



PALAVRAS CHAVES/KEY WORDS

AUTORES / AUTHORS: GENERALIZAÇÃO APRENDIZAGEM INDUTIVA  
REGRAS DE DECISÃO REPRESENTAÇÃO DE CONHECIMENTO  
INDUÇÃO

AUTORIZADA POR / AUTHORIZED BY

*Marcos Antonio Raupp*  
Diretor Geral

AUTOR RESPONSÁVEL / RESPONSIBLE AUTHOR

*Pedro Paulo B. de Oliveira*

DISTRIBUIÇÃO / DISTRIBUTION

INTERNA / INTERNAL  
 EXTERNA / EXTERNAL  
 RESTRITA / RESTRICTED

REVISADA POR / REVISED BY

*Edson L.F. Senne*

CDU/UDC

681.3.019

DATA / DATE

Agosto, 1987

TÍTULO/TITLE	PUBLICAÇÃO Nº / PUBLICATION NO INPE-4299-TDL/276
	GERAÇÃO DE MODELOS DE REGRAS DE DECISÃO: UMA ABORDAGEM CENTRADA NA APRENDIZAGEM INDUTIVA
AUTORES/AUTHORSHIP	Pedro Paulo Balbi de Oliveira

ORIGEM / ORIGIN

PG/LAC

PROJETO / PROJECT

FRH/CAP

Nº DE PAG. / NO OF PAGES

178

ULTIMA PAG. / LAST PAGE

B.9

VERSÃO / VERSION

---

Nº DE MAPAS / NO OF MAPS

---

RESUMO - NOTAS / ABSTRACT - NOTES

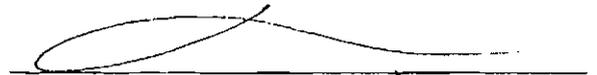
Apresenta-se uma pesquisa realizada com o objetivo de criar um mecanismo para geração de modelos (descrições) das regras de decisão constituintes de uma base de conhecimento. O mecanismo definido fundamenta-se em um processo de aprendizagem indutiva proposto e opera a partir de classes de regras definidas na base. Tais classes são formadas por regras que possuem alguma cláusula de conclusão comum e referem-se a um mesmo conjunto de objetos na premissa. O enfoque de aprendizagem adotado permite a criação de conceitos disjuntivos e conjuntivos. Prevê-se no processo que conceitos disjuntivos possam ser generalizados para um outro que os englobe, o que representa uma forma simples de aprendizagem baseada em conhecimento. Testou-se o mecanismo de aprendizagem apenas para a geração de descrições características, mas, conforme se procura mostrar, ele também poderia ser usado para a geração de descrições discriminantes. Os resultados obtidos e a própria experiência adquirida no trabalho permitem concluir que, a fim de que os modelos de regras possam melhor caracterizar o conhecimento expresso na base, são necessárias formas mais elaboradas para a definição das classes de regras.

OBSERVAÇÕES / REMARKS

Dissertação de Mestrado em Computação Aplicada, aprovada por Banca Examinadora em 11 de setembro de 1986, no Instituto de Pesquisas Espaciais - INPE.

Aprovada pela Banca Examinadora  
em cumprimento a requisito exigido  
para a obtenção do Título de Mestre  
em Computação Aplicada

Dr. Valter Rodrigues

  
Presidente

Dr. Celso de Renna e Souza

  
Orientador

Engº Edson Luiz França Senne, Mestre

  
Co-Orientador

Dr. José Armando Valente

  
Membro da Banca  
-convidado-

Dr. Paulo Ouverá Simoni

  
Membro da Banca

Candidato: Pedro Paulo Balbi de Oliveira

São José dos Campos, 11 de setembro de 1986

*"Our ultimate objective is to make  
programs that learn from experience  
as effectively as humans do."*

*John McCarthy, 1968*

*À Luara e Carmen,  
com um lamento por todo  
o transtorno que lhes  
causei.*

## AGRADECIMENTOS

Ao INPE - Instituto de Pesquisas Espaciais, pela oportunidade que me foi dada para a realização deste trabalho.

Ao *Dr. Celso* de Renna e Souza, mais que pelos comentários preciosos nas horas certas, agradeço a confiança que me foi depositada, dando-me liberdade de escolher meu próprio caminho.

Ao Edson Luiz França *Senne* pelo apoio no dia a dia e pelos erros apontados nas várias tardes que lhe roubei em nossas longas discussões.

À *Carmen* Lúcia Ruybal Santos pela crucial ajuda na edição do presente texto, por todo o apoio "logístico" nas questões familiares e pela companhia nas solitárias noites de inverno.

À IBM do Brasil pela bolsa de estudos concedida durante um período do desenvolvimento do trabalho.

Ao *Albert* Philip Close por ter permitido que eu me "refugiasse" em sua silenciosa sala durante o período mais crítico da realização deste trabalho. Ao Carlos Alberto de Oliveira "*Caverna*" pela cessão de vários de seus "preciosos" horários de utilização dos micros.

À *Mônica* Aparecida de Oliveira pelo cuidadoso serviço de datilografia, e à *Neusa* Maria Dias Bicudo pela criteriosa correção de linguagem.

Aos outros membros do INTAL (Grupo de Inteligência Artificial), que indiretamente contribuíram para este trabalho proporcionando um ambiente técnico crítico e estimulante, e aos colegas de departamento em geral, pelo ambiente amigável e agradável que sempre caracterizou o DIN.

Por fim, agradeço à *Luara*, pela sua constante alegria.

### ABSTRACT

A research for the creation of a mechanism to allow the generation of models (descriptions) of the rules which constitute a knowledge base is presented. The mechanism which was defined is based on an inductive learning process which is proposed, and works from rule classes defined in the base. Such classes are composed of rules which have some common conclusion clause and refer to a same set of objects. The learning approach allows the creation of disjunctive and conjunctive concepts. This process allows the generalization of disjunctive concepts for another that encompasses the formers, which represents a simple way of knowledge-based learning. The learning mechanism was tested only for the generation of characteristic descriptions, but, as it's tried to show, it could also be used for the generation of discriminant descriptions. The results obtained, as well as the acquired experience along the work, allow to conclude that, in order that the models may better specify the knowledge expressed in the base, more sophisticated ways of defining the rule classes are necessary.

## SUMÁRIO

	<u>Pag.</u>
LISTA DE FIGURAS .....	xv
<u>CAPÍTULO 1 - INTRODUÇÃO</u> .....	1
1.1 - Sistemas especialistas .....	1
1.2 - Aprendizagem .....	3
1.2.1 - Tipos de aprendizagem .....	4
1.3 - Organização do documento .....	6
<u>CAPÍTULO 2 - ORIGEM E OBJETIVO DO TRABALHO</u> .....	9
2.1 - O sistema DIAGNÓSTICO .....	9
2.1.1 - Introdução .....	9
2.1.2 - Representação de conhecimento utilizada .....	9
2.1.3 - A arquitetura utilizada .....	11
2.2 - O sistema TEIRESIAS .....	12
2.2.1 - Os modelos de regras .....	14
<u>CAPÍTULO 3 - APRENDIZAGEM INDUTIVA</u> .....	21
3.1 - Conceituação .....	21
3.1.1 - A aprendizagem indutiva como um processo de solução de problemas .....	21
3.1.2 - Estratégias de busca no espaço de generalizações .....	23
3.2 - Importância .....	25
3.3 - Os dois tipos: indução através de exemplos e indução através de observações .....	27
3.4 - O papel dos exemplos e dos contra-exemplos .....	28
3.5 - O papel da representação adotada: a influência do "como" representar .....	30
3.6 - O papel do conhecimento expresso: a influência de "o que" representar .....	31
3.7 - As regras sintáticas de generalização .....	33
3.7.1 - Introdução .....	33
3.7.2 - Regras seletivas de generalização .....	34
3.8 - O ponto de vista psicológico .....	42

	<u>Pág.</u>
<u>CAPÍTULO 4 - METODOLOGIAS DE APRENDIZAGEM INDUTIVA</u> .....	45
4.1 - Introdução .....	45
4.2 - O sistema SPROUTER .....	46
4.3 - Outros sistemas .....	51
<u>CAPÍTULO 5 - GERAÇÃO DE MODELOS DE REGRAS: A SOLUÇÃO PROPOSTA</u> ..	57
5.1 - A evolução do trabalho: do reconhecimento de padrões iso- lados à aprendizagem indutiva .....	57
5.1.1 - O mecanismo principal da primeira solução adotada .....	57
5.1.2 - Os testes feitos e a conclusão a que se chegou .....	58
5.1.3 - Os direcionamentos iniciais já no enfoque de aprendizagem indutiva .....	63
5.2 - Alguns aspectos sobre a representação das regras e modelos	65
5.2.1 - Modificação da representação em quádruplas .....	65
5.2.2 - As cláusulas relacionais parciais .....	69
5.2.3 - As cláusulas relacionais inversas .....	70
5.2.4 - O refinamento da representação .....	72
5.3 - Particionamento da base: geração das classes de exemplos	73
5.4 - Geração dos modelos de regras: o modelo de aprendizagem in- dutiva proposto .....	76
5.4.1 - Introdução .....	76
5.4.2 - Fase 1: formação das associações não-redundantes de obje- tos .....	78
5.4.3 - Fase 2: seleção das associações e conjuntos consistentes de máximo contexto .....	83
5.4.4 - Fase 3: geração das descrições generalizadas .....	91
5.4.5 - Fase 4: suplementação das descrições generalizadas .....	95
5.4.6 - Fase 5: generalização de conceitos disjuntivos .....	96
5.4.7 - Alternativa à fase 2: seleção das melhores associações quando são existem associações parciais no quadro de asso- ciações .....	103
5.4.8 - Justificando a geração das descrições generalizadas a partir dos conjuntos consistentes de associações .....	112

	<u>Pág.</u>
<u>CAPÍTULO 6 - OUTROS ASPECTOS DOS MODELOS DE REGRAS</u> .....	117
6.1 - Aplicação na aquisição de novas regras .....	117
6.2 - Os testes realizados .....	122
6.3 - Uma palavra sobre a geração de descrições discriminantes	127
<u>CAPÍTULO 7 - COMENTÁRIOS FINAIS</u> .....	131
REFERÊNCIAS BIBLIOGRÁFICAS .....	137
APÊNDICE A - APLICAÇÃO DO MECANISMO DE GERAÇÃO DE MODELOS DE REGRAS EM UM DOMÍNIO DA METEOROLOGIA	
APÊNDICE B - APLICAÇÃO DO MECANISMO DE APRENDIZAGEM INDUTIVA EM UM DOMÍNIO FICTÍCIO DA BIOLOGIA: GERAÇÃO DE DESCRIÇÕES CARACTERÍSTICAS DE CÉLULAS CANCEROSAS	

## LISTA DE FIGURAS

	<u>Pág.</u>
2.1 - Arquitetura do sistema DIAGNÓSTICO .....	11
2.2 - Organização dos modelos de regras do TEIRESIAS .....	15
2.3 - Exemplo de modelo de regras gerado pelo TEIRESIAS .....	16
3.1 - A aprendizagem indutiva como processo de solução de problemas: o modelo dos dois espaços .....	22
3.2 - Esquema de generalização dos métodos guiados pelas instâncias .....	24
3.3 - Exemplo de hierarquização de conceitos .....	36
3.4 - Exemplo de regra de expansão limitada no domínio .....	40
4.1 - Instâncias para ilustrar os esquemas de representação do SPROUTER .....	46
4.2 - Instâncias para comparar o desempenho dos métodos .....	46
4.3 - Busca realizada pelo SPROUTER para a obtenção das melhores associações de parâmetros, segundo as instâncias mostradas na Figura 4.2 .....	48
5.1 - Figuras geométricas que definem o primeiro conjunto de exemplos .....	59
5.2 - Figuras geométricas que definem o segundo conjunto de exemplos .....	66
5.3 - Esquema simplificado das fases do mecanismo de aprendizagem indutiva proposto .....	77
5.4 - Árvores de associações para as instâncias da Figura 4.2 ..	80
5.5 - Quadro de caracterização dos objetos presentes à Figura 4.2	82
5.6 - Árvores representativas das associações de máximo contexto para as instâncias da Figura 4.2 .....	89
5.7 - Representação da busca às associações consistentes relativas às instâncias da Figura 4.2 .....	92
5.8 - Quarta instância positiva para o conjunto de instâncias da Figura 4.2 .....	104
5.9 - Árvores de associações para as instâncias das Figuras 4.2 e 5.8 .....	106
5.10 - Representação do "grupo-base" de associações parciais para o atributo "forma", com o qual serão obtidas as cláusulas de menor número de termos disjuntivos .....	108

	<u>Pág.</u>
6.1 - Representação gráfica de uma regra supostamente sendo adquirida, para ilustrar a utilização dos modelos de regras na aquisição de conhecimento .....	119
6.2 - Conjunto de exemplos de células cancerosas · fictícias (Michalski, 1983) .....	125

## CAPÍTULO 1

### INTRODUÇÃO

#### 1.1 - SISTEMAS ESPECIALISTAS

A Inteligência Artificial é a área da Ciência da Computação que se preocupa com a criação de programas para realizar tarefas que demandam inteligência e, por isto, suas execuções têm sido vistas como atribuição exclusiva do homem. Dentre os tópicos mais abordados nesta área de pesquisa, destaca-se o estudo dos SISTEMAS ESPECIALISTAS, os quais têm produzido resultados práticos bastante úteis.

Os Sistemas Especialistas são programas computacionais que atuam como um especialista humano em uma determinada área de conhecimento. São essencialmente programas baseados em conhecimento específico de um domínio (ao invés de se proporem a tarefas gerais) e usualmente utilizam algum meta-nível de conhecimento, ou seja, conhecimento a respeito do conhecimento que ele possui, o que lhe permite justificar ou explicar as razões que o levaram a uma determinada conclusão ou ação.

Os sistemas especialistas têm sido construídos para realizar uma série de tarefas distintas (Hayes-Roth et alii, 1983):

INTERPRETAÇÃO: Análise de dados para determinar seu significado.

PREDIÇÃO: Análise de prováveis conseqüências de certas situações.

DIAGNOSE: Detecção de falhas de funcionamento em um sistema.

PROJETO: Configurações de objetos respeitando certas restrições.

PLANEJAMENTO: Projeto de uma seqüência de ações para atingir um fim.

MONITORAMENTO: Acompanhamento contínuo de um conjunto de informações.

DEPURAÇÃO: Prescrição de ações para eliminar falhas em um sistema.

REPARO: Execução de um plano a fim de eliminar uma falha.

INSTRUÇÃO: Diagnose, depuração e reparo no comportamento de estudantes.

CONTROLE: Interpretação, predição, reparo e monitoramento do comportamento de um sistema genérico.

Muitas têm sido as arquiteturas propostas para sistemas especialistas (Stefik et alii, 1982). Apesar de não haver uma arquitetura geral, pode-se identificar, no entanto, três unidades básicas: uma *Base de Conhecimento* composta dos fatos e heurísticas do domínio de aplicação; uma *Estrutura de Controle* que utiliza a base de conhecimento de alguma maneira, a fim de resolver o problema em questão; e uma *Memória de Trabalho* para acompanhar o status do programa, os dados de entrada para o problema específico e o histórico relevante do que foi feito até então. Fica claro neste esquema a total separação existente entre os dados (conhecimento) e o programa (estrutura de controle), que é a principal diferença entre os sistemas especialistas e os sistemas computacionais convencionais.

Estima-se que existam mais de 400 (Sandri, 1986) sistemas especialistas atualmente em uso no mundo, nas mais diversas áreas do conhecimento humano, tais como: Medicina, Agricultura, Genética, Economia, Simulação, Controle de Tráfego Aéreo, Prospecção Petrolífera e Mineral, Gerenciamento de Produção, Reconhecimento de Linguagem Escrita e Falada, entre muitas outras. Entre os sistemas mais famosos podem ser citados o DENDRAL (Buchanan and Feigenbaum, 1978) que infere a estrutura de compostos químicos desconhecidos a partir de dados de espectrografia de massa e ressonância magnética desses compostos; o PROSPECTOR (Duda et alii, 1978) que auxilia a prospecção mineral; o

MYCIN (Shortliffe, 1976) que faz diagnose de infecções bacteriológicas; e o HEARSAY-III (Erman et alii, 1980) que visa o entendimento de linguagem falada.

Considerando o enorme esforço necessário à construção de um sistema especialista, têm sido criadas ferramentas para auxiliar esta tarefa. Dois tipos principais têm sido desenvolvidos: *linguagens de propósito geral para processamento de conhecimento*, tais como a OPS5 (Forgy, 1981) e a RLL (Greiner and Lenat, 1980), e *sistemas do tipo arcabouço*, os quais, tendo sido na maioria dos casos derivados de sistemas especialistas já construídos, incorporam as principais características dos sistemas originais, torandó-se ambientes gerais de desenvolvimento. São mais restritivos que as linguagens de propósito geral, pois estas podem ser utilizadas com uma faixa mais ampla de estruturas de controle. Entre os mais conhecidos sistemas do tipo de arcabouço podem ser citados o EMYCIN (Van Melle, 1979) que foi derivado do MYCIN, o KAS (Duda et alii, 1978) derivado do PROSPECTOR, e o EXPERT (Weiss and Kulikowsky, 1979) que foi construído a partir do CASNET (um sistema para diagnose de casos de glaucoma).

## 1.2 - APRENDIZAGEM

A *aprendizagem* pode ser definida como um processo pelo qual uma pessoa (ou um computador) aumenta o seu conhecimento factual ou procedimental a respeito de determinada situação, de forma que o seu desempenho possa ser aperfeiçoado em alguma tarefa definida neste contexto, ou em outro a ele relacionado.

Considerando que o paradigma atual nas pesquisas em *Inteligência Artificial* são os sistemas baseados em conhecimento, e observando a profundidade com que ele se instalou, é extremamente difícil supor que esta situação possa ser alterada num horizonte de tempo visualizável. Tal fato indica a vital importância de estudar processos de aquisição, representação e utilização do conhecimento. Dentre os três, acredita o autor que os processos de aquisição são os menos

conhecidos, além de constituírem um "gargalo" mesmo para as aplicações de baixa escala; desde que tenha sido obtida uma base de conhecimento, é possível conseguir uma série de aplicações, mesmo com um processo simples e unidirecional de raciocínio operando em esquema adequado de representação.

Vários são os fatores que podem ser apontados como condiçionantes da capacidade de aprendizagem de um sistema (Dietterich, 1983). Os principais são os que relacionam ao meio ambiente onde o sistema se insere; neste contexto, deve-se ressaltar o nível de generalidade das informações que são apresentadas ao sistema, bem como o da qualidade delas, ou seja, quão "didaticamente" elas são apresentadas e com que grau de confiabilidade. A forma e o conteúdo da base de conhecimento também são fatores de influência no processo de aprendizagem; estes aspectos serão discutidos com detalhes nas Seções 3.5 e 3.6. A tarefa que vai ser realizada com o conhecimento adquirido também levanta aspectos que devem ser considerados. A complexidade da tarefa é um desses aspectos, pois, em geral, quanto maior a complexidade da tarefa, mais conhecimento deverá ser necessário, além de tornar-se mais problemática a integração, na base original, do novo conhecimento adquirido. A seguir são apresentados os tipos de aprendizagem, onde se aprofunda a influência do meio ambiente.

#### 1.2.1 - TIPOS DE APRENDIZAGEM

Conforme já foi adiantado na seção anterior, o meio ambiente é o principal responsável pela capacidade de aprendizagem de um sistema. O que se verifica é que, de acordo com o grau de generalidade da informação apresentada ao sistema, pode-se identificar quatro tipos básicos de aprendizagem: por memorização, por dedução, por indução e por analogia (Charniak and McDermott, 1985; Dietterich, 1982).

A Aprendizagem por Memorização é aquela em que o meio fornece informações exatamente ao nível necessário à tarefa que se vai

realizar, ou seja, o sistema tem que simplesmente guardar a informação. Vale observar que todo computador tem este tipo de aprendizagem na medida em que ele armazena instruções para uma determinada tarefa.

Quando a informação fornecida pelo meio é muito geral ou abstrata para a tarefa em questão, o sistema terá de operacionalizar esta informação instanciando-a (dedutivamente) a fim de que ela possa lhe ser útil. Tal situação caracteriza o processo de *Aprendizagem por Dedução*. A pesquisa nesta situação de aprendizagem tem seguido dois caminhos principais. O primeiro tenta efetivamente realizar o que foi caracterizado acima, ou seja, a automatização do processo de operacionalização de um conhecimento geral. O outro caminho tem sido no desenvolvimento de ferramentas sofisticadas (que incluem facilidades de edição e depuração da base), as quais facilitam o trabalho do especialista no que diz respeito à transformação de seu conhecimento em um certo domínio, em regras que possam efetivamente ser utilizadas pelo sistema. Nesta segunda abordagem o especialista é uma parte integrante do sistema de aprendizagem, à medida que ele detecta e diagnostica falhas na base, além de dar-lhe manutenção e refiná-la; o sistema TEIRESIAS que será apresentado na Seção 2.2 implementa este tipo de abordagem. Neste último enfoque, o aspecto dedutivo se manifesta no sentido em que é o sistema que guia o processo de aprendizagem fazendo requisições ou pedindo sugestões ao usuário, de acordo com informações concluídas dedutivamente a partir de um conjunto de informações originais. Algumas vezes este processo de aprendizagem é referido como *Aprendizagem por Aconselhamento*.

A *Aprendizagem por Analogia* refere-se àquela em que a informação provida pelo meio é relevante somente para uma tarefa análoga da situação real, o que impõe que o sistema seja capaz de descobrir a analogia existente entre elas, de modo a poder utilizar a informação dada. Poucos trabalhos têm sido desenvolvidos sobre este tipo de aprendizagem; vale citar, no entanto, o sistema ANA, conforme descrito em McDermott (1979).

Finalmente, a *Aprendizagem Indutiva* é aquela em que a informação dada pelo meio é específica demais para uma tarefa, de forma que o sistema de aprendizagem deve ser capaz de generalizar a informação fornecida. Além de este ser o principal processo de aprendizagem, é justamente o que interessa para o presente trabalho; assim ele merece um capítulo especial (Capítulo 3). Vale observar que a aprendizagem por analogia pode ser vista como um processo dependente de uma capacidade indutiva, na medida em que se pode supor que o conhecimento que vai ser aprendido deve ser "mapeado" a uma abstração do conhecimento já conhecido, e não a ele diretamente.

### 1.3 - ORGANIZAÇÃO DO DOCUMENTO

Após o Capítulo 1, onde é feita a localização do leitor com relação ao contexto mais amplo do trabalho, que são os conceitos de Sistema Especialista e de Aprendizagem, passa-se à descrição, já no Capítulo 2, dos sistemas DIAGNÓSTICO e TEIRESIAS, os quais permitem deixar claro as origens e os objetivos do trabalho.

Tendo adotado a perspectiva de um mecanismo de aprendizagem indutiva, descreve-se no Capítulo 3 uma série de aspectos concernentes a este tipo de aprendizagem. No Capítulo 4 são apresentados alguns sistemas de aprendizagem indutiva, dando-se ênfase a dois deles que mais influenciaram o presente trabalho, quais sejam, os sistemas SPROUTER e INDUCE.

No Capítulo 5 apresenta-se detalhadamente as cinco fases que compõem o mecanismo de aprendizagem desenvolvido, as quais são ilustradas com um exemplo, cujo resultado é utilizado para comparar o seu desempenho com outros modelos descritos na literatura. No Capítulo 6 discutem-se outros aspectos relativos ao modelo, tais como a possibilidade de sua utilização no auxílio a um processo de aquisição de conhecimento, e sua capacidade de gerar descrições discriminantes (generalizações que consideram tanto as instâncias positivas quanto as negativas). Além disto, são apresentados dois testes realizados com o modelo proposto.

Finalmente, no Capítulo 7 são apresentadas as conclusões a respeito da experiência adquirida, apontando direções para o prolongamento do trabalho.

## CAPÍTULO 2

### ORIGEM E OBJETIVO DO TRABALHO

#### 2.1 - O SISTEMA DIAGNÓSTICO

##### 2.1.1 - INTRODUÇÃO

O Projeto Diagnóstico é um projeto do INPE/SJC que visa a pesquisa e o desenvolvimento de ferramentas para a construção de sistemas especialistas. A primeira ferramenta concluída foi o sistema DIAGNÓSTICO (Sandri, 1986), o qual é baseado no sistema MYCIN (Shortliffe, 1976). A atual versão do DIAGNÓSTICO pretende ser uma expansão da versão anterior e já está sendo desenvolvida. A razão de descrever este sistema é que o trabalho de pesquisa que será apresentado neste documento surgiu do desejo de construir um sistema de aquisição de conhecimento mais elaborado que o atual; na busca de outras abordagens para este problema, deparou-se com o sistema TEIRESIAS que possui um interessante módulo de aquisição baseado fortemente em modelos descritivos das regras da base. No entanto, sua abordagem para a geração dos modelos é muito rudimentar, donde se lançou a tarefa de um mecanismo que permitisse gerar modelos que pudessem refletir mais profundamente o conhecimento expresso na base.

##### 2.1.2 - REPRESENTAÇÃO DE CONHECIMENTO UTILIZADA

A representação de conhecimento adotada no sistema DIAGNÓSTICO são as Regras de Produção (ou Decisão) do tipo:

SE: uma premissa é verdadeira

ENTÃO: tira-se uma ou mais conclusões, cada uma com um certo grau de crença associado.

A representação interna das regras faz-se a partir de unidades de representação chamadas *cláusulas simples*, as quais, por

sua vez, são formadas por uma quádrupla de *primitivas conceituais* que levam os nomes de atributo, objeto, função de predicado e valor de atributo (Shortliffe, 1976). O *Objeto* representa a entidade de quem se fala na cláusula; o *Atributo* refere-se a uma propriedade do objeto, propriedade esta cujo valor é alocado na primitiva *Valor de Atributo*. Finalmente, a *Função de Predicado* refere-se à relação existente entre o atributo e seu valor. Cada cláusula é, portanto, uma quádrupla do tipo:

(Atributo Objeto Função Valor-de-Atributo).

Por exemplo, a cláusula em Português:

"Há uma zona-frontal localizada na região sul.",

poderia ser representada da seguinte forma:

(localização zona-frontal é região-sul).

A premissa de uma regra é formada por uma conjunção de unidades que podem ser constituídas por uma única cláusula : simples (por exemplo, C1), ou uma disjunção delas (por exemplo, C1 v C2 v C3). A conclusão é formada apenas por uma conjunção de cláusulas simples. Um exemplo de regra válida é a mostrada a seguir:

SE : (Há uma zona frontal no Rio Grande do Sul) &  
(Velocidade da zona frontal é 20 km/h) &  
((Intensidade da zona frontal está diminuindo) ou  
(Variação de velocidade da zona frontal está diminuindo)) &  
(Condições em altos níveis em São José dos Campos são favoráveis)

ENTÃO: (É definitivo que as condições de grande escala em 24 horas em São José dos Campos são favoráveis).

É conveniente definir uma *cláusula disjuntiva* como aquela que apresenta uma disjunção de conceitos relativos a uma mesma primitiva conceitual. Este tipo de disjunção será chamada *disjunção interna*. Por outro lado, daqui para frente, uma disjunção que envolva cláusulas (simples ou disjuntivas) será referenciada como uma *disjunção externa*. Nestes termos, pode-se dizer que a premissa de uma regra admite tanto cláusulas simples quanto cláusulas disjuntivas, enquanto a conclusão admite a existência de cláusulas simples.

### 2.1.3 - ARQUITETURA UTILIZADA

O sistema Diagnóstico é constituído de quatro módulos: Inicialização, Aquisição, Consulta e Explanação. Os dois primeiros são responsáveis pela aquisição de conhecimento em um dado domínio; por intermédio destes módulos o especialista constrói a base de conhecimento sobre a qual operarão os outros dois. O módulo de Consulta constitui a máquina de inferência do sistema e o módulo de Explanação é o encarregado de fornecer informações ao usuário a respeito da consulta que foi ou que esteja sendo realizada. A arquitetura básica do sistema é mostrada na Figura 2.1.

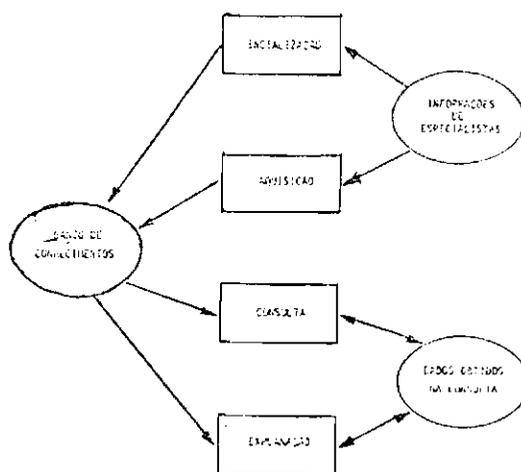


Fig. 2.1 - Arquitetura do sistema DIAGNÓSTICO.

FONTE: Sandri (1986).

O módulo de Inicialização tem o objetivo de inicializar todos os arquivos do sistema para um dado domínio de aplicação, bem como fazer a aquisição de palavras-chave, as quais são utilizadas como referência à interpretação das cláusulas em Português, para a representação interna. Para tal existe uma série de informações a respeito de cada palavra-chave, tais como: a primitiva conceitual que lhe vai ser associada; qual o tipo (lógico, de valor único ou de valores múltiplos) de atributo que ela representa, caso seja um atributo; etc.

Encerrado o processo de inicialização, através do módulo de Aquisição passa-se a adquirir as regras da base, cláusula a cláusula. Pela identificação das palavras-chaves de cada cláusula submetida pelo especialista, inicia-se um processo interativo que visa obter informações como as citadas acima, além de informações sobre as próprias regras (por exemplo, se a regra que se está recebendo é comum ou é uma meta-regra). Vale dizer que este processo é mais "amigável" que o anterior. Além disto, tem-se a capacidade de edição de regras e cláusulas. Não é possível fazer edição das palavras-chaves, o que só é permitido no módulo de inicialização.

É importante a observar com o que foi dito acima, é que todo o processo de aquisição de conhecimento é essencialmente um processo de *interpretação* das regras em Português para a representação interna em quádruplas. O sistema não possui qualquer facilidade de guia-gem da aquisição em termos do conhecimento já disponível, ou seja, o sistema não possui mecanismos para facilitar o refinamento do conhecimento já presente na base, e nem para fazer alguma sugestão a respeito de que regras deveriam ser fornecidas a fim de suprir deficiências da base. Este é um dos problemas que levaram à construção do sistema TEIRESIAS, cuja solução adotada será discutida a seguir.

## 2.2 - O SISTEMA TEIRESIAS

O sistema TEIRESIAS (Davis, 1978 e 1982) foi desenvolvido com o intuito de explorar aplicações de meta-níveis de conhecimento

em um sistema especialista, em outras palavras, explorar "o conhecimento que se possa ter a respeito do conhecimento expresso na base" de um sistema especialista. Originalmente, este sistema foi construído sobre o sistema MYCIN, apesar de não ser restrito ao domínio deste último, tendo sido testado com outras bases de conhecimento.

O TEIRESIAS deve ser considerado um sistema que procurou suplantiar uma série de deficiências do MYCIN em seus três módulos básicos (aquisição, consulta e explanação). Com relação ao módulo de inferência do MYCIN, sua contribuição básica foi a criação do conceito de uma *meta-regra*, a qual representa uma regra a respeito de regras da base. O status de uma meta-regra é de uma estratégia de raciocínio em que passa a ser possível eliminar alguns caminhos da árvore de raciocínio em uma determinada consulta, tornando o processo de inferência mais rápido. Uma interessante contribuição do TEIRESIAS no processo de explanação foi a de permitir explicações com níveis diferentes de detalhamento, de acordo com o que era requisitado pelo usuário.

Sem dúvida, no entanto, foi no que diz respeito à aquisição de conhecimento que as maiores contribuições foram feitas. Basicamente três aspectos foram explorados neste sentido: a interpretação de uma regra em Inglês para a representação em quádruplas; a aquisição de novas regras e a aquisição de novas primitivas conceituais. Esta última característica refere-se à possibilidade de expandir a capacidade de representação de conhecimento do sistema, através de um mecanismo que permite definir outras primitivas conceituais, além das quatro originais.

Quanto aos outros dois aspectos citados, eles são possíveis principalmente pela geração de *Modelos das Regras* da base. No momento, é suficiente dizer que tais modelos constituem uma caracterização da regra típica de um conjunto de regras da base. Assim, com estes modelos, o TEIRESIAS passa a ter uma estrutura conceitual que lhe permite avaliar as várias interpretações possíveis de uma cláusula e

selecionar a mais adequada. Além disto, os modelos se prestam como referência básica à aquisição de regras em dois aspectos: como indicadores de possíveis carências (em termos de regras) na base, e como indicadores de possíveis cláusulas que possam estar faltando em uma regra. Desta forma, tem-se um mecanismo de aquisição com um certo grau de guiagem do sistema. A seguir, serão detalhados alguns pontos sobre a geração dos modelos de regras.

### 2.2.1 - OS MODELOS DE REGRAS

Como já foi adiantado, os Modelos de Regras representam caracterizações da regra típica de um conjunto. Eles constituem, portanto, uma forma de conceito generalizado a partir do conjunto de regras em questão. O processo de geração dos modelos é então um mecanismo de *formação de conceitos*, se bem que muito simplificado, pois como será visto em seguida, este processo é essencialmente estatístico.

Os modelos de regras gerados pelo TEIRESIAS são compostos por quatro partes:

EXEMPLOS: o subconjunto de regras (com seus respectivos fatores de certeza (FC)) que o modelo descreve.

DESCRIÇÃO: caracterização de um membro típico deste subconjunto:

- caracterização da premissa;
- caracterização da conclusão:
  - . quais atributos aparecem tipicamente,
  - . correlações de atributos.

