado, teríamos :

$$T(carregada,s_2) \quad ,$$

que com (#) , gera uma inconsistência . Logo essa extensão se auto-elimina.

O problema na utilização dos defaults não-normais na representação deste tipo de problema é que, ao contrário do TMS, ele não responde adequadamente ao acréscimo de novas informações. Por exemplo, se no conjunto $W_a$ acima fosse colocado mais o seguinte axioma :

$$\neg S$$

quando o default $D_{a2}$ fosse aplicado, surgiria uma inconsistência , fazendo com que a extensão $E_1$ se auto-elimine, deixando a teoria default $\Delta_a' = (D_a, W_a \cup \{\neg S\})$ sem extensões (o mesmo ocorre com o YSP).

A utilização de defaults semi-normais como sugerido por REITER E CRISCUOLO(1981) traz de volta a extensão anômala. Veja que para os defaults :

$$D_1 \Leftrightarrow A : (\neg F \wedge \neg AB(A)) \ / \ \neg F$$

$$D_2 \Leftrightarrow B : (S \wedge \neg AB(B)) \ / \ S$$

ao aplicarmos $D_1$ à teoria $W_a$, junto com $S \supset F$, deduziríamos $\neg S$ , o que inibe a aplicação de $D_2$, fazendo ressurgir a extensão não-intuitiva.

Por outro lado, YOU (1989) argumentou que os problemas que surgem no YSP e no exemplo 6.1 aparecem pois estamos trabalhando com o senso comum, onde a maneira como os fatos são representados, reflete intuitivamente a forma pela qual o problema deve ser solucionado. Assim, o aparecimento de extensões anômalas seria causado pela ausência , nas lógicas não-monotônicas , de algum mecanismo que capture tal intuição.

No caso da circunscrição suportada (capítulo 4 - seção v) , que incorpora à circunscrição a noção de suporte, temos que o único meio de se obter um fato positivo é demonstrá-lo segundo a orientação das regras (o que elimina as extensões anômalas). Escrevendo (a), (b), (c), (a'), (b') e (c') como cláusulas orientadas, temos que :

(a) $\neg$ AB(carregada,esperar,$s_1$) $\rightarrow$ AB(vivo,disparar,$s_2$)

(b) $\neg$ AB(carregada,esperar,$s_1$) $\rightarrow$ T(morto,$s_3$)

(c) $\neg$ AB(vivo,disparar,$s_2$) $\rightarrow$ T(vivo,$s_3$)

(a') $\neg$ AB(B) $\rightarrow$ AB(A)

(b') $\neg$ AB(A) $\rightarrow$ $\neg$ F

(c') $\neg$ AB(B) $\rightarrow$ F

Observe que para a teoria $\tau$ composta por (a),(b) e (c), a circunscrição CIRC($\tau$;AB) possui dois modelos :

i) $M_1$ = { AB(vivo,disparar,$s_2$),T(morto,$s_3$) }, que corresponde ao modelo intuitivo ; e

ii) $M_2$ = { AB(carregada,esperar,$s_1$) }, correspondente ao modelo não-intuitivo.

Por definição, o modelo $M_1$ é suportado pelos fatos conhecidos pois existe uma cláusula orientada ((a)), na qual AB(vivo,disparar,$s_2$) é um dos literais que compõem a cabeça dessa cláusula. Já o modelo $M_2$ não é suportado pelos fatos . Logo, somente $M_1$ é um modelo da circunscrição suportada, que gera portanto apenas o modelo intuitivo.

Da mesma maneira, no caso da teoria $\tau'$ composta por (a'), (b') e (c'), apenas o modelo intuitivo $M_1$ = {AB(A),F} é obtido . Isto porque AB(B) não pode ser assumido já que nenhuma cláusula orientada a suporta. Ao contrário, podemos obter AB(A) assumindo ¬ AB(B) sem introduzir qualquer contradição. Assim, poderíamos considerar que a característica unidirecional da inferência no TMS foi transportada para a circunscrição através da noção de suporte.

O problema com a circunscrição suportada também acontece quando novos axiomas são acrescidos ao conjunto inicial. Considere o YSP representado na forma de cláusulas orientadas :

(1') T(carregada,$s_1$)

(2') T(carregada,$s_1$) ∧ ¬AB(carregada,esperar,$s_1$) →
T(carregada,$s_2$)

(3') T(vivo,$s_2$)

(4') T(vivo,$s_2$) ∧ ¬AB(vivo,disparar,$s_2$) → T(vivo,$s_3$)

(5'.1) T(carregada,$s_2$) → AB(vivo,disparar,$s_2$)

(5'.2) T(carregada,$s_2$) → T(morto,$s_3$)

Observe que se acrescentássemos o axioma :

$$(6') \ \neg T(carregada, s_2)$$

teríamos , pela circunscrição de AB na teoria (1')-(6') com T variando , um modelo que tem AB(carregada,esperar,$s_1$) mas não AB(vivo,disparar,$s_2$) , ou seja , teríamos o modelo $M_2$ visto antes, que agora deixa de ser um modelo não-intuitivo para ser intuitivo.

Entretanto, $M_2$ não é suportado pelos fatos da teoria, apesar de, nesse momento, ele ser um modelo intuitivamente correto.

Lembre que o procedimento DDB do TMS, quando o mesmo caso ocorre, é o de *criar* uma justificativa para AB(carregada,esperar,$s_1$) o que, do ponto de vista da circunscrição suportada, tornaria tal anormalidade suportada pelos fatos. Isso quer dizer que na teoria seria incluido a seguinte cláusula orientada :

$$NG \rightarrow AB(carregada, esperar, s_1)$$

onde NG é o nó (10) do TMS, que daria suporte ao modelo mínimo.

Assim, mesmo que a noção de suporte incorpore de certa forma a característica unidirecional de inferência do TMS, falta a este tipo de circunscrição alguma maneira de tratar as mudanças nos fatos suportados quando do acréscimo de novos fatos à teoria inicial.

*iii) Minimização Cronológica*

Quando aplicamos a circunscrição ao YSP, obtivemos dois modelos mínimos para a teoria, sendo um intuitivo e outro não. Além disso, a característica apresentada na construção de ambos os modelos era que uma normalidade após outra era assumida, e como consequência, uma anormalidade podia ser deduzida. A diferença entre um e outro modelo estava no sentido do tempo em que se trabalhava : no primeiro, para frente no tempo; no segundo , para trás.

Isso sugeria que o critério de seleção dos modelos, que na circunscrição é o "dos modelos mínimos", seria incorreto, e nos problemas temporais tais como o YSP, deveria ser o "dos modelos *cronologicamente* mínimos", onde as hipóteses de normalidade são feitas em ordem cronológica. Isto equivale a ter as anormalidades ocorrendo o quanto mais tarde possível. Com este novo critério, obteríamos um único modelo , justamente o que queremos.

O conceito de minimização cronológica surgiu em SHOHAM(1988), que o generalizou, construindo uma lógica modal não-monotônica. Antes de definirmos tal lógica vamos ver resumidamente o que é uma lógica modal (proposicional). Uma abordagem mais ampla sobre lógica modal pode ser encontrada em HUGHES e CRESSWELL(1968).

A lógica modal foi criada para que fosse possível distinguir num conjunto de proposições as que são obrigatoriamente verdadeiras (ou falsas) e aquelas que são ocasionalmente verdadeiras (ou falsas).

Tais noções são expressas através de dois operadores ditos *modais*, representados por $\Box$ (necessidade) e $\Diamond$

(possibilidade), que são acrescidos a lógica proposicional. Assim, sendo p uma fórmula bem formada da lógica proposicional :

$\Box p$ significa "p é necessário", ou seja, p é obrigatoriamente verdadeiro ; e

$\Diamond p$ significa "p é possível", ou seja, p não é obrigatoriamente falso .

Estes operadores são tais que $\Diamond\varphi \equiv \neg\Box\neg\varphi$ é uma fórmula válida.

Uma interpretação na lógica proposicional é uma atribuição de valores {true,false} a cada um dos símbolos proposicionais. Seu correspondente na lógica modal é chamado de *estrutura de Kripke* . Formalmente, uma estrutura de Kripke ( KRIPKE(1963) ) é uma tripla (W,M,R) onde :

- W é um conjunto não-vazio de mundos possíveis (sendo cada um deles uma interpretação clássica) ;

- M é uma função de significado $M : P \rightarrow 2^W$ , de modo que M(P) determina os mundos nos quais p é verdadeiro ; e

- R é uma relação de acessibilidade entre os mundos.

As interpretações em mundos distintos podem ser diferentes, com uma fórmula sendo verdadeira num mundo mas não em outro. Por exemplo, numa estrutura de Kripke poderíamos ter três mundos $W_1$, $W_2$, e $W_3$ tais que :

$$W_1 = \{p,q,r\} \quad ; \quad W_2 = \{p,\neg q,r\} \quad ; \quad W_3 = \{\neg p,\neg q,r\}$$

onde p, q, e r são proposições. Nesse caso, temos $M(p) = \{W_1,W_2\}$ , $M(q) = \{W_1\}$ e $M(r) = \{W_1,W_2,W_3\}$.

As estruturas de Kripke podem ser vistas como um grafo direcionado onde quaisquer dois mundos podem ser ligados em uma direção, em ambas ou em nenhuma. Esta relação entre os mundos é chamada de *relação de acessibilidade*. Ao impormos restrições sobre a relação de acessibilidade, obtemos diferentes sistemas modais. Assim, se a relação de acessibilidade é :

i) *reflexiva*, a estrutura de Kripke é dita estrutura T;

ii) *reflexiva e transitiva*, a estrutura de Kripke é dita estrutura S4;

iii) *reflexiva, transitiva e simétrica*, a estrutura de Kripke é uma estrutura S5. Neste caso, os mundos possíveis formam uma classe de equivalência e, portanto, o conjunto de todos os mundos pode ser igualado ao conjunto dos mundos acessíveis.

Quando numa estrutura de Kripke não é mencionada a relação de acessibilidade, considera-se esta estrutura como sendo S5. Este é o caso da lógica definida por SHOHAM (1988).

O fato de uma proposição p ser verdadeira numa estrutura de Kripke $K = (W,M,R)$ num mundo w é denotada por $K,w \models p$. Consequentemente :

$K,w \models p_1 \wedge p_2$ sse $K,w \models p_1$ e $K,w \models p_2$ ;

$K,w \models \neg p$ sse $K,w \not\models p$ ;

$K,w \models \Box p$ sse $K,w' \models p$ , para todo $w' \in W$ tal que $w\,R\,w'$, ou seja, para todo mundo $w'$ aos quais $w$ tem acesso .

Uma fórmula $\varphi$ é válida (resp., satisfatível) se $K,w \models \varphi$ para todos os mundos $w \in W$ (resp., em algum mundo $w \in W$).

SHOHAM(1988) utiliza uma lógica de conhecimento que nada mais é que uma lógica modal onde o símbolo $\Box$ é interpretado como "é conhecido" . Assim, $\Box p$ representará que "p é conhecido". Uma definição mais ampla de lógica do conhecimento pode ser encontrada em HALPERN e MOSES (1985).

