

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

**Aplicando Métodos de Solução de Problemas
em Tarefas de Interpretação de Rochas**

por

LUÍS ALVARO DE LIMA SILVA

Dissertação de Mestrado

Prof. Dr. Clesio Saraiva dos Santos
Orientador

Profª. Dra. Mara Abel
Co-orientadora

Porto Alegre, julho de 2001.

CIP - CATALOGAÇÃO NA PUBLICAÇÃO

Silva, Luís Alvaro de Lima

Aplicando Métodos de Solução de Problemas em Tarefas de Interpretação de Rochas / por Luís Alvaro de Lima Silva. Porto Alegre: PPGC da UFRGS, 2001.

160f.: il.

Dissertação (Mestrado) - Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR-RS, 2001. Orientador: Santos, Clesio Saraiva dos. Co-orientadora: Abel, Mara.

1. Métodos de Solução de Problemas. 2. Sistemas Baseados em Conhecimento. 3. Engenharia de Conhecimento. I. Santos, Clesio Saraiva dos; Abel, Mara. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitora: Profa. Wrana Panizzi

Pró-Reitor de Ensino: Prof. José Carlos Ferraz Hennemann

Pró-Reitor Adjunto de Pós-Graduação: Prof. Philippe Olivier Alexandre Navaux

Diretor do Instituto de Informática: Prof. Philippe Olivier Alexandre Navaux

Coordenador do PPGC: Prof. Carlos Alberto Heuser

Bibliotecária-Chefe do Instituto de Informática: Beatriz Regina Bastos Haro

Agradecimentos

Dentre muitos amigos que gostaria de agradecer, posso lembrar de alguns no instante que escrevo...

Em princípio, gostaria de agradecer aos professores Mara Abel e Clesio Saraiva dos Santos, os quais são exemplos a serem seguidos! Agradecer ao Prof. Luiz Fernando De Ros, quem disponibilizou o tempo e dados necessários para elaborar um trabalho prático em um domínio de aplicação muito interessante.

Não poderia deixar de agradecer aos alunos Laura Mastella e Rafael de Souza Lopez. Tais alunos têm um importante papel no projeto *PetroGrapher*, assim como foram fundamentais para o desenvolvimento do estudo de caso desta dissertação.

Quero prestar um agradecimento a todos os antigos membros do Grupo de Bancos de Dados Inteligentes, os quais liderados pelo Prof. Castilho demonstraram grande motivação de pesquisa. Posso considerar que tal motivação foi a fonte de inspiração para o ingresso neste curso de mestrado. Também gostaria de agradecer aos membros do grupo por mim denominado de "Grupo XML e LaTeX", os quais muitas vezes estiveram dispostos a escutar minhas idéias.

Por fim, gostaria de agradecer a Deus por seu suporte renovador e motivador!

Sumário

Lista de Abreviaturas	7
Lista de Figuras	8
Lista de Tabelas	10
Resumo	11
Abstract	12
1 Introdução	13
2 Uma visão da engenharia de conhecimento	18
2.1 As noções de ontologia e método de solução de problemas	19
2.2 Uma caracterização de métodos de solução de problemas	21
2.3 Abordagens de engenharia de conhecimento	27
2.3.1 Métodos de Limitação de Papéis	28
2.3.2 Tarefas Genéricas	29
2.3.3 CommonKADS	30
2.3.4 MIKE	33
2.3.5 Protégé-II	34
2.4 Arquitetura de métodos de solução de problemas	35
2.4.1 Competência	36
2.4.2 Especificação operacional	37
2.4.3 Requisitos / suposições	38
2.5 Desenvolvimento de métodos de solução de problemas	39
2.5.1 Enfoque baseado na teoria da competência para construção de métodos de solução de problemas	40
2.5.2 Desenvolvimento estruturado fundamentado em um sistema de eixos	43
2.5.3 Construção de métodos de solução de problemas como projetos baseados em parâmetros	48
2.6 Suposições	51
2.7 Bibliotecas de métodos de solução de problemas	55
3 Um estudo de caso de desenvolvimento de	