MAIS-GERAL: ponteiros para modelos mais gerais

MAIS-ESPECÍFICO: ponteiros para modelos mais específicos.

Já que as regras são pares do tipo premissa-conclusão, a DESCRIÇÃO contém características individuais de uma premissa e de uma conclusão típicas. Isto é feito através da especificação dos atributos e das correlações de atributos que aparecem tipicamente na premissa (ou conclusão) das regras em EXEMPLOS.

Os modelos são organizados em torno das cláusulas de ação, segundo a árvore mostrada na Figura 2.2. O modelo mais geral é o representado pela raiz da árvore e os mais específicos, pelas folhas. A razão para tal encontra-se em Davis (1982), onde é citada uma experiência informal realizada com as regras do sistema MYCIN, em que foi pedido a três médicos que, individualmente, as agrupassem em "classes de similaridade" segundo o critério que julgassem mais adequado. O resultado foi que todos eles se utilizaram das conclusões das regras como referência à classificação, e acabaram gerando praticamente as mesmas classes (exceto que, eventualmente, alguma classe era subdividida). Apenas 6 das 300 regras por eles recebidas provocaram maiores divergências sobre a que classe deveriam pertencer. A organização que se utilizou no TEIRESIAS foi sugerida pelo experimento e pode ser representada pela árvore a seguir.

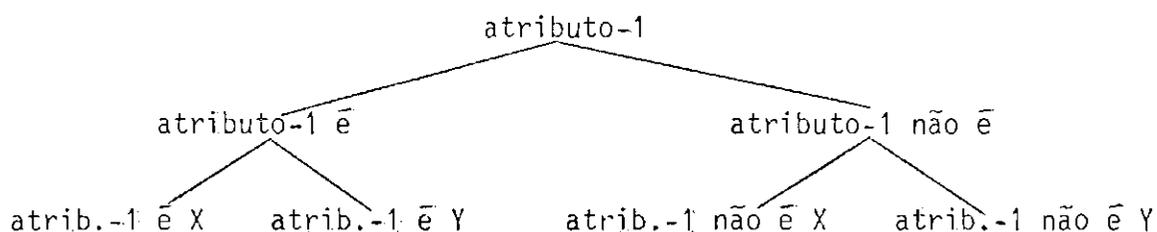


Fig. 2.2 - Organização dos modelos de regras do TEIRESIAS.

Ou seja, as regras são agrupadas segundo um atributo comum numa cláusula de conclusão, e subdivididas em dois níveis: no primeiro encontram-se aquelas que concluem positiva ou negativamente a respeito de algum

valor desse atributo, e no segundo nível estão as que concluem de forma positiva ou negativa a respeito de valores específicos (por exemplo, X e Y) do atributo.

A Figura 2.3 mostra um exemplo de modelo criado a partir de regras do MYCIN.

CATEGORY-IS

```
EXEMPLES      ((Rule116 .33) (Rule050 .70) (Rule037 .80) (Rule095 .90)
               (Rule152 1.0) (Rule140 1.0))

P-ADVICE      ((GRAM SAME NOTSAME 3.83)
               (MORPH SAME NOTSAME 3.83)
               ((GRAM SAME) (MORPH SAME) 3.83)
               ((MORPH SAME) (GRAM SAME) 3.83)
               ((AIR SAME) (NOSOCOMIAL NOTSAME SAME) (MORPH SAME)
               (GRAM SAME) 1.50
               ((NOSOCOMIAL NOTSAME SAME) (AIR SAME) (MORPH SAME)
               (GRAM SAME) 1.50)
               ((INFECTION SAME) (SITE MEMBF SAME) 1.23)
               ((SITE MEMBF SAME) (INSFECTION SAME) (PORTAL SAME) 1.23))

A-ADVICE      ((CATEGORY CONCLUDE 4.73)
               (IDENT CONCLUDE 4.05)
               (CATEGORY CONCLUDE) (IDENT CONCLUDE) 4.73))

MORE-GENL     (CATEGORY-MOD)

MORE-SPEC     NIL
```

Fig. 2.3 - Exemplo de modelo de regra gerado pelo TEIRESIAS.

FONTE: Davis (1982).

No exemplo acima podem ser identificadas duas estruturas: uma *lista simples* que indica os atributos típicos, e uma *n-upla* que referencia as correlações de atributos. As listas simples têm a forma geral

( <atributo> <função><sup>+</sup> <soma-FC> ),

onde o sinal de + "significa "1 ou mais vezes", e "soma-FC" significa a soma dos fatores de certeza associados às regras em EXEMPLOS. No exemplo aparece a lista simples

(GRAM SAME NOTSAME 3.83).

que indica que o atributo GRAM aparece tipicamente e é freqüentemente associado às funções SAME e NOTSAME. O número 3.83 indica a soma dos FCs das regras do conjunto que possuem o atributo GRAM na premissa (no caso).

As N-uplas têm a forma geral

(( <atributo> <função><sup>+</sup> )<sup>+</sup> <soma-FC> ),

sendo que, no exemplo aparece

((GRAM SAME) (MORPH SAME) 3.83),

indicando que, quando o atributo GRAM aparece na premissa (no caso) de uma regra, tipicamente também aparece o atributo MORPH.

Uma lista simples é construída para todo atributo que aparece na premissa ou conclusão de, pelo menos, 30% (este é um limiar empírico, bem como todos os outros que ainda aparecerão) das regras do conjunto. Um número suficiente de funções é então determinado de modo a cobrir, pelo menos, 75% das ocorrências do atributo.

Uma N-upla é construída quando os atributos correlacionados aparecem juntos em, pelo menos, 80% das regras do conjunto e a implicação é unidirecional. Por exemplo, a correlação expressa por ((GRAM SAME) (MORPH SAME) 3.83) indica que, pelo menos, 80% das vezes em que GRAM aparece numa regra, MORPH também aparece. Se o contrário

for verdade (e este é o caso), deve-se construir a N-upla ((MORPH SAME) (GRAM SAME) 3.83).

A principal crítica que se pode fazer ao esquema de formação de conceito acima diz respeito à própria natureza da abordagem utilizada; o enfoque estatístico que foi dado, além de ser conceitualmente fraco para este tipo de tarefa, é totalmente arbitrário, nada havendo que justifique sua aplicabilidade em outras situações. De fato esta característica dá margem a toda sorte de críticas que se queira fazer em termos de total falta de tratamento de uma série de aspectos referentes à formação de conceitos, como é o caso da detecção de aspectos estruturais entre os objetos da base, capacidade de tratamento de contra-exemplos etc.

Outra grave restrição pode ser constatada observando que a definição dos modelos não faz menção aos objetos das regras. Isto decorre do fato de que, para o domínio onde o TEIRESIAS foi desenvolvido, cada atributo está relacionado a apenas 1 objeto. Evidentemente este aspecto constitui uma forte limitação à sua utilização em domínios onde um objeto pode ter vários atributos associados, situação esta que corresponde à maioria das aplicações.

Apesar de outras críticas serem possíveis, não vale a pena insistir nelas, mesmo porque Davis (1982) deixa claro que seu objetivo não era a formação de conceitos em si, mas apenas utilizá-la como um meio de facilitar o processo de aquisição de conhecimentos. Apesar do mecanismo criado ter sido bastante rudimentar, o grande mérito conceitual do sistema TEIRESIAS, neste particular, foi o de ter conseguido mostrar a viabilidade da utilização de modelos de regras como forma de melhorar o processo de aquisição de conhecimento.

Dentro dessa perspectiva definiu-se o presente trabalho no sentido de *explorar uma forma conceitualmente mais relevante de formação de conceitos, de forma que os modelos eventualmente obtidos pudessem refletir de maneira mais realista a natureza das regras expres*

*sas na base de conhecimento.* Embora o objetivo deste processo seja a aquisição de conhecimento, adotou-se essencialmente a direção apontada acima, não se tendo explorado técnicas de utilização dos modelos no processo em si de aquisição de conhecimento.

## CAPÍTULO 3

### APRENDIZAGEM INDUTIVA

#### 3.1 - CONCEITUAÇÃO

Aprendizagem indutiva é o processo de aquisição de conhecimento em que através de fatos fornecidos direta ou indiretamente por um professor, ou pelo próprio ambiente, abstrai-se um ou mais conceitos que não estão presentes aos fatos originais.

Mais formalmente (Michalski, 1983) pode-se dizer que, a partir de um conjunto F de fatos, determina-se uma hipótese H que implica as observações realizadas (F), fracamente ou tautologicamente. Uma hipótese H implica tautologicamente F, se F é uma seqüência lógica de H, isto é, se a expressão " $H \rightarrow F$ " é sempre verdadeira. Assim, se  $H \rightarrow F$  é válido e H é verdadeiro, então pelo "modus ponens" F também é verdadeiro. Portanto, na obtenção de F a partir de H (inferência dedutiva), há preservação do valor-verdade Verdadeiro. Por outro lado, obter H a partir de F (inferência indutiva) não preserva a verdade, mas o valor-verdade Falso, ou seja, se algum fato contradiz F, então ele também contradiz H.

A condição de que H pode implicar fracamente F significa que os fatos não são certos, mas apenas conseqüências plausíveis ou parciais de H. É o caso, por exemplo, de hipóteses que representam "padrões dominantes" ou que caracterizam dados "ruidosos". Neste trabalho, somente será tratado o caso de implicação tautológica.

#### 3.1.1 - A APRENDIZAGEM INDUTIVA COMO UM PROCESSO DE SOLUÇÃO DE PROBLEMAS

É conceitualmente interessante observar que o processo de aprendizagem indutiva pode ser visto essencialmente como um processo de solução de problemas. Tal idéia foi apresentada por Simon e Lea

(1974) e merece ser aprofundada neste momento. Vale observar que os autores fazem referência à aprendizagem através de exemplos como se esta fosse a única forma de aprendizagem indutiva, o que não é verdade conforme será visto na Seção 3.3; no entanto, sua análise continua válida mesmo considerando o segundo tipo (aprendizagem através de observação).

No trabalho acima citado, descreve-se o problema de aprendizagem indutiva como um processo de utilização de instâncias de treinamento selecionadas de algum espaço de instâncias possíveis, para guiar uma busca de regras gerais. Ou seja, a partir de elementos em um *Espaço de Instâncias*, procede-se a uma busca em um *Espaço de Regras*. Além disto Simon e Lea argumentam que, para resolver ambigüidades sobre regras no espaço de regras, um programa inteligente deveria selecionar suas próprias instâncias de treinamento através de uma busca no espaço de instâncias. Assim, se o programa estivesse em dúvida quanto à validade da regra "todo verão em São Paulo é chuvoso", ele deveria proceder a uma busca no espaço de instâncias dos "dados climáticos de São Paulo", procurando encontrar referências, por exemplo, a baixos "índices pluviométricos nos meses de verão". Segundo o que foi explicado, os problemas que devem ser resolvidos são, portanto, a determinação das regras que generalizam um conjunto de instâncias e a das instâncias que eliminam a incerteza sobre a validade de uma regra. A Figura 3.1 ilustra o modelo de dois espaços para a aprendizagem indutiva.



Fig. 3.1 - A aprendizagem indutiva como processo de solução de problemas: o modelo dos dois espaços.

O esquema anterior apresenta ainda os processos de "interpretação" e de "seleção de instâncias (ou planejamento de experimento)". O primeiro refere-se ao processo de adequação das descrições das instâncias para uma forma mais relevante à busca no espaço de regras, nas situações em que a forma das descrições dos dois espaços sejam diferentes. Finalmente, o segundo processo refere-se à utilização de procedimentos com o objetivo de gerar um conjunto de instâncias a partir da hipótese (regra) de que se disponha, a fim de haver mais dados para avaliá-la.

### 3.1.2 - ESTRATÉGIAS DE BUSCA NO ESPAÇO DE GENERALIZAÇÕES

Tendo em mente que a aprendizagem indutiva pode ser vista como um processo de solução de problemas, pode-se dividir os métodos de aprendizagem em duas classes distintas, de acordo com o tipo de busca que se faça no espaço de generalizações. A primeira classe refere-se aos sistemas que são *dirigidos pelos dados*, isto é, pelas instâncias; a outra classe refere-se aos que são *dirigidos por modelos* (pode-se refinar ainda mais essa taxonomia, conforme é feito em Mitchell (1982)).

Os métodos dirigidos pelas instâncias (ou seja, os que operam de "baixo para cima") obtêm as generalizações tomando as instâncias duas a duas, como indica a Figura 3.2. No presente trabalho o esquema de generalização é diferente, apesar de ainda ser guiado pelas instâncias; neste modelo, cada generalização é obtida diretamente do conjunto de instâncias, sem se obter as generalizações intermediárias.

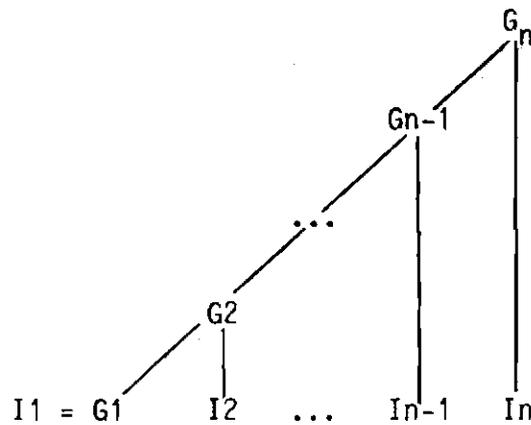


Fig. 3.2 - Esquema de generalização dos métodos guiados pelas instâncias.

No esquema acima,  $G_2$  é a generalização das instâncias  $I_1$  e  $I_2$ ; para  $k \in [3, \dots, n]$ ,  $G_k$  é a generalização de  $G_{k-1}$  e  $I_k$ . Os métodos guiados pelas instâncias têm a vantagem de poder produzir generalizações de maneira *incremental* mais facilmente que os guiados por modelos. No entanto, são bem mais sensíveis à existência de instâncias errôneas. Pode-se atenuar este problema impondo que, a cada instância considerada, a nova generalização seja obtida de forma bastante parcimoniosa, somente incorporando aspectos que, de alguma maneira, tenham sido avaliados como muito relevantes; evidentemente a rapidez de aprendizagem diminuirá com esta abordagem.

Os métodos guiados por modelos procuram um conjunto de generalizações na tentativa de achar as melhores (segundo critério de terminado pela aplicação) em relação ao conjunto de instâncias consideradas. Esses métodos usualmente apresentam melhor imunidade a "ruídos" nas instâncias, uma vez que, em geral, as instâncias são consideradas no seu conjunto. A principal desvantagem desses métodos é que cada nova generalização obtida deve ser checada com todo o conjunto de instâncias, a fim de avaliar a sua consistência e completeza.

Vale ainda observar que os métodos de aprendizagem indutiva guiados por dados impõem um número fixo de situações de comparação ( $N-1$  situações, onde  $N$  é o número total de exemplos e contra-exemplos), enquanto as abordagens guiadas por modelos provocam, no mínimo,  $N$  situações de comparação, que é o caso, bastante improvável, em que a primeira generalização obtida já cobre todos os exemplos e é consistente com todos os contra-exemplos.

### 3.2 - IMPORTÂNCIA

Em se tratando de um processo de geração de conhecimento, de forma geral pode-se dizer que a aprendizagem indutiva tem aplicação onde a aquisição de conhecimento se faz necessária, ou seja, nos sistemas baseados em conhecimento. Conforme se argumentou no Capítulo 1, pode-se mesmo dizer que a aquisição de conhecimento é o fator mais limitante para a construção de sistemas inteligentes; donde a necessidade de haver mecanismos adequados de aprendizagem indutiva. A seguir são comentadas algumas aplicações conforme Michalski (1983).

Uma dessas aplicações é na *construção automática de bases de conhecimento para sistemas especialistas*. A abordagem atual para construir bases de conhecimento envolve um processo extremamente tedioso de formalização do conhecimento de especialistas e sua codificação em algum tipo de representação, tal como as regras de produção ou as redes semânticas. Programas de aprendizagem indutiva podem fornecer tanto uma melhoria das técnicas atuais, quanto um fundamento para que se possa desenvolver métodos alternativos de aquisição de conhecimento (como é o próprio caso do TEIRESIAS).

Em pequenos domínios apropriadamente selecionados, programas que fazem indução já são capazes de determinar regras de decisão através de indução, a partir de exemplos de decisão tomadas por um especialista do domínio. Este processo simplifica enormemente a transferência de conhecimento de um especialista para a máquina. A factibilidade de um processo de aquisição como este foi demonstrada no

sistema PLANT/DS (Michalski and Chilausky, 1980). Com este sistema foi feita uma experiência interessante em que, a partir de um conjunto de descrições de plantas de soja doentes, devidamente associadas aos nomes das doenças, gerou-se automaticamente um conjunto de regras de reconhecimento da doença de uma planta de soja a partir da descrição de seu estado. Paralelamente, em entrevistas com especialistas, obteve-se um outro conjunto de regras. Curiosamente, o conjunto de regras obtidas automaticamente teve melhor desempenho na identificação das doenças de um outro conjunto de plantas do que o conjunto obtido diretamente dos especialistas.

Uma outra aplicação promissora, apesar de menos direta, é no *refinamento de bases de conhecimento inicialmente desenvolvidas por especialistas humanos*. Nesta aplicação, um programa de indução poderia ser usado para detectar e retificar inconsistências, eliminar redundâncias, preencher ausências, ou ainda simplificar as regras de decisão criadas por especialistas. Aplicando um mecanismo de aprendizagem indutiva a um conjunto de dados que consistisse de regras originais e de exemplos de resultados corretos e incorretos dessas regras em novas situações, seria possível melhorar as regras incrementalmente com um mínimo de participação do especialista. Um exemplo desta classe de aplicações é o sistema SEEK (Politakis and Weiss, 1984) que, detectando regularidades no mau desempenho de um sistema especialista no domínio de Reumatologia, sugere generalizações ou especializações em certas regras da base, refinando-as. Como uma aplicação de propósito geral, vale mencionar o sistema apresentado em Silva (1983), que cria novas regras de decisão, induzindo-as a partir de um conjunto original.

Uma outra importante aplicação de programas de aprendizagem indutiva é no *auxílio à detecção de padrões conceituais* ou na *revelação de estruturas subjacentes a um conjunto de observações*. Tal aplicação é especialmente interessante em ciências experimentais, tais como a Biologia, Química, Psicologia, Medicina, entre outras, onde há

um grande volume de dados que têm de ser analisados, e as ferramentas matemáticas e estatísticas tradicionais de análise de dados como regressão linear e taxonomia numérica podem não ser suficientemente poderosas. Para essas disciplinas é mais interessante haver disponibilidade de técnicas *conceituais* de análise, onde os resultados têm um caráter mais lógico, simbólico, do que propriamente matemático. Uma aplicação nesta linha é o sistema META-DENDRAL (Buchanan et alii, 1978) que aprende regras de clivagem molecular através de dados de espectrometria de massa. Ainda nesta linha, pode-se citar o trabalho de Michalski et alii (1981) em um sistema de propósito geral para agrupamentos de dados.

### 3.3 - OS DOIS TIPOS: INDUÇÃO ATRAVÉS DE EXEMPLOS E INDUÇÃO ATRAVÉS DE OBSERVAÇÕES

Há dois grandes processos de aprendizagem indutiva: a aprendizagem através de exemplos (ou *Aquisição de Conceito*) e a Aprendizagem através de observações (ou *Generalização Descritiva*). Na literatura de Psicologia, a aprendizagem indutiva é usualmente referenciada por *Formação (ou Identificação) de Conceitos*, ou ainda por *Abstração de Esquemas* (Elio and Anderson, 1981).

Na *aquisição de conceitos*, o objetivo é determinar uma descrição geral para os membros (objetos, situações, processos etc.) de uma classe (conceito), a partir das descrições individuais dos membros da classe. O conceito generalizado pode ser visto como uma regra de classificação de conceito, de tal forma que se um objeto satisfaz esta regra, então ele é uma instância positiva (exemplo) deste conceito. Por exemplo, uma regra de reconhecimento do conceito "inverno em São José dos Campos" poderia ser: Se o início da manhã é bastante frio e enevoado, o meio do dia e o começo da tarde são quentes, o fim da tarde e o começo da noite são frios, a madrugada é bem fria e úmida, e quase não chove, então é inverno em São José dos Campos". Vale mencionar que os trabalhos de classificação estatística de padrões constituem formas de aquisição de conceitos.

Na *generalização descritiva* procura-se obter uma descrição geral (lei, teoria etc.) para um conjunto de observações ou fatos. Por exemplo, a observação de que a Lua cheia de cada mês do primeiro semestre deste ano provocou instabilidades climáticas, permite generalizar que nas próximas fases da Lua cheia o tempo deverá ficar instável. Assim, em contraste com a aquisição de conceito, na qual são produzidas descrições para classificar objetos em classes a partir de suas propriedades, na generalização descritiva são produzidas descrições através da especificação de propriedades de objetos que pertencem a uma certa classe. O presente trabalho tratará apenas de "aquisição de conceitos".

Como se percebe pelas definições acima, a diferença entre os dois tipos de indução é sutil; o ponto fundamental é que, conforme as instâncias estão originalmente disponíveis, isto é, como exemplos ou como observações, os procedimentos de generalização devem ser definidos de acordo.

#### 3.4 - O PAPEL DOS EXEMPLOS E DOS CONTRA-EXEMPLOS

Os exemplos (instâncias positivas) e os contra-exemplos (instâncias negativas) constituem, respectivamente, o conjunto de fatos que pertencem ou não a uma determinada categoria, ou seja, que são instâncias ou não de um determinado conceito. A existência de apenas exemplos, ou de exemplos e contra-exemplos permite dividir os processos de aquisição de conceito em dois tipos (Michalski, 1983): a aprendizagem de *Descrições Características* e a aprendizagem de *Descrições Discriminantes*.

Antes de definir essas duas formas de aprendizagem, deve-se observar as duas condições de validade de uma asserção indutiva, a *Condição de Completeza* e a *Condição de Consistência*. A primeira estabelece que esta asserção deve cobrir (isto é, implicar logicamente) todas as instâncias positivas que se conhece do conceito, enquanto a segunda impõe que a descrição induzida não deverá cobrir qualquer uma das instâncias negativas conhecidas.

A aprendizagem de descrições características significa a especificação de um conjunto de propriedades (características) que devem ser comuns a todos os membros de uma classe, ou seja, é tipicamente a conjunção de um conjunto de propriedades referentes aos elementos dessa classe. Por isto as descrições características são também conhecidas por *Generalizações Conjuntivas*. Assim, pode-se dizer que estas asserções indutivas são as que satisfazem apenas a condição de completeza. O objetivo destas descrições é o de servir como regras que permitam discriminar a classe em questão de todas as outras classes. Conseqüentemente, quanto maior o número de elementos da descrição, mais interessante será. Em NASA (1982), este tipo de descrição é chamada de "Abdução".

A aprendizagem de descrições discriminantes significa a especificação de um conjunto de propriedades dos membros de uma classe, de forma que ela possa ser distingüida de um número limitado de outras classes. Desta forma, pode-se dizer que as descrições discriminantes são as que devem satisfazer tanto a condição de completeza, quanto a de consistência. Conseqüentemente, tanto mais interessante sob o ponto de vista das aplicações, quanto menor for o número de elementos componentes da descrição; as melhores constituem as chamadas *descrições discriminantes mínimas*.

Os contra-exemplos mais importantes são os *contra-exemplos com diferenças próximas* ("near misses") (Winston, 1975), cuja característica é a de serem bastante semelhantes às instâncias positivas. Como conseqüência, eles fornecem muita informação ao processo de generalização, informação esta que se traduz na imposição de fortes restrições à descrição dos exemplos.

Como será mostrado no Capítulo 6, apesar de se ter adotado uma perspectiva básica de geração de descrições características, o modelo de aprendizagem desenvolvido pode ser também aplicado na geração de descrições discriminantes.

### 3.5 - O PAPEL DA REPRESENTAÇÃO ADOTADA: A INFLUÊNCIA DO "COMO" REPRESENTAR

Muitas formas de representação têm sido usadas em sistemas de aprendizagem, entre as quais se encontram os vetores de características, o cálculo de predicados, as redes semânticas, as regras de produção etc. Essas representações variam entre si segundo quatro aspectos: sua expressividade, a facilidade de inferência dentro da representação, sua capacidade de ser estendida, e sua capacidade de permitir alterações na base de conhecimento.

A *expressividade* refere-se à capacidade de exprimir com facilidade o conhecimento relevante a uma determinada aplicação, sejam os fatos, sejam as descrições gerais procuradas. Assim, se está sendo utilizado o cálculo de predicados, tem-se uma maior flexibilidade de representação do que se os vetores de características estivessem sendo usados. Por exemplo, nesta última representação não seria possível exprimir aspectos estruturais de um conjunto de objetos. Um outro exemplo é o fato de que se a disjunção não é permitida na representação, então não é possível descobrir uma descrição que a envolva, como a seguinte: "em todas as grandes tempestades deste ano a Lua estava na fase minguante *ou* na quarto-crescente". Deve-se observar que a medida que se enriquece a representação (por exemplo com a existência da disjunção), ganha-se em expressividade, mas perde-se na própria complexidade do processo de busca a generalização, já que o espaço de regras é aumentado significativamente. Outros comentários sobre este aspecto serão vistos na Seção 5.2.

Quanto à *facilidade de inferência dentro da representação*, pode-se observar, por exemplo, que a verificação de igualdade entre duas descrições é bem mais custosa em termos computacionais quando se está utilizando o cálculo de predicados, do que na situação em que a representação se baseia nos vetores de características.

Em geral pode-se dizer que as representações declarativas, por ser modulares, permitem com facilidade que a base de conheci

mento possa ser modificada. No entanto, mesmo utilizando estas representações, o aspecto da *modificação da base de conhecimento* pode se tornar bastante complexo quando o novo conhecimento aprendido levantar problemas de integração das novas informações com as já presentes na base.

Finalmente, vale ressaltar que sendo desejável que um sistema de aprendizagem tenha a *capacidade de estender sua própria representação*, esta deve oferecer facilidades para tal. Apesar de este aspecto não ter sido ainda muito explorado, alguns resultados interessantes já foram obtidos como é o caso do sistema AM (Lenat, 1982) que gera conceitos em matemática; este sistema permite a definição de novos conceitos a partir de outros já conhecidos. No próprio sistema TEIRESIAS, o processo de aquisição de novas primitivas conceituais é uma forma de extensão da representação original.

### 3.6 - O PAPEL DO CONHECIMENTO EXPRESSO: A INFLUÊNCIA DE "O QUE" REPRESENTAR

Apesar de algum conhecimento relevante a uma tarefa de aprendizagem indutiva poder estar imbutido no próprio mecanismo de generalização, ou possa ter sido recebido do ambiente (professor) juntamente com as instâncias positivas e negativas, de forma geral pode-se dizer que a maior fonte de conhecimento relevante são as próprias instâncias positivas e negativas que lhe são apresentadas. Portanto, é bastante importante determinar quais serão os itens de informação (ou seja, as cláusulas das descrições das instâncias) que serão fornecidos ao sistema de indução.

Sob a ótica de relevância das cláusulas das instâncias para uma tarefa de indução, Michalski (1983) distingue três níveis de relevância: total, parcial e indireta, as quais serão descritas a seguir. Estes níveis de relevância requerem, na ordem em que foram citados anteriormente, cada vez menos trabalho da pessoa que esteja definindo a situação de aprendizagem, isto é, formulando o problema sobre o qual o sistema atuará.

No caso de *relevância total*, todas as cláusulas das instâncias são consideradas diretamente relevantes à tarefa de aprendizagem. Nesta situação, o papel do sistema de aprendizagem é formular uma asserção indutiva em uma forma matemática ou lógica, suficientemente geral, a fim de levar em consideração todas as cláusulas das instâncias, sem poder desprezar qualquer uma delas. Os primeiros trabalhos de sistemas de controle adaptativo (Dietterich, 1983) foram desenvolvidos sob esta abordagem, onde era freqüente a utilização de técnicas de regressão polinomial como mecanismo de geração da descrição generalizada, isto é, da curva que se desejava ajustar sobre um conjunto de pontos em um sistema de coordenadas.

Na situação em que há *relevância parcial* das instâncias, estas podem conter um conjunto de cláusulas irrelevantes ou redundantes, juntamente com um conjunto de cláusulas relevantes. A tarefa do sistema de aprendizagem neste caso é selecionar as cláusulas mais relevantes e, a partir delas, construir uma asserção indutiva apropriada. Este tipo de abordagem é chamada "*Aprendizagem Indutiva Seletiva*" (Michalski, 1980). A maior parte das pesquisas realizadas em aprendizagem indutiva (entre as quais o próprio modelo desenvolvido) refere-se a sistemas que operam sob essa perspectiva. No próximo capítulo serão discutidos alguns.

Finalmente, na situação de *relevância indireta* das instâncias, apesar de que todas as suas cláusulas componentes possam não ser diretamente relevantes à tarefa de aprendizagem, existem algumas que podem ser usadas para construir cláusulas que são diretamente relevantes. Assim, o papel do sistema de aprendizagem é construir essas novas cláusulas e formular uma asserção indutiva adequada. Por exemplo, esta situação é criada quando uma informação relevante é a pressão atmosférica de uma certa região da Terra, mas dispõe-se apenas da informação da altitude desta região com relação ao nível do mar. Este tipo de mecanismo é conhecido por "*Aprendizagem Indutiva Construtiva*" (Michalski, 1980). Pouca pesquisa tem sido feita sob esta perspectiva,

sem dúvida porque este é o problema mais amplo que se coloca no âmbito da aprendizagem indutiva. Os trabalhos que têm sido produzidos sobre generalização construtiva referem-se a domínios específicos como é o caso do sistema AM citado, e do sistema BACON (Dietterich, 1983) que descobre leis simples da física clássica. O único sistema de propósito geral que oferece alguma facilidade de generalização construtiva é o sistema INDUCE (e suas várias versões), o qual será comentado no próximo capítulo.

### 3.7 - AS REGRAS SINTÁTICAS DE GENERALIZAÇÃO

#### 3.7.1 - INTRODUÇÃO

Várias são as regras de generalização que têm sido referenciadas na literatura. Conforme as regras de generalização estejam sendo usadas em um contexto de aprendizagem seletiva ou construtiva, elas podem ser classificadas como de um destes dois tipos. As regras "construtivas" de generalização são regras que geram atributos não-referenciados nas descrições originais. Exemplos desta classe de regras são: a regra de contagem do número de objetos presentes a cada instância; a regra de contagem do número de objetos com determinada especificação de atributos; a regra de especificação da posição de um objeto em seqüência ordenada, e a regra de determinação de interdependência entre valores de atributos (Michalski, 1983).

As regras que serão apresentadas a seguir são todas do tipo "seletivo" (as descrições geradas apresentam apenas atributos já presentes nas descrições originais), já que somente este enfoque de aprendizagem foi explorado no presente trabalho. Tais regras podem ser consideradas as mais importantes, uma vez que outras não-mencionadas não passam de casos particulares dessas. É interessante observar que todas as regras foram obtidas direta ou indiretamente das regras de inferência da lógica simbólica (Copi, 1967). Um exemplo de aplicação direta é a "regra de eliminação de condição", obtida a partir da regra de inferência conhecida por "simplificação"; um exemplo de

aplicação indireta é a "regra de hierarquização conceitual", obtida indiretamente a partir da regra de inferência conhecida por "adição". Esses exemplos serão detalhados na próxima seção e correspondem às duas primeiras regras de generalização.

### 3.7.2 - REGRAS SELETIVAS DE GENERALIZAÇÃO

Para todo o resto da seção considere-se que P, Q, R e S são proposições da lógica simbólica. Já que por simples inspeção nas regras de inferência é possível visualizar como cada uma delas se relaciona com as suas respectivas regras de generalização, não se dará maiores detalhes a respeito deste interrelacionamento (com exceção das duas primeiras regras). As regras serão apresentadas a partir da sua regra de inferência de origem.

#### i) SIMPLIFICAÇÃO

Esta regra de inferência estabelece que:

$$P \ \& \ Q \ \rightarrow \ P$$

##### 1) Regra de Eliminação de Condição

Esta regra estabelece que uma descrição qualquer pode ser generalizada, bastando que lhe seja retirada (eliminada) uma das cláusulas componentes. Conforme foi citado anteriormente, esta regra é obtida diretamente da "regra de simplificação". Como esta última pode ser expressa por  $P \ \& \ Q \ \rightarrow \ P$ , basta que se interprete o antecedente e o conseqüente como se fossem descrições relativas a uma classe K para ser obtida a regra de eliminação de condição, conforme definida anteriormente. Assim,  $P \ \& \ Q$  levando a K pode ser generalizado para P levando ao mesmo K. Por exemplo, a cláusula:

"Dias quentes e chuvosos são dias de verão",

que na realidade são duas cláusulas expressas de forma compacta, poderia ser generalizada para:

"Dias chuvosos são dias de verão".

ii) ADIÇÃO

Esta regra de inferência diz que:

$P \rightarrow P \text{ ou } Q.$

2) Regra de Hierarquização Conceitual

Refere-se à troca de um grupo de conceitos, em um conjunto de cláusulas de instâncias, por um outro mais geral que os englobe (Figura 3.3). Esta regra é obtida indiretamente da "regra de adição". Para tal, deve-se primeiramente aplicar a regra da adição, acrescentando-se elementos semanticamente relacionados ao original. Este processo pode ser representado por:

$P \rightarrow P \text{ ou } P' \text{ ou } P''.$

Supondo a existência de um conceito  $P^*$  que engloba os três anteriores, obtêm-se finalmente a regra de hierarquização conceitual interpretando o conseqüente da implicação acima e o conceito  $P^*$  como descrições relativas a uma mesma classe  $K$ . Assim,  $P$  ou  $P'$  ou  $P''$  levando a  $K$  pode ser generalizado para  $P^*$  levando ao mesmo  $K$ . Por exemplo, o conjunto de cláusulas:

"A zona frontal está localizada no *Rio Grande do Sul*" e

"A zona frontal está localizada em *Santa Catarina*",

pode ser generalizado para:

"A zona frontal está localizada na *Região Sul*".