A lógica da Ignorância Cronológica (CI - Chronological Ignorance), é uma lógica modal não-monotônica, obtida associando-se à uma lógica de conhecimento temporal TK, uma ordem sobre as suas estruturas de Kripke .

O conjunto de fórmulas bem formadas (f.b.f.) de TK é definido indutivamente da seguinte maneira : dados

P : o conjunto de proposições primitivas ;

TC : o conjunto de símbolos dos pontos de tempo $\{\ldots,-1,0,1,\ldots\}$ ;

TV : um conjunto de variáveis temporais ;

$U = TC \cup TV$ e $\angle$ um símbolo de relação binária ;

temos que :

(1) se $u_1$ , $u_2 \in U$ então $u_1=u_2$ e $u_1 \angle u_2$ são f.b.f. ;

(2) se $u_1$ , $u_2 \in U$ e $p \in P$ então $TC(u_1,u_2,p)$ é uma f.b.f. ;

(3) se $\varphi_1$ e $\varphi_2$ são f.b.f. então $\varphi_1 \wedge \varphi_2$, $\neg\varphi_1$ e $\Box\varphi_1$ também são ;

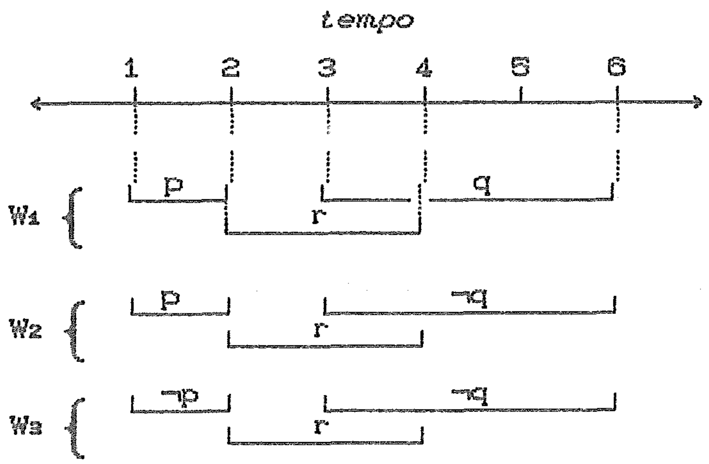
(4) se $\varphi$ é uma f.b.f. e $v \in TV$, então $\forall v\ \varphi$ também é .

Valem as definições usuais de $\vee$, $\supset$, $\equiv$, $\exists$ e $\hat{\nabla}$.

Usaremos as seguintes abreviações :

$$\Box T(t_1,t_2,p) \text{ por } \Box(t_1,t_2,p)\ ;$$
$$\Box\neg T(t_1,t_2,p) \text{ por } \Box(t_1,t_2,\neg p)\ ;$$
$$\hat{\nabla} T(t_1,t_2,p) \text{ por } \hat{\nabla}(t_1,t_2,p)\ ;$$
$$\hat{\nabla}\neg T(t_1,t_2,p) \text{ por } \hat{\nabla}(t_1,t_2,\neg p)\ ;$$
$$T(t,t,p) \text{ por } T(t,p)\ .$$

Nas estruturas de Kripke K = (W,M,R) em TK, uma mesma interpretação de tempo é partilhada pelos mundos pertencentes à W, fazendo com que cada um deles descreva um possível curso do universo. A estrutura de tempo será isomorfa aos números inteiros. Por exemplo, considere uma estrutura formada pelos mundos $W_1$, $W_2$ e $W_3$, representada pela figura abaixo :




Nela, temos T(2,4,r) nos três mundos, e portanto, esta estrutura satisfaz $\Box(2,4,r)$. As proposições p e q não são conhecidas pois em alguns mundos (resp. $W_3$ e $W_2$,$W_3$)

T(1,2,p) e T(3,6,q) são falsos.

Assim, uma interpretação de Kripke é uma tripla (W,M,R) onde :

- W é um universo não-vazio de possíveis mundos ;
- M é uma função de significado M : P → $2^{W \times N \times N}$ (N denota os inteiros com ≤) , ou seja , M(p) determina em cada mundo w os intervalos de tempo nos quais p é verdadeiro ( vamos considerar apenas os intervalos de tempo $(t_1,t_2)$ tais que $t_1 \leq t_2$ ) ; e
- R é uma relação de acessibilidade reflexiva, simétrica e transitiva (estrutura S5) e que por essa razão será omitida daqui para frente.

No caso da estrutura representada pela figura acima temos :

$W = \{ W_1, W_2, W_3 \}$

$M(p) = \{ (W_1,1,2) , (W_2,1,2) \}$

$M(q) = \{ (W_1,3,6) \}$

$M(r) = \{ (W_1,2,4) , (W_2,2,4) , (W_3,2,4)\}$

Definindo uma função de <u>atribuição de variáveis temporais</u> VA : TV → N e a função VAL(u) que para u ∈ U é definida como sendo VA(u) , se u ∈ TV ; e a interpretação padrão de u se u ∈ TC, temos que uma interpretação KI = (W,M) e um mundo w ∈ W <u>satisfazem</u> $\varphi$ sob a atribuição de variáveis VA ( ou seja, KI,w $\models \varphi$ [VA] ) sob as seguintes condições :

$KI, w \models u_1 = u_2$ [VA] sse $VAL(u_1) = VAL(u_2)$ ;

$KI, w \models u_1 \angle u_2$ [VA] sse $VAL(u_1) \leq VAL(u_2)$ ;

$KI, w \models T(u_1, u_2, p)$ [VA] sse $(w, VAL(u_1), VAL(u_2)) \in M(p)$ ,

onde $VAL(u_1) \leq VAL(u_2)$;

$KI, w \models (\varphi_1 \wedge \varphi_2)$ [VA] sse $KI, w \models \varphi_1$ [VA] e

$KI, w \models \varphi_2$ [VA] ;

$KI, w \models \neg\varphi$ [VA] sse $KI, w \not\models \varphi$ [VA] ;

$KI, w \models \Box\varphi$ [VA] sse $KI, w' \models \varphi$ [VA] , $\forall w' \in W$ ;

$KI, w \models \forall v\ \varphi$ [VA] sse $KI, w \not\models \varphi$ [VA'] , $\forall$ VA' que concorde com VA em toda parte exceto possivelmente em v .

Se , para alguma atribuição de variáveis VA, $KI, w \models \varphi$ [VA] então a interpretação de Kripke $KI = (W, M)$ e um mundo $w \in W$ são um *modelo* para uma fórmula $\varphi$ (ou seja, $KI, w \models \varphi$). Uma f.b.f. $\varphi$ é satisfatível se $\varphi$ tem um modelo e é válida se $\neg\varphi$ não possui nenhum modelo (ou seja, $\neg\varphi$ é insatisfatível) .

A lógica de conhecimento temporal TK é monotônica. A não-monotonicidade é obtida definindo-se uma ordem parcial sobre as estruturas de Kripke de TK. Para isto, é necessário que os seguintes conceitos sejam estabelecidos.

**DEFINIÇÃO 6.2**

As f.b.fs. *básicas* são aquelas que não contém ocorrência de operadores modais.

**EXEMPLO 6.3**

As fórmulas :

(a) $T(3,r) \supset \forall v\ T(4,v,\neg r)$

(b) $T(1,\neg p)$

são exemplos de fórmulas básicas . Já as fórmulas :

(c) $\Box(4,p)$

(d) $\hat{\vee}\ [T(3,r) \vee T(7,r)]$

não são fórmulas básicas ■

**DEFINIÇÃO 6.4**

O *último ponto de tempo* (*u.p.t.*) de uma fórmula básica é o último (cronologicamente) ponto de tempo que a fórmula menciona. Formalmente :

(1) O u.p.t. de $T(t_1,t_2,p)$ é o maior entre $t_1$ e $t_2$;

(2) O u.p.t. de $\varphi_1 \wedge \varphi_2$ é o maior entre os u.p.t. de $\varphi_1$ e $\varphi_2$;

(3) O u.p.t. de $\neg\varphi$ é o u.p.t. de $\varphi$;

(4) O u.p.t. de $\forall v\ \varphi$ é o menor entre todos os u.p.t. de todas as sentenças resultantes da substituição da variável quantificada v por uma constante. Caso não exista este menor u.p.t. , o u.p.t. é $-\infty$ .

**EXEMPLO 6.5**

As fórmulas (a) e (b) do exemplo 6.3 têm como seus u.p.t., respectivamente , 3 e 1. No caso de (a), o u.p.t. de $\forall v\ T(4,v,\neg r)$ é igual a 4 ■

A definição seguinte estabelece a relação de ordem parcial sóbre as estruturas de Kripke de TK :

DEFINIÇÃO 6.6

Uma interpretação de Kripke $M_2$ é <u>cronologicamente mais ignorante</u> que a interpretação $M_1$ ( $M_1 \subset_{ci} M_2$ ) se existe um tempo to tal que :

(1) para toda sentença básica $\varphi$ cujo u.p.t. $\leq$ to, se $M_2 \models \square\varphi$, então $M_1 \models \square\varphi$ ; e

(2) existe alguma sentença básica $\varphi$ cujo u.p.t. é to tal que $M_1 \models \square\varphi$, mas $M_2 \not\models \square\varphi$ .

DEFINIÇÃO 6.7

Um modelo M que satisfaz uma sentença $\varphi$ é chamado <u>ignorante cronológico maximal</u> (c.m.i. - chronologically maximally ignorant) se $M \models \varphi$ e não existe outro modelo M' tal que $M' \models \varphi$ e $M \subset_{CI} M'$ . Isto será representado por $M \models_{Cci} \varphi$ .

Assim, associando a lógica de conhecimento temporal TK com a relação $\subset_{CI}$ obtem-se a lógica da ignorância cronológica CI.

Dentro da lógica da CI, Shoham sugeriu que o YSP poderia ser representado pelo seguinte conjunto de axiomas:

(1) $\square$(1,carregada)

(2) $\square$(3,disparar)

(3) $\forall$t [$\square$(t,carregada) $\wedge$ $\hat{\diamond}$(t,$\neg$disparar) $\wedge$ $\hat{\diamond}$(t,$\neg$descarregar) $\supset$ $\square$(t+1,carregada)]

(4) $\forall$t [$\square$(t,carregada) $\wedge$ $\square$(t,disparar) $\wedge$ $\hat{\diamond}$(t,$\neg$defeito) $\supset$ $\square$(t+1,morto)]

onde (1) representa que sabemos que no tempo 1 , a arma está carregada ; (2) representa que sabemos que no tempo 3, a arma é disparada; (3) representa o axioma "frame" de que se sabemos que a arma está carregada em t e neste instante ignoramos que ela tenha sido disparada ou descarregada, então saberemos que ela permanecerá carregada em t+1; (4) representa que se sabemos que a arma está carregada e sabemos que ela foi disparada em t e ignoramos que a arma esteja defeituosa em t ,então saberemos que Fred morrerá em t+1 .