métodos de solução de problemas	58
3.1 Interpretação de rochas	60
3.2 Os passos de desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos	62
3.2.1 Aquisição de conhecimento	63
3.2.2 Detalhamento do problema	64
3.2.3 Reutilização de bibliotecas de PSMs	65
3.2.4 Definição de um esquema de inferência básico	65
3.2.5 Adaptação / refinamento	67
3.2.5.1 <i>Estabelecimento de suposições</i>	68
3.2.5.2 <i>Detalhamento operacional</i>	70
3.2.6 Validação das especificações	70
3.2.7 Abstração dos PSMs	71
3.3 Uma análise do desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos	72
4 Um esquema de raciocínio para interpretação de rochas	74
4.1 O PSM Interpretação de Rochas	74
4.1.1 A competência do PSM Interpretação de Rochas	77
4.1.2 Uma especificação operacional para o PSM Interpretação de Rochas	77
4.1.3 Um esquema de domínio típico do PSM Interpretação de Rochas	80
4.1.4 Alguns requisitos/suposições do PSM Interpretação de Rochas	81
4.2 Um modelo de conhecimento de domínio	85
4.2.1 Um modelo de conhecimento inferencial de domínio	88
4.3 O PSM Interpretação de Ambientes Diagenéticos	94
4.3.1 A competência do PSM Interpretação de Ambientes Diagenéticos	97
4.3.2 Uma especificação operacional para o PSM Interpretação de Ambientes Diagenéticos	97
4.3.3 Um esquema de domínio típico do PSM Interpretação de Ambientes Diagenéticos	101
4.3.4 Alguns requisitos/suposições do PSM Interpretação de Ambientes Diagenéticos	102
4.4 Uma classificação para os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos	108
5 Um mecanismo de inferência para interpretação de rochas	109

5.1 A arquitetura do sistema <i>PetroGrapher</i>	109
5.2 Uma arquitetura de inferência para interpretação de rochas	111
5.3 Um mapeamento do modelo de frames especificado para um modelo de dados	115
5.4 Um mapeamento do modelo de conhecimento inferencial de domínio especificado para um modelo de dados	119
5.5 Um algoritmo de inferência para interpretação de ambientes diagenéticos	123
6 Conclusões	128
Anexo 1 Modelo de Frames da Aplicação <i>PetroGrapher</i>	131
Anexo 2 Modelo de Grafos de Conhecimento da Aplicação <i>PetroGrapher</i>	139
Bibliografia	153

Lista de Abreviaturas

BD	- Banco de Dados
CML	- <i>Conceptual Modeling Language</i>
CLOS	- <i>Common Lisp Object System</i>
CRLM	- <i>Configurable Role-Limiting Methods</i>
DFD	- Diagrama de Fluxo de Dados
GT	- <i>Generic Tasks</i>
KBS	- <i>Knowledge Based System</i>
KE	- <i>Knowledge Engineering</i>
KL	- <i>Knowledge Level</i>
MIKE	- <i>Model-Based and Incremental Knowledge Engineering</i>
PSM	- <i>Problem-Solving Methods</i>
RLM	- <i>Role-Limiting Methods</i>
SGBD	- Sistema de Gerência de Banco de Dados
SQL	- <i>Structured Query Language</i>
UPML	- <i>Unified Problem-Solving Method Development Language</i>