Ou seja, foi feita a seguinte hierarquização de conceitos:

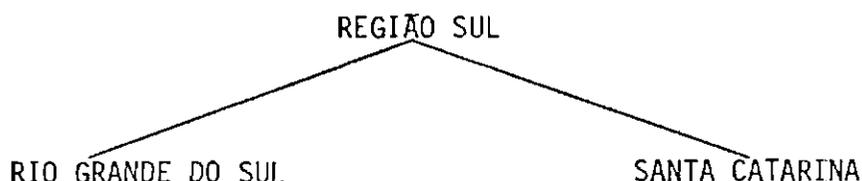


Fig. 3.3 - Exemplo de hierarquização de conceitos.

### 3) Regra de Adição de Alternativa

Consiste na extensão do escopo de validade de um certo atributo de objeto, através da utilização do conectivo lógico de disjunção. Serve tanto para valores numéricos quanto para valores simbólicos. Por exemplo, a cláusula:

"A zona frontal se desloca para o norte",

poderia ser generalizada para:

"A zona frontal se desloca para norte *ou* nordeste".

### 4) Regra de Especificação de Intervalo Numérico

Refere-se à extensão dos valores numéricos que uma propriedade de algum objeto possa assumir; através da troca destes valores numéricos por um intervalo que os englobe.

Dois casos são possíveis: O "Fechamento de Intervalo" e a "Extensão de Intervalo" propriamente dita. Inicialmente considere-se que:

$N, N_1, N_2, \longrightarrow$  números quaisquer;

$I_i \longrightarrow$  um intervalo numérico qualquer, aberto ou fechado em qualquer um dos extremos;

$N_{\min_i} \longrightarrow$  extremo inferior do intervalo  $I_i$ ;

$N_{\max_i} \longrightarrow$  extremo superior do intervalo  $I_i$ .

Embora outras situações sejam possíveis na "Extensão de Intervalo", duas delas são mostradas a seguir: uma em que os valores de atributo envolvidos constituem um intervalo e um número, e a outra em que todos os valores são intervalos. Na primeira situação, a regra estabelece que as cláusulas:

(Atributo Objeto =  $N$ ) e

(Atributo Objeto  $\in I_1$ )

poderiam produzir a generalização:

(Atributo Objeto  $\in [N \dots N_{\max_1}]$ ) se  $N \leq N_{\min_1}$ , ou

(Atributo Objeto  $\in [N_{\min_1} \dots N]$ ) se  $N \geq N_{\max_1}$ .

Na situação em que os valores de atributos são intervalos, tem-se que as cláusulas:

(Atributo Objeto  $\in I_1$ ) e

(Atributo Objeto  $\in I_2$ ),

poderiam ser generalizados para:

(Atributo Objeto  $\in [N_{\min} \dots N_{\max}]$ ),

onde:

$$N_{\min} = \text{Min} \{N_{\min 1}, N_{\min 2}\},$$

$$N_{\max} = \text{Max} \{N_{\max 1}, N_{\max 2}\}.$$

O caso de "Fechamento de Intervalo" estabelece que as cláusulas:

(Atributo Objeto =  $N_1$ ) e

(Atributo Objeto =  $N_2$ )

poderiam ser generalizadas para:

(Atributo Objeto  $\in [N_1 \dots N_2]$ ) se  $N_1 < N_2$ , ou

(Atributo Objeto  $\in [N_2 \dots N_1]$ ) se  $N_2 < N_1$ .

Por exemplo, a partir das cláusulas:

"A precipitação no início de janeiro foi de 680 mm", e

"A precipitação no fim de janeiro foi de 850 mm",

seria possível a generalização:

"A precipitação em janeiro esteve entre 680 e 850 mm".

##### 5) Regra de Troca de Constantes por Variáveis

Refere-se à troca de uma série de nomes específicos de objetos, que têm uma ou mais propriedades comuns, por uma variável, o que significa que se generalizou esta propriedade para todos os outros objetos do domínio. Por exemplo, o conjunto de cláusulas:

"O Estado do Amazonas é chuvoso",

"O Estado do Acre é chuvoso", e

"O Estado de Rondônia é chuvoso",

poderia ser generalizado (mais adequado seria dizer "sobregeneralizado", pois levaria a um fato não-verdadeiro) para

"Todos os Estados são chuvosos".

#### 6) Regra de Expansão Limitada no Domínio

Consiste em expandir o valor de determinado atributo em uma cláusula, respeitando o limite do domínio de validade do atributo, conforme imposto pelo conhecimento expresso em outra cláusula. Considere-se que "c1 e c2" são cláusulas de instâncias, cuja diferença é a presença de diferentes valores de atributos; sejam  $V(c1)$  e  $V(c2)$  estes valores de atributo. Considere-se também que K e não-K são, respectivamente, os conceitos a que levam as duas instâncias. Assim, a regra de expansão limitada no domínio pode ser entendida, genericamente, da seguinte forma:

SE:

$V(c1) \rightarrow K$

$V(c2) \rightarrow \text{não-K},$

ENTÃO:

$V(c1) \rightarrow K$

$\{\forall V/(V \neq V(c1))\} \rightarrow \text{não-K},$

Qu:

$V(c2) \rightarrow \text{n\~{a}o-K},$

$\{\forall V/(V \neq V(c2))\} \rightarrow K.$

Por exemplo, supondo que haja duas instâncias que caracterizam um volume elevado ou não de chuva, e a partir das cláusulas

"A estação do ano é verão", e

"A estação do ano é inverno",

chega-se ao esquema mostrado na Figura 3.4, que ilustra mais claramente a regra de expansão limitada. Vale ressaltar que podem ser obtidas duas descrições generalizadas a partir das duas cláusulas acima.

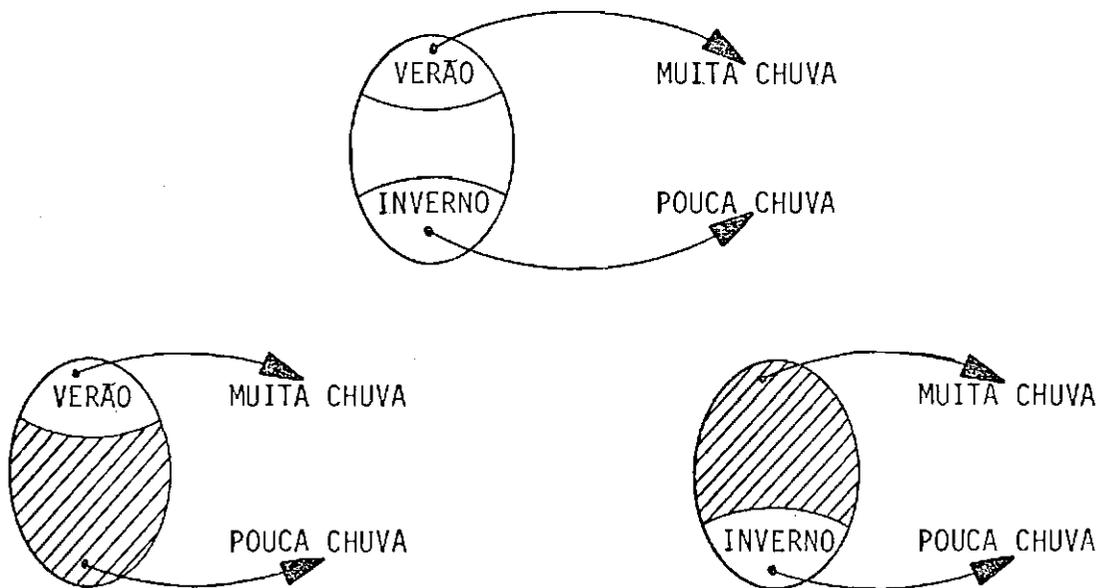


Fig. 3.4 - Exemplo da regra de expansão limitada no domínio.

iii) SIMPLIFICAÇÃO E ADIÇÃO

7) Regra da Mudança de Conjunção em Disjunção

Como o próprio nome já sugere, consiste em mudar o relacionamento conjuntivo que possa haver entre um grupo de cláusulas para um relacionamento disjuntivo. Assim, a mesma cláusula do exemplo anterior pode ser generalizada para:

"Todos os verões são quentes *ou* chuvosos".

iv) DILEMA CONSTRUTIVO

Esta regra de inferência estabelece que:

SE:

$(P \rightarrow Q) \& (R \rightarrow S)$

P ou R,

ENTÃO:

Q ou S.

Se R é equivalente a não-P ( $\sim P$ ), tem-se o conhecido *princípio da resolução*, o qual pode ser expresso da seguinte forma:

SE:  $(P \rightarrow Q) \& (\sim P \rightarrow S)$ ,

ENTÃO: Q ou S.

### 8) Regra de Resolução Indutiva

Sejam  $C_1$  e  $C_2$  duas cláusulas de duas instâncias positivas do conceito  $K$ . Seja ainda  $*C$  uma cláusula da instância que contém  $C_1$ , de tal forma que  $\text{não-}*C$  seja uma cláusula da outra instância. A regra da Resolução Indutiva estabelece que:

SE:

$$*C \ \& \ C_1 \quad \longrightarrow \ K,$$

$$\text{não-}*C \ \& \ C_2 \quad \longrightarrow \ K,$$

ENTÃO:

$$C_1 \ \vee \ C_2 \quad \longrightarrow \ K.$$

Por exemplo, as regras:

"Verão úmido provoca chuva", e

"Estações diferentes de verão na Lua cheia provocam chuva",

podem ser generalizadas para a regra:

"Umidade *ou* Lua cheia provoca chuva".

### 3.8 - O PONTO DE VISTA PSICOLÓGICO

Uma vez que o problema de aprendizagem indutiva não está fechado dentro da inteligência artificial, é interessante observar rapidamente o que se diz no âmbito da psicologia cognitiva. Neste sentido vale considerar a distinção feita por Elio e Anderson (1981) e Anderson (1979), a respeito dos vários modelos da aprendizagem indutiva humana.

Segundo esses autores, pode-se distinguir três tipos básicos de teoria sobre o processo humano de "abstração de esquemas". A primeira classe de teorias, conhecida como a dos *modelos de protótipo*, propõe que os seres humanos constroem uma caracterização única (protótipo) sobre a tendência central do conceito considerado. Considera-se freqüentemente que o protótipo para este conceito é uma das próprias instâncias positivas. Outras instâncias são membros desta categoria, à medida que elas se assemelham a este protótipo. Esta classe de teorias inclui, por exemplo, os trabalhos de Franks e Bransford (1971) e Rosch e Mervis (1975).

O segundo tipo de teorias constitui os chamados *modelos apenas de instâncias*, em que se propõe que os seres humanos armazenam apenas instâncias individuais e fazem seus julgamentos de categorização com base na similaridade que possa existir entre a instância de teste e as instâncias armazenadas. Um representante desta classe é a teoria proposta por Medin e Schaffer (1978).

Finalmente, o terceiro tipo corresponde aos *modelos de reforço ou freqüência*. Nestes propõe-se que a representação do conceito é resultado da freqüência de ocorrência de características individuais das instâncias e de combinações destas características. Em particular, poderia ser armazenada uma única abstração (que caracteriza a primeira classe), ou várias instâncias isoladas (que caracteriza a segunda); as duas classes consideradas acima podem ser vistas, portanto, como casos particulares deste terceiro tipo. Destacam-se nesta classe de teorias, as várias versões do sistema ACT (Anderson, 1983; Anderson and Kline, 1979; e Anderson et alii, 1979)), além do modelo proposto em Hayes-Roth e Hayes-Roth (1977). Esta abordagem vem se fixando como a mais adequada em termos de explicar resultados obtidos experimentalmente. Vale observar que o modelo que será apresentado e o próprio sistema TEIRESIAS constituem instâncias desta classe.

## CAPÍTULO 4

### METODOLOGIAS DE APRENDIZAGEM INDUTIVA

#### 4.1 - INTRODUÇÃO

Dois sistemas de aprendizagem indutiva tiveram uma forte influência sobre o presente trabalho. Sob o ponto de vista teórico, ou seja, do embasamento necessário para penetrar na área e começar a entender os seus problemas, a principal influência veio dos trabalhos desenvolvidos por Michalski e colaboradores. Vale adiantar que tais trabalhos acabaram por gerar uma metodologia denominada INDUCE, sobre a qual será feito um rápido comentário na Seção 4.3.

Como ficará patente nos próximos capítulos, sob o ponto de vista prático, ou seja, a respeito do mecanismo em si que foi desenvolvido, as idéias contidas no sistema SPROUTER (Hayes-Roth and McDermott, 1977 e 1978; Dietterich and Michalski, 1979 e 1981) foram fundamentais. Pela sua importância, e a fim de que se possa entender melhor o próprio mecanismo que será apresentado posteriormente, o sistema SPROUTER será suficientemente detalhado neste capítulo.

A seguir apresentam-se as Figuras 4.1 e 4.2, a partir das quais será apresentado o sistema SPROUTER. A Figura 4.1 servirá para apresentar a linguagem utilizada para a representação das instâncias, e a Figura 4.2 mostra o desempenho do sistema. A Figura 4.2 é particularmente importante pois foi utilizada no estudo comparativo entre métodos de aprendizagem indutiva descrito em Dietterich and Michalski (1979 e 1981). Além do sistema SPROUTER, será apresentado (na Seção 4.3) o desempenho de outros três sistemas a partir desse mesmo conjunto de instâncias, isto é, o da Figura 4.1. De posse de todos estes resultados, será possível ter uma boa visão do desempenho do mecanismo proposto, pois, no Capítulo 5, também ele será testado no referido conjunto de instâncias.



Fig. 4.1 - Instâncias para ilustrar os esquemas de representação do SPROUTER.

FONTE: Dietterich e Michalski (1979 e 1981).

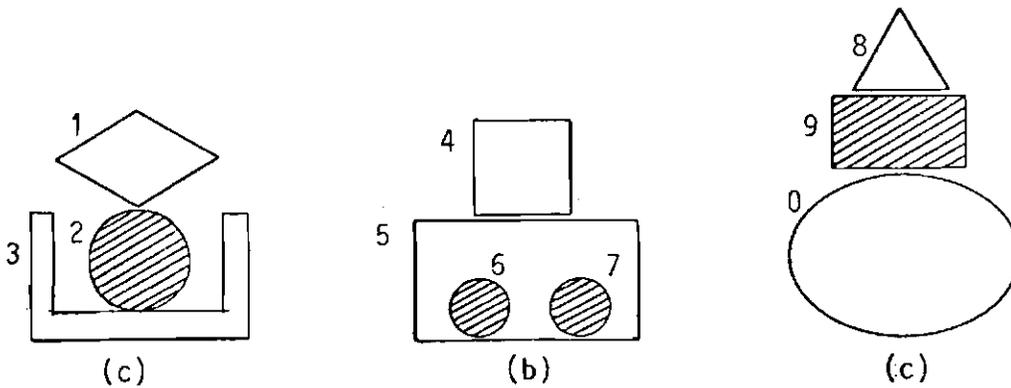


Fig. 4.2 - Instâncias positivas para comparar o desempenho dos métodos.

FONTE: Dietterich e Michalski (1979 e 1981).

#### 4.2 - O SISTEMA SPROUTER

O sistema SPROUTER é um sistema de aprendizagem indutiva guiado pelas instâncias, cujo objetivo é a busca de generalizações conjuntas a partir de um conjunto de exemplos, as quais são aqui chama

das *abstrações máximas* ou *casamentos por interferência* ("interference matches"). Uma abstração máxima de duas instâncias I1 e I2 é toda generalização conjuntiva destas instâncias que não esteja propriamente contida em qualquer outra generalização.

As chamadas *Representações Estruturais Parametrizadas* (REP) constituem o esquema de representação, tanto para as instâncias originais quanto para suas generalizações. A REP para os dois eventos da Figura 4.1 são:

Instância (a): {{círculo:a} {quadrado:b} {pequeno:a}  
                  {pequeno:b} {sobre:a, sob:b}}

Instância (b): {{círculo:c} {quadrado:d} {círculo:e}  
                  {pequeno:c} {grande:d} {pequeno:e}  
                  {sobre:c, sob:d} {dentro:e, fora:d}}

Expressões tais como {pequeno:a} são *quadros de caso* ("case frames"), construídos a partir de *rótulos de caso* ("case labels") (pequeno, círculo, etc.) e *parâmetros* (a, b, c, etc.).

O algoritmo de indução opera de baixo para cima, exatamente como no esquema da Figura 4.3 e seu objetivo é encontrar o mais longo casamento um-a-um de parâmetros e quadros de caso entre as instâncias, isto é, a mais longa subexpressão comum a elas. Isto é realizado em dois passos. Primeiro, os quadros de casos relativos às instâncias (a) e (b) são casados em todas as formas possíveis para obter o conjunto "A". Dois quadros de caso casam se todos seus rótulos de caso também casam. Cada elemento de "A" é constituído de um rótulo de caso e uma lista de correspondências de parâmetros, a qual representa a condição para que ambos os eventos sejam casados, com relação ao referido rótulo do caso:

```
A = {{c{circulo}:(a/c) (a/e)}}
      {quadrado:(b/d)},
      {pequeno:(a/c) (b/c) (a/e) (b/e)},
      {sobre, sob:(a/c b/d)}}.
```

O segundo passo envolve a seleção de um subconjunto de correspondências de parâmetros em "A", tal que todos os parâmetros possam ser associados de forma consistente. Isto é realizado por meio de uma busca em largura no espaço das possíveis associações, sendo então podados os nós não-promissores. A busca pode ser vista como um processo de criação de nós (e de fato é isto que acontece, pois estes vão sendo criados em uma rede chamada ACORN, cuja função é servir de estrutura de reconhecimento para futuras instâncias). A Figura 4.3 apresenta o exemplo de um grafo de busca já podado.

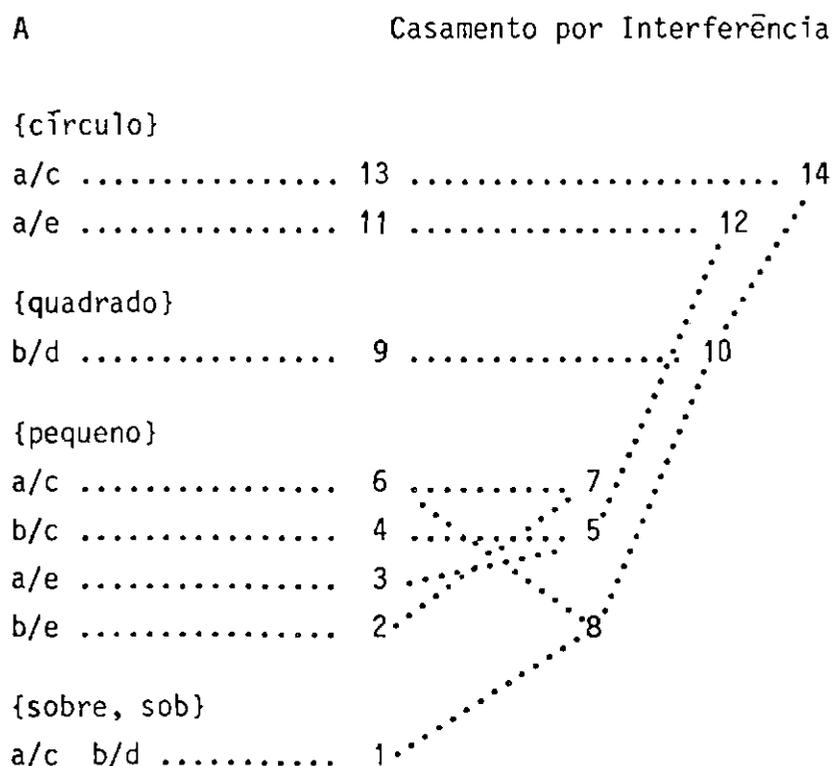


Fig. 4.3 - Busca realizada pelo SPROUTER para a obtenção das melhores associações de parâmetros, segundo as instâncias mostradas na Figura 4.1.

FONTE: Dietterich e Michalski (1981).

Os nós são numerados na ordem de sua geração. Um de cada vez, cada par de parâmetros correspondentes é selecionado de "A", e um novo nó é criado para ele. Este nó é então comparado com todos os outros nós previamente gerados. Nós adicionais são criados para cada caso em que ele pode ser consistentemente associado a um nó previamente existente. No grafo de busca da Figura 4.3, quando a associação de parâmetro (a/c) relativa ao rótulo de caso "pequeno" é selecionada, o nó 6 é criado. O nó 6 é então comparado aos nós de 1 a 5 e dois novos nós são criados: o nó 7, que é criado ligando o nó 6 (a/c) ao nó 2 (b/e), e o nó 8 que é criado pela associação do nó 6 (a/c) ao nó 1 (a/c b/d). O nó 6 não pode ser ligado ao nó 3, por exemplo, porque o parâmetro "a" seria inconsistentemente ligado a ambos os parâmetros "c" e "e".

Quando a busca é completada, os nós 7, 12 e 14 representam associações que conduzem a generalizações conjuntivas. O nó 14, por exemplo, associa "a" a "c" (dando "v1") e "b" a "d" (dando v2), produzindo a conjunção:

{c:rculo:v1} {quadrado:v2} {pequeno:v1} {sobre:v1, sob:v2}}.

Duas heurísticas são invocadas durante o processo de construção dos nós. A primeira delas tem o objetivo de manter o espaço de busca em um tamanho razoável, através da poda automática de todas as conjunções que provavelmente não farão parte de uma abstração máxima. Isto é possível por meio do cálculo de uma função de utilidade para cada nó. Quando se ultrapassa um limite estipulado do número de nós que se permite considerar durante a busca, utiliza-se esta função podando os nós que têm "baixa" utilidade (vale dizer que nos artigos já citados, relativos ao SPROUTER, não é feita qualquer referência a qual valor ou a quais critérios se deve obedecer para a especificação desse número de nós). A função de utilidade de um nó é uma função crescente do número de propriedades cobertas pelo nó, e uma função decrescente do número de parâmetros distintos necessários a sua validação.

A outra heurística fornece uma direção à busca, indicando qual correspondência de parâmetros deverá ser utilizada no próximo ciclo de construção, a fim de atingir mais rapidamente o caminho da melhor generalização. Inicialmente, estabelece-se um fator de amostragem ("atualmente" 20%) para a proporção de associações ainda não-utilizadas, sendo que no mínimo se deve usar três delas. Para cada uma destas associações amostradas, determina-se um limite superior da utilidade de todos os nós que poderiam ser construídos através da correspondência da associação amostrada com todas as outras restantes que ainda estão sendo consideradas. A associação que produzir o nó com a maior utilidade "potencial" será selecionada.

O algoritmo descrito opera somente para um par de instâncias. Se mais instâncias forem disponíveis, deverá ser tomada uma das generalizações obtidas para o par original, com a qual o processo poderá prosseguir. Com este procedimento, generalizações bastante relevantes poderão estar sendo perdidas. Este tipo de problema será referido como *erro de propagação* das generalizações, e pode ser visto como decorrente da *incapacidade do sistema de tratar associações de parâmetros  $N$  a  $1$* .

Deve-se notar que o algoritmo de casamento por interferência tem de ser heurístico, pois a *natureza combinatorial de seu processo de busca* impõe um custo computacional muito caro ao processo de generalização. Se por um lado as heurísticas tornam o processo mais eficiente, por outro introduzem, evidentemente, a possibilidade de não se chegar às generalizações procuradas.

De fato, pode acontecer que o número estipulado, relativo à quantidade máxima de nós que podem ser considerados numa dada aplicação não tenha sido suficiente para evitar que nós relevantes fossem podados ao longo do processo; um dos nós podados poderia inclusive estar no caminho da melhor generalização. Mesmo chegando-se ao fim do processo, pode ser que simplesmente não "caibam" (de acordo com o número máximo estipulado) na ACORN todos os nós que tenham sido obtidos, e alguns deles tenham de ser eliminados.

O algoritmo descobre as seguintes generalizações conjuntivas para as instâncias mostradas na Figura 4.2.

1) {{sobre:v1, sob:v2} {médio:v1} {branco:v1}}

Hã um objeto médio e branco sobre alguma coisa.

2) {{sobre:v1, sob:v2} {médio:v1} {grande:v2} {branco:v2}}

Hã um objeto de tamanho médio sobre um outro branco e grande.

3) {{médio:v1} {branco:v1} {grande:v3} {branco:v3} {preto:v2}}

Hã um objeto branco e de tamanho médio, um outro objeto que é grande e branco, e um outro objeto preto.

Deve-se observar ainda que o algoritmo descrito *não tem a capacidade de gerar asserções disjuntivas*, além de ser *incapaz de gerar descrições que envolvam contra-exemplos*. Resta dizer que no modelo que será apresentado, serão propostas soluções para todos os problemas levantados nesta seção com relação ao SPROUTER.

#### 4.3 - OUTROS SISTEMAS

Alguns sistemas de aprendizagem indutiva têm sido apresentados na literatura seja com objetivos específicos seja com propósito geral. Também no problema da aprendizagem indutiva, tem sido constatado um maior desempenho dos sistemas que possuem capacidade de manipular conhecimentos específicos ao seu domínio de aplicação, fato este que está começando a levar os novos sistemas a tentar incorporar esta facilidade. Conhecimento específico ao domínio pode entrar tanto para impor restrições às generalizações que se possa fazer, quanto para permitir outras (generalização construtiva). Finalmente, vale ressaltar a falta de uniformidade tanto na representação utilizada para instâncias e gene

realizações, quanto na terminologia em geral concernente aos vários temas do problema. A seguir, apresentam-se outros sistemas de aprendizagem indutiva encontrados na literatura.

Inicialmente vale comentar, ainda que ligeiramente, a metodologia desenvolvida em decorrência dos trabalhos de Michalski e colaboradores. Tal metodologia apresentou duas versões básicas: o sistema INDUCE 1.2 (Michalski, 1980; Dietterich and Michalski, 1979, 1981; Michalski and Chilausky, 1980), e a versão atual chamada STAR, implementada no sistema INDUCE (Michalski, 1983).

Ambos os sistemas são sistemas de propósito geral guiados por modelos. Utilizam-se basicamente das regras de generalização, definidas no Capítulo 3, para lançar hipóteses de generalizações válidas para o conjunto de instâncias. Ambos possuem facilidades de generalização construtiva. Suas diferenças principais são:

- . em ambos os casos as linguagens de representação são extensões do cálculo de predicados de primeira ordem, porém na versão mais recente a linguagem é mais poderosa;
- . o sistema INDUCE 1.2 não permite obter generalizações discriminantes, isto é, ele não trata de contra-exemplos;
- . as hipóteses (generalizações a serem testadas) geradas no INDUCE 1.2 vão sendo podadas a partir das mais específicas, enquanto no INDUCE elas são podadas a partir das mais gerais.

O algoritmo de aprendizagem utilizado no INDUCE 1.2 é baseado na "observação" de que as informações de especificação de estruturas são responsáveis pela complexidade computacional do processo de generalização de descrições estruturais; daí a idéia de separar as informações de especificação de atributo, fazendo as generalizações em separado.

No sistema INDUCE, lançam-se hipóteses a partir de cada instância positiva, as quais são testadas e, se não forem consistentes, são especializadas segundo um determinado mecanismo. Dentre as que forem consistentes, são mantidas as melhores segundo um certo critério (que envolve, por exemplo, as que cobrem o maior número de instâncias positivas). Para cada conjunto de hipóteses lançadas, procede-se a uma nova poda, até que, considerados todos os exemplos, as hipóteses válidas até este ponto constituam as generalizações discriminantes procuradas.

Utilizando apenas regras seletivas de generalização, o sistema INDUCE 1.2, quando aplicado ao conjunto de instâncias da Figura 4.2, fornece, entre outras, as seguintes generalizações (Dietterich and Michalski, 1979 e 1981):

1. [sobre(p1, p2)] [tamanho(p1) = médio]  
[forma(p1) = círculo, quadrado, retângulo]  
[tamanho(p2) = grande]  
[forma(p2) = caixa, retângulo, elipse]  
[cor(p2) = branca]

Há um círculo, retângulo ou quadrado de tamanho médio em cima de uma caixa, de um retângulo ou de uma elipse branca e grande.

2. [sobre(p1, p2)] [tamanho(p1) = médio]  
[forma(p1) = polígono] [cor(p1) = branca]  
[tamanho(p2) = médio, grande]  
[forma(p2) = retângulo, círculo]

Há um polígono branco de tamanho médio em cima de um círculo ou retângulo médio ou grande.

3. [sobre(p1, p2)] [tamanho(p1) = médio]  
[forma(p1) = polígono]  
[tamanho(p2) = médio, grande]  
[forma(p2) = retângulo, elipse, círculo]

Há um polígono de tamanho médio em cima de um retângulo, de uma elipse ou de um círculo de tamanho médio ou grande.

4. [tamanho(p1) = pequeno, médio]  
[forma(p1) = círculo, retângulo]  
[cor(p1) = preta]

Há um objeto preto de tamanho médio ou pequeno, e de forma circular ou retangular.

Permitindo regras construtivas de generalização, o sistema INDUCE 1.2 ainda é capaz de gerar a seguinte generalização:

5. [ $\#p = 3,4$ ] [ $\#p$  de cor branca = 2] &  
[mais-de-cima(p1)] [sobre(p1,p2)] [tamanho(p1) = médio]  
[forma(p1) = polígono] [cor(p1) = branca]  
[tamanho(p2) = médio, grande]  
[forma(p2) = círculo, retângulo]

Há 3 ou 4 objetos em cada instância. Exatamente 2 deles são brancos. O objeto mais de cima é um polígono branco de tamanho médio, que está em cima de um círculo ou retângulo de tamanho médio ou grande.

Antes do início do processo de indução, o usuário deverá fornecer todas as informações disponíveis que possam ser usadas pelo sistema, tais como o domínio dos atributos envolvidos (aqui chamados "descritores") e as hierarquizações conceituais do tipo mostrado na Figura 3.3. Devido a uma informação deste último tipo, é que se torna possível a geração automática do conceito "polígono" nas descrições acima.

Finalmente, deve-se salientar que as versões INDUCE representam o resultado do maior esforço teórico e prático em aprendizagem indutiva que se pode encontrar na literatura.

Outro sistema que vale ser citado é o programa THOTH (Vere, 1978; Dietterich and Michalski, 1979 e 1981)). Este sistema também é de propósito geral, sendo porém guiado pelas instâncias. As descrições geradas pelo THOTH são apenas conjuntivas. Como o INDUCE 1.2, ele também não é capaz de produzir generalizações discriminantes.

Uma vantagem que ele apresenta com relação ao SPROUTER é a sua capacidade de tratar associações N a 1 entre os objetos referenciados; vale no entanto a ressalva de que isto é possível ao custo da geração de alguns termos muito pouco significativos nas generalizações. Por exemplo, uma das descrições que ele é capaz de gerar para as instâncias da Figura 4.2 é a mostrada a seguir (Dietterich and Michalski, 1979 e 1981, onde se verificam algumas informações muito pouco relevantes:

- 1) (em-cima 1 2) (branca 1) (médio 1) (9 1) (5 3 4)  
(preta 3) (7 3) (6 3) (branca 4) (grande 4) (8 4)

Há um objeto branco e médio em cima de outro objeto e há dois objetos relacionados de alguma maneira, tal que um é preto e o outro é grande e branco.

Em Vere (1978) são descritas interessantes aplicações desse sistema, em situações de aprendizagem onde o conceito procurado é na realidade uma regra de mudança de estado (e os exemplos são os próprios estados). Em uma dessas aplicações (obtenção da regra de formação da voz passiva em inglês a partir de três frases nas duas vozes, isto é, na ativa e na passiva, o sistema SPROUTER também havia sido testado, sendo que o THOTH se mostrou bem mais eficiente que o outro.

Observando que o espaço das generalizações forma uma "ordem parcial" (Dietterich, 1983), Mitchell (1979 e 1982) propõe-se um algoritmo denominado "algoritmo de eliminação de candidato" para explorar esta característica. Para Mitchell o espaço de generalizações é chamado "espaço das versões". Partindo do conjunto de todas as instâncias (que pode ser visto como o conjunto das generalizações mais específicas, chamado "S") e de um conjunto onde todos os termos são variáveis (que é o conjunto das generalizações mais gerais, chamado "H"), passa-se a considerar sucessivamente cada uma das instâncias com as quais se vai generalizando o conjunto S e especializando o conjunto H. Quando os conjuntos se igualam, seus elementos fornecem as generalizações procuradas. Este método não gera conceitos disjuntivos e, apesar de ser guiado por instâncias, ele pôde ser adaptado para operar em ambiente ruidoso.

Vale ainda citar o sistema de aprendizagem de regras de produção desenvolvido por Hedrick (1976) e um outro do mesmo tipo desenvolvido por Weiss et alii (1979); o sistema META-DENDRAL (Buchanan and Mitchell, 1978; Buchanan and Feigenbaum, 1978) que infere regras de clivagem de moléculas a partir de dados de espectrometria de massa; o sistema desenvolvido por Lemerle-Loisel e Kodratoff (1983) destinado a exemplos com estrutura interna bastante definida; e o sistema desenvolvido por Winston (1973) que utilizava redes semânticas para representar as instâncias, o qual teve um papel fundamental para as pesquisas posteriores na área.