Observe que existem dois tipos de declarações na teoria acima : no primeiro (ax. 1 e 2), onde as sentenças são da forma "sabemos que o fato ... é veradeiro no tempo ...", representando o conhecimento a respeito de uma sentença atômica ; e o segundo , onde as declarações são da forma "sabemos que os fatos ... são verdadeiros no tempo ... , e não sabemos que outros fatos tais como ... são verdadeiros no instante de tempo ... , então também sabemos que os fatos ... são verdadeiros em um instante de tempo futuro ...", onde o conhecimento e a ignorância a cerca de sentenças atômicas passadas implicam no conhecimento de sentenças atômicas futuras. As sentenças do primeiro tipo são um caso especial do segundo, onde o antecedente da implicação é identicamente verdadeiro.

Assim, as sentenças podem ser escritas na forma :

$$\phi \wedge \theta \supset \Box\varphi \qquad (*)$$

onde $\varphi$ é uma sentença atômica básica (s.a.b.), positiva ou negativa, tendo $t_1$ como o seu u.p.t. ; $\phi$ é uma conjunção de sentenças $\Box\varphi_i$, onde $\varphi_i$ é uma s.a.b. (positiva ou negativa) cujo seu u.p.t. $t_i$ é tal que $t_i < t_1$ ; e $\theta$ é uma conjunção de sentenças $\hat{\nabla}\varphi_j$ , onde $\varphi_j$ é uma s.a.b. (positiva ou negativa) cujo seu u.p.t. $t_j$ é tal que $t_j < t_1$. Quando $\phi$ é vazio, a sentença é chamada de *condição de limite*; nos outros casos, elas são chamadas de *regras causais*.

Teorias compostas por sentenças da forma (*) são chamadas de teorias causais. Este tipo de teoria tem as seguintes propriedades :

(a) existe um ponto de tempo $t_0$ tal que se $\theta \supset \Box\varphi$ é uma condição de limite na teoria $\psi$ e u.p.t. de $\varphi$ é $t_1$, então $t_0 < t_1$ ;

(b) não existem duas sentenças na teoria $\psi$ tais que para qualquer $p, t_1, t_2$ , uma contém $\hat{\nabla}(t_1, t_2, p)$ e a outra $\hat{\nabla}(t_1, t_2, \neg p)$ ;

(c) se $\phi_1 \wedge \theta_1 \supset \Box(t_1, t_2, p)$ e $\phi_2 \wedge \theta_2 \supset \Box(t_1, t_2, \neg p)$ estão na teoria $\psi$, então $\phi_1 \wedge \theta_1 \wedge \phi_2 \wedge \theta_2$ é inconsistente ;

Este tipo de teoria possui uma característica importante : ela pode ter vários modelos c.m.i., mas em todos eles as mesmas sentenças básicas são conhecidas. No caso do YSP, isto significa que as seguintes sentenças básicas são conhecidas : T(1,carregada), T(2,carregada), T(3,carregada), T(3,disparar) e T(4,morto) (além de suas consequências tautológicas) .

Formalmente, este resultado é expresso pelo :

**TEOREMA 6.7**

Para qualquer teoria causal $\psi$, temos :

(1) $\psi$ possui um modelo c.m.i. ;

(2) se $M_1$ e $M_2$ são ambos modelos c.m.i. de $\psi$ , e $\varphi$ é qualquer sentença básica, então $M_1 \models \Box\varphi$ se, e somente se, $M_2 \models \Box\varphi$ .

A construção do modelo c.m.i. relativo ao YSP (e, pelo teorema 6.7, para teorias causais em geral) é feita indutivamente (no tempo) .

O ponto de partida $t_0$ deve ser tal que preceda o u.p.t. de todas as condições de limite da teoria envolvida. Note que a propriedade (a) das teorias causais garante a existência deste ponto. Considerando a condição de limite $\Box$(1,carregada) (axioma (1)), temos que $t_0 = 0$. Observe que este ponto precede o u.p.t. da condição de limite $\Box$(3,disparar).

A partir daí, define-se uma <u>estrutura de Kripke limitada pelo tempo $M/t_0$</u>. Este tipo de estrutura de Kripke é de tal maneira que são atribuídos valores verdade às proposições atômicas e sentenças em geral que tenham os seus u.p.t. menores ou iguais a $t_0$. Assim, para uma sentença $\varphi$ tal que o seu u.p.t. $\leq t_0$, escrevemos $M/t_0 \models \varphi$ (resp. $M/t \models \varphi$ , para um $t$ qualquer) quando $\varphi$ é satisfeita por $M/t_0$ (resp. por $M/t$). Além disso, dizemos que uma estrutura $M/t$ <u>satisfaz</u> <u>parcialmente</u> uma teoria $\Psi$ se $M/t$ satisfaz todos os membros de $\Psi$ cujo u.p.t. $\leq t$.

Observe que no caso do YSP, $M/0$ satisfaz $\Box\varphi$ ( $M/0 \models \Box\varphi$) , para qualquer tautologia básica $\varphi$ cujo u.p.t. $\leq 0$.

Para qualquer outra f.b.f. básica $\varphi'$ cujo u.p.t. $\leq$ to, teremos M/O $\not\models$ $\Box\varphi'$. Note que M/O satisfaz parcialmente os axiomas (1) e (2), uma vez que os seus u.p.t. são maiores que O, e os axiomas (3) e (4), pois quando os seus u.p.t. são menores ou iguais a O, os seus antecedentes são falsos (não são tautologias). Logo, M/O satisfaz parcialmente $\Psi$.

Para qualquer t, obter M/t+1 a partir de M/t, representa aumentar o nosso conhecimento com os fatos obtidos a partir do que é conhecido e ignorado até o instante t, e que está expresso em M/t. Assim, devemos definir o conjunto $CONSEQ_{t+1}$ da seguinte maneira :

$CONSEQ_{t+1}$ = $\{\Box(t',t+1,x) : \phi \wedge \theta \supset \Box(t',t+1,x) \in \Psi$ e $M/t \models \phi \wedge \theta \}$, ou seja, é o conjunto de conhecimentos que são consequências daquilo que era conhecido e ignorado em tempos passados (instantes de tempo menores ou iguais a t).

Para obter M/t+1, fazemos todas as f.b.f. do conjunto $CONSEQ_{t+1}$ (além de suas consequências tautológicas) verdadeiras, e qualquer outra f.b.f. básica $\varphi'$, cujo u.p.t. = t+1, falsa. No caso do YSP, temos que :

$$CONSEQ_1 = \{ \Box(1,carregada) \},$$

pois o axioma (1) pertence a teoria $\Psi$ e M/O $\models$ true. Assim teremos, M/1 $\models$ $\Box(1,carregada)$, e para qualquer outra f.b.f. básica $\varphi'$ que tenha u.p.t. $\leq$ 1 (e que não seja consequência tautológica), M/1 $\not\models$ $\Box\varphi'$.

Prosseguindo no tempo, temos que :

$$CONSEQ_2 = \{\Box(2,carregada)\} ,$$

já que o axioma (3) pertence a teoria e $M/1 \models \Box(1,carregada) \wedge \hat{\nabla}(1,\neg disparar) \wedge \hat{\nabla}(1,\neg descarregar)$. Note que $\hat{\nabla}(1,\neg disparar) \equiv \neg\Box(1,disparar)$, e assim, $M/1 \models \hat{\nabla}(1,\neg disparar)$ se, e somente se,

$$M/1 \not\models \Box(1,disparar) \quad (*)$$

o que ocorre pois $T(1,disparar)$ não é uma consequência tautológica de $CONSEQ_1$ e nem pertencem a $CONSEQ_1$. O mesmo ocorre para $\hat{\nabla}(1,\neg descarregar)$. Além disso, $\Box(2,morto)$ não pertencerá a $CONSEQ_2$ pela expressão (*).

No instante de tempo t=3, o conjunto $CONSEQ_3$ é composto por :

$\Box(3,carregada)$, caso semelhante a $\Box(2,carregada)$ ; e $\Box(3,disparar)$, pois o axioma (2) pertence a teoria $\Psi$ e $M/2 \models true$.

Finalmente, $CONSEQ_4 = \{\Box(4,morto)\}$, já que o axioma (4) pertence a teoria $\Psi$ e $M/3 \models \Box(3,carregada) \wedge \Box(3,disparar) \wedge \hat{\nabla}(3,\neg defeito)$.

Observe que $\Box(4,carregada)$ não pertence a $CONSEQ_4$ pois neste caso deveríamos ter : $M/3 \not\models \Box(3,disparar)$.

Assim, obtivemos um modelo c.m.i. M no qual sabemos exatamente aquilo que queríamos : $T(1,carregada)$, $T(2,carregada)$, $T(3,carregada)$, $T(3,disparar)$ e $T(4,morto)$.

Este modelo c.m.i. do YSP é único. Vamos supor que exista um modelo M' cronologicamente mais ignorante que o modelo M, de forma que o ponto to da definição 6.6 é igual a 4, ou seja, M e M' só diferem na maneira de interpretar a sentença $\Box(4,\text{morto})$. Logo,

(1) para toda sentença básica $\varphi$ que tenha o u.p.t. menor ou igual a 4, se M' $\models \Box\varphi$ então M $\models \Box\varphi$ ; e

(2) existe uma sentença básica $\varphi'$ (que é T(4,morto)) que tem o u.p.t. igual a 4 e é tal que : M $\models \Box(4,\text{morto})$ , mas M' $\not\models \Box(4,\text{morto})$.

Observe que pela construção de M, como M $\models \Box(4,\text{morto})$, então existe $\phi \wedge \theta \supset \Box(4,\text{morto})$ que pertence a teoria $\Psi$ tal que os u.p.t. em $\phi$ e $\theta$ são menores ou iguais a 4 e M $\models \phi \wedge \theta$. De fato,

$$\phi \wedge \theta \equiv ( \Box(3,\text{carregada}) \wedge \hat{\nabla}(3,\neg\text{disparar}) \wedge \hat{\nabla}(3,\neg\text{descarregar})).$$

Como M e M' satisfazem as mesmas sentenças básicas que tenham o u.p.t. menor ou igual a 4, deveremos ter :

$$M' \models \phi \wedge \theta \,.$$

Logo, M' $\models \Box(4,\text{morto})$ , o que contradiz a nossa hipótese.

A forma como é construido o modelo M do teorema 6.7 (semelhante ao caso do YSP) sugere um procedimento que permite determinar todas as sentenças básicas comuns a todos os modelos c.m.i. . Nele, começaríamos com alguma sentença atômica básica anterior ao u.p.t. da primeira condição de limite. Então, prosseguindo no tempo, acrescentamos somente as sentenças atômicas básicas que são obtidas a partir do nosso conhecimento e ignorância passada. Tal algorítmo está descrito no apêndice B.

Outras propostas de solução que envolvem o conceito de minimização cronológica foram apresentadas. KAUTZ(1986) e LIFSCHITZ(1987) introduziram um novo predicado no axioma de circunscrição de forma a permitir que os fatos fossem minimizados cronologicamente.