Lista de Figuras

FIGURA 2.1 - Estrutura de inferência do PSM Classificação Heurística	22
FIGURA 2.2 - PSM Diagnóstico Sistemático: um exemplo para diagnóstico de falhas mecânicas em automóveis	23
FIGURA 2.3 - PSM Diagnóstico Sistemático: um exemplo para diagnóstico médico.....	24
FIGURA 2.4 - Um PSM de Diagnóstico	25
FIGURA 2.5 - Esquema de domínio típico de um PSM de Diagnóstico	26
FIGURA 2.6 - A relação entre conhecimento de domínio e conhecimento de inferência	26
FIGURA 2.7 - Uma representação em Diagrama de Fluxo de Dados	27
FIGURA 2.8 - Um exemplo de uma estrutura de tarefas para diagnóstico modelada como Tarefas Genéricas	30
FIGURA 2.9 - As camadas do modelo de conhecimento	31
FIGURA 2.10 - Um Exemplo para o Modelo de Conhecimento	32
FIGURA 2.11 - Ontologias e PSMs em Protégé-II	35
FIGURA 2.12 - A arquitetura de PSMs	36
FIGURA 2.13 - O enfoque baseado na teoria de competência para construção de PSM	41
FIGURA 2.14 - As três dimensões da descrição e desenvolvimento de PSMs	44
FIGURA 2.15 - Os componentes da especificação de um PSM	48
FIGURA 2.16 - Um diagrama de estados para o controle do método de configuração de PSMs	51
FIGURA 2.17 - As lacunas que isolam um PSM	52
FIGURA 2.18 - Os efeitos das suposições	53
FIGURA 4.1 - Estrutura de inferência do PSM Interpretação de Rochas	75
FIGURA 4.2 - Especificação de controle do PSM Interpretação de Rochas	79
FIGURA 4.3 - Esquema de domínio típico do PSM Interpretação de Rochas	80
FIGURA 4.4 - Conhecimento de domínio requerido pelos passos de inferência do PSM Interpretação de Rochas	82
FIGURA 4.5 - Partes de um caso de descrição de amostra de rocha	85
FIGURA 4.6 - Modelo de frames do domínio	87
FIGURA 4.7 - Um exemplo de grafo de conhecimento: interpretação e expressões de domínio	91
FIGURA 4.8 - Um exemplo de expressão de domínio modelada	91

FIGURA 4.9 - Exemplo de expressão de domínio modelada	92
FIGURA 4.10 - PSM Interpretação de Ambientes Diagenéticos	95
FIGURA 4.11 - Especificação de controle do PSM Interpretação de Ambientes Diagenéticos	99
FIGURA 4.12 - Esquema de domínio típico do PSM Interpretação de Ambientes Diagenéticos	102
FIGURA 5.1 - Arquitetura do sistema <i>PetroGrapher</i>	110
FIGURA 5.2 - Uma arquitetura de inferência para interpretação de rochas	112
FIGURA 5.3 - Mapeamento do modelo de frames para o modelo de dados	116
FIGURA 5.4 - Modelo de dados do domínio de aplicação	119
FIGURA 5.5 - Base de conhecimento inferencial de domínio	120
FIGURA 5.6 - Estrutura simbólica de representação de grafos de conhecimento	124
FIGURA 5.7 - Estrutura simbólica de representação de expressões de domínio	124
FIGURA 5.8 - Especificação simbólica da expressão de domínio "caulinização" ...	125
FIGURA 5.9 - Consultas SQL geradas a partir da especificação simbólica da expressão de domínio "caulinização"	125
FIGURA 5.10 - Estrutura simbólica de representação de grafos de conhecimento validados	126

Lista de Tabelas

TABELA 4.1 - Um exemplo do modelo de frames: frame composição-diagenética	88
TABELA 4.2 - Significado dos pesos modelados	90
TABELA 4.3 - Grafos de conhecimento e expressões de domínio modelados	93
TABELA 5.1 - Tabela grafos: exemplos de grafos de conhecimento armazenados no BD	121
TABELA 5.2 - Tabela expressõesDomínio: exemplos de expressões de domínio armazenadas no BD	121
TABELA 5.3 - Tabela conceitosDomínio: exemplos de conceitos de domínio armazenados no BD	122

Resumo

A Engenharia de Conhecimento (*Knowledge Engineering* - KE) atual considera o desenvolvimento de Sistemas Baseados em Conhecimento (*Knowledge-Based Systems* - KBSs) como um processo de modelagem baseado em modelos de conhecimento reusáveis. A noção de Métodos de Solução de Problemas (*Problem-Solving Methods* - PSMs) desempenha um importante papel neste cenário de pesquisa, pois representa o conhecimento inferencial de KBSs em um formalismo explícito. Não menos importante, PSMs também facilitam a compreensão do processo de raciocínio desenvolvido por humanos.

PSMs são descritos em um formalismo abstrato e independente de implementação, facilitando a análise do conhecimento inferencial que muitas vezes é obscurecido em grandes bases de conhecimento. Desta forma, este trabalho discute a noção de PSMs, avaliando os problemas de pesquisa envolvidos no processo de desenvolvimento e especificação de um método, como também analisando as possibilidades de aplicação de PSMs.