## CAPÍTULO 5

### GERAÇÃO DOS MODELOS DE REGRAS: A SOLUÇÃO PROPOSTA

#### 5.1 - A EVOLUÇÃO DO TRABALHO: DO RECONHECIMENTO DE PADRÕES ISOLADOS À APRENDIZAGEM INDUTIVA

##### 5.1.1 - O MECANISMO PRINCIPAL DA PRIMEIRA SOLUÇÃO ADOTADA

O ponto de partida para a solução apresentada neste trabalho envolvia, basicamente, a idéia de fazer uma cobertura das regras segundo o que elas tivessem de comum, considerando as cláusulas "isoladamente". Tal procedimento estava mais próximo à abordagem de formação de conceitos utilizada no sistema TEIRESIAS do que às abordagens tradicionais de aprendizagem indutiva, o que é natural considerando a origem do trabalho. Os aspectos comuns às regras seriam buscados seletivamente, de acordo com os "*tipos de padrões*" definidos pelo usuário, ou seja, pelos arranjos das primitivas conceituais (atributo, objeto, função e valor) usadas na representação, que fossem de interesse para o usuário numa certa aplicação. Os padrões obtidos nessa busca definiriam o conteúdo dos modelos. Por exemplo, se fosse definido o tipo de padrão AOV, relativo ao arranjo de atributo-objeto-valor, como a referência para a construção do modelo de um conjunto de regras, somente seriam buscadas nas regras ocorrências de padrões do tipo AOV, sem ter de se considerar o campo de função das cláusulas das regras.

Depois do particionamento do conjunto das regras, o qual definiria classes a partir daquelas que compartilhassem uma cláusula de ação, iniciar-se-ia a busca das ocorrências comuns de acordo com os tipos de padrões especificados e de forma a se conseguir padrões os mais amplos possíveis, isto é, aqueles que cobrissem o maior número de regras da classe considerada. Assim, partir-se-ia da tentativa de obtenção da melhor solução, a qual seria um padrão comum a todas as regras, ou seja, um padrão que cobrisse o conjunto totalmente. No entanto, se esse padrão não existisse, proceder-se-ia então à busca de padrões refe

rentes a agrupamentos de regras, agrupamentos estes com cardinalidade igual ao número total de regras do conjunto, menos uma unidade. Se os padrões obtidos nesta etapa cobrissem todas as regras, já se teria uma solução; caso contrário, tentar-se-ia cobrir as regras ainda não cobertas, procurando padrões em agrupamentos das regras contendo as regras não-cobertas em parte ou em todo, sendo agora a cardinalidade desses agrupamentos igual ao número total de regras do conjunto menos duas. Novamente, se todas as regras pudessem ser cobertas nesta etapa, ter-se-ia chegado a uma solução; caso contrário, o processo continuaria analogamente aos passos anteriores, considerando, a cada etapa, agrupamentos de regras com cardinalidade sucessivamente decrementada de uma unidade. Por exemplo, partindo de um conjunto de 10 regras, tentar-se-ia obter um padrão que envolvesse todas elas; se ele não existisse, tentar-se-ia cobrir as regras por meio de padrões relativos a 9 regras. Se ainda isto não fosse possível, tentar-se-ia a cobertura com padrões de 8 regras, e assim sucessivamente. Se, somente quanto se considerassem agrupamentos de 6 regras algumas delas (por exemplo, 3 regras) fossem cobertas, as 7 restantes poderiam ser cobertas em agrupamentos de no máximo 5 regras.

#### 5.1.2 - OS TESTES FEITOS E A CONCLUSÃO A QUE SE CHEGOU

O que se deseja enfatizar com a explicação dada na seção anterior é o caráter essencialmente voltado para um reconhecimento de padrões isolados nas regras, ou seja, a ocorrência de características fragmentadas, sem se cogitar de qualquer aspecto integrado, contextual, entre estas características. De fato, ao se fazer alguns testes simples do programa implementado para realizar o processo descrito, ficou bastante nítida a inadequação do processo.

A seguir serão apresentados os testes realizados, os quais se basearam na representação de três configurações no mundo dos blocos, conforme a Figura 5.1. A regra associada a cada configuração geométrica é obtida interpretando a sua descrição como um conjunto de premissas

que, se satisfeitas, levam à conclusão de que esta configuração é instância de um conceito genérico. Considere-se que os nomes dos objetos expressos por letras representam variáveis, as quais podem ser instanciadas para números representando nomes específicos.

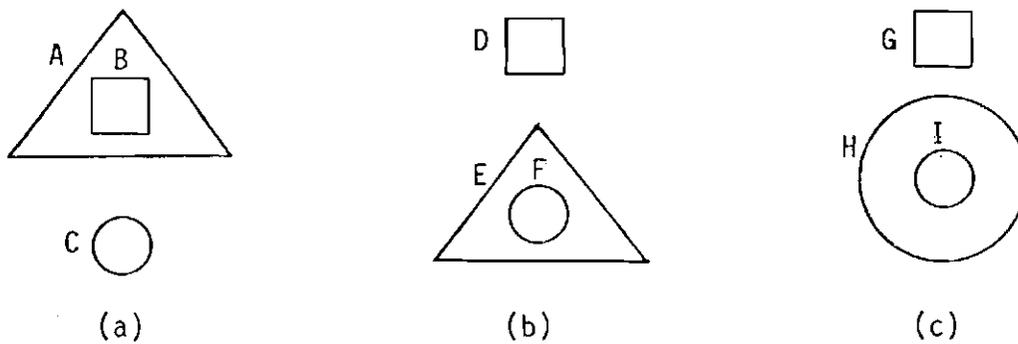


Fig. 5.1 - Figuras geométricas que definem o primeiro conjunto de exemplos.

As figuras geométricas que definem a Figura 5.1 estão descritas, respectivamente, pelos seguintes conjuntos de cláusulas:

Representação para a Figura 5.1 (a)

- (forma A é triângulo)
- (forma B é quadrado)
- (forma C é círculo)
- (tamanho A é grande)
- (tamanho B é pequeno)
- (tamanho C é pequeno)
- (tamanho A é por-fora)
- (posição B é dentro)
- (posição A é acima)
- (posição B é acima)
- (posição C é abaixo)

Representação para a Figura 5.1 (b)

(forma D é quadrado)  
(forma E é triângulo)  
(forma F é círculo)  
(tamanho D é pequeno)  
(tamanho E é grande)  
(tamanho F é pequeno)  
(posição E é por-fora)  
(posição F é dentro)  
(posição D é acima)  
(posição E é abaixo)  
(posição F é abaixo)

Representação para a Figura 5.1 (c)

(forma G é quadrado)  
(forma H é círculo)  
(forma I é círculo)  
(tamanho G é pequeno)  
(tamanho H é grande)  
(tamanho I é pequeno)  
(posição H é por-fora)  
(posição I é dentro)  
(posição G é acima)  
(posição H é abaixo)  
(posição I é abaixo)

Tomando o tipo de padrão AOFV (atributo - objeto - função - valor) como referência, rodou-se o programa para dois conjuntos de nomes de objetos. No primeiro teste, para o conjunto:

A = 1	D = 2	G = 2
B = 2	E = 1	H = 1
C = 3	F = 3	I = 3

obteve-se o seguinte padrão:

(forma 2 é quadrado)  
(forma 3 é círculo)  
(tamanho 2 é pequeno)  
(tamanho 1 é grande)  
(tamanho 3 é pequeno)  
(posição 1 é por-fora)  
(posição 2 é acima)  
(posição 3 é abaixo).

No segundo teste, para o conjunto de nomes:

A = 1	D = 4	G = 7
B = 2	E = 5	H = 8
C = 3	F = 6	I = 9

nenhum padrão do tipo AOFV foi obtido, como era de se esperar, já que cada objeto tinha um nome diferente.

O primeiro ponto que se deseja deixar claro é que o padrão que se abstrairia das regras era "dependente do nome que se atribuisse aos objetos", fato este inadmissível, pois os padrões devem depender apenas da natureza (expressa pelos atributos) dos objetos em si, e não de seus nomes. Se, para resolver este problema, não se consideram os objetos para a busca dos padrões, o que equivale a uma mudança do

tipo de padrão AOFV para o AFV, obtém-se uma informação do padrão existente, mas perde-se toda a informação associadas aos objetos, seja sobre as relações existentes entre elas, seja sobre as suas características individuais, através dos atributos que a eles se referem. Este é o segundo ponto que se deseja salientar. A título de exemplo, rodou-se o programa para o tipo de padrão AFV e obteve-se um padrão que é o mesmo padrão anterior, excetuando-se o campo de objeto.

Nesse padrão obtido, foi possível descobrir que existia algum objeto com a forma quadrada e que algum outro era pequeno, mas não se conseguiu descobrir que se tratava do mesmo objeto, ou seja, perdeu-se a coerência contextual, relativa aos itens de informação descobertos. Com esta constatação, resolveu-se dar um salto qualitativo com relação à abordagem cogitada inicialmente, o que redirecionou a busca da solução na linha dos modelos usuais de aprendizagem indutiva, nos quais se objetiva uma análise integrada dos objetos nos exemplos e não a utilização de unidades fragmentadas de informação.

Os fatos observados nos testes realizados forneceram a direção ao "salto", que era a especificação de um mecanismo que garantisse a independência dos padrões com relação aos nomes dos objetos, sem que estes tivessem de ser desconsiderados do processo. A idéia de fazer "*associações de objetos*", conforme implementada no sistema SPROUTER e explicada na Seção 4.3, pareceu adequada e foi adotada. Lembrando, a idéia é considerar associados os objetos (de regras diferentes) que compartilhem alguma propriedade. Por exemplo, considerando o segundo conjunto de nomes de objetos especificados nesta seção, os objetos 1,5 e 8 podem ser considerados associados, pois compartilham o fato de serem grandes e de terem um outro objeto em seus interiores. Na Seção 5.4.1 será discutida a criação e utilização das associações de objetos.

### 5.1.3 - OS DIRECIONAMENTOS INICIAIS JÁ NO ENFOQUE DE APRENDIZAGEM INDUTIVA

A julgar pela descrição do sistema SPROUTER em Hayes-Roth e McDermott (1977, 1978), onde as associações de objetos foram utilizadas de forma explícita, naturalmente se é levado a pensar que esta idéia representa apenas uma etapa do mecanismo de aprendizagem implementado. No entanto, em seu contexto mais amplo, o processo de associar objetos deve ser interpretado não como uma técnica particular, mas como uma etapa fundamental dentro de qualquer processo de aprendizagem indutiva, qualquer que seja a estratégia de busca no espaço de generalizações. Tal afirmação é decorrente do fato de que a aprendizagem indutiva deve, em alguma altura do processo, permitir a comparação entre duas descrições quaisquer, o que só é possível através de uma referência comum, que é justamente o conjunto dos objetos associados.

Se o mecanismo de indução é guiado pelos dados, existe diretamente a necessidade de associar os objetos para comparar os exemplos, a fim de poder criar as generalizações. Por exemplo, isto se verifica tanto no sistema SPROUTER quanto no sistema THOTH, através da criação, respectivamente, do "conjunto dos rótulos de caso ("case labels") com parâmetros associados" e do "conjunto de pares de literais associados". Por outro lado, se o mecanismo de indução é guiado por modelos, a criação das generalizações não depende de fazer associações dos objetos (embora estas possam ser utilizadas), havendo necessidade, porém, no processo de validação das generalizações, seja com relação à completude, seja quanto à consistência. Por exemplo, o sistema INDUCE 1.2, na etapa de determinação dos atributos que farão parte das generalizações, utiliza diretamente a associação de objetos na geração do espaço de valores de atributos. Além disto, na verificação de completude, tanto das descrições generalizadas puramente estruturais quanto dos vetores de atributos, "deverão" estar sendo feitas associações entre os objetos das descrições parciais generalizadas e os objetos dos exemplos, apesar deste fato não ser citado em Michalski (1980, 1983), nem em Dietterich

e Michalski (1979, 1981). O sistema INDUCE 1.2 apresenta, portanto, duas situações em que a associação de objetos é realizada: uma em que se visa obter as generalizações, e outra em que se pretende validá-las.

Como será mostrado nas seções posteriores, o modelo definido é guiado pelos dados, isto é, pelos exemplos (no presente caso considera-se a inexistência de contra-exemplos; sobre este aspecto, ver Seção 6.2). A guiagem pelos dados deve-se principalmente à característica de "*incrementalidade*" (conforme a Seção 3.1.2) desejável em uma aplicação à aquisição de regras de uma base de conhecimento, em que o modelo das regras deverá ser alterado a cada nova regra incorporada à base.

Visando levantar subsídios para a formulação de um modelo de aprendizagem indutiva guiado pelos exemplos, implementou-se uma versão modificada do mecanismo usado no sistema SPROUTER e foram feitos alguns testes. A principal modificação realizada foi a implementação de um mecanismo simplificado para permitir alguma disjunção interna nas cláusulas induzidas. A evolução deste mecanismo gerou o processo que acabou sendo usado no modelo final e que será descrito nas Seções 5.4.1 e 5.4.3.

Outras modificações incluíram:

- . A anexação de uma outra heurística ao processo de busca dos melhores conjuntos de associações de objetos. Esta heurística pressupunha que as associações de objetos estivessem em uma lista ordenada, decrescentemente, segundo o número de cláusulas que seriam geradas devido à associação em questão, caso ela fosse escolhida no processo de busca. A heurística consistia na especificação do número N que era interpretado como uma referência às N melhores associações que deveriam ser consideradas no processo de busca, ou seja, o algoritmo tomava as N primeiras associações da lista ordenada. Esta heurística apresentou resultados

positivos mas não foi utilizada, pois o que foi aproveitado no modelo final a respeito do mecanismo a que ela se vinculava não impunha a necessidade de uma busca heurística. Este fato ficará mais claro na Seção 5.4.3.

A consideração das regras de modo global na definição de um único espaço de busca, em vez de utilizar a definição de vários espaços incrementalmente. Tal procedimento foi adotado na tentativa de contornar o problema da propagação da melhor generalização ao longo dos exemplos, conforme explicado na Seção 4.3. Se por um lado o objetivo foi alcançado, por outro o espaço de busca aumentou muito, tornando o processo ineficiente.

Apesar de somente o mecanismo de disjunção ter dado frutos diretamente para o modelo de aprendizado finalmente definido, houve ganho conceitual como, por exemplo, na constatação de que seria difícil conseguir contornar o problema de propagação da melhor generalização, sem se alterar de forma mais drástica o mecanismo de aprendizagem do SPROUTER. Além disto, houve um ganho "colateral", resultado de um dos testes realizados, conforme é explicado a seguir.

## 5.2 - ALGUNS ASPECTOS SOBRE A APRENDIZAGEM DAS REGRAS E MODELOS

### 5.2.1 - MODIFICAÇÃO DA REPRESENTAÇÃO EM QUÁDRUPLAS

Um dos testes realizados para explicar o mecanismo do sistema SPROUTER envolveu as duas configurações no mundo dos blocos, conforme a Figura 5.2.

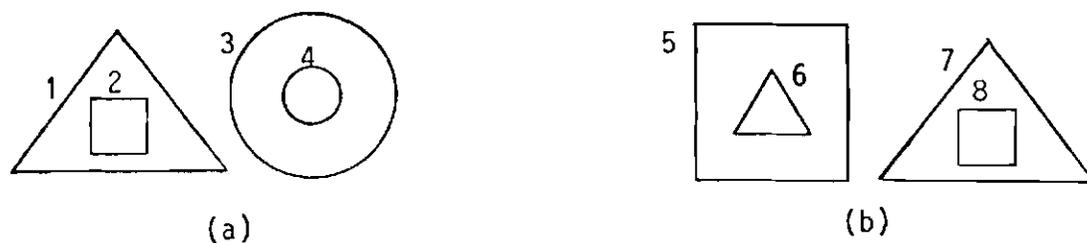


Fig. 5.2 - Figuras geométricas que definem o segundo conjunto de exemplos.

Considerando as figuras como instâncias de um exemplo genérico, resultam as seguintes representações:

Representação para a Figura 5.2 (a):

(forma 1 é triângulo)

(forma 2 é quadrado)

(forma 3 é círculo)

(forma 4 é círculo)

(tamanho 1 é grande)

(tamanho 2 é pequeno)

(tamanho 3 é grande)

(tamanho 4 é pequeno)

(posição 1 é por-fora)

(posição 2 é dentro)

(posição 3 é por-fora)

(posição 4 é dentro)

Representação para a Figura 5.2 (b):

(forma 5 é quadrado)  
(forma 6 é triângulo)  
(forma 7 é triângulo)  
(forma 8 é quadrado)  
(tamanho 5 é grande)  
(tamanho 6 é pequeno)  
(tamanho 7 é grande)  
(tamanho 8 é pequeno)  
(posição 5 é por-fora)  
(posição 6 é dentro)  
(posição 7 é por-fora)  
(posição 8 é dentro).

A generalização obtida, consideradas apenas as cláusulas sem disjunção interna, é a seguinte:

Generalização obtida para a Figura 5.2:

(forma 17 é triângulo)  
(tamanho 17 é grande)  
(posição 17 é por-fora)  
(forma 28 é quadrado)  
(tamanho 28 é pequeno)  
(posição 28 é dentro)  
(tamanho 35 é grande)  
(posição 35 é por-fora)  
(tamanho 46 é pequeno)  
(posição 46 é dentro).

Em texto corrente, a descrição generalizada indica que os exemplos apresentam um quadrado pequeno dentro de alguma figura, uma figura pequena dentro de alguma outra, um triângulo grande contendo uma

figura e uma figura grande contendo uma outra. No entanto, não se consegue descobrir, por exemplo, que a figura que está dentro do triângulo grande é o quadrado pequeno. Esta deficiência decorre da impossibilidade de explicitar numa mesma cláusula todos os objetos envolvidos em uma relação. Desta forma há necessidade de alterar o esquema da representação de modo a sanar este problema.

A modificação será no sentido de acrescentar um nodo campo semântico às "*cláusulas relacionais*" (por exemplo: (posição objeto-1 é dentro)), onde deverão ser representados os objetos que se relacionam ao objeto principal da cláusula. As "*cláusulas de propriedade*" (por exemplo: (tamanho objeto-1 é grande)) não precisam sofrer modificação. Daqui para frente, tratar-se-á por "*campo do objeto principal*" o campo de objeto já existente na representação, e por "*campo dos objetos secundários*" o novo campo recém-criado. Exemplificando, a representação de que "um objeto grande contém dois outros objetos pequenos" poderia ser a seguinte:

(tamanho objeto-1 é grande)  
(tamanho objeto-2 é pequeno)  
(tamanho objeto 3 é pequeno)  
(posição objeto-1 é por-fora (objeto-2 objeto-3)).

O fato de se ter alterado a representação das regras não suscita maiores problemas com relação ao sistema Diagnóstico (ao qual o presente trabalho se vincula), já que o processo de aprendizagem indutiva em si, que será apresentado, é compatível com o esquema original de representação, muito embora as generalizações possíveis possam vir a apresentar o problema mostrado no início da seção.

No exemplo anterior, a lista de dois elementos que aparece no campo secundário de objetos da cláusula relacional não reflete a existência de uma relação ternária entre os objetos 1, 2, 3, mas uma relação binária entre o objeto 1 e o conjunto dos objetos 2 e 3. De fa

to, a modificação proposta pra a representação é no sentido de permitir a explicitação apenas dos objetos envolvidos em uma *relação binária*.

Essa restrição na representação das regras é resultado de uma solução de compromisso entre a capacidade expressiva da representação e o aumento da complexidade em sua utilização. O ponto crucial é que o processo de inferência que se espera implementar na versão atual do sistema DIAGNÓSTICO ficaria muito complexo e, portanto, ineficiente, se fosse permitida a representação explícita de relações N-árias. Deve-se observar, no entanto, que a solução adotada não é tão restritiva quanto possa parecer, já que qualquer relação N-ária pode ser decomposta em N-1 relações binárias, se bem que ao custo de um pouco da naturalidade expressiva da representação original.

#### 5.2.2 - AS CLÁUSULAS RELACIONAIS PARCIAIS

Apesar de as cláusulas relacionais das regras deverem ser expressas especificando o campo secundário de objetos, não se deve impor este aspecto às generalizações obtidas. Desta forma, se não for possível induzir o fato de que "o objeto 1 contém os objetos 2 e 3", pelo menos será possível descobrir que "o objeto-1 contém algum outro objeto", apesar de não se conhecer quem é este último. Este tipo de cláusula, daqui para frente referida como "*cláusula relacional parcial*", é decorrente da "*capacidade de casamento parcial entre cláusulas*", sem ter de alterar o esquema de representação. Evidentemente esta capacidade será possível com alguma perda de eficiência no processo, mas será compensada com um ganho significativo no desempenho.

É interessante observar que a alteração no esquema de representação do SPROUTER, criando a chamada "PSR uniforme" (Hayes-Roth and McDermott, 1978), decorre principalmente do fato de se querer atingir a característica acima descrita. No entanto, esta mudança na representação ocasiona alguns problemas. O principal é a diminuição da eficiência do processo de indução, já que a PSR uniforme provoca um aumento do número de elementos das descrições dos exemplos, o espaço de bus

ca de generalizações torna-se significativamente maior. Outro problema é uma considerável perda de naturalidade na representação das instâncias. Por causa destes problemas, foi adotada no SPROUTER a estratégia de deixar que o usuário identifique as situações em que a utilização da PSR uniforme é necessária; se por um lado este procedimento contorna o problema de eficiência, por outro cria uma situação bastante incomfortável para o usuário.

Considerando que a capacidade de geração de cláusulas parciais atua no SPROUTER em uma situação combinatorial (aumento do espaço de busca devido ao aumento do número possível de associações de objetos) e como o modelo que será apresentado na Seção 5.4 não apresenta este problema, conseguiu-se manter a existência da referida capacidade a um ganho relativo (ao SPROUTER) de eficiência. Este fato ficará mais claro na Seção 5.4.1. Ainda vale dizer que o sistema THOTH também apresenta a capacidade de casamento parcial entre cláusulas, se bem que até em excesso, como se pode verificar em algumas cláusulas (literais) extremamente vagas que são induzidas, conforme mostrado na Seção 4.3. Quanto ao sistema INDUCE, a partir das referências consultadas (Michalski, 1980, 1983; Dietterich and Michalski, 1979, 1981) não é conclusivo dizer que este sistema não apresenta a referida capacidade, apesar de que aparentemente esta seja a conclusão.

### 5.2.3 - AS CLÁUSULAS RELACIONAIS INVERSAS

Como consequência de permitir a geração de cláusulas parciais, passa a ser interessante a explicitação das relações inversas já representadas. Por exemplo, supondo que um conjunto de regras apresente a relação "contém", seria interessante que estas regras também tivessem cláusulas que apresentassem a relação inversa "está contido". Desta forma, pelo menos seria possível inferir que "um objeto conhecido está contido em algum outro", apesar de não se saber qual é este outro, e mesmo que, nenhuma cláusula com a relação "contém" tenha sido induzida. Deve-se ressaltar que ao se falar em "relação inversa" deve

-se pensar na realidade, na "inversa do valor de atributo" da cláusula. Assim, com relação ao exemplo acima, ter-se-ia de fato o atributo "posição" (nome da relação), o qual estaria associado aos valores "dentro" e "por-fora" (relações propriamente ditas). Daqui para frente, uma cláusula que contiver uma relação inversa a uma outra será, com respeito a esta, uma "*cláusula relacional inversa*", ou simplesmente, uma "*cláusula inversa*".

Além do custo associado ao processamento das novas cláusulas, um outro problema com a representação das cláusulas inversas é a criação de uma redundância semântica na regra. Uma forma de resolver este último seria permitir que, interativamente, o sistema se encarregasse de requisitar do usuário o fornecimento das cláusulas referentes às relações inversas. Para tal, seria necessário que o sistema identificasse cada cláusula relacional e, para cada uma delas, pedisse ao usuário que lhe fornecesse a relação inversa à existente na cláusula, ou seja, o valor de atributo inverso ao existente nela; a partir desta informação o próprio sistema geraria as cláusulas inversas. Evidentemente a informação do usuário estaria vinculada à existência da relação inversa. Por exemplo, se o sistema se deparasse com uma regra que tivesse a cláusula

(posição objeto-1 é por-fora (objeto-2 objeto-3),

o usuário seria requisitado a fornecer a relação inversa à "por-fora", de forma a relacionar os objetos 2 e 3 com o objeto 1. Com este procedimento o usuário não teria de se preocupar com as relações inversas na criação das regras. Evidentemente o sistema deveria ter uma forma de armazenar as informações de associação entre as relações e suas inversas, de modo a não ter que refazer perguntas ao usuário. Na atual versão do sistema este procedimento interativo não está implementado.

#### 5.2.4 - O REFINAMENTO DA REPRESENTAÇÃO

O último ponto que se deseja comentar envolve a Figura 5.1 e as suas representações associadas. Observando-as, constata-se que um único atributo foi usado para expressar ambos os aspectos de posição relativa entre os objetos (a posição na vertical, e a relação de pertinência). No entanto, poderiam ter sido usados dois atributos: por exemplo, o atributo "posição" referente à posição na vertical, e o atributo "pos-pert" referente à relação de pertinência entre os objetos. A diferença entre as duas representações está no "*grau de refinamento dos atributos*"; na primeira representação considera-se todo o conjunto de valores dentro de um mesmo contexto, enquanto na segunda dá-se ênfase à ligação semântica mais específica existente entre os valores "dentro e por-fora" e entre os valores "acima e abaixo", considerados isoladamente.

Com relação ao esquema de aprendizagem que será apresentado, deve-se dizer que o grau de refinamento da representação tem influência direta sobre o seu desempenho e sua eficiência no que diz respeito à obtenção de cláusulas disjuntas. Em decorrência do maior número de valores que estão associados a um determinado atributo quando a representação é menos refinada, a possibilidade de existência de uma cláusula disjuntiva generalizada é facilitada, conseguindo-se, portanto, um aumento potencial de desempenho; por outro lado, o custo para obtê-las tende a aumentar. Ao longo da Seção 5.4 estas afirmações ficarão mais claras.

Finalizando, vale observar que o conceito de "refinamento" empregado está associado ao número de valores de um mesmo atributo, os quais podem ser associados a um objeto em uma mesma regra. Assim, o máximo grau de refinamento definível, é aquele em que cada objeto só pode estar associado uma única vez a cada atributo presente na regra.

### 5.3 - PARTICIONAMENTO DA BASE: GERAÇÃO DAS CLASSES DE EXEMPLOS

Toda regra de uma base de conhecimento pode ser vista como um conjunto de descrições (o conjunto das cláusulas de premissas), as quais, caracterizando uma certa situação, associam a esta um determinado conceito (a conclusão da regra). É segundo esta interpretação que se pode pensar em aplicar um mecanismo de formação de conceitos para a geração de modelos das regras.

Interpretando uma regra como exemplo de um determinado conceito, será necessário definir as classes que farão os papéis dos conjuntos de exemplos sobre os quais serão induzidos os conceitos desejados, isto é, os conjuntos sobre os quais serão gerados os modelos.

Duas são as etapas adotadas no processo de particionamento da base de regras, a fim de gerar as classes de exemplos. A primeira etapa consiste em formar grupos de no mínimo duas regras que compartilham um conjunto de cláusulas de conclusão. A idéia de organizar as classes a partir da informação das cláusulas de conclusão das regras foi tirada do próprio sistema TEIRESIAS, conforme explicações dadas na Seção 2.2.1. As classes geradas não precisam ser disjuntas. Por exemplo, se a base for constituída de cinco regras, tais que:

R1 conclui C1  
R2 conclui C2 e C4  
R3 conclui C3, C1 e C4  
R4 conclui C4 e C1  
R5 conclui C5, C2 e C3,

serão geradas as seguintes classes:

Classe C1       $\longrightarrow$       R1, R3 e R4  
Classe C2       $\longrightarrow$       R2 e R5  
Classe C3       $\longrightarrow$       R3 e R5



A *segunda etapa* para o particionamento da base de regras é considerar os conjuntos de regras obtidos no passo anterior, e particioná-los gerando grupos de pelo menos duas regras que referenciam na premissa o mesmo conjunto de objetos. Os grupos assim formados se rão as chamadas "*classes de exemplos*". Tomando por exemplo a classe C4 (obtida no exemplo do passo anterior), definida pelas regras R2, R3 e R4, se:

R4 referencia objeto-1 e objeto-2

R3 referencia objeto-1, objeto-2 e objeto-3

R4 referencia objeto-1 e objeto-2,

então será obtida a classe de exemplos C4', formada pelas regras R2 e R4, e classe C4", formada pela regra R3.

O que está por trás desse procedimento é o fato de que diferentes conjuntos de objetos que levam a uma mesma conclusão podem ser interpretados como um indício de que os contextos que envolvem esses conjuntos devem ser considerados em separado, o que reflete linhas causais distintas. Em outras palavras, está sendo suposto que diferentes linhas de raciocínio estejam sendo representadas na base, porém todas levando a uma mesma conclusão; daí considerá-las independentemente. Tal abordagem representa, portanto, um refinamento conceitual bastante significativo com relação ao esquema de particionamento utilizado no TEIRESIAS.

É interessante observar que a identificação (melhor se ria dizer "tentativa" de identificação, já que o mecanismo para tal é extremamente simples) de linhas causais na base de conhecimento está para o particionamento determinado somente pela igualdade das cláusulas de conclusão, assim como a abordagem de geração dos modelos de regras, que utiliza um mecanismo de aprendizagem indutiva, está para a primeira solução (o reconhecimento isolado de padrões, conforme cita

do na Seção 5.1.1) imaginada para a geração dos modelos. Ou seja, da mesma forma que ao partir para o enfoque de aprendizagem indutiva, consegue-se atribuir aos modelos uma informação mais integrada das unidades de informação (cláusulas) das regras, ao adotar a perspectiva de identificação de linhas causais na base, consegue-se atribuir aos modelos uma informação mais integrada das unidades de informação (regras) da base.

#### 5.4 - GERAÇÃO DOS MODELOS DE REGRAS: O MODELO DE APRENDIZAGEM INDUTIVA PROPOSTO

##### 5.4.1 - INTRODUÇÃO

Uma vez que as regras já estejam agrupadas em classes, isto é, já tenham sido definidos os conjuntos de exemplos, pode-se proceder à criação dos modelos propriamente ditos. Para cada conjunto de exemplos, o mecanismo de aprendizagem indutiva é acionado a fim de criar uma ou mais generalizações desses exemplos, as quais adquirem o status de modelo das regras (exemplos) em questão.

O ponto de partida para o mecanismo de aprendizagem de de desenvolvido foi a observação de que sistemas, tais como o SPROUTER e o THOTH, são guiados unicamente pelos exemplos. A informação sobre "on de" se deseja chegar no processo de generalização é muito genérica e pouco aproveitável. Na busca de uma maior especificação dessa informação, chegou-se ao conceito de "associações de máximo contexto" que passou a fornecer uma maior orientação à busca das generalizações.

Basicamente, a generalização de um conjunto de exemplos envolve cinco fases. Na *primeira*, são geradas todas as associações de objetos possíveis para os exemplos considerados. Poderá haver "*associa*ções *totais ou parciais*", dependendo se uma associação contiver ou não objetos de todas as regras. Na *segunda fase*, a partir de uma anãlise da informação relativa aos objetos que definem cada uma das associações geradas no passo anterior, são selecionadas as *associações e*

*conjuntos consistentes de máximo contexto*, conforme serão definidos na Seção 5.4.3, as quais passam a constituir um "foco" em torno do qual o modelo das regras é construído. Esta fase poderá se apresentar sob duas formas distintas, dependendo se as associações geradas na fase 1 representarem tanto cláusulas simples quanto disjuntivas, ou se representarem apenas cláusulas disjuntivas. Na *terceira fase*, são obtidas uma ou mais descrições, chamadas *descrições generalizadas*, as quais representam efetivamente a informação que foi obtida pela generalização dos exemplos, e que agora pode ser apresentada ao usuário de forma inteligível. Na *quarta fase*, se for possível, suplementa-se a informação presente nas descrições generalizadas já disponíveis. Finalmente, na *quinta fase*, algumas cláusulas disjuntivas podem ser ainda generalizadas através da troca de uma disjunção de conceitos por um outro mais geral que os englobe.

A seguir apresenta-se a Figura 5.3, que resume o que foi comentado acima; maiores detalhes serão vistos ao longo de toda a Seção 5.4.

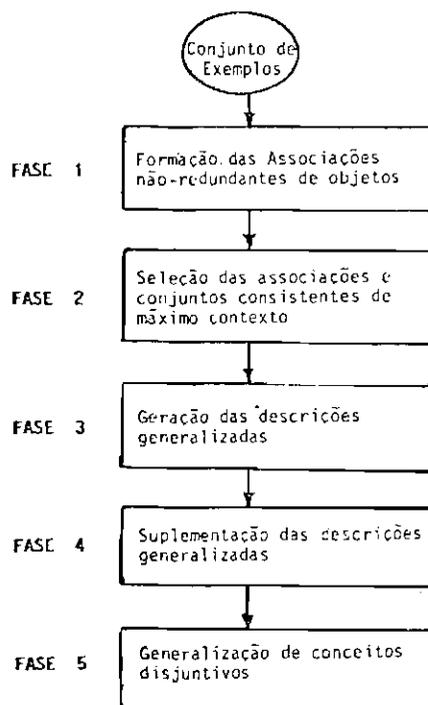


Fig. 5.3 - Esquema simplificado das fases do mecanismo de aprendizagem indutiva proposto.

Como a fase de seleção de associações se diferencia dependendo se elas referenciam apenas cláusulas disjuntivas ou não, a apresentação do modelo de aprendizagem será feita em duas etapas. Inicialmente, a partir do conjunto de exemplos mostrado na Figura 4.2 (que foi utilizado para a apresentação dos modelos de aprendizagem descritos no Capítulo 4), ilustra-se o caso em que é possível obter tanto cláusulas simples quanto cláusulas disjuntivas nas descrições generalizadas. Posteriormente, na Seção 5.4.7, acrescentando mais uma instância positiva ao conjunto original de instâncias, será possível explicitar o caso em que só é possível cláusulas generalizadas disjuntivas para o conjunto de exemplos.

#### 5.4.2 - FASE 1: FORMAÇÃO DAS ASSOCIAÇÕES NÃO-REDUNDANTES DE OBJETOS

Este passo consiste em gerar todas as associações totais e parciais de objetos que podem ser definidas para o conjunto de exemplos considerados. Cada associação indica um conjunto de objetos que tem alguma propriedade comum. A notação utilizada para representar as associações envolve a referência a esses objetos.