O problema com todas as propostas que utilizam este tipo de minimização é que elas são restritas a problema do tipo "o que vai acontecer". Se , por exemplo, soubessemos que Fred está vivo no instante 4 e tentássemos descobrir o que ocorreu para que isto acontecesse (ou seja, em algum momento entre 1 e 3 a arma foi descarregada), deveríamos construir uma outra lógica , na qual o critério de minimalidade fosse o "cronológico" de trás para frente. Naturalmente que o melhor seria trabalhar com uma lógica que, no caso de problemas temporais, permita tanto "prever" os fatos quanto "explicá-los".

*iv) Minimização Causal*

Esta proposta de solução para o YSP foi desenvolvida por LIFSCHITZ(1987), e a sua idéia básica é reformular a descrição original do YSP de modo que seja possível dar um tratamento axiomático a causalidade, permitindo que esta noção fique explicitamente estabelecida através de axiomas dentro da representação do problema.

Nas Teorias Formais da Ação (Formal Theories of

Action), LIFSCHITZ(1987) procura através do acréscimo de novos predicados à linguagem , estabelecer uma nova interpretação para o que significam modelos intuitivos e não-intuitivos : no primeiro, todas as mudanças nos valores dos fluentes são causados por eventos; no segundo isso não acontece.

Os predicados acrescentados à linguagem original do YSP (representação de **HANKS** e **McDERMOTT** (1987)) são :

CAUSA(e,f,v) que expressa que o evento e faz com que o fluente f assuma o valor v (true ou false) ;

PRECOND(f,e) que expressa que o fato f é uma pré-condição para a execução bem sucedida do evento e .

Assim , Lifschitz representou o YSP através dos seguintes grupos de axiomas :

*GRUPO 1 :* os que descrevem a situação inicial :

(1) T(vivo,so), representando que Fred está vivo no estado so ;

(2) ¬ T(carregada,so), representando que a arma não está carregada em so;

*GRUPO 2 :* que descrevem como os fatos são afetados pelos eventos e quando um evento pode ser bem-sucedido:

(3) CAUSA(carregar,carregada,true), que representa que o evento carregar faz com que a arma fique carregada;

(4) CAUSA(disparar,carregada,false), representa que o evento disparar faz com que a arma fique descarregada;

(5) CAUSA(disparar,vivo,false), representa que o evento <u>disparar</u> faz com que Fred deixe de estar vivo;

(6) PRECOND(carregada, disparar), representa que o evento <u>disparar</u> é bem-sucedido somente quando a arma estiver carregada .

*GRUPO 3* : que descreve como o valor de um fluente muda depois de um evento que o afete ocorre (7) e que descreve que se um fluente não é afetado então seu valor permanece o mesmo (8) :

(7) $SUCESSO(e,s) \wedge CAUSA(e,f,v) \supset (T(f,r(e,s)) \equiv (v=true))$ , onde

$SUCESSO(e,s) \Leftrightarrow \forall f( PRECOND(f,e) \supset T(f,s) )$ (7.1);

(8) $\neg AFETA(e,f,s) \supset (T(f,r(e,s)) \equiv T(f,s))$ , onde

$AFETA(e,f,s) \Leftrightarrow (SUCESSO(e,s) \wedge \exists v\ CAUSA(e,f,v))$

**OBSERVAÇÃO 2 :** De forma a simplificar a discussão, vamos considerar que a representação axiomática acima considera apenas modelos de Herbrand. No caso mais geral , há a necessidade de se acrescentar mais axiomas.

A seguir , circunscrevemos os predicados CAUSA e PRECOND na teoria composta pelos axiomas (1)-(8) com T variando, o que faz com que as causas e pré-condições estabelecidas pela teoria sejam as únicas conhecidas. Logo, ao aplicarmos a circunscrição vamos obter que as únicas causas verdadeiras são aquelas expressas por (3)-(5) e a única pré-condição é a que está expressa por (6) .

Logo,

(9) $\forall e,f,v$ [ CAUSA(e,f,v) $\equiv$ (e=carregar $\wedge$ f=carregada $\wedge$ v=true) $\vee$
(e=disparar $\wedge$ f=carregada $\wedge$ v=false) $\vee$
(e=disparar $\wedge$ f=vivo $\wedge$ v=false)]

(10) $\forall$ e,f [ PRECOND(f,e) $\equiv$ (f=carregada $\wedge$ e=disparar) ]

Assim , a teoria circunscrita fica composta pelos axiomas (1)-(8) mais (9) e (10). O que queremos verificar é se desta teoria podemos concluir que :

$\neg$ T(vivo,r(disparar,r(esperar,r(carregar,so)))) ou seja,
$\neg$ T(vivo,s3)

onde s3 = r(disparar,r(esperar,r(carregar,so))).

Suponha, por absurdo, que Fred esteja vivo na situação s3, resultante da ocorrência do evento <u>disparar</u> no estado s2 :

T(vivo,r(disparar,s2)) (*)

Como o axioma (5) nos diz que o evento <u>disparar</u> faz com que o fluente <u>vivo</u> passe a ter o valor <u>false</u>, dele, do axioma (7) e mais (*), obtemos :

SUCESSO(disparar,s2) é falso , ou seja,
$\forall$f (PRECOND(f,disparar) $\supset$ T(f,s2)) é falso.

Logo, pela expressão (10), temos :

$$\neg T(carregada, s_2) \qquad (**)$$

Por outro lado, sabemos do axioma (7) que a arma está carregada em $s_1$, pois o evento carregar é bem-sucedido no estado $s_0$ (não existem pré-condições para esse evento - expressão 10), além de fazer com que o fluente carregada fique *true* (axioma 3).

Pela expressão (9), o evento esperar não muda o valor do fluente carregada ( $\forall v\ \neg CAUSA(esperar, carregada, v)$ ) e, portanto , o axioma (8) nos dá que :

$$T(carregada, s_2) \qquad (***)$$

o que contradiz (**).

O que difere este método dos outros que também tentaram aplicar a circunscrição ao cálculo situacional é que os predicados circunscritos neste caso não possuem entre seus argumentos situações . Desta forma, o conflito entre minimizar em instantes de tempo mais cedo ou mais tarde não aparece.

Este método é problemático pois a utilização do predicado CAUSA impõe uma grande restrição sobre o nosso poder de expressão. Por exemplo, se acrescentarmos o fluente morto à nossa teoria junto com uma restrição de domínio nos dizendo que se Fred está não vivo, ele está morto :

$$T(morto, s) \equiv \neg T(vivo, s)$$

seria interesante que pudéssemos concluir da teoria circunscrita resultante que, como o evento disparar faz com que Fred esteja não vivo, ele também deveria causar a sua morte , ou seja :

CAUSA(disparar,morto,true)

Infelizmente isso não ocorre pois , para essa nova teoria existe um outro modelo mínimo onde o evento carregar causa a morte de Fred (CAUSA(carregar,morto,true)). Nesse caso, não temos nenhuma situação na qual a arma está carregada enquanto Fred está vivo, e portanto o evento disparar não vai afetar o fluente morto (axioma (8)). Logo, CAUSA(disparar,morto,true) não poderá estar neste modelo, e será necessário acrescentar mais axiomas a teoria para eliminar este novo modelo anômalo.

v) *Sistema Temporal Não-Monotônico*

A lógica temporal não-monotônica apresentada por **MORGENSTEIN e STERN (1988)** , procura solucionar o problema da projeção temporal baseada na idéia de que os eventos só acontecem quando eles têm motivação para tal, ou seja, eles acontecem somente quando têm que acontecer. Esse sistema não só determina o que é válido após a ocorrência de uma sequência de eventos mas também permite determinar quais os prováveis eventos que podem ter ocorrido para se chegar a uma certa situação.

Nesta lógica, a estrutura do tempo é isomorfa a estrutura dos números inteiros. As sentenças são da forma :

$$T(f, j)$$

onde <u>f</u> representa um fluente e <u>j</u> um ponto no tempo. Os fluentes usados na representação do YSP são :

vivo : significa que Fred está vivo;

carregada : significa que a arma está carregada;

oc(e) : significa que o evento <u>e</u> ocorre. Os eventos podem ser : <u>carregar</u> , <u>disparar</u> e <u>descarregar</u>.

Assim , a sentença T(oc(descarregar),3) representa que no instante de tempo 3, o evento <u>descarregar</u> ocorre . Já a sentença T(¬carregada,4) representa que no instante de tempo 4 , a arma não está carregada .

**OBSERVAÇÃO 3 :** $T(\neg f, j) \equiv \neg\, T(f, j)$ .

Uma <u>*teoria*</u> <u>*instanciada*</u> TI ( Theory Instantiation ) é composta pela união de uma <u>*teoria*</u> $\tau$ e uma <u>*descrição histórica*</u> CD ( Chronicle Description ). A teoria $\tau$ é formada por dois tipos de regras que procuram governar o comportamento de alguns aspectos do mundo. São elas : as *causais* e as de *persistência* .

As <u>*regras causais*</u> são sentenças da forma :

$$\alpha \wedge \beta \supset \gamma \ ,$$

onde : $\alpha$ é um conjunto não-vazio de sentenças da forma T(oc(e),j) ; $\beta$ é uma conjunção de sentenças que estabelecem as pré-condições do evento ; e $\gamma$ descreve os resultados da ocorrência do evento .

As *regras de persistência* são da forma :

$$T(p,j) \wedge \beta \supset T(p,j+1) \ ,$$

onde $\beta$ inclue uma conjunção de declarações da forma $T(\neg\ oc(e),j)$ .

A *descrição histórica* CD descreve alguns fatos que são verdadeiros em determinados instantes de tempo.

O YSP é representado nesta lógica através da seguinte descrição histórica CD e teoria $\tau$ :

Descrição histórica CD :

(cd1) $T(vivo,1)$ ;

(cd2) $T(oc(carregar),1)$ ;

(cd3) $T(oc(disparar),5)$ .

Teoria $\tau$ :

i) Regras Causais :

(rc1) $T(oc(carregar),j) \supset T(carregada,j+1)$ ;

(rc2) $T(oc(disparar),j) \wedge T(carregada,j) \supset$
$T(\neg\ vivo,j+1)$ ;

(rc3) $T(oc(disparar),j) \supset T(\neg\ carregada,j+1)$ ;

(rc4) $T(oc(descarregar),j) \supset T(\neg\ carregada,j+1)$ ;

ii) Regras de Persistência :

(rp1) $T(vivo,j) \wedge [\ T(\neg\ oc(disparar),j) \vee$
$T(\neg\ carregada,j)\ ] \supset T(vivo,j+1)$ ;

(rp2) $T(carregada,j) \wedge T(\neg\ oc(disparar),j) \wedge$
$T(\neg\ oc(descarregar),j) \supset T(carregada,j+1)$ .