No caso de *uma associação total*, todas as regras estão representadas na associação por algum modelo. Por exemplo, para as instâncias da Figura 4.2, o identificador 350 representa uma associação total válida e indica que os objetos 3, 5 e 0, respectivamente, pertencentes às instâncias positivas (a), (b) e (c), possuem pelo menos uma *subcláusula AFV* comum, ou seja, que os conteúdos de seus campos de atributo, função e valor são iguais. Assim, a existência de uma associação total representa a garantia de que pelo menos uma cláusula simples será possível no modelo das regras.

No caso de *uma associação parcial*, não há representantes de todas as instâncias positivas na associação. Para as instâncias da Figura 4.2, tem-se, por exemplo, o identificador 26\*, que representa uma associação parcial válida, o qual indica que os objetos 2 e 6, respectivamente, pertencentes às instâncias (a) e (b), possuem pelo menos uma

subcláusula AFV comum, não possuindo, porém, qualquer subcláusula em comum com algum objeto da instância (c).

Para gerar as associações de objetos, todas as instâncias positivas de entrada são consideradas sucessivamente e, varrendo a representação de cada uma, monta-se o *quadro de associações de objetos* que caracteriza o dado conjunto de instâncias positivas. Esta estrutura é formada por um conjunto de *árvores de associações de objetos*, as quais nada mais são do que representações das subcláusulas AFV existentes no *conjunto* de exemplos, juntamente com os objetos para os quais a subcláusula é verdadeira. Cada ramo das árvores representa uma subcláusula AFV que caracteriza um aspecto das associações especificadas neste ramo.

Para as instâncias positivas da Figura 4.2, resultam as seguintes árvores de associação (Figura 5.4).

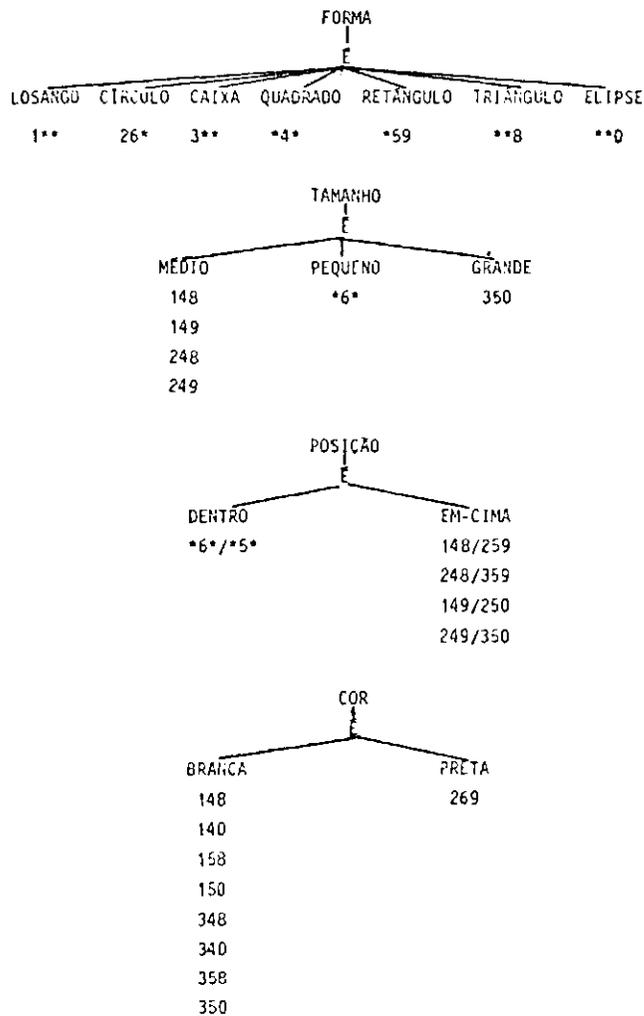


Fig. 5.4 - Árvores de associações para as instâncias da Figura 4.2.

Cada árvore de associação poderá ter várias funções de predicado e vários valores de atributo, mas somente um atributo, o qual acaba por defini-la. Como será visto adiante, se dois exemplos têm alguma informação em comum, no mínimo eles têm o mesmo atributo.

Na árvore relativa ao valor EM-CIMA, as associações separadas por uma barra são devidas ao fato de estarem no contexto de uma cláusula *relacional*. Assim, a associação 148/259, indica que os objetos 1,4 e 8 estão, respectivamente, EM-CIMA dos objetos 2, 5 e 9, eventos estes que ocorrem, respectivamente, nas instâncias positivas (a), (b) e (c).

À medida que cada instância vai sendo considerada para a construção do quadro de associações, deve-se fazer uma comparação entre os seus objetos a fim de detectar a possível presença de objetos idênticos, o que poderia levar à geração de *associações redundantes por igualdade de objetos*, isto é, associações com nomes diferentes mas que geram descrições exatamente iguais. Dois objetos são considerados idênticos se suas descrições forem as mesmas e se seus relacionamentos com outros objetos também forem os mesmos, ou seja, se em toda cláusula que um deles ocorrer houver, na mesma instância, uma outra em que o outro ocorra, tal que ambas possuam as mesmas ocorrências em todos os outros campos conceituais da representação. Nestes casos, para cada subcláusula AFV escolhida, novas associações são estabelecidas, sem que as já existentes devam ser atualizadas (expandidas) para cada um dos objetos envolvidos na redundância. O que deve ser feito é atualizá-las seqüencialmente, sem que se considere o mesmo objeto duas vezes consecutivas. Por exemplo, com relação ao exemplo mostrado, verifica-se que os objetos 6 e 7 são idênticos. Se ambos fossem considerados para a geração das associações relativas subcláusulas, as associações totais 269 e 279 teriam sido geradas redundantemente para a subcláusula (cor é preta), além de várias outras associações parciais. Como para este caso só havia uma única associação para as várias subcláusulas, apenas o objeto 6 foi construído na atualização das associações.

À medida que é criado o quadro de associações de objetos, vai-se computando a "*freqüência de ocorrência*" de cada associação, a qual é anexada ao identificador que define a associação. Para o cômputo deste parâmetro, basta que se incremente a freqüência de ocorrência da associação cada vez que ela for gerada em uma das árvores. Este parâmetro é fundamental para a escolha das melhores associações, conforme ficará claro na próxima seção.

Uma outra organização da informação levantada durante a geração do quadro de associações constitui o chamado *quadro de caracterização individual dos objetos* presentes ao conjunto de instâncias, que

são listas de todas as subcláusulas AFV que definem cada objeto do conjunto de instâncias positivas. As informações presentes neste quadro é que permitirão a construção das cláusulas disjuntivas que comporão as descrições generalizadas. Para as instâncias da Figura 4.2, tem-se que os objetos em questão são caracterizados conforme indica a Figura 5.5.

- 1 --- R1 ---- (forma é losango)  
(cor é branca)  
(tamanho é médio)  
(posição é em-cima 2)
- 2 --- R1 ---- (forma é círculo)  
(cor é preta)  
(tamanho é médio)  
(posição é em-cima 3)
- 3 --- R1 ---- (forma é caixa)  
(tamanho é grande)
- 4 --- R2 ---- (forma é quadrado)  
(cor é branca)  
(tamanho é médio)  
(posição é em-cima 5)
- 5 --- R2 ---- (forma é retângulo)  
(cor é branca)  
(tamanho é grande)
- 6 --- R2 ---- (forma é círculo)  
(cor é preta)  
(tamanho é pequeno)  
(posição é dentro 5)
- 7 --- R2 ---- (forma é círculo)  
(cor é preta)  
(tamanho é pequeno)  
(posição é dentro 5)
- 8 --- R3 ---- (forma é triângulo)  
(cor é branca)  
(tamanho é médio)  
(posição é em-cima 9)
- 9 --- R3 ---- (forma é retângulo)  
(tamanho é médio)  
(posição é em-cima 0)
- 0 --- R3 ---- (forma é elipse)  
(cor é branca)  
(tamanho é grande)

Fig. 5.5 - Quadro de caracterização dos objetos presentes à Figura 4.2.

#### 5.4.3 - FASE 2: SELEÇÃO DAS ASSOCIAÇÕES E CONJUNTOS CONSISTENTES DE MÁXIMO CONTEXTO

Conforme já foi comentado anteriormente, em se tratando de obter generalizações conjuntivas (descrições características), como é o presente caso, deve-se objetivar descrições as mais longas possíveis, ou seja, que tenham o maior número possível de elementos. Tais descrições constituem as chamadas *generalizações máximas*. Em geral este tem sido o objetivo de vários sistemas descritos na literatura.

Fazer referência às generalizações máximas simplesmente como "descrições mais longas" (como nas várias versões do INDUCE) não revela aspectos mais relevantes da natureza delas, os quais podem vir a ser usados de alguma maneira no processo indutivo. O ponto é: quanto mais clara e precisa a definição das generalizações máximas de modo a se poder enxergar melhor onde se quer chegar, tanto mais "focalizadamente" pode-se-ã orientar o processo de busca no espaço de generalizações. Além desta vantagem, vale observar que a precisão na conceituação das generalizações máximas permite enxergar mais claramente a orientação da da a um ou outro trabalho que se encontra na literatura. Neste sentido pode-se assinalar a referência feita em Dietterich e Michalski (1979, 1981), onde se comenta que tanto o sistema SPROUTER quanto o INDUCE 1.2 têm o mesmo objetivo em termos do tipo de descrição procurada, apenas com a diferença de que, em um caso e em outro, dão-se nomes diferentes a esses tipos. Como ficará claro em seguida, este fato não é verdadeiro, pois a caracterização do SPROUTER em termos das "abstrações máximas" representa um subconjunto das descrições consideradas no INDUCE 1.2.

A princípio não se pode dizer que a conceituação de "abstração máxima" utilizada no sistema SPROUTER tenha ocorrido dentro da ótica de tentar ir além da especificação, simplesmente, em termos do tamanho das descrições; no entanto, é inegável que ela representa um avanço neste sentido. De fato, o pensamento que norteia a obtenção das generalizações no SPROUTER pode ser entendido em termos não de buscar

simplesmente as descrições de maior número de elementos, mas as *mais longas, tal que nenhuma delas esteja contida em qualquer outra*. Assim, mesmo que um conjunto de associação de objetos possa gerar uma descrição relativamente longa, ela só será mantida se não houver uma outra que a contenha. Consegue-se com isto eliminar associações ou conjuntos de associações redundantes; já que isto representa uma redução do espaço de busca de generalizações, o processo de indução tende a ser facilitado.

No entanto, essa idéia de desconsiderar as descrições que estejam contidas em outras não é bem aproveitada no SPROUTER. De fato, somente se identifica neste sistema a *redundância trivial* que ocorre quando a diferença entre dois conjuntos de associações de objetos é a ausência, em uma delas, de uma ou mais associações que estão presentes na outra (é evidente que neste caso nem se faz necessária a verificação de quais cláusulas estão representadas por uma ou outra associação). Por exemplo, a associação representada pelo identificador 350-148 é trivialmente redundante com a associação representada por 350-148-269 e, portanto, seria desprezada pelo SPROUTER.

Se o conceito de generalização máxima é colocado de maneira ainda mais precisa, de modo que fique especificado de forma mais operacional o significado de uma descrição "estar contida em outra", passa a ser possível enxergar alguns aspectos que antes não se conseguia ver. Assim, em última instância, pode-se dizer que uma descrição está contida em outra se, e somente se, existir uma correspondência um a um entre as associações de objetos de cada um dos dois conjuntos de associações em questão, de tal forma que, estabelecida a correspondência, consiga-se reproduzir uma descrição a partir da outra apenas através da eliminação de uma ou mais cláusulas de uma das descrições. Neste contexto, o caso de redundância trivial é, evidentemente aquele em que a correspondência das associações é feita entre "mesmas" associações. Para o presente trabalho, as generalizações máximas significam, portanto, o conjunto das descrições generalizadas mais longas, tal que elas sejam diferentes de qualquer outra, e nenhuma dessas a contenha (segundo a conceitualização de pertinência acima).

Vale observar na Figura 4.3 que os nós 7 e 12 do grafo de busca gerado pelo SPROUTER apresentam as seguintes descrições conjuntas:

Nó 7 -----> {{ pequeno : ((a/c) (b/e)) }}

Existe um objeto pequeno e um outro também pequeno.

Nó 12 -----> {{ pequeno : ((a/e) (b/c)) }, { círculo : ((a/e)) }}

Existe um círculo pequeno e um outro objeto também pequeno.

Evidentemente, pela descrição em português, é fácil perceber que a descrição do nó 7 está contida na descrição do nó 12. De forma mais operacional, pode-se perceber este fato bastando que se faça, por exemplo, as correspondências

(a/c) <-----> (a/e)

(b/e) <-----> (b/c),

e se elimine o segundo termo da descrição do nó 12. Se no SPROUTER já se estivesse pensando nestes termos, eventualmente o nó 7 não poderia ter sido descartado.

Se uma descrição apresenta uma subcláusula AFV que uma segunda não possui, este tipo é suficiente para garantir que a primeira não está contida na segunda. Se a referida subcláusula ocorre dentro do seu maior "contexto de caracterização", isto é, acompanhada do maior número possível de outras subcláusulas, tal que todas referenciam uma mesma associação de objetos, este fato é suficiente para garantir que a descrição referente a esta associação não estará contida em qualquer outra que se possa obter. Em outras palavras, terá sido conseguida uma associação com a qual poderá certamente ser construída uma generalização máxima.

A determinação do maior contexto de caracterização relativo a uma subcláusula AFV qualquer, traduz-se, na realidade, na busca

da associação de objetos que referencia a subcláusula em questão, juntamente com o maior número de outras, evidentemente esta busca se processa no espaço das associações válidas para a subcláusula considerada. Se for aplicado este raciocínio a cada subcláusula AFV para a qual exista uma associação válida relativamente ao conjunto de instâncias positivas consideradas, tem-se a criação de um grupo de associações que garantem a possibilidade de criação de um conjunto de generalizações máximas. Exemplificando, considerem-se as instâncias da Figura 4.1. Se é tomada a subcláusula (tamanho é pequeno) como referência, por inspeção constata-se que a associação "ae" permite que esta subcláusula ocorra no seu maior contexto de caracterização, qual seja, o de ocorrer juntamente com a subcláusula (forma é círculo). Desta forma, garante-se a possibilidade de construção de uma generalização máxima, a qual deverá ter a associação "ae", além de uma outra consistente com ela; no caso, a única associação consistente é a associação "bc". Vale observar que a generalização máxima resultante do conjunto de associações "ae-bc" é a própria descrição relativa do nó 12 apresentada anteriormente.

Feitas todas essas colocações já deve ter ficado claro que essa é a idéia central explorada no mecanismo de aprendizagem que se apresenta. Na fase 2 do processo, a seleção das associações mais importantes refere-se justamente à obtenção do conjunto de associações que definem os contextos de caracterização mais amplos para cada uma das subcláusulas AFV existentes.

Genericamente, pode-se dizer que em cada ramo das árvores de associação em que haja associações totais, serão tomadas as associações que apresentam as maiores frequências de ocorrência; desta forma, cada subcláusula AFV estará sendo tomada no seu contexto mais amplo. No entanto, deve-se identificar duas situações. A primeira refere-se ao caso em que uma determinada subcláusula AFV caracteriza objetos em todas as instâncias, mas apenas 1 objeto em pelo menos uma delas. Neste caso, deve-se simplesmente tomar a associação de maior frequência, a qual será chamada *associação de máximo contexto de caracterização* para a dada subcláusula. Por exemplo, para as instâncias da Figura 4.2, a

subcláusula (tamanho é médio) só referencia o objeto 4 na instância (b); assim, pode-se tomar imediatamente a associação 148 que é a mais frequente (valor 3).

A outra situação define-se quando uma subcláusula caracteriza mais de 1 objeto em cada uma das instâncias. Desta forma existem conjuntos consistentes de associações que podem ser obtidos de modo a considerar simultaneamente as várias ocorrências da subcláusula em questão. Nesta situação, além da associação de máximo contexto definida acima, deve-se tomar também o *conjunto de associações consistentes de máximo contexto de caracterização*, que envolve as associações mais frequentes consideradas isoladamente, ou seja, toma-se o conjunto consistente, cuja soma das frequências das associações isoladas é a maior. Por exemplo, para as instâncias representadas na Figura 5.4, a subcláusula (cor é branca) existem os seguintes conjuntos de associações:

148-350; 140-358; 158-340; 150-348,

entre as quais é escolhido o primeiro, já que sua frequência ( $3+2=5$ ) é a maior de todos eles. Para esta subcláusula a associação de máximo contexto é a já selecionada "148", que apresenta frequência 3.

É interessante observar que os conceitos de "associações e conjuntos consistentes de associações de máximo contexto" constituem entidades distintas, embora conceitualmente próximas. Assim, eles permitem definir classes disjuntas de generalizações máximas. Isto ocorre, pois não necessariamente a associação de máximo contexto para uma certa subcláusula corresponde a uma das associações do conjunto consistente de contexto máximo dessa subcláusula.

Para determinar quantas serão as associações necessárias a fim de formar os conjuntos consistentes a uma dada subcláusula AFV, deve-se tomar o mínimo número de objetos diferentes que esta subcláusula referencia em cada instância. Isto acontece, pois uma associação con

sistente significa uma associação "1 a 1" entre os objetos. Por exemplo, a subcláusula (cor é branca) referencia nas instâncias (a), (b) e (c), respectivamente, 2, 2 e 2 objetos; donde conclui-se que as associações consistentes deverão ter 2 ( $\min \{2,2,2\}$ ) objetos.

Se uma associação escolhida ocorre em alguma cláusula relacional, pelo menos mais uma associação referenciada na cláusula também é escolhida. Se a associação escolhida é secundária (isto é, refere-se ao campo dos objetos secundários), a principal (isto é, a que se refere ao campo do objeto principal) também é escolhida. Por outro lado, se a associação escolhida é a principal, também são tomadas as associações que definem o conjunto consistente de maior frequência de ocorrência (a frequência de ocorrência do conjunto é a soma das frequências individuais). A quantidade de associações do conjunto é igual ao número mínimo de associações secundárias que ocorrem nas instâncias, considerando cada um dos objetos referenciados na associação principal. Ou seja, a quantidade de associações é obtida com o mesmo procedimento do parágrafo anterior. Esse procedimento permite que se possa desde já caminhar na direção do conjunto consistente relativo à generalização máxima procurada. Por exemplo, para o conjunto de instâncias já referido, quando a associação 350 é escolhida, ela leva consigo uma única ( $\min \{1,1,1\}$ ) associação. No caso, esta associação é a 249, sendo que ambas se relacionam através da subcláusula (posição é em-cima).

Quando há mais de uma associação ou mais de um conjunto consistente de associações com a mesma frequência de ocorrência, deve-se comparar as cláusulas simples que eles podem gear. Para tal, as associações ou conjuntos consistentes devem ser agrupados com outros que gerem as mesmas cláusulas simples. Para cada um destes agrupamentos, seleciona-se um representante, dando-se preferência à associação ou conjunto consistente que ocorrer no campo secundário de alguma subcláusula relacional.

O resultado da seleção das associações e conjuntos consistentes de máximo contexto, realizada no quadro mostrado na Figura 5.4,

gera as árvores apresentadas a seguir. Deve-se ressaltar que o quadro de associações de objetos "não é destruído". A Figura 5.6 a seguir pode ser vista como uma "máscara" sobre o quadro de associações.

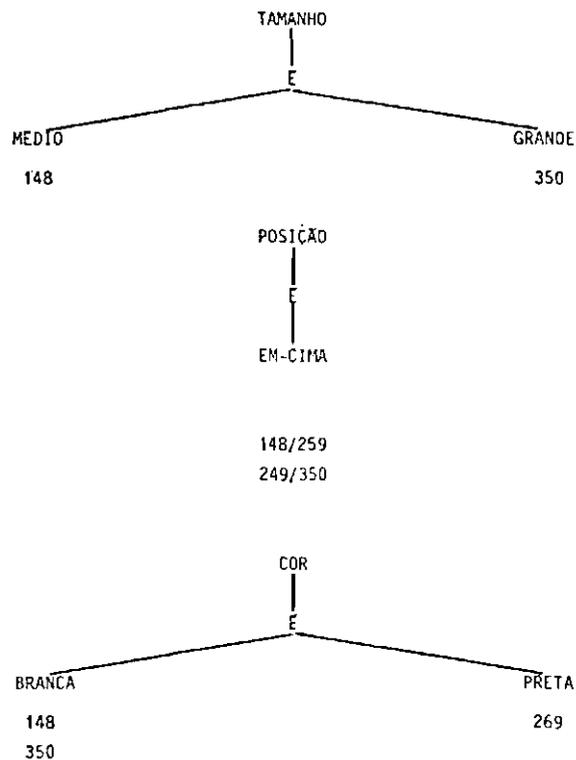


Fig. 5.6 - Árvores representativas das associações de máximo contexto para as instâncias da Figura 4.2.

Sobre o processo que acaba de ser apresentado, vários comentários devem ser feitos. Antes de mais nada, deve-se dizer que a fase 2 que acaba de ser descrita pode ser vista como um processo de "filtragem" do conjunto de associações válidas, que visa diminuir significativamente o número de associações com as quais é realizado o processo de busca dos conjuntos consistentes de associações, já que este mecanismo é muito caro em termos de processamento. No SPROUTER, a fim de contornar este problema combinatorial, optou-se pela eliminação de várias associações a cada etapa do processo, já que apenas uma das associações consistentes encontradas será mantida para a comparação com uma nova instância. Evidentemente, introduz-se um erro significativo de propagação da generalização, conforme já comentado na Seção 4.2.

Para o presente trabalho, a manutenção do quadro de associações visa justamente a eliminação desse erro de propagação. Assim, a cada instância positiva que o mecanismo recebe, as novas informações são incorporadas ao quadro, e a partir desse novo conjunto de associações são obtidas as associações de máximo contexto. É claro que este procedimento tem um custo. Embora ele não seja nada desprezível, não se espera que ele seja muito elevado pelas seguintes razões: em primeiro lugar, as buscas são realizadas em espaços distintos (relativos aos vários ramos da árvore de associações), cada um deles representando um subconjunto do conjunto total de associações válidas, ou seja, as associações são procuradas em um escopo combinatorial mais reduzido. Em segundo lugar, a busca em alguns desses espaços não passa da obtenção da associação de maior frequência, o que não chega a constituir um problema combinatorial. Finalmente, a busca do conjunto mais freqüente de associações consistentes (processo este que realmente tem natureza combinatorial) constitui uma forma mais eficiente de busca do que a realizada no SPROUTER, pois, como no presente caso se sabe exatamente de quantas associações deve ser composto cada conjunto consistente, a busca fica mais direcionada.

Uma outra característica da seleção de associações é que o conjunto selecionado não constitui um conjunto consistente, embora possa sê-lo. Em outras palavras, o processo de seleção permite o tratamento de associações  $N$  a  $1$ , ou seja, que vários ( $N$ ) objetos possam ser associados simultaneamente a um outro comum. Vale lembrar que no SPROUTER somente são permitidas as associações  $1$  a  $1$  que caracterizam as associações consistentes; por outro lado, no THOTH é possível fazer associações  $N$  a  $1$ , porém ao custo de provocar a criação de algumas descrições muito pouco significativas.

Até então argumentou-se apenas em termos da natureza das associações e conjuntos consistentes de contexto máximo, nada tendo sido ainda comentado a respeito da "qualidade" das descrições que estas associações podem gerar. Um ponto é definitivo: ambos os conceitos acima

podem conduzir a generalizações máximas significativas. No entanto, não se pode afirmar nem que a melhor generalização seja decorrente de uma associação ou conjunto consistente de máximo contexto, nem que essas associações sempre geram as melhores generalizações máximas. Tais constatações decorrem do fato de que a natureza das associações de máximo contexto e, em menor grau, a dos conjuntos consistentes conduzem o processo na direção de um certo isolamento das associações envolvidas, com relação a outras que são consistentes com elas.

Apesar dessa "tendência ao isolamento" ser uma realidade, ela não é de todo verdadeira. De fato, o isolamento completo só ocorre na situação em que a subcláusula considerada refere-se a uma cláusula de propriedade que caracterize apenas um objeto em pelo menos uma das instâncias. como, por exemplo, é a situação em que se procura a associação de máximo contexto para a subcláusula (tamanho é médio), no quadro de associações da Figura 5.3. Lembrando que foi imposto que toda associação escolhida "carrega" com ela as associações que a ela se relacionam via uma cláusula relacional qualquer, mesmo a situação que acaba de ser descrita não implica necessariamente isolamento total da associação de máximo contexto que tenha sido determinada. De fato, no exemplo acima, a associação 148 relativa à subcláusula expressa faz com que a associação 259 também seja selecionada.

Considerados os comentários feitos no parágrafo anterior, conclui-se que a geração de descrições generalizadas, baseadas na determinação das associações e conjuntos consistentes de máximo contexto, constitui um processo aproximado para atacar o problema, mas pode dar bons resultados devido à significância desses conceitos. De fato, os resultados empíricos têm confirmado esta expectativa.

#### 5.4.4 - FASE 3: GERAÇÃO DAS DESCRIÇÕES GENERALIZADAS

A geração das descrições generalizadas é conseguida através de processo semelhante ao descrito para o sistema SPROUTER. A partir das associações, procede-se a uma busca em largura objetivando os

conjuntos de associações consistentes; a diferença com relação ao processo usado no SPROUTER é que agora a busca será "não-heurística". Cada conjunto de associações consistentes representa uma descrição generalizada. Este processo será referido como *busca dos conjuntos de associações consistentes*. Para as associações e conjuntos consistentes de máximo contexto, obtidos na fase 2, obtêm-se três conjuntos de associações consistentes (Figura 5.7).

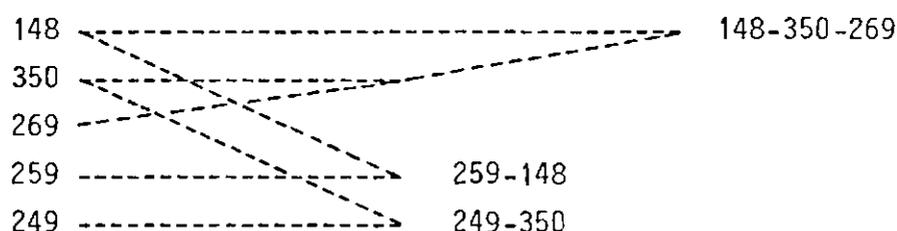


Fig. 5.7 - Representação da busca às associações consistentes relativas às instâncias da Figura 4.2.

Para a geração das descrições a partir dos conjuntos consistentes, uma observação deve ser feita com respeito às cláusulas relacionais. Quando a associação principal destas cláusulas se relacionar com uma associação secundária que não pertence ao conjunto consistente considerado, a associação secundária não aparecerá na cláusula. Em seu lugar, aparecerá uma disjunção de associações pertencentes ao conjunto consistente, se esse referenciar todos os objetos da associação secundária original; caso isto não seja possível, tem-se a criação de uma cláusula relacional (conforme a Seção 5.2.2), na qual não é possível explicitar qualquer associação secundária da cláusula. Conforme será visto a seguir, para o conjunto consistente 148-269-350, a associação 259 (que é secundária na cláusula que especifica a posição da 148) será substituída pela disjunção "269 ou 350".

Muitos conjuntos poderão ser gerados nesta fase. Assim, para fins de utilização das descrições generalizadas, é conveniente selecionar alguns dos conjuntos de associações gerados. Um critério razoável é tomar, para cada associação envolvida, o melhor conjunto consis

tente que a cubra, isto é, o conjunto consistente que a referencie e que, ao mesmo tempo, gere o maior número de cláusulas simples.

As três descrições generalizadas resultantes das associações consistentes geradas na Figura 5:7 são:

148-259 -----> (cor 148 é branca)  
(tamanho 148 é médio)  
(forma 148 é (losango ou quadrado ou triângulo))  
(posição 148 é em-cima 259)  
(forma 259 é (retângulo ou círculo))  
(tamanho 259 é (médio ou grande))  
(cor 259 é (branca ou preta))

148-269-350 -----> (cor 148 é branca)  
(tamanho 148 é médio)  
(posição 148 é em-cima (269 ou 350))  
(forma 148 é (losango ou quadrado ou triângulo))  
(cor 269 é preta)  
(forma 269 é (círculo ou retângulo))  
(tamanho 269 é (pequeno ou médio))  
(posição 269 é (em-cima ou dentro) 350)  
(cor 350 é branca)  
(tamanho 350 é grande)  
(forma 350 é (caixa ou retângulo ou elipse))

249-350 -----> (tamanho 249 é médio)  
(forma 249 é (círculo ou quadrado ou retângulo))  
(cor 249 é (preta ou branca))  
(posição 249 é em-cima 350)  
(cor 350 é branca)  
(tamanho 350 é grande)  
(forma 350 é (caixa ou retângulo ou elipse))

Vale observar que s̄o foi possível obter a cláusula disjun  
tiva

(posição 269 é (em-cima ou dentro) 350)

devido ao grau de refinamento (ver Seção 5.2.4) empregado na representa  
ção das instâncias, no que diz respeito ao atributo "posição".

Nota-se ainda que a disjunção "branca ou preta" é uma tautologia, contendo portanto pouca informação. Além do fato de que sua geração foi decorrente de uma situação particular, deve-se observar que em termos de uma aplicação em aquisição de regras, mesmo a disjunção tautológica constitui informação relevante, já que se traduz na forma de expectativa sobre qual a informação poderá chegar a uma nova regra (este fato ficará mais claro na Seção 6.1).

A descrição relativa a uma associação qualquer é obtida tomando como base os atributos comuns aos objetos que aparecem na associação; neste processo são utilizadas as informações que definem o quadro de caracterização dos objetos. Se, ao considerar um dado atributo, tanto as funções de predicado quanto os valores de atributo forem iguais para todos os objetos envolvidos, então obtêm-se uma clásula simples. Se, no entanto, houver mais de uma função ou mais de um valor diferente, obtêm-se uma clásula disjuntiva.

É interessante perceber uma consequência do fato de que as associações que aparecem como os objetos das descrições generalizadas podem receber quaisquer nomes, apesar de que têm sido usados nomes que referenciam os objetos originais. Desta forma, estas associações podem ser interpretadas como "variáveis" que estão representando os objetos originais. Tem-se, portanto, que este mecanismo de abstração dos objetos originais nada mais é do que uma forma dissimulada de utilização da regra de generalização da "troca de constantes por variável", como foi vista na Seção 3.7.2.

No mesmo sentido acima, isto é, da aplicação implícita de uma regra de generalização, deve-se notar que a existência de cláusulas simples nas descrições generalizadas pode ser interpretada como decorrente da utilização implícita da regra de "eliminação de condição". Segundo esta interpretação, as condições eliminadas são as cláusulas simples que ocorriam na mesma instância positiva, juntas com as cláusulas que foram "selecionadas" para a descrição generalizada.

#### 5.4.5 - FASE 4: SUPLEMENTAÇÃO DAS DESCRIÇÕES GENERALIZADAS

A penúltima fase do esquema de aprendizagem indutiva refere-se à *suplementação* das descrições já disponíveis. Este processo de suplementação ocorre em duas etapas consecutivas, conforme descritas a seguir.

A *primeira etapa* corresponde à anexação, aos conjuntos consistentes já determinados, de um conjunto de associações que se referem a pelo menos uma cláusula simples, ao mesmo tempo que sejam consistentes com os conjuntos existentes. Por exemplo, supondo que não tenha sido obtido o conjunto 148-269-350, mas apenas o conjunto 148-350, verifica-se que duas associações (269 e 279) são consistentes com o conjunto 148-350, ao mesmo tempo que se referem a uma cláusula simples (que é o caso da informação de que "a cor da associação 269 (ou da 279) é preta"). Assim, seriam gerados os conjuntos consistentes 148-269-350 e 148-279-350, em substituição ao conjunto 148-350. Em particular, como as associações 269 e 279 são redundantes, bastaria manter apenas uma delas (por exemplo, a 269), o que implicaria manter apenas o conjunto 148-269-350. Evidentemente, para executar o processo descrito, é necessário que sejam conhecidos os objetos que ocorrem em cada uma das cláusulas, informação esta obtida durante a fase 1 já descrita.

A *segunda etapa* de suplementação ocorre por conta da permissão de que algum atributo que não tenha sido referenciado em uma cláusula simples de uma descrição generalizada, pelo menos, possa sê-lo em uma cláusula disjuntiva. Esse processo é executado para cada uma das

descrições generalizadas obtidas, ou seja, cada descrição é suplementada por um de seus atributos independentemente das outras. Considerando cada uma delas sucessivamente, determinam-se inicialmente os atributos existentes nos quadros de associações que a descrição em questão não possui. Em seguida, a partir das associações parciais relativas a cada um desses atributos, determinam-se as associações totais consistentes com o conjunto referente à descrição considerada. Entre estas, selecionam-se as que tiverem o menor número de termos em disjunção.

À medida que as associações totais vão sendo formadas, vai-se computando a frequência de ocorrência de cada uma. Com as associações já obtidas, procede-se a uma busca do conjunto consistente de maior frequência. O número de associações deste conjunto deve ser no máximo igual ao número de atributos ausentes na descrição. Finalmente, as associações do conjunto obtido são acrescentadas aos conjuntos já existentes, bem como suas descrições associadas.

Resta dizer que o procedimento de geração das associações totais a partir das parciais será descrito na Seção 5.4.7, pois o contexto é mais apropriado, uma vez que para as descrições obtidas na fase 3 nenhuma suplementação deste tipo pode ser feita.

#### 5.4.6 - FASE 5: GENERALIZAÇÃO DE CONCEITOS DISJUNTIVOS

Observando que algumas das regras sintáticas de generalização (conforme a Seção 3.7) são criadas a partir da disjunção lógica, resolveu-se utilizar este fato a fim de tentar eliminar, na medida do possível, as disjunções presentes nas descrições generalizadas. A idéia é supor que toda disjunção de conceitos possa ter um conceito mais geral que a englobe.

Originalmente, em todos os trabalhos de Michalsky já referenciados, utilizam-se as regras efetivamente operando em um conjunto de conceitos primitivos, a fim de obter um conceito generalizado.

Nesta abordagem, o usuário tem o encargo de especificar tanto os conceitos primitivos quanto o generalizado, ou seja, ele é forçado a prever situações em que as generalizações poderiam ser feitas e deixá-las já explicitadas, a fim de que a regra de generalização possa operar. No enfoque utilizado no presente trabalho, estas situações são determinadas automaticamente quando é obtida uma cláusula disjuntiva em uma descrição qualquer; nestes casos, interpreta-se a ocorrência da disjunção como uma situação em que pode ser possível criar um conceito generalizado. Desta forma, tem-se a vantagem de o próprio programa ser capaz de conduzir o processo de obtenção do conceito generalizado.

Se o programa é feito sensível tanto às disjunções simbólicas quanto às numéricas, pode-se interpretar as suas ações de guiagem das generalizações como se ele estivesse realizando generalizações por "hierarquização de conceitos", ou por "especificação de intervalo numérico". Deve-se salientar, no entanto, que estas regras somente existem "conceitualmente".

Em termos da regra de hierarquização de conceitos, se o conceito generalizado já é conhecido, ele é submetido ao usuário, juntamente com os conceitos primitivos para uma avaliação, a partir da qual é feita ou não a troca da disjunção pelo conceito mais geral. Por outro lado, se é conhecida apenas a disjunção primitiva, o usuário é requisitado a fornecer o conceito mais geral. Para o presente caso considera-se válida apenas a hierarquização de valores de atributo.

Em termos da regra de especificação de intervalo numérico, a generalização será sempre proposta ao usuário. Para tanto, as situações possíveis (conforme descritas na Seção 3.7.2) já devem estar previstas internamente. Por exemplo, supondo as cláusulas

C1 : (atributo objeto = N1)

C2 : (atributo objeto = N2),

pode-se propor a seguinte generalização ao usuário:

(atributo objeto  $\in$  [min ... Nmax])

onde:  $N_{min} = \min \{ N1, N2 \}$

$N_{max} = \max \{ N1, N2 \}$ .

Resta dizer que deve ser mantido um registro das hierarquizações simbólicas que forem sendo adquiridas durante o processo, bem como das disjunções que não forem generalizáveis. Desta forma, evita-se fazer perguntas que já foram feitas em outra ocasião.

Permitindo-se a hierarquização de conceitos nas descrições obtidas na fase anterior, as descrições poderão ser generalizadas ainda mais. Supondo que seja relevante para o exemplo em questão a constatação de que "quadrados, triângulos ou losangos" são instâncias do conceito "polígono", então as descrições generalizadas obtidas transformam-se finalmente nas seguintes:

148-259 -----> (cor 148 é branca)  
(tamanho 148 é médio)  
(forma 148 é *polígono*)  
(posição 148 é em-cima 259)  
(forma 259 é (retângulo ou círculo))  
(tamanho 259 é (médio ou grande))  
(cor 259 é (branca ou preta))

"Há um polígono médio ou branco em cima de um retângulo ou círculo, cujo tamanho é médio ou grande e cuja cor é branca ou preta".

148-269-350 -----> (cor 148 é branca)  
(tamanho 148 é médio)

(posição 148 é em-cima (269 ou 350))  
(forma 148 é *polígono*)  
(cor 269 é preta)  
(forma 269 é (círculo ou retângulo))  
(tamanho 269 é (pequeno ou médio))  
(posição 269 é (em-cima ou dentro ) 350)  
(cor 350 é branca)  
(tamanho 350 é grande)  
(forma 350 é (caixa ou retângulo ou elipse))

"Há um círculo ou retângulo preto de tamanho pequeno ou médio, que está em cima ou dentro de um objeto grande e branco, cuja forma é um retângulo, elipse ou tipo-caixa; há ainda um polígono médio e branco que está em cima de um dos dois objetos anteriores".

249-350 -----> (tamanho 249 é médio)  
(forma 249 é (*polígono* ou círculo))  
(cor 249 é (preta ou branca))  
(posição 249 é em-cima 350)  
(cor 350 é branca)  
(tamanho 350 é grande)  
(forma 350 é (caixa ou retângulo ou elipse))

"Há um círculo ou um polígono de cor branca ou preta, que está em cima de uma caixa, retângulo ou elipse grande e branca".

O que se chamará *modelo de regras* daqui para frente refere-se ao conjunto das descrições generalizadas, obtidas após a generalização dos conceitos disjuntivos, como é o caso do conjunto das três descrições acima.

O que se deseja com a generalização de conceitos disjuntivos é que cláusulas disjuntivas se transformem em cláusulas simples (ou, pelo menos, que diminua o número de termos em disjunção), tornando ainda mais precisa a informação expressa nas descrições generalizadas. Este ganho na capacidade expressiva das descrições é significativo, em particular, na situação em que as cláusulas envolvidas referenciam associações que não ocorrem em qualquer cláusula simples (o que significa que as associações foram geradas por suplementação). Pode acontecer também a situação em que a disjunção não é transformada, mas apenas diminui o número de termos na disjunção. Nestes casos, apesar de não se conseguir transformar as disjunções em cláusulas simples (o que seria realmente desejável), ainda se tem uma situação relevante em termos de aplicação das generalizações, dado que se conseguiu obter descrições comuns às instâncias disponíveis.

A respeito do tratamento que se está dando à disjunção no presente trabalho, o esquema é operacional, apesar das restrições existentes; vale lembrar ainda que sistemas como o SPROUTER e o THOTH nem chegam a considerar a disjunção. O ponto principal que envolve o tratamento da disjunção é o problema de descobrir uma forma de "freá-la", ou seja, de limitar sua ação. Este ponto é essencial, pois a disjunção pura e simples de todos os elementos de um conjunto qualquer de instâncias positivas (a qual poderia ser chamada disjunção trivial), constitui uma descrição que, apesar de válida para o conjunto, não é relevante.

No modelo apresentado, a existência da disjunção é limitada, antes de mais nada, pelo fato de não ser permitida a disjunção no campo de atributo. Além disto, as disjunções são amarradas às cláu

sulas simples, já que elas somente são permitidas depois de terem sido procuradas as associações que geram cláusulas simples. Determinadas as cláusulas simples, as cláusulas disjuntivas são criadas mais no sentido de complementar a informação relativa a elas, seja para aumentar as informações sobre uma associação total, seja para definir uma associação parcial consistente com as totais já existentes. Somente quando não tiverem sido encontradas cláusulas simples é que as cláusulas disjuntivas adquirem um caráter mais autônomo, a partir do qual são criados conjuntos de associações consistentes parciais.

Ainda sobre a disjunção, deve-se observar que a utilização das regras de generalização da maneira como foram usadas sugerem que outras também possam ser usadas. De fato, a regra de "expansão limitada no domínio" (ver Seção 3.7.2) poderia ser usada, por exemplo, a partir da existência de duas cláusulas: uma que concluísse positivamente a respeito do valor V1 de um atributo genérico, e outra que concluísse negativamente sobre o valor V2. Supondo que o domínio para este atributo fosse { V1, V2, V3, V4, V5 }, o usuário poderia impedir a criação da disjunção.

(atributo objeto (é ou não-é) (V1 ou V2),

se fosse válida para o caso em questão a criação da cláusula simples

(atributo objeto  $\in$  { V1, V2, V3, V4, V5 }),

onde  $\in$  representa a função de predicado "pertence ao conjunto".

Neste momento é interessante comparar as descrições obtidas nesta última fase, com as obtidas no SPROUTER e no INDUCE 1.2, conforme mostradas no Capítulo 4. O primeiro ponto a observar é que, todas as associações consistentes geradoras de cláusulas simples que foram obtidas segundo a presente abordagem, também foram determinadas pelo SPROUTER, com a diferença que no presente caso as descrições geradas são significativamente mais relevantes.

Com relação ao INDUCE 1.2, das quatro descrições geradas por técnica seletiva, três delas (descrições 1, 2 e 4) correspondem às associações consistentes obtidas na fase 3, sendo que as descrições geradas nesta fase são mais completas que as obtidas por aquele sistema. A terceira descrição gerada pelo INDUCE 1.2 corresponde ao conjunto de associações consistentes 149-250, o qual não foi obtido pelo esquema mostrado neste capítulo, dado que este conjunto não é de máximo contexto. No entanto, isto não representa uma perda significativa, pois, enquanto no conjunto 149-250 as cláusulas simples descobertas permitem dizer que "existe um polígono médio em cima de outro objeto", no conjunto de associações 148-259 obtido descobre-se que "existe um polígono médio e *branco* em cima de outro objeto".

Com respeito à quinta descrição descoberta pelo INDUCE 1.2, para a qual foram usadas regras construtivas de generalização, a descrição obtida é realmente mais relevante. No entanto, deve-se observar que, apesar de nas descrições obtidas neste capítulo não aparecem cláusulas sobre o número de objetos com algum atributo (que constitui a regra construtiva mais frequentemente utilizada pelas versões INDUCE) ou o número de objetos presentes às instâncias, tais aspectos estão implicitamente expressos nos conjuntos de associações obtidos, devido à sua própria natureza. Por exemplo, para o conjunto consistente 148-269-350 seria possível anexar à descrição em português, já disponível, a informação de que "existem pelo menos 3 objetos em cada instância, sendo que pelo menos dois deles são brancos". Para que estas informações possam ter utilidade prática, deve-se ter um procedimento de geração de cláusulas que contenham essas informações, de modo que as cláusulas geradas sejam anexadas às descrições de origem. A restrição de que as informações obtidas são sempre na forma de um "mínimo" de ocorrência decorre do fato de que não se pode afirmar apenas pela descrição generalizada que não existem outros objetos com a propriedade considerada. Para alterar esta característica, deve-se ter procedimentos que permitam atuar diretamente nas regras, o que para o presente caso bastaria considerar os quadros de associações.

5.4.7 - ALTERNATIVA À FASE 2: SELEÇÃO DAS MELHORES ASSOCIAÇÕES QUANDO SÕ EXISTEM ASSOCIAÇÕES PARCIAIS NO QUADRO DE ASSOCIAÇÕES

Como a obtenção de uma descrição generalizada está condicionada à existência de uma associação total, algum procedimento deve ser definido no sentido de poder contornar o problema que aparece em situações de aprendizagem para as quais os quadros de associações obtidos somente apresentam associações parciais. A inexistência de associações totais nos quadros não indica que elas não possam existir, mas sim que qualquer uma eventualmente estabelecida representa uma descrição generalizada, formada apenas por cláusulas disjuntivas. É para estas situações que será definido um procedimento alternativo para a seleção das associações totais mais importantes (a fase 2 do esquema de aprendizagem), conforme adiantou-se na Seção 5.4.1. Evidentemente, como essas situações são exclusivas com relação àquelas consideradas na fase 2 (em que se pode obter generalizações formadas tanto por cláusulas simples quanto por cláusulas disjuntivas) os dois procedimentos que as caracterizam também o são.

O processo que se descreverá a seguir envolve *dois passos*: no primeiro, obtêm-se um conjunto de associações totais a partir das associações parciais presentes no quadro de associações; no segundo, procede-se à seleção propriamente dita das melhores associações, com as quais poderão ser definidos os conjuntos de associações consistentes.

Na presente situação, o desejável é obter descrições generalizadas "as mais longas e menos disjuntivas" possíveis, isto é, que apresentem o máximo número de cláusulas com o mínimo número de termos disjuntivos. A solução adotada é aproximada nos dois passos citados no parágrafo anterior: no primeiro através da desconsideração de todas as associações totais que não sejam relativas às cláusulas disjuntivas de menor número de termos, e no segundo, pela seleção apenas das associações que apresentam a maior frequência de ocorrência. A seleção das associações segundo o critério acima se justifica pelas

razões apresentadas na Seção 5.4.3. Quanto à geração das associações totais, algum tipo de critério de poda no espaço de busca dessas associações se faz necessário para viabilizar o processo na prática; o critério utilizado é o mais natural e tem dado bons resultados.

Deve-se observar que na atual situação, a existência de uma associação total está relacionada à existência de apenas um atributo comum às instâncias, e não de toda uma subcláusula AFV. Desta forma, deve-se esperar que as situações como a presente envolvam um conjunto muito maior de associações totais do que as situações já consideradas nas seções anteriores; em outras palavras, o espaço de busca para a seleção das associações totais mais relevantes tende a ser significativamente maior.

Para ilustrar o procedimento, considere-se que às instâncias da Figura 4.2 uma quarta instância positiva é acrescentada, a qual é mostrada na Figura 5.8.

É suposto que, para esta figura, o hachurado representa a cor verde.

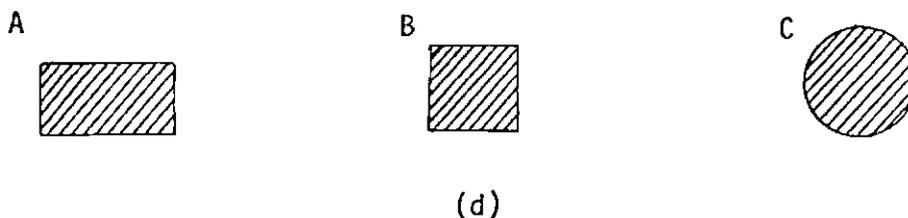


Fig. 5.8 - Quarta instância positiva para o conjunto de instâncias da Figura 4.2.

Observando a instância (b) da Figura 4.2, verifica-se que poderia ser representado o fato de que o objeto 6 está à esquerda do objeto 7; no entanto, isto não foi feito para que a comparação do resultado do esquema de aprendizagem proposto com os resultados dos sistemas SPROUTER e INDUCE 1.2 pudesse ser mais coerente, já que para estes sistemas não foi representada a relação entre os objetos citados. Não por este motivo, mas por uma mera questão de uniformidade, também para a instância acima a representação não envolverá o valor de atributo "à-esquerda", conforme pode ser visto a seguir:

Representação para a instância (d):

(forma A é retângulo)  
(tamanho A é pequeno)  
(cor A é verde)  
(forma B é quadrado)  
(tamanho B é pequeno)  
(cor C é verde)  
(forma C é círculo)  
(tamanho C é pequeno)  
(cor C é verde).

O quadro de associações resultante para as quatro instâncias das Figuras 4.2 e 5.8 está representado na Figura 5.9 a seguir.

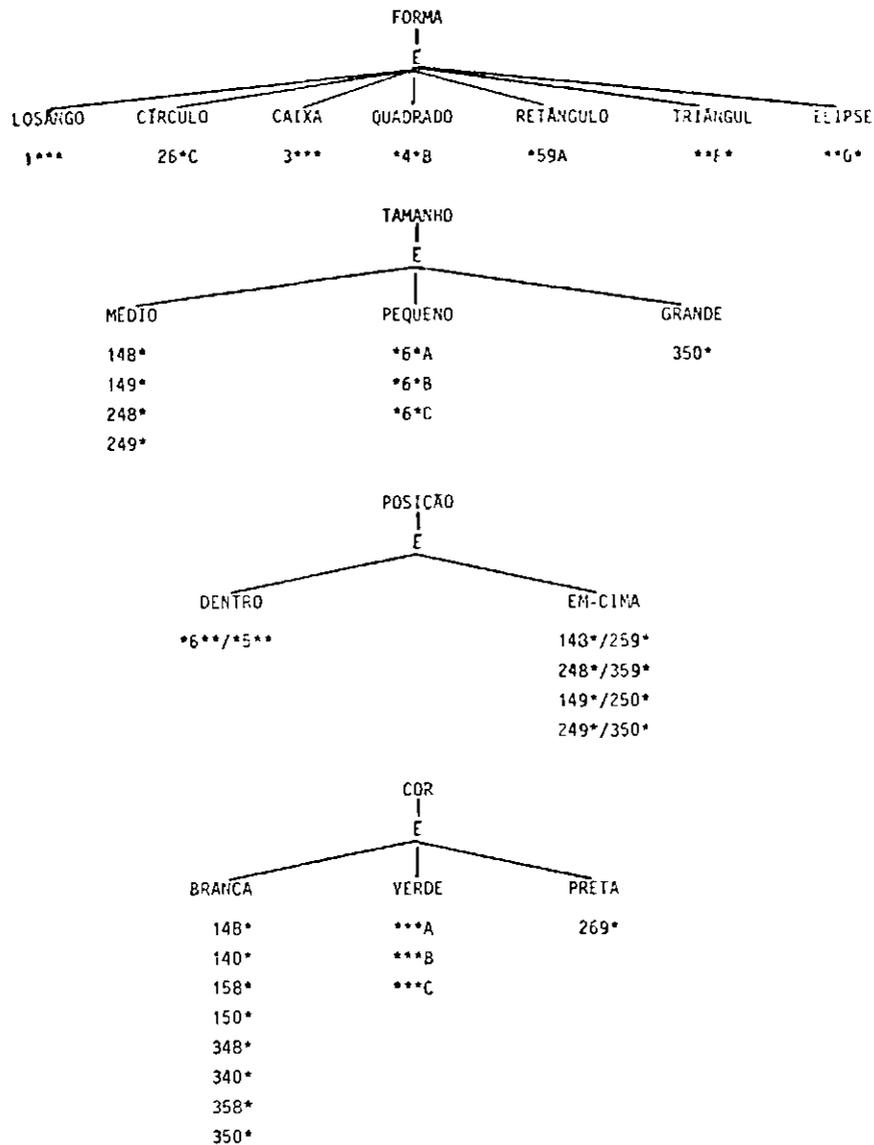


Fig. 5.9 - Árvores de associações para as instâncias das Figuras 4.2 e 5.8.

Para obter as associações totais a partir das associações parciais disponíveis (*primeiro passo* do processo), bastaria ir fazendo combinações entre estas últimas; no entanto, este procedimento geraria muitas associações, entre as quais várias certamente seriam irrelevantes. O critério utilizado para a determinação das associações totais é obter, para cada um dos atributos existentes, aquelas que geram *cláusulas disjuntivas com o menor número de termos em disjunção*, dando preferência às cláusulas com disjunção apenas no campo de valor. Este pro

cesso é realizado para cada atributo, e pode ser visto tendo *duas etapas*: uma em que se determina quais cláusulas disjuntivas de menor número de termos podem-se definir para o atributo considerado, e uma outra etapa em que a partir dessas cláusulas são obtidas todas as associações totais que podem gerá-las.

É importante notar que, além de seu status natural no contexto de alternativa à fase 2, o mecanismo que será descrito a seguir representa essencialmente o processo básico que permite a segunda etapa do esquema de suplementação das descrições generalizadas. Como se adiantou na Seção 5.4.5, o mecanismo que se descreve a seguir deve ser considerado como o próprio mecanismo que caracteriza a segunda etapa do processo de suplementação.

Para descobrir quais as cláusulas disjuntivas que apresentam o menor número de termos em disjunção, focalizam-se inicialmente as associações parciais existentes no contexto de cada par atributo-função (subcláusula AF) isoladamente, ou seja, inicialmente procura-se obter o menor número de termos de uma cláusula com disjunção apenas no campo de valor de atributo. Caso este tipo de cláusula não seja possível, passa-se a considerar cláusulas com disjunção no campo de função, ou seja, passa-se a utilizar todas as associações parciais existentes no quadro de associações. Através deste esquema de prioridade está sendo dada *ênfase às cláusulas disjuntivas em que a disjunção ocorra apenas no campo de valor*, mesmo que seja possível obter uma outra com menos termos em disjunção, mas em que a disjunção ocorra em campos de função e valor, ou apenas no de função. As cláusulas com disjunção apenas no campo de valor são preferíveis, pois nem sempre as funções componentes de uma disjunção (por exemplo: " $=$  ou  $\neq$ ", " $>$  ou  $<$ ") permitem a transformação da cláusula disjuntiva em uma cláusula simples, enquanto, pelo menos em princípio, isto é sempre possível para as disjunções apenas de valor.

Em ambos os casos citados acima, parte-se da formação de um *grupo-base* de associações parciais, o qual se define pelo agrupa

pamento de uma associação parcial relativa a cada subcláusula AF ou AFV, conforme estejam sendo procuradas disjunções apenas de valor, ou não. No grupo-base buscam-se os membros conjuntos de associações parciais que cobrem individualmente todas as instâncias positivas. Se o contexto é de disjunção apenas de valor, tais conjuntos de associações parciais já constituem os conjuntos que serão referidos como *conjuntos mínimos de cobertura*. Por outro lado, se o contexto é o de disjunção de função, então os conjuntos de cobertura são obtidos pela seleção, entre os recém-formados, daqueles que estão associados às cláusulas de menor número de termos em disjunção. Uma vez que cada conjunto mínimo de cobertura define uma cláusula disjuntiva para o atributo considerado, tem-se a possibilidade de criação de associações totais relativas a essas cláusulas. Considere-se, por exemplo, o atributo "FORMA" da Figura 5.9; tomando a primeira associação parcial para cada uma das subcláusulas AFV, obtêm-se o seguinte grupo-base de associações, o qual está acompanhado de suas respectivas subcláusulas AFV (Figura 5.10).

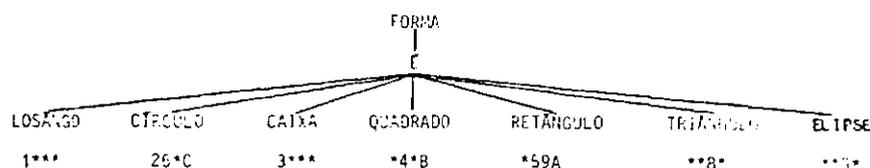


Fig. 5.10 - Representação do "grupo-base" de associações parciais para o atributo "FORMA", com o qual serão obtidas as cláusulas de menor número de termos disjuntivos.

Para a obtenção dos conjuntos mínimos de cobertura relativos ao grupo-base de associações parciais, deve-se, inicialmente, tentar formá-los a partir das associações do grupo-base, tomando-as 2 a 2; se não for possível, então elas serão tomadas 3 a 3, e assim por diante. Para o grupo-base acima, ao considerar suas associações 2 a 2, já é possível obter seus conjuntos mínimos de cobertura. Cada um desses conjuntos refere-se, portanto, à existência de cláusulas disjuntivas com dois termos em disjunção. Os conjuntos mínimos de cobertura relativos ao grupo-base e suas correspondentes cláusulas disjuntivas são:

- 1\*\*\*-\*59A ----> (forma objeto é (losango ou retângulo))
- 3\*\*\*-\*59A ----> (forma objeto é (caixa ou retângulo))
- 26\*C-\*59A ----> (forma objeto é (círculo ou retângulo))
- 26\*C-\*\*8\* ----> (forma objeto é (círculo ou triângulo))
- 26\*C-\*\*0\* ----> (forma objeto é (círculo ou elipse)).

Deve-se observar que, no caso geral, outros conjuntos mínimos de cobertura são possíveis, já que foi tomada apenas 1 associação parcial de cada subcláusula AFV para compor o grupo-base de associações. Como a existência de outros conjuntos mínimos de cobertura está relacionada à existência de outras associações totais, esses outros conjuntos também devem ser considerados. Deve-se notar que a existência deles é determinável diretamente pelas cláusulas disjuntivas obtidas: todas as associações geradoras de uma cláusula simples expressa na disjunção definem conjuntos mínimos de cobertura com as associações geradoras das outras cláusulas simples também expressas na disjunção. Como o presente caso não permite uma exemplificação, suponha-se que houvesse a associação parcial 2W\*C no mesmo ramo onde existe a 26\*C. Pelo que foi explicado, como 26\*C-\*59A constitui um conjunto mínimo de cobertura, 2W\*C-\*59A também constituirá.

Para obter as associações totais a partir das associações parciais, deve-se considerar sucessivamente cada uma das cláusulas disjuntivas obtidas anteriormente. Tomando os conjuntos mínimos de cobertura que se definem a partir dessas cláusulas, basta combinar os objetos referenciados em suas associações parciais. Para o atributo "FORMA", obtêm-se várias associações totais, as quais são mostradas abaixo, juntamente com as cláusulas disjuntivas que elas geram.

- 159A -----> (forma objeto é (losango ou retângulo))
- 269C; 269A; 259C; 259A ----> (forma objeto é (círculo ou retângulo))
- 359A -----> (forma objeto é (caixa ou retângulo))
- 268C -----> (forma objeto é (círculo ou triângulo))
- 260C -----> (forma objeto é (círculo ou elipse)).

Deve-se observar que as associações obtidas são todas derivadas dos conjuntos mínimos de cobertura relativos ao grupo-base de associações, já que neste caso não existem outros conjuntos de cobertura que se possa definir.

Pela razão já apresentada, o número de associações totais tende a ser bem maior na situação presente que na situação já discutida nas seções anteriores. De fato, para os outros atributos expressos na Figura 5.9, foram geradas as seguintes associações totais:

Atributo "TAMANHO" ----->	148A	148B	148C
	168A	168B	168C
	149A	149B	149C
	169A	169B	169C
	248A	148B	148C
	268A	268B	268C
	249A	249B	249C
	269A	269B	269C
	350A	350B	350C
	360A	360B	360C
Atributo "COR" ----->	148A	148B	148C
	140A	140B	140C
	158A	158B	158C
	150A	150B	150C
	348A	348B	348C
	340A	340B	340C
	358A	358B	358C
	350A	350B	350C
	269A	269B	269C.

Uma vez obtidas para cada atributo todas as associações totais responsáveis pela existência de cláusulas disjuntivas de número mínimo de termos, deve-se proceder finalmente à escolha das mais relevantes (que é o *segundo passo* do processo). O critério de seleção é tomar as associações que tiverem maior

freqüência de ocorrência para cada atributo existente; em caso de haver mais de uma com a mesma freqüência, todas elas são selecionadas. Para que o critério possa ser aplicado, há necessidade de computar as freqüências de ocorrência das associações a medida que elas forem sendo geradas no passo anterior. No presente exemplo todas as associações tais selecionadas apresentam freqüência 2, e são as seguintes:

269C; 269A; 350A; 350B; 350C; 148A; 148B; 148C.

Com essas associações pode-se obter as descrições generalizadas e, a partir das disjunções obtidas, passa-se à generalização dos conceitos disjuntivos. Pela própria natureza da etapa descrita, torna-se desnecessária a suplementação das descrições geradas. Observe-se que, diferentemente da situação descrita na Seção 5.4.3, não são selecionados aqui os conjuntos de associações consistentes de máximo contexto. Esta opção é decorrente do problema combinatorial que surgiria se esses conjuntos fossem permitidos, já que o espaço de busca na presente situação é muito grande. Este procedimento certamente imporá uma queda de desempenho ao processo de indução, queda esta que poderá ser contornada na fase de suplementação das descrições generalizadas.

Como ilustração, é mostrada a seguir a descrição mais longa e menos disjuntiva que se pôde (e se pode) obter para as instâncias consideradas. Um fato curioso que pode ser notado com esta generalização diz respeito à situação que foi criada pelo fato de ser permitida a generalização conceitual de termos disjuntivos: tendo partido da situação em que somente seria possível obter cláusulas disjuntivas, acabou-se obtendo uma cláusula simples, conforme se pode constatar a seguir:

148B-269C-350A -----> (forma 148B é polígono)  
(tamanho 148B é (médio ou pequeno))  
(cor 148B é (branca ou verde))  
(forma 269C é (círculo ou retângulo))

(tamanho 269C é (médio ou pequeno))  
(cor 269C é (verde ou preta))  
(forma 350A é (caixa ou retângulo ou elipse))  
(forma 350A é (pequeno ou grande))  
(cor 350A é (branca ou verde)).

"Existe um polígono médio ou pequeno, cuja cor é branca ou verde; existe ainda um círculo ou retângulo médio ou pequeno, de cor verde ou preta; e existe uma caixa, retângulo ou elipse pequena ou grande, cuja cor é branca ou verde".

Deve-se observar que o mecanismo apresentado pode levar à geração de uma descrição generalizada que é a própria disjunção trivial; numa situação de aplicação caberá ao usuário a conveniência de usá-la ou não.

Vale dizer finalmente que a generalidade da informação contida em descrições básica ou totalmente disjuntivas não constitui um fator de desmerecimento a estas descrições. De fato, a possibilidade de obtê-las pode adquirir um papel essencial, principalmente no caso de se estar operando com instâncias a respeito das quais não se pode garantir a existência de cláusulas simples em comum (como é o caso quando se opera sobre regras de uma base de conhecimento).

#### 5.4.8 - JUSTIFICANDO A GERAÇÃO DAS DESCRIÇÕES GENERALIZADAS A PARTIR DOS CONJUNTOS CONSISTENTES DE ASSOCIAÇÕES

Um último comentário que se deve fazer refere-se à justificativa do fato de se ter utilizado o procedimento que caracteriza a fase 3. Observando o resultado da fase 2, constata-se que as associações selecionadas já constituem por si sô descrições generalizadas das instâncias originais. Desta forma "não haveria" a necessidade de proceder à busca dos conjuntos consistentes, processo que caracteriza a fase 3, o que implicaria um ganho considerável de eficiência.

Inicialmente o que se pretende nesta seção é argumentar que essa observação não é verdadeira, sendo porém aceitável no caso em que o domínio das instâncias consideradas é "pouco estruturado". Por *estruturação do domínio* deve-se entender a existência de relações entre os objetos desse domínio. Um domínio pouco estruturado se caracteriza (informalmente), portanto, pelo fato de seus objetos serem definidos basicamente por atributos de propriedade, praticamente não havendo a necessidade de atributos relacionais; um exemplo deste tipo de domínio será apresentado no próximo capítulo, onde cada instância possui pelo menos oito objetos, e apenas dois ou três deles se relacionam em cada uma das instâncias. O domínio de figuras geométricas com o qual vem se trabalhando neste capítulo representa um domínio bastante estruturado, ou seja, onde os atributos relacionais despenham um papel bastante importante para a caracterização dos objetos.

Em um domínio qualquer é possível detectar, nas descrições generalizadas, três formas possíveis de relacionamento (estruturação) entre os objetos. A mais comum delas é a que ocorre em uma cláusula simples, por exemplo a que ocorre na cláusula (posição 249 é em-cima 350). A segunda situação é aquela em que dois ou mais objetos podem se relacionar em uma cláusula disjuntiva na qual a disjunção ocorre no campo de função, ou no campo de valor de atributo; por exemplo, com disjunção no campo de valor tem-se a cláusula (posição 269 é (em-cima ou dentro) 350). Finalmente, tem-se a situação de uma cláusula disjuntiva na qual a disjunção ocorre no campo dos objetos secundários; um exemplo para esta situação é a cláusula (posição 148 é em-cima (269 ou 350)). Vale levantar neste momento uma restrição às versões INDUCE que não ocorre no presente trabalho; nesse sistema somente é possível gerar cláusulas com o primeiro tipo de relacionamento entre os objetos.

Considerando o caso extremo de um domínio totalmente não-estruturado, se para uma tarefa qualquer de reconhecimento de novas instâncias forem utilizadas as descrições generalizadas, obtidas diretamente das associações selecionadas na fase 2, não haverá quaisquer pro

blemas de reconhecimento, dado que cada uma das descrições poderá ser utilizada isoladamente para reconhecer uma ou mais propriedades de um dado objeto. No entanto, se o domínio apresenta algum grau de estruturação, o que representa sem dúvida a situação geral, o reconhecimento terá de ocorrer a partir de conjuntos consistentes de associações. Neste ponto já se teria a necessidade de criar um procedimento especial que permitisse a formação de conjuntos consistentes de números quaisquer de associações, de modo que fosse possível reconhecer quaisquer graus de relacionamentos entre os objetos (por exemplo, "obj-1 está em cima do obj-2, o qual está em cima do obj-3 que está dentro do obj-4 etc.").

Apesar desse novo procedimento que teria de ser definido, a princípio não haveria maiores problemas. No entanto, essa abordagem de utilizar as descrições diretamente das associações selecionadas apresenta uma limitação quanto ao fato de que a terceira forma de estruturação de objetos (conforme citada acima) jamais poderá aparecer em uma das descrições. Isto se deve à própria forma com que surgem os relacionamentos existentes nessas descrições, isto é, através dos conjuntos consistentes de máximo contexto, ou através de uma associação que impõe a seleção de outras associações consistentes com aquela (conforme o que foi explicado na Seção 5.4.3); em qualquer dos dois casos não é possível aparecer a disjunção no campo do objeto secundário.

O que se perde com a incapacidade de geração de cláusulas, com o tipo de disjunção acima, é a possibilidade de uma maior especificidade na caracterização de um objeto envolvido em uma determinada relação. Por exemplo, tomando por base as descrições da Seção 5.4.6, observa-se para o conjunto de associações 148-259 que o objeto 259 (o qual está em baixo do objeto 148) é caracterizado totalmente por cláusulas disjuntivas; no entanto, para o conjunto 148-269-350, verifica-se que o objeto que está em baixo do 148 (representado pela disjunção das associações 269 e 350) está melhor caracterizado, já que os dois objetos envolvidos apresentam também cláusulas simples que os caracterizam. Esta possibilidade de referenciar um objeto de forma mais específica é

particularmente desejável em termos de uma aplicação em aquisição de novas instâncias, a partir das descrições que já se possui, conforme será visto na Seção 6.1. Com os comentários feitos, conclui-se que a obtenção das descrições de forma mais direta (a partir das associações selecionadas na fase 2) não conduzem, necessariamente, às descrições mais precisas a respeito das instâncias; evidentemente, quanto menos estruturado for o domínio, menor será a perda de desempenho, eventualmente possibilitando essa abordagem.