Vamos considerar agora os dois modelos do YSP : o intuitivo $M_1$ , onde a arma está carregada no instante 5 e

Fred está morto no instante 6 ; e o não-intuitivo $M_2$ , onde o evento <u>descarregar</u> ocorre em algum instante entre 2 e 4,e nesse caso , Fred está vivo em 6 . Veja nas figuras abaixo quais os fluentes verdadeiros em cada instante de tempo :

| 1 | 2 | 3 | 4 |
|---|---|---|---|
| vivo<br>oc(carregar) | vivo<br>carregada | vivo<br>carregada | vivo<br>carregada |

| 5 | 6 |
|---|---|
| oc(disparar)<br>vivo<br>carregada | ¬ vivo<br>¬ carregada |

Fig. 6.4 : modelo $M_1$

| 1 | 2 | 3 | 4 |
|---|---|---|---|
| vivo<br>oc(carregar) | vivo<br>carregada<br>oc(descarregar) | vivo<br>¬ carregada | vivo<br>¬ carregada |

| 5 | 6 |
|---|---|
| oc(disparar)<br>vivo<br>¬ carregado | vivo<br>¬ carregada |

Fig 6.5 : modelo $M_2$, quando o evento descarregar ocorre em 2

**OBSERVAÇÃO 4 :** Sem perda de generalidade, assumimos que o evento <u>disparar</u> no modelo $M_2$ ocorre no instante 2. Como veremos adiante, tal evento poderia ocorrer também nos instantes

3 ou 4.

Observe que ambos são modelos da teoria instanciada TI e, apresentados dessa maneira, nada nos leva a preferir um ou outro. Sendo o sistema em questão baseado na idéia de que os eventos que acontecem são exatamente aqueles que têm que acontecer , o critério de preferência entre os modelos utilizado por esta lógica será definido em termos de eventos motivados. Assim, a estratégia utilizada pelo sistema é o de minimizar os eventos que não são motivados, conhecidos como eventos <u>imotivados</u> .

Vamos denotar os modelos da teoria instanciada TI por M(TI). Logo, $M_1$ e $M_2$ passam a ser representados daqui para frente, respectivamente, por $M_1$(TI) e $M_2$(TI). Dizemos que uma sentença $\varphi$ é <u>*fortemente motivada*</u> com respeito a TI se $\varphi$ está em todos os modelos de TI . A sentença $\varphi$ é dita ser <u>*fracamente motivada*</u> em M(TI) se existir em TI uma regra causal ou de persistência da forma $\alpha \wedge \beta \supset \varphi$ , onde $\alpha$ é (fraca ou fortemente) motivada em M(TI) e M(TI) $\models \beta$ .

No caso do YSP , é fácil ver que cd1, cd2 e cd3 são sentenças fortemente motivadas pois, por pertencerem à CD, devem ser verdadeiras em todos os modelos. Já a sentença T($\neg$vivo,6) é fracamente motivada em $M_1$(TI) pois a regra causal rc2 :

$$T(oc(disparar),5) \wedge T(carregada,5) \supset T(\neg vivo,6)$$

é tal que T(oc(disparar),5) (pertence a CD) é fortemente motivada e $M_1$(TI) $\models$ T(carregada,5).

**DEFINIÇÃO 6.8** *(Sentença Motivada)*

Dada uma teoria instanciada TI = T ∪ CD , dizemos que uma sentença $\varphi$ é motivada em M(TI) se ela é forte ou fracamente motivada em M(TI).

O critério de minimalidade será de tal maneira que um modelo será um dos *preferidos* de TI se ele possuir tão poucos eventos imotivados quanto possível. Formalmente, a relação de preferência sobre os modelos é definida por :

**DEFINIÇÃO 6.9** *(Modelo Preferível)*

Dizemos que $M_i(TI)$ é *preferível* a $M_j(TI)$ , denotado por $M_i(TI) \leq_{TI} M_j(TI)$ se para todo $\varphi$ da forma T(oc(e),j), se $M_i(TI) \models \varphi$ e $M_j(TI) \not\models \varphi$ , então $\varphi$ é motivada em $M_i(TI)$ .

Logo, $M_i(TI)$ é preferível a $M_j(TI)$ se qualquer evento que ocorre em $M_i(TI)$ e não ocorre em $M_j(TI)$, for motivado em $M_i(TI)$. Vamos dizer que os modelos $M_i(TI)$ e $M_j(TI)$ são *equipreferíveis* caso $M_i(TI) \leq_{TI} M_j(TI)$ e $M_j(TI) \leq_{TI} M_i(TI)$, que representaremos por $M_i(TI) =_{TI} M_j(TI)$. Dizemos que um modelo M é o *preferido* se ele for um elemento mínimo sob $\leq_{TI}$ , ou seja, se $M'(TI) \leq_{TI} M(TI)$ então $M'(TI) =_{TI} M(TI)$.

Assim , no caso do YSP, temos que o modelo $M_1(TI)$ é preferível a $M_2(TI)$. De fato, note que cd1, cd2, cd3 e T(carregada,2) são fortemente motivados pois devem ser verdadeiros em todos os modelos de TI. Além disso, temos que T(¬ vivo,6) é fracamente motivada pois existe uma regra causal (rc2) tal que o evento <u>disparar</u> é motivado ( T(oc(disparar),5) ) e $M_1(TI) \models$ T(carregada,5) . Já no modelo $M_2$ , a ocorrência do evento <u>descarregar</u> no instante 2 é imotivada pois ela não é (fraca nem fortemente) motivada uma vez que não existe nenhuma regra causal ou de persistência da forma :

$$\alpha \wedge \beta \supset T(oc(descarregar),2) ,$$

e nem T(oc(descarregar),2) está em $M_1$. Como tal evento não ocorre em $M_1$ , pela definição 6.8, não temos que $M_2(TI)$ seja preferível a $M_1(TI)$.

Note que o fato da arma estar carregada no instante 5 e Fred estar morto no instante 6 , são verdadeiros em todos os modelos preferidos de TI . Isto porque , num modelo onde a arma estiver descarregada no instante 5, ou o evento <u>disparar</u> ou <u>descarregar</u> deve ocorrer entre 2 e 5. Mas ambos os eventos são imotivados. Logo, os fatos projetados por TI serão T(carregada,5) e T(¬vivo,6).

De maneira igualmente satisfatória, essa lógica apresentada manipula a projeção para trás no tempo. No caso do YSP, considere TI' = TI ∪ {T(vivo,6)}, onde TI' é uma teoria instanciada resultante de acrescentarmos à descrição histórica CD de TI o fato de que Fred está vivo no instante 6.

Como sabemos que o evento <u>disparar</u> ocorre em 5, teremos por (rc2) que a arma não pode estar carregada em 5. Entretanto, sabemos que a arma estava carregada no instante 2 (cd2 e rc1) e portanto , ela deve ter sido descarregada entre os instantes 2 e 5.

**OBSERVAÇÃO 5** : O evento <u>disparar</u> pode também ser responsável pela arma ficar descarregada (rc3), porém isso também causaria a morte de Fred (rc2), que persistiria até o instante 6 , onde vale que T(vivo,6). Logo, em todos os modelos de TI' o evento que deve ocorrer é <u>descarregar</u>.

Assim, considere os seguintes modelos de TI' :

i) $M_1$(TI') onde <u>descarragar</u> ocorre em 2 e a arma está descarregada em 3, 4 e 5 ;

ii) $M_2$(TI') onde <u>descarragar</u> ocorre em 3 e a arma está descarregada em 4 e 5 , e carregada em 3;

iii) $M_3$(TI') onde <u>descarregar</u> ocorre em 4 e a arma está descarregada em 5 , e carregada em 3 e 4.

Os três modelos são equipreferíveis já que :

$$\text{(i)} \qquad M_1(TI') \leq_{TI} M_2(TI') \text{ ,}$$

$$\text{(ii)} \qquad M_2(TI') \leq_{TI} M_3(TI') \text{ , e}$$

$$\text{(iii)} \qquad M_3(TI') \leq_{TI} M_1(TI')$$

De fato,

(i) $M_1(TI') \leq_{TI} M_2(TI')$ pois $M_1(TI') \models T(oc(descarregar),2)$ e $M_2(TI') \not\models T(oc(descarregar),2)$ .

Como, de (rp2) temos :

$$T(carregada,2) \wedge T(\neg carregada,3) \wedge T(\neg oc(disparar),j) \supset T(oc(descarregar),2) \qquad (i.1)$$

onde $M_1(TI') \models T(carregada,2) \wedge T(\neg carregada,3)$ e $T(\neg oc(disparar),j)$ é motivada (pela rc2) , teremos por definição que :

$$T(oc(descarregar),2)$$

é fracamente motivada em $M_1(TI')$. Logo, pela definição 6.8, (i) é válida.

De maneira análoga, (ii) e (iii) podem ser provados.

Além disso, $M_1(TI')$, $M_2(TI')$ e $M_3(TI')$ são modelos preferíveis para TI'. Vejamos no caso de $M_1(TI')$ : se $M_1(TI')$ é preferido para TI', então para todo modelo $M(TI')$, se $M(TI') \leq_{TI} M_1(TI')$ então $M(TI') =_{TI} M_1(TI')$.

Supondo que não tenhamos $M_1(TI') \leq_{TI} M(TI')$, por definição , existirá uma sentença $\varphi$ da forma $T(oc(e),j)$, tal que :

(A) $M_1(TI') \models \varphi$ ,

(B) $M(TI') \not\models \varphi$ e

(C) $\varphi$ não é motivada em $M_1(TI')$.

Mas (C) é falso, pois todas as sentenças $\varphi$ em $M_1(TI')$ são motivadas ( $T(oc(carregar),1)$ e $T(oc(disparar),5)$ fortemente ; e $T(oc(descarregar),2)$ fracamente ).

Os casos $M_2(TI')$ e $M_3(TI')$ são análogos a esse.

Assim, tudo o que TI' pode "*imaginar*" que ocorreu é a disjunção :

$$T(oc(descarregar),2) \vee$$
$$T(oc(descarregar),3) \vee$$
$$T(oc(descarregar),4)$$

que é exatamente o que esperávamos.

*vi) Nova Política de Circunscrição*

Na formulação original do YSP, HANKS e McDERMOTT(1987) circunscreveram o predicado AB, com o predicado T variando e a função r(e,s) fixa. Essa política de circunscrição determinava, para as diversas situações fixas r(e,s), os fluentes que eram válidos ou não em cada uma delas. Como vimos, isso acarreta que a teoria circunscrita obtida possui dois modelos mínimos, um intuitivo e outro não, permitindo dessa maneira que somente conclusões fracas, tais como T(vivo,s3) ∨ T(morto,s3), sejam obtidas.

BAKER(1989) sugeriu uma mudança na política de circunscrição utilizada no YSP : ao invés do predicado T variar quando AB é circunscrito, teríamos a função r(e,s) como variável e T fixo. Assim, para cada combinação de fluentes, assumimos alguma possível situação na qual eles são válidos (o que nos obrigará a ter todas as situações consistentes explícitas no nosso domínio de discurso).

Considere a seguinte versão simplificada do YSP :

(1) $\neg AB(f,e,s) \supset (T(f,r(e,s)) \equiv T(f,s))$

(2) $T(carregada,s) \supset \neg T(vivo,r(disparar,s))$

(3) $T(carregada,s_0)$

(4) $T(vivo,s_0)$

Vamos considerar $t, t_1, t_2, \ldots$ como variáveis de tempo e definir a função tempo(s) que mapeia uma situação s em seu tempo. A estrutura do tempo é isomorfa aos números inteiros.