Um último ponto que se deseja comentar refere-se a uma característica interessante que resulta do fato de se utilizar descrições a partir de conjuntos consistentes e não diretamente a partir das associações selecionadas. Uma vez que os conjuntos consistentes representam associações "1 a 1" entre os objetos das instâncias, passa a ser possível a obtenção de informações que dependam exclusivamente desta característica, mas que não tenham sido explicitadas nas instâncias originais. Ou seja, está se falando de "generalização construtiva", nos termos em que foi definida no Capítulo 3.

As informações que estão implícitas nas descrições são, por exemplo, o número de objetos com determinados atributos, o número de objetos que se relaciona a um outro de determinada forma, tal que possua determinados atributos, ou ainda a posição que um certo objeto ocupa em uma seqüência ordenada por alguma relação (por exemplo, se "obj-1 está em cima do obj-2 que, por sua vez, está em cima do obj-3", pode-se querer determinar qual é o objeto que está em cima de todos os outros). Para que isto possa ocorrer, deve-se definir procedimentos especiais que, operando sobre as descrições induzidas, geram as cláusulas com as informações procuradas. O interessante é que exatamente os três tipos de informações acima constituem todas as regras de generalização construtiva utilizadas no INDUCE (Michalski, 1980 e 1983). A diferença entre as utilizações é que neste último os procedimentos devem operar diretamente a partir das instâncias, gerando os atributos desejados, os quais passam a ser considerados no processo de generalização, augmentan

do, no entanto, o espaço de busca. Apesar desta vantagem, a idéia apresentada não possui a generalidade do esquema utilizado no INDUCE. Se no presente trabalho fossem definidos os referidos procedimentos, teria sido possível obter, para o exemplo desenvolvido, as informações de que "existe pelo menos um objeto médio (conforme já se mencionou na Seção 5.4.6) e pelo menos dois objetos brancos, ou que existem um ou dois objetos médios e exatamente dois objetos brancos, ou ainda que o objeto mais de cima é o representado pela associação 148.

## CAPÍTULO 6

### OUTROS ASPECTOS DOS MODELOS DE REGRAS

#### 6.1 - APLICAÇÃO NA AQUISIÇÃO DE NOVAS REGRAS

Uma vez que o problema que originou este trabalho foi a aquisição de conhecimento, vale observar como os modelos de regras poderiam ser aplicados com este fim.

Nos poucos trabalhos sobre aquisição de conhecimento relatados na literatura percebe-se a ênfase em fazer a aquisição baseada em modelos descritivos das regras da base, como é o caso do TEIRESIAS e do MORE (Nowlan et alii, 1985). Aliás, como já foi comentado, o presente trabalho visa justamente caminhar nesta direção através da utilização efetiva de um mecanismo de aprendizagem indutiva. Vale observar que a utilização de meta-níveis de representação no processamento do conhecimento em geral tem sido vista cada vez mais como uma característica essencial ao comportamento inteligente de seres humanos e computadores (Barr, 1979).

Como os modelos representam alguma informação existente na base de conhecimento sobre um conjunto de regras, ao se deparar com uma nova regra pertencente a este conjunto durante a aquisição de conhecimento, as informações do modelo podem se traduzir na geração de *expectativas* sobre o conteúdo desta nova regra. Por sua vez, estas *expectativas* podem se traduzir em *sugestões* para que o usuário acrescentasse à nova regra alguma informação que eventualmente esteja faltando.

No presente trabalho tais sugestões podem se referir a todo um conjunto de cláusulas que descrevem um objeto ausente na nova regra, ou a um conjunto de cláusulas, sejam relacionais ou de propriedade, que complementam a descrição de um objeto nesta regra. No TEIRESIAS a expectativa de presença de alguma informação refere-se à ocorrência de um ou mais atributos na nova regra, dado que outros tenham

ocorrido, ou seja, apenas é possível o caso de geração de expectativa a respeito de propriedades de um objeto. Nenhuma das outras duas situações caracterizadas acima são possíveis. Além dessa vantagem com relação ao TEIRESIAS, no que diz respeito à aquisição de conhecimento, o presente mecanismo de geração de modelos de regras apresenta ainda uma vantagem, uma vez que a sugestão dada ao usuário pode ser *mais refinada* (referenciando "valores de atributos", por exemplo), já que a representação utilizada nos modelos é mais rica que a usada naquele sistema.

Como não foi explicitado no capítulo anterior, deve-se observar neste momento que, a cada uma das *classes de regras* obtidas pelo particionamento da base de conhecimento, são acrescentados os nomes de cada uma de suas regras componentes, o conjunto de objetos comuns referenciados nas suas premissas e o modelo das regras propriamente dito.

Para verificar qual é o modelo correspondente a uma nova regra, deve-se mapeá-la a uma das classes de regras existentes, tomando-se então o seu modelo correspondente. Para selecionar a classe, deve-se tomar aquela que for definida pelo maior número de cláusulas (de conclusão), tal que as cláusulas de conclusão da nova regra sejam englobadas e todos os objetos da premissa desta regra sejam referenciados no modelo correspondente à classe. O mapeamento deve ser para a classe de maior número de cláusulas, para que seja possível a criação de expectativa (apesar de bastante rudimentar) a respeito da presença de outras cláusulas na conclusão da nova regra. Por exemplo, se a conclusão da regra é formada pela cláusula "C1" e a classe "C1-C2" pôde ser referenciada, pode-se considerar que falta a cláusula "C2" na regra, o que permite que ela seja requisitada ao usuário. Se este não julgar válido o acréscimo de qualquer outra cláusula à conclusão da regra, deve-se selecionar a classe definida diretamente pelas cláusulas da regra em questão. Por outro lado, se alguma das cláusulas sugeridas for acrescentada, a classe definida deve levar em conta o novo conjunto de cláusulas de conclusão.

Observando-se que a seleção do modelo para uma nova regra envolve tanto as cláusulas de conclusão quanto as de premissa (através da referência a objetos comuns ao modelo e à nova regra), verifica-se também neste aspecto uma vantagem da presente abordagem com relação ao TEIRESIAS, já que neste último apenas se consideram as cláusulas de conclusão.

Para exemplificar o que foi comentado acima, considere-se que as três instâncias da Figura 4.2 são regras que definem uma classe de exemplos e que uma nova regra vai ser adquirida, de tal modo que esta regra (representada na Figura 6.1) leve ao mesmo conceito que as três anteriores. Como todas as descrições envolvidas neste exemplo não são realmente regras de uma base de conhecimento, e portanto não estiveram sujeitas à imposição que elas devem referenciar os mesmos objetos nas premissas, este aspecto não é obedecido. Apesar disto, as descrições são válidas para os pontos que se deseja mostrar.

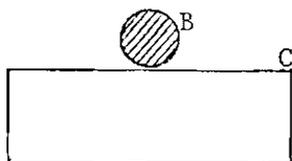


Fig. 6.1 - Representação gráfica de uma regra supostamente sendo adquirida, para ilustrar a utilização dos modelos de regras na aquisição de conhecimento.

Representação para a figura acima

(forma B é círculo)

(cor B é preta)

(tamanho B é pequeno)  
(posição B é em-cima)  
(forma C é retângulo)  
(tamanho C é grande).

Para que o modelo possa ser usado, deve-se fazer um mapeamento entre as associações presentes ao modelo e os objetos referenciados na nova regra. Isto pode ser conseguido diretamente pela execução dos procedimentos definidos para as fases 1 e 2 do mecanismo de aprendizagem. No entanto, como a expectativa que se quer criar tem de depender apenas do modelo de regras conveniente, não há necessidade de executar essas fases totalmente, já que elas envolverão muitas associações que não pertencem ao modelo.

Assim, o mapeamento entre as associações existentes e os objetos da nova regra pode ser feito considerando apenas as associações referenciadas no modelo. No entanto, mais do que todas as associações, somente serão consideradas aquelas que tenham pelo menos uma ocorrência em alguma cláusula simples. Quando não houver associações que ocorram em cláusulas simples, todas elas serão consideradas para o mapeamento.

Para a regra da Figura 6.1 e com base no exposto no parágrafo anterior, os mapeamentos possíveis são os seguintes:

<u>Associação</u>	<u>Objetos</u>
148	B, C
269	B
350	C
249	B

onde, por exemplo, o primeiro mapeamento indica que a associação 148 apresenta em comum com os objetos B e C as subcláusulas AFV (posição é em cima) e (cor é branca), respectivamente.

Para que sejam possíveis as sugestões sobre eventuais faltas de cláusulas de propriedades ou de cláusulas relacionais referentes a algum objeto, basta que sejam considerados sucessivamente cada um dos mapeamentos possíveis e comparados às cláusulas relativas à associação e ao novo objeto. Se a associação possuir uma cláusula que não aparece na descrição do novo objeto, está criada uma expectativa a respeito desta cláusula, a qual se traduzirá em uma sugestão ao usuário para que ele a forneça. Esta ação se justifica uma vez que a ocorrência de uma regra com essa descrição na base tipicamente apresenta a cláusula requisitada. Para o caso acima, apenas a associação 350 permite que se crie uma expectativa, sendo que a cláusula sugerida é

(cor C é branca),

que representa uma característica da associação 350.

Para que sejam possíveis as sugestões sobre a descrição completa de um objeto supostamente ausente, deve-se tomar sucessivamente todos os conjuntos consistentes referentes ao modelo, tais que eles sejam relativos a um número de associações maior que o número de objetos referenciados na nova regra. Para cada um desses conjuntos, estabelecem-se os mapeamentos consistentes possíveis, tais que eles levem em conta o maior número de associações do conjunto. Por exemplo, para as regras consideradas, o maior conjunto consistente possui três associações e é único (148-269-350); os possíveis mapeamentos consistentes a partir deste conjunto consistente são:

148,C	e	269,B
148,B	e	350,C
269,B	e	350,C.

De posse dos conjuntos consistentes de mapeamentos possíveis, deve-se selecionar o "melhor" deles. Para tal, a cada conjunto deve-se associar um número cujo objetivo é refletir quanto especificamente o conjunto referencia a nova regra. Será escolhido o mapeamento que

computar o maior número de pontos. Para a contagem dos pontos, deve-se considerar que uma cláusula simples de propriedade vale 1 ponto, e uma relacional simples ou relacional com disjunção apenas no campo de objeto secundário vale 2 pontos. A contagem de pontos para as cláusulas disjuntivas é obtida dividindo o valor básico da cláusula (1 ou 2) pelo número de termos em disjunção. O número de pontos para os conjuntos de mapeamentos acima são:

$$\begin{array}{llll} 148,C & \text{e} & 269,B & \longrightarrow (1 + 1) & + (1 + 1/2 + 1/2) = 4 \\ 148,B & \text{e} & 350,C & \longrightarrow (2) & + (1 + 1 + 1/3) = 4,3 \\ 269,B & \text{e} & 350,C & \longrightarrow (1 + 1/2 + 1/2 + 2/2) & + (1 + 1 + 1/3) = 5,3. \end{array}$$

Como o melhor conjunto de mapeamentos é o 269, B e 350,C, este é o escolhido. Assim, a sugestão que se pode fazer ao usuário é quanto ao acréscimo de mais um objeto, tal que ele tenha as características da associação 148. Observe-se que, se a abordagem de utilização das associações diretamente da fase 2 (como explicada na Seção 5.4.8) estivesse sendo usada, a possível ausência da associação 148 não teria sido notada, uma vez que o conjunto "148-269-350" não estaria disponível. De fato, uma outra desvantagem desta abordagem é exatamente a aplicação em aquisição de novas regras, pois uma vez que ela tem intrinsecamente a dificuldade de criação de conjuntos consistentes mais longos, torna-se mais difícil a obtenção de conjuntos que englobem outros totalmente. Na presente situação, além do referido conjunto já poder existir devido à própria característica da fase 3, foi possível criar uma cláusula que relaciona os três objetos, que é a cláusula

(posição 148 é em-cima (269 ou 350)).

## 6.2 - OS TESTES REALIZADOS

Foram realizados dois testes para ilustrar o que foi exposto no Capítulo 5. O primeiro visou uma aplicação real de geração dos modelos, tendo tomado como base o conjunto de regras de decisão sobre previsão pluviométrica em São José dos Campos, conforme apresen

tadas em Sandri (1986). O segundo visou uma aplicação especificamente em aprendizagem indutiva, a partir de um conjunto de instâncias que definem um grande espaço de busca para as descrições generalizadas; este exemplo foi tomado de Michalski (1983) e diz respeito a um conjunto de instâncias positivas de células cancerosas fictícias. Como o programa que implementa as idéias apresentadas ainda está em desenvolvimento (na linguagem LISP (Winston, 1981)), parte de ambos os testes foi realizada em computador e parte, manualmente.

Os resultados obtidos para as regras no domínio de Meteorologia, bem como as próprias regras utilizadas, estão mostrados no Apêndice A. Vale aqui chamar a atenção para alguns pontos. O principal deles é que, apesar de os modelos serem obtidos por um processo de aprendizagem indutiva, as descrições obtidas não devem ser consideradas como novas regras (induzidas) válidas para o domínio. Se bem que isto possa eventualmente acontecer, não se tem no entanto esta pretensão. Isto decorre do fato de que, para induzir novas regras para a base de conhecimento, é necessário considerar efetivamente o conhecimento do domínio, e não simplesmente observar as representações de algumas unidades (regras) desse conhecimento. O status real dos modelos é o de indicar quais informações aparecem tipicamente nas regras de determinada classe.

Outro aspecto importante diz respeito às generalizações de conceitos disjuntivos. Tal processo é realizado com o objetivo de criar descrições as "menos" disjuntivas possíveis, o que torna as descrições mais precisas. Quando se está operando normalmente a partir de um conjunto de instâncias, sempre que for possível o usuário deverá generalizar as disjunções. No entanto, quando se opera sobre uma base de regras, nem toda disjunção deve ser generalizada, pois poderá se perder alguma informação útil ao processo de aquisição de conhecimento. É o caso das cláusulas

(verão SJC (igual ou não-igual) T), e

(condições-de-pressão-nas-últimas-3-horas SJC (favoráveis ou não-favoráveis) T),

que se referem, respectivamente, às classes 4 e 2. Apesar de haver a tendência de eliminar essas cláusulas já que são tautológicas, isto não deve ser feito, pois seria perdida, por exemplo, a informação de que na classe 4 "existe uma referência ao atributo "verão" relativamente ao objeto SJC".

Com relação às classes de 6 a 12, que apresentam valores numéricos, os resultados poderiam ser mais naturais e precisos se a representação das regras envolvidas também o fosse; ou seja, se a representação dessas regras tivesse sido realizada em termos de especificação de intervalos (através da função de predicado  $\epsilon$  (pertence)), em vez de serem utilizadas as funções *maior-igual* e *menor-igual*.

Como a regra 6 (na realidade uma meta-regra) é composta de três cláusulas em disjunção, ela foi desmembrada em três outras (conforme a Seção 5.3), cada uma delas com uma das cláusulas disjuntivas, e definiram integralmente a classe 17. Por isto, o modelo da classe é igual à cláusula de conclusão que a caracteriza.

Finalmente, vale ressaltar que os modelos mais interessantes foram os relativos às classes 1 e 3, as quais se referem às regras de maior número de cláusulas. Em resumo, pelas colocações feitas conclui-se que a qualidade dos modelos reflete a qualidade das regras componentes da base.

Para o outro exemplo citado, o de geração de descrições características de um conjunto de instâncias positivas, usou-se o grupo de células cancerosas fictícias, mostrado na Figura 6.2.

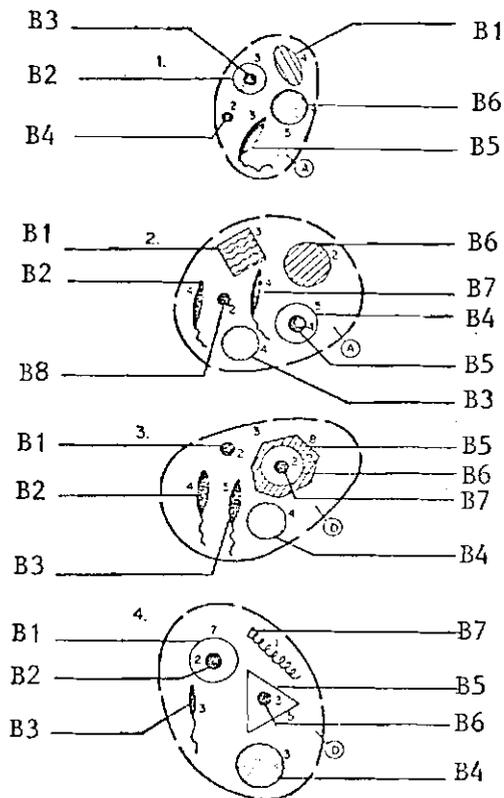


Fig. 6.2 - Conjunto de exemplos de células cancerosas fictícias.

FONTE: Michalski (1983).

Uma característica desse conjunto de instâncias é a sua pouca estruturação, conforme se pode constatar pelas descrições no Apêndice B. No entanto, ele possui uma característica interessante que são as várias disjunções que se pode obter, as quais permitem a generalização dos termos disjuntivos. Deve-se observar que uma mesma disjunção pode ser transformada em várias outras, cabendo unicamente ao usuário a escolha da mais adequada.

Várias foram as descrições características obtidas entre as quais algumas são apresentadas no Apêndice B. Vale observar que todas as descrições apresentadas em Michalski (1983) e obtidas pelo sistema INDUCE, também foram aqui obtidas, além de outras.

Entre as descrições mais importantes vale mostrar as seguintes, já em uma possível interpretação em português:

Conjunto consistente de associações

protoplasma-célula-5723-4812-6344-3576-2455.

Descrição

"Cada célula possui de 6 a 10 membranas, 6 ou mais organelas, e o seu protoplasma é do tipo A ou D.

Em cada célula existe:

.uma organela preta em forma de barco que está orientada para norte ou nordeste, possui uma cauda e tem peso 3 ou 4;

.uma organela preta de peso 2 e de forma circular;

.uma organela sombreada de peso maior ou igual a 2, e de forma circular;

.uma organela branca ou listrada de peso maior ou igual a 3, que apresenta forma circular ou poligonal, além de conter uma organela preta e circular de peso maior ou igual a 3."

Conjunto consistente de associações

protoplasma-célula-5723-2461-6344-3572-4816.

Descrição

"Cada célula possui de 6 a 10 membranas, 6 ou mais organelas e o seu protoplasma é do tipo A ou D.

Em cada célula existe:

- .uma organela preta em forma de barco que está orientada para norte ou nordeste, possui uma cauda e tem peso 3 ou 4;
- .uma organela sombreada de peso maior ou igual a 2, e de forma circular;
- .uma organela preta e circular, de peso 2 ou 3;
- .uma organela branca e circular, cujo peso é um número ímpar, que contém uma organela preta e circular, cujo peso é um número par."

Vale ainda dizer que as associações representadas pelos identificadores *protoplasma* e *célula* foram obtidas através da execução da segunda etapa do processo de suplementação das descrições.

### 6.3 - UMA PALAVRA SOBRE A GERAÇÃO DE DESCRIÇÕES DISCRIMINANTES

Embora toda a apresentação do modelo tenha sido em termos da geração de descrições características (isto é, as baseadas somente em instâncias positivas), é possível adaptar o esquema descrito, de modo que ele passe a gerar também descrições discriminantes, ou seja, descrições baseadas em instâncias positivas e negativas. Deve-se ressaltar, no entanto, que ainda não foi possível fazer testes reais com este objetivo.

Basicamente, para que isso seja possível deve-se acrescentar mais uma fase ao processo, cuja função é fazer uma *pré-seleção* das associações, em vista da existência de contra-exemplos do conceito considerado, antes da seleção propriamente dita que caracteriza a fase 2. O que deve ser feito é comparar as descrições de cada uma das associações obtidas na fase 1 com as descrições dos objetos referenciados na instância negativa considerada, e, se houver uma associação que cubra algum dos objetos, as cláusulas comuns deverão ser eliminadas da associação. Se a cláusula eliminada é simples, a associação em questão

deve ser eliminada do quadro de associações, no ramo correspondente à cláusula. Se a cláusula é disjuntiva, ela apenas não poderá mais se referenciar àquela associação. A partir desse procedimento, as outras fases seguem o seu curso normal. Na situação em que só existem associações parciais (conforme a Seção 5.4.7), a fase de pré-seleção ocorrerá entre as fases de geração das associações totais e a de seleção das menos disjuntivas. Deve-se observar que, enquanto não houver instâncias positivas, as negativas deverão permanecer armazenadas para ser usadas posteriormente, quando o primeiro exemplo for disponível.

A geração de descrições discriminantes para um conjunto de instâncias é possível simplesmente através do que foi comentado acima. No entanto, a utilização desse processo no contexto de um conjunto de regras de uma base de conhecimento apresenta um problema adicional. Pela natureza da representação utilizada para as regras (e instâncias), não se tem, a princípio, uma forma de saber se duas funções de predicado são opostas, o que permitiria caracterizar duas cláusulas de conclusão como opostas, e, conseqüentemente, perceber um contra-exemplo de uma regra. Por exemplo, apesar de poder parecer à primeira vista que os predicados "favoráveis" e "não-favoráveis" (que ocorrem na base de regras do exemplo de meteorologia apresentado) são opostos, apenas através das regras não é possível obter esta informação.

Para resolver essa questão, duas soluções poderiam ser adotadas. A primeira é dispor de um conjunto de pares de predicados opostos, a fim de que se possa automaticamente detectar as instâncias positivas e negativas. A outra solução (conforme relatada em Sandri, 1986) é supor a existência de um novo campo na representação, cuja função é modificar (por exemplo, através do modificador "não") a semântica da função de predicado; neste caso, duas cláusulas opostas seriam facilmente identificadas, bastando que elas possuíssem a mesma função de predicado, e que o modificador de função para uma delas fosse o modificador "não". Evitando alterar o esquema de representação, e por uma questão de naturalidade expressiva (a fim de garantir, por exemplo,

que o oposto de "maior-ou-igual" seja simplesmente "menor"), pretende-se adotar a primeira solução quando, para a geração dos modelos de regras, forem considerados os contra-exemplos.

## CAPITULO 7

### COMENTÁRIOS FINAIS

Inicialmente, vale observar ainda uma outra possibilidade de aplicação do mecanismo de aprendizagem, que é no aprendizado de conceitos representados por uma regra de mudança de estado. Por exemplo, em Vere (1978) descreve-se uma aplicação no domínio de blocos, onde os exemplos são pares de configurações de blocos, tais que a diferença existente entre as configurações de uma mesma instância se deve ao deslocamento de um dos blocos. Desta forma, foi possível ao sistema de aprendizagem descobrir regras de empilhamento e desempilhamento de blocos, e de transferência um um bloco de uma pilha à outra. Ainda neste trabalho, apresenta-se uma aplicação de obtenção da regra de transformação de uma frase (em Inglês) na voz ativa para a voz passiva, a partir de três pares de frases em uma representação profunda.

Uma forma aparentemente válida para que o tipo de aprendizagem acima seja possível, segundo a presente abordagem, implica inicialmente obter uma generalização do conjunto de instâncias de aplicação da regra. Para tal, deve-se generalizar separadamente os dois conjuntos de descrições que definem os dois lados de cada instância, utilizando o procedimento normal. Em seguida, de posse desta "instância generalizada de aplicação da regra", que constitui um par ordenado de descrições, deve-se obter a generalização destas duas descrições; isto também é realizado pelo esquema normal de generalização. Finalmente, com base nesta generalização eliminam-se das descrições que definem o par ordenado as cláusulas que são cobertas pela generalização. O novo par ordenado de descrições obtido constitui os dois lados da regra procurada. Além do teste do processo de geração de descrições discriminantes, pretende-se testar proximo o processo descrito acima.

Recentemente, conforme é mostrado em Mitchell et alii (1986), começaram a ser propostos os chamados *métodos de generalização baseada em explicação*, que representam uma nova abordagem para o pro

blema da aprendizagem através de exemplos e têm permitido explorar uma série de aspectos relativos à aprendizagem que não vinham sendo atacados nos métodos tradicionais de indução (como é o caso de todos os métodos apresentados no presente texto), os quais têm sido chamados *métodos baseados em similaridade*. No momento, a principal deficiência dos métodos baseados em explanação refere-se ao fato de eles terem utilizado somente uma única instância como base para a generalização. A geração de processos mistos baseados nos dois métodos é bastante desejável e já começa a ser tentada, mas ainda não forneceu resultados muito significativos. Assim, um interessante tópico a ser pesquisado no prolongamento do presente trabalho seria a investigação da possibilidade de integração do mecanismo proposto com algum método de generalização baseado em explanação.

Um outro comentário que vale ser feito diz respeito ao status das cláusulas disjuntivas. Inicialmente, deve-se observar que o condicionamento da disjunção à conjunção, conforme foi realizado, deveu-se a uma experiência informal feita com algumas pessoas, em que se percebeu uma certa tendência dessas pessoas em não formular generalizações a respeito de um conjunto de instâncias que lhes era mostrado quando não havia um padrão diretamente identificável. Não se pretende justificar a solução adotada com esta constatação; no entanto, dado que a disjunção ainda não foi suficientemente tratada na literatura de Inteligência Artificial, parece ser uma boa estratégia, neste problema em particular, que se busque subsídios na literatura de Psicologia, com vistas aos refinamentos que certamente deverão ocorrer no modelo proposto.

Ainda sobre a disjunção, deve-se ressaltar o aspecto de "aprendizagem baseada em conhecimento" que ela atribui ao processo, devido à possibilidade de criação de um novo conceito que elimine os termos disjuntivos, mesmo que este conceito não tenha sido representado originalmente no conjunto de instâncias. A regra geral para este processo é: toda disjunção pode ser submetida ao usuário, a fim de que ele possa eventualmente transformá-la. É interessante observar que

todas as regras seletivas de generalização utilizadas no sistema INDUCE são redutíveis em termos de desempenho, ao esquema utilizado. Deve-se observar que o prosseguimento natural para este tipo de abordagem é a automatização da obtenção do conhecimento necessário à transformação da disjunção, através de uma base de conhecimento preexistente.

Com relação à abordagem em si que foi utilizada para o esquema de aprendizagem, deve-se ainda fazer outros testes de avaliação para acertar eventuais detalhes imprevistos e melhorar os algoritmos atuais. Neste aspecto, tem-se a idéia de melhorar o algoritmo de determinação dos conjuntos consistentes de máximo contexto, através de uma busca inicial do subespaço definido pelas associações de máximo contexto que já tenham sido determinadas. Ainda com vistas à melhoria do processo de aprendizagem, um objetivo que vale ser perseguido é o de obter as generalizações de forma totalmente incremental, isto é, que, para obter a generalização válida em uma certa etapa, não se tenha de considerar toda a "história" de associações até aquele ponto; o problema principal que tem de ser contornado é não cair em erro de propagação.

Embora o processo careça de uma análise formal de eficiência, pode-se dizer que, apesar de se ter conseguido diminuir substancialmente o espaço de busca de generalizações com o processo de seleção das associações, situações práticas ainda poderão inviabilizar a utilização do mecanismo de aprendizagem. Assim, eventualmente poderia se pensar em fazer a busca da fase 3 de forma heurística, possivelmente utilizando as próprias heurísticas definidas no SPROUTER. Em uma atitude mais drástica que implicaria uma mudança de enfoque, pode-se pensar em eliminar a fase 3, passando-se a operar diretamente com as associações selecionadas, segundo o esquema citado na Seção 5.4.8; nessa situação, uma queda de desempenho é esperada, já que não seria mais possível a detecção do tipo de estruturação de objetos caracterizada por uma disjunção no campo de objeto secundário.

O último comentário que se deseja fazer refere-se ao processo em si de geração dos modelos de regras. Uma vez que o objetivo básico dos modelos é a aquisição de conhecimento, eles devem refletir o mais fielmente possível o conhecimento expresso na base. A utilização de um mecanismo de aprendizagem indutiva, a fim de obter os modelos diretamente das regras, parece ser realmente um passo necessário. No entanto, a questão fundamental que se coloca é qual a melhor forma de definir as classes de regras sobre as quais o mecanismo de aprendizagem indutiva vai atuar? Em outras palavras, de que forma as regras devem ser agrupadas, a fim de que elas possam realmente refletir a *estrutura do conhecimento* expresso, e não simplesmente a *estrutura de sua representação* em termos dos conceitos explicitamente expressos nas regras? Foi a partir desta perspectiva que se adotou o critério de considerar também as premissas no particionamento da base de regras, impondo que somente regras com um mesmo conjunto de objetos na premissa fossem consideradas na mesma classe. Embora esta posição tenha uma justificativa coerente, a medida que tenta distinguir entre diferentes processos causais que conduzem a uma mesma conclusão, ela ainda é muito simples.

Uma idéia que se teve e que vale ser explorada na continuidade do presente trabalho de pesquisa resultou de algumas idéias expressas em Kahn et alii (1985), onde é descrito o sistema MORE de aquisição de conhecimento. O ponto de interesse nesse artigo é a descrição de oito estratégias de aquisição de conhecimento, identificadas pelos autores para uma tarefa de diagnose que, segundo se argumenta, são suficientes para adquirir o conhecimento necessário a esta tarefa. Cada uma dessas estratégias é caracterizada em termos de um relacionamento bem definido entre entidades, tais como sintomas, hipóteses causadoras e situações condicionantes de sintomas. Observando que a existência dessas estratégias faz com que a base possa ser vista como se estivesse particionada em até oito classes fundamentais, tem-se um critério básico de particionamento realmente decorrente da estrutura do conhecimento expresso, conforme desejado. Dentro de uma certa classe poderia haver ainda subclasses relativas a sintomas específicos, sintomas quais

quer associados a situações condicionantes específicas etc. Resta observar que este tipo de abordagem implica uma forma de generalização construtiva, à medida que as entidades que definem as estratégias de vem ser identificadas a partir das cláusulas componentes das regras originais da base.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ANDERSON, J.R. *The architecture of cognition*. Cambridge, Harvard University, 1983.
- ANDERSON, J.R.; KLINE, P.J. A learning system and its psychological implications. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 6., Tokyo, Aug. 20-23, 1979. *Proceedings*. Palo Alto, CA, USA, William Kaufmann, 1979. v. 1., p. 16-21.
- ANDERSON, J.R.; KLINE, P.J.; BEASLEY JR., C.M. A general learning theory and its application to schema abstraction. In: BOWER, G.H., ed. *The psychology of learning and motivation*. New York, NY, Academic, 1979. v. 13, p. 277-318.
- BARR, A. Meta-knowledge and cognition. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 6., Tokyo, Aug. 20-23, 1979. *Proceedings*. Palo Alto, CA, USA, William Kaufmann, 1979, v. 1, p. 31-33.
- BUCHANAN, B.G.; FEIGENBAUM, E.A. DENDRAL and META-DENDRAL: their applications dimension. *Artificial Intelligence* 11(1-2):5-24, 1978.
- BUCHANAN, B.G.; MITCHELL, T.M. Model-directed learning of production rules. In: WATERMAN, D.A.; HAYES-ROTH, F., ed. *Pattern-directed inference systems*. New York, NY, Academic, 1978, p. 297-312.
- CHARNIAK, E.; McDERMOTT, D. Learning. In: ----- *Introduction to artificial intelligence*. Reading, MA, Addison-Wesley, 1985, cap. 11, p. 609-662.
- COPI, I.M. *Symbolic logic*. Toronto, Macmillan, 1967.
- DAVIS, R. Knowledge acquisition in rule-based systems - Knowledge about representations as a basis for system construction and maintenance. In: WATERMAN, D.A.; HAYES-ROTH, F., ed. *Pattern-directed inference systems*. New York, NY, Academic, 1978. p. 99-134.
- DAVIS, R. TEIRESIAS: applications of meta-level knowledge. In: DAVIS, R.; LENAT, D. *Knowledge-based systems in artificial intelligence*. New York, NY, McGraw-Hill, 1982. p. 227-490.

- DIETTERICH, T.G. Learning and inductive inference. In: COHEN, P.R.; FEIGENBAUM, E.A., ed. *The Handbook of Artificial Intelligence*, Stanford, CA, Heuristic, 1982, v. 3, cap. 14, p. 233-511.
- DIETTERICH, T.G.; MICHALSKI, R.S. Learning and generalization of characteristic descriptions: evaluation criteria and comparative review of selected methods. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 6., Tokyo, Aug. 20-23, 1979. *Proceedings*. Palo Alto, CA, William Kaufmann, 1979. v. 1, p. 223-231.
- DIETTERICH, T.G.; MICHALSKY, R.S. Inductive learning of structural descriptions: evaluation criteria and comparative review of selected methods. *Artificial Intelligence*, 16(3):257-294, 1981.
- DUDA, R.O.; HART, P.E.; BARRET, P.; GASCHNIG, J.; KONOLIGE, K.; REBOH, R.; SLOCUM, J. *Development of the PROSPECTOR consultation system for mineral exploration*. Menlo Park, CA, Artificial Intelligence Center, SRI International, 1978. (Final Rept SRI Projects 5821 and 6415).
- ELIO, R.; ANDERSON, J.R. The effects of category generalizations and instance similarity on schema abstraction. *Journal of Experimental Psychology: Human Learning and Memory*, 7(6):397-417, Nov. 1981.
- ERMAN, L.D.; HAYES-ROTH, F.; LESSEE, V.R.; REDDY, D.R. The HEARSAY-III speech understanding system: integrating knowledge to resolve uncertainty. *Computing Surveys*, 12(2):213-253, 1980.
- FORGY, C.L. *The OPS5 user's manual*. Pittsburgh, PA. Computer Science Department, Carnegie-Mellon University, 1981 (Tech. Rept. CMU-CS-81-135).
- FRANKS, J.J.; BRANSFORD, J.D. Abstraction of visual patterns. *Journal of Experimental Psychology*, 90(1):65-74; 1971. Apud Anderson (1983).
- GREINER, R.; LENAT, D. *RLL-1: A representational language language*. Stanford, CA, Heuristic Programming Project, Computer Science Department, Stanford University, 1980 (Rep. no. HPP-80-9).