Aos axiomas (1)-(4) que descrevem o YSP, devem ser acrescidos mais três axiomas :

(5) $tempo(r(e,s)) = tempo(s) + 1$ , representando que os eventos incrementam o tempo ;

(6) $esperar \neq disparar$ , que é o axioma do nome único ;

(7) $\forall t\ [\ \exists s\ (tempo(s) = t \wedge T(vivo,s) \wedge T(carregada,s))] \wedge$
$\forall t\ [\ \exists s\ (tempo(s) = t \wedge T(vivo,s) \wedge \neg T(carregada,s))] \wedge$
$\forall t\ [\ \exists s\ (tempo(s) = t \wedge \neg T(vivo,s) \wedge T(carregada,s))] \wedge$
$\forall t\ [\ \exists s\ (tempo(s) = t \wedge \neg T(vivo,s) \wedge \neg T(carregada,s))]$

representando que para todo ponto de tempo e toda possível combinação de fluentes, existe alguma situação nesse instante de tempo onde os fluentes são válidos. Esta expressão é chamada de *axioma de existência das situações*.

Para simplificar a discusão, vamos trabalhar com o axioma (7), ressaltando que para problemas mais complexos, a utilização de um axioma desse tipo torna-se inviável. BAKER(1989) apresenta uma formulação alternativa, onde um predicado de anormalidade de situações e uma função que mapeia um tempo e um fluente em alguma situação são acrescidos à linguagem. Esta nova anormalidade vai introduzir um novo axioma de circunscrição ( além daquele onde AB é circunscrito com r(e,s) variando ) com todos os outros predicados e funções variando.

Considere que T seja a teoria composta por (1)-(7) e $T' \Leftrightarrow CIRC(T;AB;r)$. Sejam as figuras 1 e 2 a representação dos modelos de T :

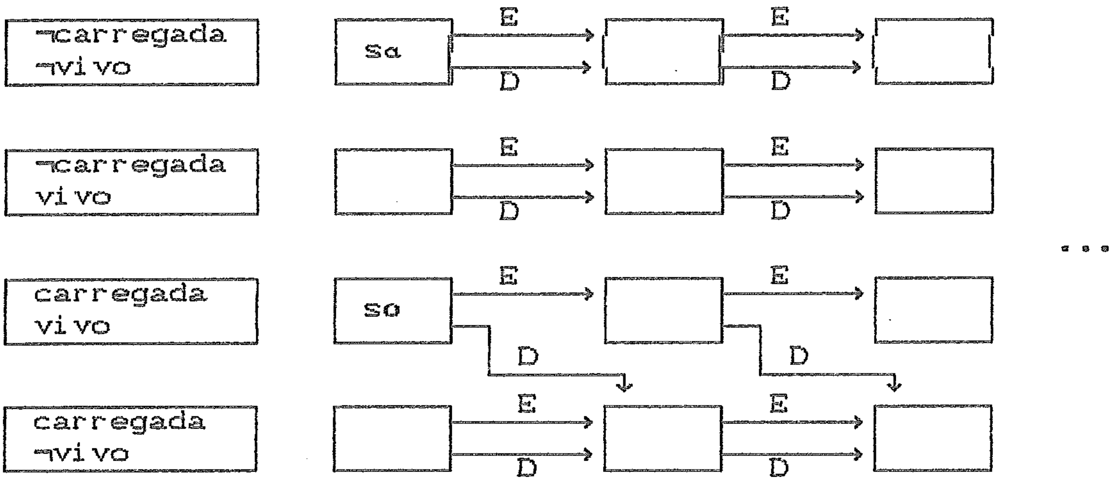



FIG. 6.6

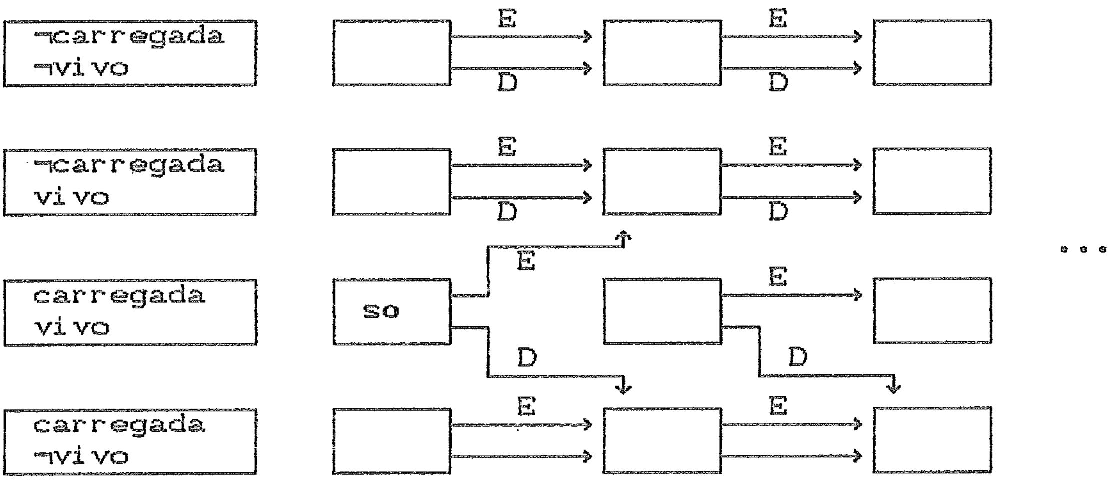



FIG. 6.7

Nas figuras acima, na coluna mais à esquerda temos representadas todas as combinações possíveis de fluentes no YSP. Cada um dos retângulos à direita dos fluentes representam um conjunto de situações num certo estado de tempo no qual os fluentes valem ou não. Por exemplo, sa (FIG. 6.6) pertence ao conjunto de situações nas quais os fluentes <u>carregada</u> e <u>vivo</u> não valem. Já so pertence ao conjunto de situações nas quais aqueles fluentes valem. O axioma (7) nos garante a existência de pelo menos uma situação em cada retângulo.

O tempo flui horizontalmente e as setas E e D mostram o resultado, respectivamente, dos eventos <u>esperar</u> e <u>disparar</u>. As setas em diagonal representam eventos que mudam pelo menos um fluente, e por isso são associados a pelo menos uma anormalidade.

A figura 1 representa um modelo onde as setas diagonais D que partem de situações onde os fluentes <u>vivo</u> e <u>carregar</u> são válidos (entre elas, so) , levam a situações nas quais o fluente <u>vivo</u> deixa de ser válido. As anormalidades associadas à essas setas (do tipo AB(vivo,disparar,s)) são as únicas anormalidades desse modelo, sendo este portanto um modelo mínimo de T e, consequentemente um modelo de T'. Este corresponde ao modelo intuitivo do YSP.

Já a figura 2 corresponde ao modelo não-intuitivo do YSP, pois a nova anormalidade (AB(carregada,esperar,s)) não elimina qualquer das outras. Logo, este modelo não é mínimo.

Observe que em todos os modelos mínimos de T, a arma permanece carregada e Fred morre após o seu disparo. Isto porque num modelo onde o evento <u>esperar</u> em so fizer com que a arma fique descarregada, deveremos ter que :

$$\neg\, T(carregada, r(esperar, so)) \wedge AB(carregada, esperar, so) \qquad (*)$$

Mas, como o axioma (7) nos garante a existência de pelo menos uma situação para cada combinação de fluentes, podemos variar a função <u>r(e,s)</u> de forma a manter a arma carregada, e assim eliminar a anormalidade (*) sem introduzir qualquer outra. Logo :

$$T' \models \neg T(vivo, r(disparar, r(esperar, so))).$$

BAKER e GINSBERG(1989-b) produziram um trabalho que pode ser considerado como uma consequência natural da proposta que acabamos de examinar. Nele, é feita uma distinção entre uma *situação* (sequência de ações), tal como so, e um *estado* , que será descrito simplesmente em termos dos valores dos fluentes. Assim, se um fluente f muda quando um evento e ocorre na situação s, pensaremos nas anormalidades não como uma função da situação s, mas como uma função do estado à ela correspondente.

Para formalizar tal idéia, introduzimos um subscrito no axioma *frame* :

$$\neg AB_s(e,f,s) \supset [\ T(f,s) \equiv T(f,\langle e|s\rangle)\ ] \qquad (1)$$

para indicar que é uma situação que está sendo considerada anormal, onde $\langle e|s\rangle$ é o resultado do evento e na situação s ($\langle\ \rangle$ corresponde a situação inicial).

Afim de estabelecer uma relação entre as situações e os vetores de estado que os descrevem, introduzimos o predicado :

$$DESC(s,v)$$

que pretende capturar a idéia de que a situação s é descrita pelo vetor de estado v. A intenção aqui é que v contenha informação suficiente para determinar o valor atribuido por s à qualquer fluente particular f. No YSP, v contém duas entradas, correspondendo aos valores dos fluentes vivo e carregada respectivamente; o valor de qualquer outro fluente , tal como morte ($\neg$vivo) , pode ser computado destas duas entradas.

Assim, podemos dizer que situações anormais correspondem a vetores de estado anormais :

$$AB_s(e,f,s) \wedge DESC(s,v) \supset AB_v(e,f,v) \qquad (2)$$

Necessitamos também um axioma dizendo que para qualquer vetor de estado v, existe alguma situação s(v) que é descrita por ele :

$$DESC(s(v),v) \qquad (3)$$

(Observe que (3) diz somente que alguma situação é descrita pelo vetor de estado v; de fato várias situações podem ser).

Completando o conjunto de axiomas temos :

$$\left.\begin{array}{ll} (4) & T(vivo,\langle\ \rangle) \\ (5) & T(carregada,\langle\ \rangle) \end{array}\right\} \text{condições iniciais}$$

(6) $T(carregada,s) \supset \neg\, T(vivo,\langle disparar \mid s\rangle)$

(7) $DESC(s,[vivo,carregada]) \equiv T(vivo,s) \wedge T(carregada,s)$

(8) $DESC(s,[vivo,descarregada]) \equiv T(vivo,s) \wedge \neg T(carregada,s)$

(9) $DESC(s,[morto,carregada]) \equiv \neg T(vivo,s) \wedge T(carregada,s)$

(10) $DESC(s,[morto,descarregada]) \equiv \neg T(vivo,s) \wedge \neg T(carregada,s)$

Suponha que denotássemos por so a situação s[vivo,carregada]. De (3), temos:

$$DESC(s_0,[vivo,carregada]) \quad (a)$$

De (7) e (a) : $T(vivo,s_0) \wedge T(carregada,s_0)$ (b).

De (6) e (b) : $\neg T(vivo,\langle disparar|s_0\rangle)$ (c).

De (1), (b) e (c) : $AB_s(disparar,vivo,s_0)$ (d).

De (d), (2) e (a) :

$$AB_v(disparar,vivo,[vivo,carregada])$$

Nenhum outro vetor de estado ou evento será anormal se ABv tem sua extensão mínima, portanto concluímos que :

$$\neg\ AB_v(esperar,carregada,[vivo,carregada]) \ ,$$

e daí segue que :

$$\neg AB_s(esperar,carregada,\langle\ \rangle) \ ,$$

onde a situação inicial < > é dada por (4) e (5). Segue que a arma permanece carregada durante o evento <u>esperar</u>, e Fred morre após o evento <u>disparar</u> .

Ambas as soluções apresentadas funcionam basicamente pelo mesmo motivo que a minimização causal : a minimização não envolve situações e portanto não gera conflitos entre minimizar mais cedo ou mais tarde no tempo. Além disso, as formalizações apresentadas não sofrem das restrições impostas pela utilização do predicado CAUSA.

*vii) Conclusão*

Neste capítulo, examinamos a deficiência apresentada quando da utilização de métodos não-monotônicos na formalização de problemas que envolvam o raciocínio do senso-comum. Vimos que tal deficiência se refere a derivação de fatos dentro desses sistemas uma vez que, por gerarem modelos conflitantes (intuitivo e não-intuitivo), somente fatos "fracos" podem ser derivados. Este foi o caso dos exemplos vistos neste capítulo, quando as teorias envolvidas só nos permitiram derivar fatos como : "Fred está vivo ou morto", "Tweety voa ou não" e "O bloco A está sobre B ou B está sobre A" .

Visando solucionar este problema foram apresentadas uma série de soluções. De formas distintas, todas as soluções buscam obter, dentre os vários modelos gerados, aquele considerado intuitivo,ou seja, aquele correspondente ao senso-comum.

Para alcançar tal objetivo, as várias abordagens sugerem a utilização de algum critério de seleção entre os modelos gerados. Como vimos no caso específico do YSP, cada proposta de solução apresentada utiliza um critério de seleção distinto :

- *na circunscrição suportada, são selecionados os modelos* suportados pelos fatos ;

- *na minimização cronológica, são selecionados os modelos* nos quais as anormalidades ocorrem o mais tarde possível;

- *na minimização causal e no sistema temporal*, são selecionados os modelos nos quais os fatos mudam de valor verdade por algum motivo (ou causa);

- através de uma nova política de circunscrição, quando os modelos "desejados" são obtidos mudando-se a tupla de predicados Z.

Todas as soluções apresentadas não são totalmente satisfatória , uma vez que elas ou apresentam problemas quando novos fatos são acrescentados à teoria original; ou restringem o poder de expressão , exigindo que mais axiomas sejam utilizados para eliminar eventuais modelos anômalos; ou ainda são soluções restritas ao caso de situações temporais.

# CAPÍTULO VII

## CONCLUSÃO

Neste trabalho, analisamos a utilização de métodos não-monotônicos na formalização do senso-comum.

Para isto, apresentamos dois métodos não-monotônicos : a lógica de default e a circunscrição. O primeiro, busca ampliar os conhecimentos expressos através de uma teoria de primeira-ordem , utilizando para tal regras defaults. O outro, procura explicitar hipoteses que estejam implícitas numa dada teoria, através de um axioma.

Apresentamos também dois sistemas de manutenção da verdade : o TMS, que trata de maneira adequada os exemplos utilizados; e o ATMS, que sugere uma implementação para determinar se um fato segue ou não de uma teoria circunscrita.

A partir de um problema (YSP) vimos que tais métodos geravam modelos contraditórios (intuitivo e não-intuitivo), e isto só nos permitia derivar fatos "fracos" das teorias envolvidas. Tentando solucionar tal problema , uma série de propostas foram apresentadas, cujo objetivo comum era introduzir um critério de seleção , de forma que o modelo obtido era aquele compatível com o senso-comum .

A introdução destes critérios também é feita de maneira distintas nas várias abordagens : umas constrõem sistemas que incorporam o critério desejado como parte integrante do método (caso do sistema de SHOHAM (1988) e o sistema de MORGENSTEIN e STERN (1988)), outras introduzem em sistemas já existentes o critério desejado (como no caso da circuncrição suportada).

O problema de se estabelecer um critério de seleção entre modelos parece ser algo bastante complexo . Utilizar critérios específicos para cada tipo diferente de problema, obriga a criação de vários sistemas, cada um sendo responsável por um caso particular. Uma alternativa seria escolher um critério bem amplo, que servisse para uma classe mais abrangente de problema, o que parece ser mais difícil.

Vamos considerar a circunscrição suportada que sugere um critério de seleção mais abrangente. Neste caso, o critério é utilizado para selecionar, dentre os modelos da teoria circunscrita, aqueles suportados pelos fatos . Isto leva em conta que a circunscrição é incapaz de ela própria obter os modelos desejados.

Observe que a solução apresentada por BAKER (1989) consegue o modelo desejado apenas com a alteração da política de circunscrição. A primeira vista isto apresenta dois aspectos positivos : (1) não é mais necessário estabelecer um critério de seleção dos modelos; (2) elimina a necessidade da criação de novos sistemas onde o critério de seleção venha incorporado ao método . Portanto, somente com a circunscrição poderíamos modelar o raciocínio do senso-comum.

Porém, não é exatamente isto o que ocorre. Note que na circunscrição, os modelos mínimos da teoria que é circunscrita são obtidos através da relação $\leq$ mod(P,Z), que por sua vez é determinada pelas tuplas de predicados P e Z.

Logo, podemos considerar que os modelos mínimos de uma teoria <u>são *selecionados pela política de circunscrição* que *é utilizada*</u>. Deste ponto de vista, os aspectos positivos que pareciam existir desaparecem : (1) o critério de seleção de modelos é determinado pela política de circunscrição adotada; (2) a circunscrição é um sistema que traz incorporado em si o critério de seleção de modelos.

Novamente voltamos à questão levantada no capítulo V: pouco se estudou a respeito de quais predicados de uma determinada teoria devem compor as tuplas P e Z. A solução de BAKER (1989) sugere que a partir de uma política adequadamente escolhida somente os modelos desejados são obtidos.

O problema é que a escolha da política certa parece ser tão complexa quanto escolher um critério de seleção de modelos. Na verdade, como dissemos antes, podemos encarar a política de circunscrição como um critério de seleção de modelos.

Assim, a partir desta visão do problema, trabalhos futuros poderiam ser desenvolvidos visando estabelecer uma relação entre a política de circunscrição adotada e os modelos de uma teoria que ela seleciona, e como tal política poderia ser determinada numa implementação.

Outro caminho de pesquisa seria tentar aperfeiçoar os sistemas apresentados, ou criar novos sistemas, de forma que seja possível tratar adequadamente a formalização do raciocínio do senso-comum por intermédio de métodos não-monotônicos.

# APÊNDICE A

## *ALGORÍTMO 1*

Seja (CP C IH OH) uma justificativa CP válida , onde C é o nó consequente , IH é a lista de *hipóteses* in e OH é a lista de *hipóteses* out :

PASSO 1 : Marque cada um dos nós em IH e OH com as marcas E (examinada) e S (subordinada) .

PASSO 2 : Faça o seguinte procedimento para C :

(a) i : Se o nó possui a marca E , retorne. Caso contrário, marque o nó com E e utilize recurssivamente o PASSO 2 para cada um dos antecedentes do nó corrente ;

ii : Caso quaisquer dos nós antecedentes seja marcado com S, marque o nó corrente com S. Retorne.

PASSO 3 : Usando um processo recursivo semelhante ao PASSO2-(a), remova as marcas E dos fundamentos de C .

PASSO 4 : Marque as hipóteses (nós em IH e OH) com E.

PASSO 5 : Faça o seguinte procedimento para C :

(b) i : Se o nó possui a marca E, retorne. Caso contrário, marque o nó com uma marca E.

ii : Se o nó tem uma marca S, utilize recursivamente o PASSO 5-(b) para cada um dos antecedentes do nó corrente . Caso contrário, coloque o nó corrente na lista IS, quando o seu estado for IN e na lista OS, quando o seu estado for OUT.

Retorne.

PASSO 6 : Remover todas as marcas E e S de todos os nós, e retorne a justificativa (SL IS OS) .

*ALGORITMO 2*

Procedimento DDB (Dependency-Directed Backtracking) :

PASSO 1 : Determinar o conjunto de suposições máximas S = {$A_1$,...,$A_n$} com relação ao nó contradição C. Uma suposição A ∈ S se, e somente se, A é um fundamento de C e não existe nenhuma outra suposição B nos fundamentos de C tal que A pertence aos fundamentos de B .

PASSO 2 : Se não houve nenhuma tentativa anterior de determinar S a partir de C , cria-se um novo nó. Este novo nó é representado por NG (nogood), e serve para representar a inconsistência existente no conjunto S. O nó NG é justificado por :

(CP C ($A_1$,...,$A_n$) ( )) .

PASSO 3 : Selecionar, aleatoriamente, algum $A_i$ ∈ S . Se $D_1$ ,...,$D_k$ são nós OUT da *lista* out de $A_i$ , escolhe-se algum desses $D_j$ e o justificamos com :

(SL (NG $A_1$ ... $A_{i-1}$ $A_{i+1}$ ... $A_n$) ($D_1$ ... $D_{j-1}$ $D_{j+1}$ ... $D_k$))

PASSO 4 : Se o TMS encontrar outros argumentos de modo que o nó contradição C permaneça no estado IN após acrescentarmos a justificativa para $D_j$, repita o procedimento.

Caso o nó contradição passe para o estado OUT , o procedimento termina .

Se nenhuma suposição for descoberta nos fundamentos do nó contradição C, o TMS notifica o programa principal que foi encontrado uma contradição que não pode ser analisada.

*ALGORITMO 3*

Acréscimo de novas justificativas :

PASSO 1 : O ATMS recebe uma nova justificativa do RP para um certo nó n. Supondo que todo nó antecedente possua um rótulo completo e correto, um novo nó é computado para $\gamma_n$ . Este novo rótulo é composto por ambientes Ej, para j=1,...,k , tal que Ej é a união de uma combinação de ambientes dos nós antecedentes.

Após remover os ambientes inconsistentes e que tenham algum outro ambiente como subconjunto , o rótulo é correto e mínimo.

PASSO 2 : Se o rótulo recentemente computado é o mesmo que o antigo, terminamos com o nó .

PASSO 3 : Se o nó for $\gamma_\perp$ , cada ambiente do rótulo obtido é acrescido à uma base de dados de ambientes nogood , além de serem removidos dos rótulos de todos os outros nós nos quais eles aparecem .

PASSO 4 : Para qualquer outro nó que não $\gamma_\perp$ , o processo recomeça para atualizar outros nós que têm justificativas que mencionam o nó cujo rótulo mudou devem ser atualizados.

**OBSERVAÇÃO** : Os rótulo não são garantidos serem consistentes e completos até o algorítmo terminar.

# APÊNDICE B

## *ALGORÍTMO 4*

*Organiza todas as subfórmulas de $\psi$ em ordem crescente do u.p.t. e então caminha ao longo delas um tempo por vez, determinando para cada uma se ou ele ou sua negação devem ser conhecidas.*

PASSO 1 : Seja a lista T de todas as sentenças em $\psi$.
Forme a lista S com todas as s.a.b. que aparecem em T, omitindo os sinais de negação e eliminado as sentenças duplicadas.

PASSO 2 : Ordene T pelo argumento de tempo mais tarde do consequente.
Ordene S da mesma forma .
Marque todas os membros de S com NM (Não-Marcado).