- HAYES-ROTH, B.; HAYES-ROTH, F. Concept learning and the recognition and classification of exemplars. *Journal of Verbal Learning and Verbal Behavior*. 16(3):321-338, June, 1977. Apud Anderson (1983).
- HAYES-ROTH, F.; McDEMOTT, J. Learning structured patterns from examples. In: INTERNATIONAL JOINT CONFERENCE ON PATTERN RECOGNITION, 3., Coronado, CA, Nov. 8-11, 1976. *Proceedings*. Piscataway, NJ, IEEE Computer Society, 1977. p. 419-423.
- HAYES-ROTH, F.; McDERMOTT, J. An interference matching technique for inducing abstractions. *Communications of the ACM*, 21(5):401-411, May 1978.
- HAYES-ROTH, F.; WATERMAN, D.A.; LENAT, D.B. ed. *Building Expert Systems*. Reading, MA, Addison-Wesley, 1983.
- HEDRICK, C.L. Learning production systems from examples. *Artificial Intelligence*, 7(1):21-49, 1976.
- KAHN, G.; NOWLAN, S.; McDEMOTT, J. Strategies for knowledge acquisition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(5):511-522, Sept. 1985.
- LEMERLE-LOISEL, R.; KODRATOFF, Y. Learning complex structural descriptions from examples. In: INTERNATIONAL JOINT CONFERENCE ON PATTERN RECOGNITION, 6., Munich, Oct. 19-22, 1982. *Proceedings*. Piscataway, NJ, IEEE Computer Society, 1983. p. 1056-1058.
- LENAT, D. AM: Discovery in mathematics as heuristic search. In: DAVIS, R.; LENAT, D. *Knowledge-based systems in artificial intelligence*. New York, Mc Graw-Hill, 1982. p. 1-225.
- McCARTHY, J. Programs with common sense. In: MINSKY, M. ed. *Semantic Information Processing*. Cambridge, MA, MIT, 1968. p. 403-409.
- McDERMOTT, J. Learning to use analogies. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 6., Tokyo, Aug. 20-23, 1979. *Proceedings*. Palo Alto, CA, William Kaufmann, 1979. v. 1, p. 568-576.

- MEDIN, D.L.; SCHAFFER, M.M. A context theory of classification learning. *Psychological Review*. 1978. Apud Anderson (1983).
- MICHALSKI, R.S. Pattern recognition as rule-guided inductive inference. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(4):349-361, July 1980.
- MICHALSKI, R.S. A theory and methology of inductive learning. *Artificial Intelligence*, 20(2):111-161, 1983.
- MICHALSKI, R.S.; CHILAUSSKY, R.L. Learning by being told and learning by examples. *International Journal of Policy Analysis and Information Systems*, 4(2):125-161, 1980.
- MICHALSKI, R.S.; STEPP, R.; DIDAY, E. A recent advance in data analysis: clustering objects into classes characterized by conjunctive concepts. In: KANAL, L.; ROSENFELD, A., ed. *Progress in Pattern Recognition*. Amsterdam, North-Holland, 1981.
- MITCHELL, T.M. An analysis of generalization as a search problem. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 6., Tokyo, Aug. 20-23, 1979. *Proceedings*. Palo Alto, CA, William Kaufmann, 1979, v. 1, p. 577-582.
- MITCHELL, T.M. Generalization as search. *Artificial Intelligence*, 18(2):203-226, 1982.
- MITCHELL, T.M.; KELLER, R.M.; KEDAR-CABELLI, S.T. Explanation-Based Generalization: à Unifying Approach. *Machine Learning*, 1(1):47-80, 1986.
- NASA. *Advanced automation for space missions*. Santa Clara, CA, 1982. Cap. 3, p. 60-64. (NASA CP-2255).
- POLITAKIS, P.; WEISS, S.M. Using empirical analysis to refine expert system knowledge based. *Artificial Intelligence*, 22(1):23-48, 1984.
- ROSCH, E.; MERVINS, C.B. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*. 7(4):573-605, Oct., 1975. Apud Anderson (1983).
- SANDRI, S. *Uma ferramenta para a construção de sistemas especialistas*. Dissertação de Mestrado em Computação Aplicada, São José dos Campos, INPE, maio, 1986. (INPE-3890-TDL/222).

- SILVA, O.O. *Indução de regras de decisão nebulosas*. Tese de Doutorado em Computação Aplicada. São José dos Campos, INPE, jul. 1983. 145 p. (INPE-2855-TDL/143).
- SIMON, H.A.; LEA, G. Problem solving and rule induction: a unified view. In: GREGG, L. *Knowledge and cognition*. Hillsdale, NJ, Lawrence Erlbaum, 1974. p. 105-127.
- SHORTLIFFE, E. *Computer-based medical consultations: MYCIN*. New York, NY, American Elsevier, 1976.
- STEFIK, M.; AIKINS, J.; BALZER, R.; BENOIT, J.; BIRBAUM, L.; HAYES-ROTH, F.; SACERDOTI, E. The organization of expert systems: a tutorial. *Artificial Intelligence*, 18(2):135-173, 1982.
- VAN MELLE, W. A domain-independent production rule system for consultation programs. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 6., Tokyo, 1979. *Proceedings*. Stanford, CA, Stanford University, 1979. v. 2, p. 923-925.
- VERE, S.A. Inductive learning of relational productions. In: WATERMAN, D.A.; HAYES-ROTH, F., ed. *Pattern-directed inference systems*. New York, NY, Academic, 1978, p. 281-295.
- WEISS, S.M.; KULIKOWSKI, C.A. EXPERT: a system for developing consultation models. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 6., Tokyo, 1979. *Proceedings*. Stanford, CA, Stanford University, 1979. v.2, p. 942-947.
- WINSTON, P.H. Learning structural descriptions from examples. In: -----, ed. *The psychology of computer vision*. New York, McGRAW-HILL, 1975, p. 157-209.
- WINSTON, P.H.; HORN, B.K.P. *LISP*. Reading, MA, Addison-Wesley, 1981.

## APÊNDICE A

### APLICAÇÃO DO MECANISMO DE GERAÇÃO DE MODELOS DE REGRAS EM UM DOMÍNIO DA METEOROLOGIA

A seguir apresentam-se as regras constituintes da base de conhecimento. Cada regra apresenta três elementos: o seu número, uma conjunção de cláusulas de premissa e uma conjunção de cláusulas de conclusão. Cada cláusula de premissa possui um número e pode ser formada por uma disjunção de subcláusulas. Cada cláusula de conclusão possui um número, e o fator de certeza (multiplicado por 10) com que ela é concluída.

(1

- (1 (local zona-frontal há Paranã)
- 2 (velocidade zona-frontal igual 10 km/h)
- 3 (atividade zona-frontal está T)
- 4 (desenvolvendo-se-continuamente zona-frontal está T)
- 5 (condições-em-altos-níveis SJC favoráveis T))
- (1 8 (condições-de-grande-escala-em-24-horas SJC favoráveis T)))

(2

- (1 (local zona-frontal há Rio-Grande-Sul)
- 2 (velocidade zona-frontal igual 20 km/h)
- 3 (atividade zona-frontal está T)
- 4 (desenvolvendo-se-continuamente zona-frontal está T)
- 5 (condições-em-altos-níveis SJC favoráveis T))
- (1 6 (condições-de-grande-escala-em-24-horas SJC favoráveis T)))

(3

- (1 (local zona-frontal há Rio-Grande-Sul)
- 2 (velocidade zona-frontal igual 20km/h)
- 3 ((intensidade zona-frontal está diminuindo) ou (variação-de-velocidade zona-frontal está diminuindo))
- 4 (condições-em-altos-níveis SJC favoráveis T))
- (1 2 (condições-de-grande-escala-em-24-horas SJC favoráveis T)))

(4

- (1 (proximidade sistema-sinótico há T)
- 2 (intensidade sistema-sinótico há pouca)
- 3 (jatos-em-altos-níveis SJC há T)
- 4 (convergência-de-umidade SJC há T))
- (1 7 (condições-de-grande-escala-em-24-horas SJC favoráveis T)))

- (5)  
 (1 (local sistema-sinótico hã sul)  
 2 (bloqueio-não-escoamento SJC hã T)  
 3 (condições-em-altos-níveis SJC não-favoráveis T))  
 (1 9 (condições-de-grande-escala-em-24-horas SJC não-favoráveis T)))
- (6)  
 (1 (local sistema-sinótico hã Paranã)  
 (local sistema-sinótico hã Sta-Catarina)  
 (local sistema-sinótico hã Rio-Grande-Sul))  
 (1 10 (local sistema-sinótico hã sul)))
- (7)  
 (1 (condições-de-pressão-nas-últimas-24-horas SJC favoráveis T)  
 2 (depressão-do-bulbo-úmido SJC menor 3gc))  
 (1 8 (condições-locais-em-24-horas SJC favoráveis T)))
- (8)  
 (1 (condições-de-pressão-nas-últimas-24-horas SJC favoráveis T)  
 2 (depressão-do-bulbo-úmido SJC maior 3gc))  
 (1 6 (condições-locais-em-24-horas SJC favoráveis T)))
- (9)  
 (1 (condições-de-pressão-nas-últimas-24-horas SJC não-favoráveis T)  
 2 (depressão-do-bulbo-úmido SJC menor 3gc))  
 (1 4 (condições-locais-em-24-horas SJC favoráveis T)))
- (10)  
 (1 (condições-de-pressão-nas-últimas-24-horas SJC não-favoráveis T)  
 2 (depressão-do-bulbo-úmido SJC maior 3gc))  
 (1 7 (condições-locais-em-24-horas SJC não-favoráveis T)))
- (11)  
 (1 (condições-de-pressão-nas-últimas-3-horas SJC favoráveis T)  
 2 (depressão-do-bulbo-úmido SJC menor 3gc))  
 (1 8 (condições-locais-para-menos-de-24horas SJC favoráveis T)))
- (12)  
 (1 (condições-de-pressão-nas-últimas-3-horas SJC favoráveis T)  
 2 (depressão-do-bulbo-úmido SJC maior 3gc))  
 (1 4 (condições-locais-para-menos-de-24-horas SJC favoráveis T)))

(13

- (1 (condições-de-pressão-nas-últimas-3-horas SJC não-favoráveis T)
- 2 (depressão-do-bulbo-úmido menor 3gc))
- (1 4 (condições-locais-para-menos-de-24-horas SJC favoráveis T)))

(14

- (1 (condições-de-pressão-nas-últimas-3-horas SJC não-favoráveis T)
- 2 (depressão-do-bulbo-úmido SJC maior 3gc))
- (1 7 (condições-locais-para-menos-de-24-horas SJC não-favoráveis T)))

(15

- (1 (proximidade cavado há T)
- 2 (intensidade cavado está amplificado))
- (1 8 (condições-em-altos-níveis SJC favoráveis T)))

(16

- (1 (proximidade nuvem-convectiva igual T)
- 2 (desenvolvendo-na-parte-da-manhã nuvem-convectiva igual T))
- (1 5 (condições-locais-para-menos-de-24-horas SJC favoráveis T)))

(17

- (1 (jato-em-altos-níveis SJC há T))
- (1 7 (condições-em-altos-níveis SJC favoráveis T)))

(18

- (1 (proximidade vórtice-ciclônico há T))
- (1 8 (condições-em-altos-níveis SJC favoráveis T)))

(19

- (1 (local zona-frontal há S.Paulo)
- 2 (atividade zona-frontal está T)
- 3 (condições-em-altos-níveis SJC favoráveis T))
- (1 9 (condições-de-grande-escala-para-menos-de-24-horas SJC favoráveis T)))

(20

- (1 (local zona-frontal há Paraná)
- 2 (atividade zona-frontal está T)
- 3 (condições-em-altos-níveis SJC favoráveis T))
- (1 7 (condições-de-grande-escala-para-menos-de-24-horas SJC favoráveis T)))

(21

- (1 (condições-locais-para-menos-de-24-horas SJC favoráveis T)
- 2 (verão SJC igual T)
- (1 7 (precipitação-em-menos-de-24-horas SJC igual T)))

(22

- (1 (mês-atual SJC menor 3)
- (mês-atual SJC maior 10))
- (1 10 (verão SJC igual T)))

(23

- (1 (condições-locais-para-menos-de-24-horas SJC favoráveis T)
- 2 (verão SJC não-igual T))
- (1 4 (precipitação-em-menos-de-24-horas SJC haverã T)))

(24

- (1 (condições-locais-em-24-horas SJC favoráveis T)
- 2 (verão SJC igual T))
- (1 6 (precipitação-em-24-horas SJC haverã T)))

(25

- (1 (condições-locais-em-24-horas SJC favoráveis T)
- 2 (verão SJC não-igual T))
- (1 5 (precipitação-em-24-horas SJC haverã T)))

(26

- (1 (condições-de-grande-escala-para-menos-de-24-horas SJC
- favoráveis T)
- 2 (verão SJC igual T))
- (1 9 (precipitação-em-menos-de-24-horas SJC haverã T)))

(27

- (1 (condições-de-grande-escala-para-menos-de-24-horas SJC
- favoráveis T)
- 2 (verão SJC não-igual T))
- (1 8 (precipitação-em-menos-de-24-horas SJC haverã T)))

(28

- (1 (condições-de-grande-escala-em-24-horas SJC favoráveis T)
- 2 (verão SJC igual T))
- (1 8 (precipitação-em-24-horas SJC haverã T)))

(29

- (1 (condições-de-grande-escala-em-24-horas SJC favoráveis T)
- 2 (verão SJC não-igual T))
- (1 7 (precipitação-em-24-horas SJC haverã T)))

(30

- (1 (hora-atual SJC maior-igual 10)
- 2 (hora-atual SJC menor-igual 16))
- (1 10 (pressão-esperada SJC deveria-estar diminuindo)
- 2 10 (pressão-esperada SJC não-deveria-estar aumentando)))

(31

- (1 (hora-atual SJC maior-igual 0)
- 2 (hora-atual SJC menor-igual 4))
- (1 10 (pressão-esperada SJC deveria-estar diminuindo)
- 2 10 (pressão-esperada SJC não-deveria-estar aumentando)))

(32

- (1 (hora-atual SJC maior-igual 4)
- 2 (hora-atual SJC menor-igual 10))
- (1 10 (pressão-esperada SJC deveria-estar aumentando)
- 2 10 (pressão-esperada SJC não-deveria-estar diminuindo)))

(33

- (1 (hora-atual SJC maior-igual 16))
- (1 10 (pressão-esperada SJC deveria-estar aumentando)
- 2 10 (pressão-esperada SJC não-deveria diminuir)))

(34

- (1 (valor-na-carta-isalobārica SJC menor -6)
- 2 (valor-na-carta isalobārica SJC maior-igual -10))
- (1 8 (condições-de-pressão-nas-últimas-24-horas SJC favoráveis T)))

(35

- (1 (valor-na-carta-isalobārica SJC menor -2)
- 2 (valor-na-carta-isalobārica SJC maior-igual -6))
- (1 5 (condições-de-pressão-nas-últimas-24-horas SJC favoráveis T)))

(36

- (1 (valor-na-carta-isalobārica SJC menor 0)
- 2 (valor-na-carta-isalobārica SJC maior-igual -2))
- (1 2 (condições-de-pressão-nas-últimas-2-horas SJC favoráveis T)))

(37

- (1 (pressão SJC está diminuindo)
- 2 (pressão-esperada SJC deveria-estar diminuindo)
- 3 (tendência-da-pressão SJC menor -3))
- (1 6 (tendência-pressão-nas-últimas-3-horas SJC favoráveis T)))

(38

- (1 (pressão SJC está aumentando)
- 2 (pressão-esperada SJC deveria-estar aumentando)
- 3 (tendência-da-pressão SJC maior -3))
- (1 6 (condições-de-pressão-nas-últimas-3-horas SJC não-favoráveis T)))

(39

- (1 (pressão SJC está estacionária)
- 2 (pressão-esperada SJC deveria-estar diminuindo))
- (1 6 (condições-de-pressão-nas-últimas-3-horas SJC não-favoráveis T)))

(40

- (1 (mês-atual SJC maior-igual 3)
- 2 (mês-atual SJC menor-igual 10))
- (1 10 (verão SJC não-igual T))).

A seguir são apresentados os resultados da aplicação do mecanismo de geração de modelos de regras. Cada classe de regras é caracterizada por uma ou mais cláusulas de conclusão. Quando alguma regra referenciada em uma classe apresenta um número seguido por uma letra, tem-se que essa regra foi obtida da regra referenciada por esse número, através do desmembramento de uma disjunção de subcláusulas de premissa.

Os modelos gerados para cada classe estão apresentados sob duas formas: os modelos com generalização de conceitos disjuntivos (MODELO c/ GCD), e os modelos sem esse tipo de generalização (MODELO s/ GCD).

<u>CLASSE 1</u>	: (condições-de-grande-escala-em-24-horas SJC favoráveis T)
REGRAS	: 1, 2, 3, 3A
MODELO s/GCD	: (localização zona-frontal hã (Paraná ou Rio-Grande-Sul)) (velocidade zona-frontal igual (10km/h ou 20 km/h)) (condições-em-altos-níveis SJC favoráveis T)
MODELO s/ GCD	: (localização zona-frontal hã região-sul) (velocidade zona-frontal e [10 km/h ... 20 km/h]) (condições-em-altos-níveis SJC favoráveis T)

- A.7 -

<u>CLASSE 2</u>	: (condições-locais-para-menos-de-24-horas SJC favoráveis T)
REGRAS	: 11, 12, 13
MODELO s/ GCD	: (condições-de-pressão-nas-últimas-3-horas SJC (favoráveis ou não-favoráveis) T) (depressão-do-bulbo-úmido SJC (menor ou maior) 3)
MODELO c/ GCD	: (condições-de-pressão-nas-últimas-3-horas SJC (favoráveis ou não-favoráveis) T) (depressão-do-bulbo-úmido SJC não-igual 3)
<u>CLASSE 3</u>	: (condições-de-grande-escala-para-menos-de-24-horas SJC favoráveis T)
REGRAS	: 19, 20
MODELO s/ GCD	: (localização zona-frontal hã (S.Paulo ou Paraná )) (atividade zona-frontal está T) (condições-em-altos-níveis SJC favorável T)
MODELO c/ GCD	: (localização zona-frontal hã região-sul) (atividade zona-frontal está T) (condições-em-altos-níveis SJC favorável T)
<u>CLASSE 4</u>	: (precipitação-em-menos-de-24-horas SJC haverá T)
REGRAS	: 23, 26, 27
MODELO s/ GCD	: (verão SJC (igual ou não-igual) T)
<u>CLASSE 5</u>	: (precipitação-em-24-horas haverá T)
REGRAS	: 24, 25, 28, 29
MODELO s/GCD	: (verão SJC (igual ou não-igual) T)
<u>CLASSE 6</u>	: (pressão-esperada SJC não-deveria-estar aumentando)
REGRAS	: 30, 31
MODELO s/ GCD	: (hora-atual SJC maior-igual (0 ou 10)) (hora-atual SJC menor-igual (4 ou 16))
MODELO c/ GCD	: (hora-atual SJC maior-igual 0) (hora-atual SJC menor-igual 16)

CLASSE 7 : (pressão-esperada SJC deveria-estar diminuindo)  
REGRAS : 30, 31  
MODELO s/ GCD : (hora-atual SJC maior-igual (0 ou 10))  
                  . (hora-atual SJC menor-igual (4 ou 16))  
MODELO c/ GCD : (hora-atual SJC maior-igual 0)  
                  (hora-atual SJC menor-igual 16)

CLASSE 8 : (pressão-esperada SJC não-deveria-estar aumentando)  
                  (pressão-esperada deveria-estar diminuindo)  
REGRAS : 30, 31  
MODELO s/ GCD : (hora-atual SJC maior-igual (0 ou 10))  
                  (hora-atual SJC menor-igual (4 ou 16))  
MODELO c/ GCD : (hora-atual SJC maior-igual 0)  
                  (hora-atual SJC menor-igual 16)

CLASSE 9 : (pressão-esperada SJC deveria-estar aumentando)  
REGRAS : 32, 33  
MODELO s/ GCD : (hora-atual SJC maior-igual (4 ou 16))  
MODELO c/ GCD : (hora-atual SJC maior-igual 4)

CLASSE 10 : (pressão-esperada SJC não-deveria-estar diminuindo)  
REGRAS : 32, 33  
MODELO s/ GCD : (hora-atual SJC maior-igual (4 ou 16))  
MODELO c/ GCD : (hora-atual SJC maior-igual 4)

CLASSE 11 : (pressão-esperada SJC deveria-estar aumentando)  
                  (pressão-esperada SJC não-deveria-estar diminuindo)  
REGRAS : 32, 33  
MODELO s/ GCD : (hora-atual SJC maior-igual (4 ou 16))  
MODELO c/ GCD : (hora-atual SJC maior-igual 4)

<u>CLASSE 12</u>	: (condições-de-pressão-nas-últimas-24-horas SJC favoráveis T)
REGRAS	: 34, 35, 36
MODELO s/ GCD	: (valor-na-carta-isalobárica SJC menor (-6 ou -2 ou 0)) (valor-na-carta-isalobárica SJC maior-igual (-10 ou -6 ou -2))
MODELO c/ GCD	: (valor-na-carta-isalobárica SJC menor 0) (valor-na-carta-isalobárica SJC maior-igual -10)
<u>CLASSE 13</u>	: (tendência-da-pressão-nas-últimas-24-horas SJC favoráveis T)
REGRAS	: 37, 40
MODELO s/ GCD	: (pressão SJC está (diminuindo ou estacionária)) (pressão-esperada SJC deveria-estar (diminuindo ou aumentando))
MODELO c/ GCD	: (pressão SJC está (diminuindo ou estacionária)) (pressão-esperada SJC deveria-estar variando)
<u>CLASSE 14</u>	: (condições-de-pressão-nas-últimas-3-horas SJC não-favoráveis T)
REGRAS	: 38, 39
MODELO s/ GCD	: (pressão SJC está (aumentando ou estacionária)) (pressão-esperada SJC deveria estar (aumentando ou diminuindo))
MODELO c/GCD	: (pressão SJC está (aumentando ou estacionária)) (pressão-esperada SJC deveria-estar variando)
<u>CLASSE 15</u>	: (precipitação-em-menos-de-24-horas não-haverá T)
REGRAS	: 46, 47, 48
MODELO s/ GCD	: (verão SJC (igual ou não igual) T)

- CLASSE 16 : (precipitação-em-24-horas SJC não-haverã T)  
REGRAS : 41, 42, 43, 44  
MODELO s/ GCD : (verão SJC (igual ou não-igual) T)
- CLASSE 17 : (local sistema-sinótico hã sul)  
REGRAS : 6, 6A, 6B  
MODELO s/ GCD : (local sistema-sinótico hã (Paraná ou Sta-Catarina ou Rio-Grande-Sul))  
MODELO c/ GCD : (local sistema-sinótico hã região- sul)
- CLASSE 18 : (condições-locais-em-24-horas SJC favoráveis T)  
REGRAS : 7, 8, 9  
MODELO s/ GCD : (condições-de-pressão-nas-últimas-24-horas SJC (favoráveis ou não-favoráveis) T)  
(depressão-do-bulbo-úmido SJC (menor ou maior) 3)  
MODELO c/ GCD : (condições-de-pressão-nas-últimas-24-horas SJC (favoráveis ou não-favoráveis) T)  
(depressão-do-bulbo-úmido SJC não-igual 3)

## APÊNDICE B

### APLICAÇÃO DO MECANISMO DE APRENDIZAGEM INDUTIVA EM UM DOMÍNIO FICTÍCIO DA BIOLOGIA: GERAÇÃO DE DESCRIÇÕES CARACTERÍSTICAS DE CÉLULAS CANCEROSAS

A seguir são apresentadas as descrições das instâncias positivas conforme os desenhos mostrados no Capítulo 6.

#### Instância 1:

(tipo protoplasma é A)  
(número-de-organelas célula é 6)  
(número-de-membranas célula é 8)  
(forma B1 é elipse)  
(textura B1 é listrada)  
(peso B1 é 4)

(forma B2 é círculo)  
(textura B2 é branca)  
(peso B2 é 3)  
(posição B2 é por-fora B3)

(forma B3 é círculo)  
(textura B3 é preta)  
(peso B3 é 2)

(forma B4 é círculo)  
(textura B4 é preta)  
(peso B4 é 2)

(forma B5 é barco)  
(número-de-caudas é 1)  
(textura B5 é preta)  
(peso B5 é 3)  
(direção B5 é NE)

(forma B6 é círculo)  
(textura B6 é sombreado)  
(peso B6 é 5)

Instância 2:

(tipo protoplasma é A)  
(número-de-organelas célula é 8)  
(número-de-membranas célula é 10)  
(forma B1 é quadrado)  
(textura B1 é ondulada)  
(peso B1 é 3)

(forma B2 é barco)  
(número-de-caudas é 1)  
(textura B2 é preta)  
(peso B2 é 4)  
(direção B2 é N)

(forma B3 é círculo)  
(textura B3 é sombreado)  
(peso B3 é 4)

(forma B4 é círculo)  
(textura B4 é branca)  
(peso B4 é 5)  
(posição B4 é por-fora B5)

(forma B5 é círculo)  
(textura B5 é preta)  
(peso B5 é 4)

(forma B6 é círculo)  
(textura B6 é listrada)  
(peso B6 é 2)

(forma B7 é barco)  
(número-de-caudas é 1)  
(textura B7 é preta)  
(peso B7 é 4)  
(direção B7 é N)

(forma B8 é círculo)  
(textura B8 é preta)  
(peso B8 é 2).

Instância 3:

(tipo protoplasma  $\bar{e}$  D)  
(número-de-organelas célula  $\bar{e}$  7)  
(número de-membranas célula  $\bar{e}$  6)  
(forma B1  $\bar{e}$  círculo)  
(textura B1  $\bar{e}$  preta)  
(peso B1  $\bar{e}$  2)

(forma B2  $\bar{e}$  barco)  
(número-de-caudas  $\bar{e}$  1)  
(textura B2  $\bar{e}$  preta)  
(peso B2  $\bar{e}$  4)  
(direção B2  $\bar{e}$  N)

(forma B3  $\bar{e}$  barco)  
(número-de-caudas  $\bar{e}$  1)  
(textura B3  $\bar{e}$  preta)  
(peso B3  $\bar{e}$  5)  
(direção B3  $\bar{e}$  N)

(forma B4  $\bar{e}$  círculo)  
(textura B4  $\bar{e}$  sombreada)  
(peso B4  $\bar{e}$  4)

(forma B5  $\bar{e}$  heptágono)  
(textura B5  $\bar{e}$  listrada)  
(peso B5  $\bar{e}$  8)  
(posição B5  $\bar{e}$  por-fora (B6 B7))

(forma B6  $\bar{e}$  círculo)  
(textura B6  $\bar{e}$  branca)  
(peso B6  $\bar{e}$  3)  
(posição B6  $\bar{e}$  por-fora B7)

(forma B7  $\bar{e}$  círculo)  
(textura B7  $\bar{e}$  preta)  
(peso B7  $\bar{e}$  2)

Instância 4:

(tipo protoplasma é D)  
(número-de-organelas célula é 7)  
(número-de-membranas célula é 8)  
(forma B1 é círculo)  
(textura B1 é branca)  
(peso B1 é 7)  
(posição B1 é por-fora B2)

(forma B2 é círculo)  
(textura B2 é preta)  
(peso B2 é 2)

(forma B3 é barco)  
(número-de-caudas é 1)  
(textura B3 é preta)  
(peso B3 é 3)  
(direção B3 é N)

(forma B4 é círculo)  
(textura B4 é sombreada)  
(peso B4 é 3)

(forma B5 é triângulo)  
(textura B5 é branca)  
(peso B5 é 5)  
(posição B5 é por-fora B6)

(forma B6 é círculo)  
(textura B6 é preta)  
(peso B6 é 3)

(forma B7 é mo1a)  
(peso B7 é 1)))

Para que as descrições características mostradas a seguir fossem criadas, deve-se ressaltar que as seguintes generalizações de conceitos foram tomadas como válidas:

(... é (3 ou 4 ou 5)) -----> (... maior-igual 3)  
(... é (3 ou 5 ou 8)) -----> (... maior-igual 3)  
(... é (2 ou 3 ou 4)) -----> (... maior-igual 2)  
(... é (3 ou 5 ou 7)) -----> (... é ímpar)  
(... é (2 ou 4)) -----> (... é par)  
(... é (6 ou 7 ou 8)) -----> (... maior-igual 6)  
(... é (6 ou 8 ou 10)) -----> (... é [6...10]).

Todos os conjuntos de associações representados a seguir possuem, além das associações mostradas, as associações

"célula-célula-célula-célula", e

"protoplasma-protoplasma-protoplasma-protoplasma".

Como são comuns a todos os conjuntos e por uma questão de clareza elas serão omitidas dos mesmos. Além disso, como todos os objetos nelas referenciados são iguais, eles podem ser fundidos em um único.

5723-4812-6344-3576-2455 ----> (forma 5723 é barco)  
(textura 5723 é preta)  
(peso 5723 é (3 ou 4))  
(número-de-caudas 5723 é 1)  
(direção 5723 é (N ou NE))  
  
(forma 4812 é círculo)  
(textura 4812 é preta)  
(peso 4812 é 2)  
  
(forma 6344 é círculo)  
(textura 6344 é sombreada)  
(peso 6344 maior-igual 3)

- B.6 -

(forma 3576 é círculo)  
(textura 3576 é preta)  
(peso 3576 maior-igual 2)

(forma 2455 é (círculo ou polígono))  
(textura 2455 é (branca ou listrada))  
(peso 2455 maior-igual 3)  
(posição 2455 é por-fora 3576)

(tipo *protoplasma* é (A ou D))

(número-de-organelas *celula* maior-igual 6)  
(número-de-membranas *celula* e [6...10])

5723-4812-2461-6344-3576 --->(forma 5723 é barco)  
(textura 5723 é preta)  
(peso 5723 é (3 ou 4))  
(número-de-caudas 5723 é 1)  
(direção 5723 é (N ou NE))

(forma 4812 é círculo)  
(textura 4812 é preta)  
(peso 4812 é 2)

(forma 2461 é círculo)  
(textura 2461 é branca)  
(peso 2461 é ímpar)  
(posição 2461 é por-fora (3576 ou 4812))

(forma 6344 é círculo)  
(textura 6344 é sombreada)  
(peso 6344 maior-igual 3)

(forma 3576 é círculo)  
(textura 3576 é preta)  
(peso 3576 maior-igual 2)

(tipo *protoplasma* é (A ou D))

(número-de-organelas *celula* maior-igual 6)  
(número-de-membranas *celula* e [6...10])

- B.7 -

- 5723-3812-6435-4576-2164 ---> (forma 5723 é barco)  
(textura 5723 é preta)  
(peso 5723 é (3 ou 4))  
(número-de-caudas 5723 é 1)  
(direção 5723 é (N ou NE))
- (forma 3812 é círculo)  
(textura 3812 é preta)  
(peso 3812 é 2)
- (forma 6435 (círculo ou barco ou triângulo))  
(textura 6435 é (sobreada ou branca ou preta))  
(peso 6435 é 5)
- (forma 4576 é círculo)  
(textura 4576 é preta)  
(peso 4576 maior-igual 2)
- (forma 2164 é (círculo ou quadrado))  
(textura 2164 é (branca ou ondulada  
ou sombreada))  
(peso 2164 é 3)
- (tipo *protoplasma* é (A ou D))
- (número-de-organelas *celula* maior-igual 6)  
(número-de-membranas *celula* e [6...10])
- 4812-6344-5163-3576-2465 ---> (forma 4812 é círculo)  
(textura 4812 é preta)  
(peso 4812 é 2)
- (forma 6344 é círculo)  
(textura 6344 é sombreada)  
(peso 6344 maior-igual 3)
- (forma 5163 é (barco ou quadrado ou círculo))  
(textura 5163 é (preta ou ondulada ou  
branca))  
(peso 5163 é 3)
- (forma 3576 é círculo)  
(textura 3576 é preta)  
(peso 3576 maior-igual 2)

- B.8 -

(forma 2465 é (círculo ou triângulo)  
(textura 2465 é branca)  
(peso 2465 é (3 ou 5))  
(posição 2465 é por-fora 3576))

(tipo *protoplasma* é (A ou D))

(número-de-organelas *celula* maior-igual 6)  
(número-de-membranas *celula* e [6...10])

5723-2166-3812-6344

----> (forma 5723 é barco)  
(textura 5723 é preta)  
(peso 5723 é (3 ou 4))  
(número-de-caudas 5723 é 1)  
(direção 5723 é (N ou NE))

(forma 2166 (círculo ou quadrado))  
(textura 2166 é (branca ou ondulada ou preta))  
(peso 2166 é 3)

(forma 3812 é círculo)  
(textura 3812 é preta)  
(peso 3812 é 2)

(forma 6344 é círculo)  
(textura 6344 é sombreada)  
(peso 6344 maior-igual 3)

(tipo *protoplasma* é (A ou D))  
(número-de-organelas *celula* maior-igual 6)  
(número-de-membranas *celula* e [6...10])

5723-2461-6344-3572-4816

----> (forma 5723 é barco)  
(textura 5723 é preta)  
(peso 5723 é (3 ou 4))  
(número-de-caudas 5723 é 1)  
(direção 5723 é (N ou NE))

(forma 2461 é círculo)  
(textura 2461 é branca)  
(peso 2461 é ímpar)  
(posição 2461 é por-fora 3572)

(forma 3572 é círculo)  
(textura 3572 é preta)  
(peso 3572 é par)

- B.9 -

(forma *4816* é círculo)  
(textura *4816* é preta)  
(peso *4816* é (2 ou 3))

(forma *6344* é círculo)  
(textura *6344* é sombreada)  
(peso *6344* maior-igual 3)

(tipo *protoplasma* é (A ou D))

-(número-de-organelas *célula* maior-igual 6)  
(número-de-membranas *célula* ∈ [6...10])