PASSO 3 : Se T é vazio, pare. As s.a.b. (positivas e negativas) que são conhecidas nos modelos c.m.i. de $\psi$ são exatamente aquelas marcadas com SIM em S e a negação daqueles marcados NÃO em S.

PASSO 4 : Remova o primeiro elemento de T. Suponha que ele seja $\phi \wedge \theta \supset \Box(t_1, t_2, [\neg]p)$ . Para cada elemento $\Box(t_{i1}, t_{i2}, [\neg]p_i)$ de $\phi$ e cada elemento $\hat{\nabla}(t_{i1}, t_{i2}, [\neg]p_i)$ de $\theta$ , determine como $T(t_{i1}, t_{i2}, p_i)$ está marcado em S.
Se uma das seguintes condições é verdadeira :

(a) $\square(t_{i1}, t_{i2}, p_i)$ é um elemento de $\phi$ e $T(t_{i1}, t_{i2}, p_i)$ não é marcada SIM em S ;

(b) $\square(t_{i1}, t_{i2}, \neg p_i)$ é um elemento de $\phi$ e $T(t_{i1}, t_{i2}, p_i)$ não é marcada NÃO em S ;

(c) $\hat{\vee}(t_{i1}, t_{i2}, p_i)$ é um elemento de $\theta$ e $T(t_{i1}, t_{i2}, p_i)$ é marcada NÃO em S ;

(d) $\hat{\vee}(t_{i1}, t_{i2}, \neg p_i)$ é um elemento de $\theta$ e $T(t_{i1}, t_{i2}, p_i)$ é marcada SIM em S ;

então vá para o passo 3.

Caso contrário, rotule $T(t_1, t_2, p)$ em S com SIM ou NÃO, dependendo se o consequente é $\square(t_1, t_2, p)$ ou $\square(t_1, t_2, \neg p)$. Vá para o passo 3.

## REFERÊNCIAS BIBLIOGRÁFICAS

[1] BAKER, A.B. e GINSBERG, M.L. (1989-a), " A Theorem Prover for Prioritized Circumscription", Proceedings of the 11th IJCAI-89, pp. 463-467.

[2] BAKER, A.B. e GINSBERG, M.L., (1989-b) "Temporal Projection and Explanation", Proceedings of the 11th IJCAI-89, pp. 906-911.

[3] BAKER, A.B., (1989), "A Simple Solution to the Yale Shooting Problem", Proceedings of the 1st Conference on Principles of Knowledge Representation and Reasoning, pp. 11-20.

[4] CASANOVA, M. A. ; GIORNO, F. A. C. e FURTADO, A. L. (1988) , "Programação Lógica e a Linguagem PROLOG" , Editora Edgar Blüncher Ltda.

[5] CHANG, C. L. e LEE, R.C.T. (1971) , "Symbolic Logic and Mechanical Theorem Proving" , Academic Press Inc.

[6] DAVIS, M. (1980), "The Mathematics of Non-Monotonic Reasoning", Artificial Intelligence , vol. 13, pp. 73-80.

[7] DE KLEER, J. (1986-a), "An Assumption-Based TMS", Artificial Intelligence , vol. 28, pp. 127-162.

[8] DE KLEER, J. (1986-b), "Extending the ATMS", Artificial

Intelligence , vol. 28, pp. 163-192.

[9] DE KLEER, J. (1988), "A General Labeling Algorithm for Assumption-Based Truth Maintenance" , Proceedings AAAI-88 , pp. 188-192.

[10] DOYLE, J. (1979), "A Truth Maintenance System", Artificial Intelligence , vol. 12, pp. 231-272.

[11] ETHERINGTON, D.W. ; MERCER, R.E. e REITER, R. (1985), "On the Adequacy of Predicate Circumscription for Closed-World Reasoning" , Computational Intelligence , vol. 1 , pp. 11-15.

[12] ETHERINGTON, D.W. (1987), "Formalizing Nonmonotonic Reasoning Systems", Artificial Intelligence , vol. 31, pp. 41-85.

[13] ETHERINGTON, D.W., FORBUS, K.D., GINSBERG, M.L., ISRAEL, D., LIFSCHITZ, V. (1989), "Critical Issues in Nonmonotonic Reasoning", Proceedings of the 1st Conference on Principles of Knowledge Representation and Reasoning, pp. 500-504.

[14] GELFOND, M. e LIFSCHITZ, V. (1988) , "Compiling Circumscription Theories into Logic Programs : Preliminary Report", Proceedings AAAI-88, pp. 455-459.

[15] GELFOND, M. ; PRZYMUSINSKA, H. e PRZYMUSINSKI, T. (1989) , "On the Relationship between Circunscription and

Negation as Failure", Artificial Intelligence , vol. 38, pp. 75-94.

[16] GINSBERG, M.L. e SMITH, D.E. (1988-a) , "Reasoning about Action I : A Possible Worlds Approach" , Artificial Intelligence , vol. 35, pp. 165-195.

[17] GINSBERG, M.L. e SMITH, D.E. (1988-b) , "Reasoning about Action II : The Qualification Problem " , Artificial Intelligence , vol. 35, pp. 311-342.

[18] GINSBERG, M.L. (1989), "A Circumscriptive Theorem Prover", Artificial Intelligence , vol. 39, pp. 209-230.

[19] HALPERN, J.Y. e MOSES, Y. (1985) , "A Guide to the Modal Logics of Knowledge and Belief : Preliminary Draft" , Proceedings of IJCAI-85 , pp. 480-490.

[20] HANKS, S. e McDERMOTT, D. (1987), "Nonmonotonic Logic and Temporal Projection", Artificial Intelligence , vol. 33, pp. 379-412.

[21] HAUGH, B.A. (1988), "Simple Causal Minimizations for Temporal Persistence and Projection", Proceedings AAAI-87, pp. 218-223.

[22] HUGHES, G. E. e CRESWELL, M. J. (1968) , "An Introduction to Modal Logic" , Methuen and Co LTD.

[23] IMIELINSKI, T. (1987) , "Results on Translating

Defaults to Circumscription" , Artificial Intelligence , vol. 32 , pp. 131-146.

[24] KAUTZ, H.A. (1986), "The Logic of Persistence", Proceedings AAAI-86, pp. 401-405.

[25] KOLAITIS, P.G. e PAPADIMITRIOU, C.H. (1989) , "Some Computational Aspects of Circumscription" , Journal of ACM, vol. 37 , n. 1 , pp. 1-14.

[26] LIFSCHITZ, V. (1984) , "Some results on Circumscription" , Workshop on Nonmonotonic Reasoning , pp. 151-164.

[27] LIFSCHITZ, V. (1985-a), "Computing Circumscription", Proceedings of the 9th IJCAI-85, pp. 121-127.

[28] LIFSCHITZ, V. (1985-b), "Closed-Word Databases and Circumscription" , Artificial Intelligence , vol. 27, pp. 229-235.

[29] LIFSCHITZ, V. (1986), "On the Satisfiability of Circumscription" Artificial Intelligence , vol. 28, pp. 17-27.

[30] LIFSCHITZ, V. (1987-a), "Pointwise Circumscription", Readings in Nonmonotonic Logic, Morgan Kaufmann Publishers, pp. 179-193.

[31] LIFSCHITZ, V. (1987-b), "Formal Theories of Action",

Readings in Nonmonotonic Logic, Morgan Kaufmann Publishers, pp. 410-432.

[32] LIFSCHITZ, V. (1989), "Miracle in Formal Theories of Action", Artificial Intelligence , vol. 38, pp. 225-237.

[33] LUKASZEWICZ, W. (1985), "Two Results on Default Logic", Proceedings of the 9th IJCAI-85, pp. 459-461.

[34] McCARTHY, J. e HAYES, P.J. (1969), "Some Philosophical Problems from the Standpoint of Artificial Intelligence" , Machine Intelligence , vol. 4 , pp. 463-502.

[35] McCARTHY, J. (1980), "Circumscription: A Form of Non-Monotonic Reasoning", Artificial Intelligence , vol. 13, pp. 27-39.

[36] McCARTHY, J. (1986), "Applications of Circumscription to Formalizing Common-Sense Knowledge", Artificial Intelligence , vol. 28, pp. 89-116.

[37] McDERMOTT, D. e DOYLE, J. (1980) , "Non-monotonic Logic I" , Artificial Intelligence , vol. 13, pp. 41-72.

[38] McDERMOTT, D. (1982-a) , "Nonmonotonic Logic II : Nonmonotonic Modal Theories" , Journal of ACM , vol. 29 , pp. 33-57.

[39] McDERMOTT, D. (1982), "A Temporal Logic for Reasoning About Processes and Plans", Cognitive Science , vol. 6,

pp. 101-155.

[40] McDERMOTT, D. (1987), "A Critique of Pure Reason", Computational Intelligence, vol. 3, pp. 149-150.

[41] MORGENSTERN, L. e STEIN, L.A. (1988), "Why Things go Wrong: A Formal Theory of Causal Reasoning", Proceedings AAAI-88, pp. 518-523.

[42] MORRIS, P.H. (1988), "The Anomalous Extension Problem in Default Reasoning", Artificial Intelligence, vol. 35, pp. 383-399.

[43] PRZYMUSINSKI, T.C. (1988), "On the Relationship Between Logic Programming and Nonmonotonic Reasoning", Proceedings AAAI-88, pp. 444-448.

[44] PRZYMUSINSKI, T.C. (1989), "An Algorithm to Compute Circumscription", Artificial Intelligence, vol. 38, pp. 49-73.

[45] REITER, R. (1980), "A Logic for Default Reasoning", Artificial Intelligence, vol. 13, pp. 81-132.

[46] REITER, R. e CRISCUOLO, G. (1981), "On Interacting Defaults", Proceedings of the 7th IJCAI-81, pp. 270-276.

[47] REITER, R. e DEKLEER, J. (1987), "Foundations of Assumption-Based Truth Maintenance Systems: Preliminary Report", Proceedings AAAI-87, pp. 183-188.

[48] REITER, R. (1988) , "Nonmonotonic Reasoning" , Exploring Artificial Intelligence : Survey Talks from the National Conference on A.I. , pp. 439-481.

[49] SHOHAM , Y. (1985) , "Ten Requirements for a Theory of Change" , New Generation Computing , vol. 3 , pp. 467-477.

[50] SHOHAM , Y. (1988) , "Reasoning about Change : Time and Causation from the Standpoint of Artificial Intelligence" , MIT Press.

[51] WINSLETT, M. (1989) , "Sometimes Updates are Circumscription" , Proceedings of the 11th IJCAI-89 , pp. 859-863.

[52] YOU, J.H. (1989), "Supported Circumscription and Its Relation to Logic Programming With Negation", Logic Programming-Proceedings of the North American Conference , pp. 291-309.

[53] YUAN, L.Y. e WANG, C.H. (1988), "On Reducing Parallel Circumscription", Proceedings AAAI-88, pp. 460-464.

[54] RABINOV, A. (1989) , "A Generalization of Collapsible Cases of Circumscription" , Artificial Intelligence , vol. 38 , pp. 111-117.