O trabalho apresenta a descrição e análise de um estudo de caso sobre o processo de desenvolvimento, especificação e aplicação de um PSM Interpretação de Rochas. As tarefas de interpretação de rochas são desenvolvidas por petrógrafos especialistas e correspondem a um importante passo na caracterização de rochas-reservatório de petróleo e definição de técnicas de exploração, permitindo que companhias de petróleo reduzam custos de exploração normalmente muito elevados. Para suportar o desenvolvimento de KBSs neste domínio de aplicação, foram desenvolvidos dois PSMs novos: o PSM Interpretação de Rochas e o PSM Interpretação de Ambientes Diagenéticos. Tais métodos foram especificados a partir de uma análise da perícia em Petrografia Sedimentar, como também a partir de modelos de conhecimento e dados desenvolvidos durante o projeto *PetroGrapher*.

O PSM Interpretação de Rochas e o PSM Interpretação de Ambientes Diagenéticos são especificados conceitualmente em termos de competência, especificação operacional e requisitos/suposições. Tais definições detalham os componentes centrais de um esquema de raciocínio para interpretação de rochas. Este esquema é empregado como um modelo de compreensão e análise do processo de raciocínio requerido para orientar o desenvolvimento de uma arquitetura de raciocínio para interpretação de rochas. Esta arquitetura é descrita em termos de requisitos de armazenamento e manipulação de dados e conhecimento, permitindo projetar e construir um algoritmo de inferência simbólico para uma aplicação de bancos de dados inteligentes denominada *PetroGrapher*.

Palavras-chave: métodos de solução de problemas, sistemas baseados em conhecimento, engenharia de conhecimento

TITLE: "APPLYING PROBLEM-SOLVING METHODS IN ROCK INTERPRETATION TASKS"

Abstract

The recent view of Knowledge Engineering (KE) considers the Knowledge-Based Systems (KBSs) development as a process of modeling based on reusable knowledge models. The notion of Problem-Solving Methods (PSMs) plays an important role in this research scene by representing the inferential knowledge of KBSs in an explicit formalism. Not less important, PSMs also contributes with the understanding of the reasoning process developed by human beings.

PSMs are described in an abstract and implementation-independent formalism, facilitating the analysis of the inferential knowledge, which is frequently overcastted in large knowledge bases. Thus, this essay discusses the PSMs notion, evaluates the research problems involved in the development process and specification of a method, and analyzes the possibilities of the PSMs application.

This work presents the description and analysis of a case study about the development process, specification and application of a rock interpretation PSM. The rock interpretation tasks are developed by petrology experts and correspond to an important step in petroleum rock-reservoir characterization and exploration technique definition, allowing petroleum companies to reduce costs normally very raised. In order to support the KBSs development in this application domain, two new PSMs were developed: the Rock Interpretation PSM and the Diagenetic Environment Interpretation PSM. Such methods were specified from an analysis of the Sedimentary Petrography skill, as also from knowledge and data models developed during the *PetroGrapher* project.

The Rock Interpretation PSM and the Diagenetic Environment Interpretation PSM are conceptually specified in competence, operational specification and requirement/assumption terms. Such definitions detail the central components of an interpretation of rocks reasoning schema. This schema is used as a comprehension and analysis model of the reasoning process required in guiding a rock interpretation reasoning architecture development. This architecture is described in term of data and knowledge storage and manipulation requirements, allowing the project and development of a symbolic inference algorithm for the intelligent database application called *PetroGrapher*.

Keywords: problem-solving methods, knowledge-based systems, knowledge engineering

1 Introdução

Nos últimos anos, um considerável esforço de pesquisa na área de Engenharia de Conhecimento (*Knowledge Engineering* - KE) tem sido realizado, assim buscando caracterizar os princípios metodológicos requeridos para a construção sistemática de bases de conhecimento e métodos de inferência reusáveis. O processo de KE que permite construir esses componentes centrais de Sistemas Baseados em Conhecimento (*Knowledge-Based Systems* - KBSs) é compreendido atualmente como um processo de especificação de um conjunto de modelos, os quais sistematizam o processo de aquisição de conhecimento e desenvolvimento de KBSs.

A área de pesquisa da KE atual que estuda Métodos de Solução de Problemas (*Problem-Solving Methods* - PSMs) tem buscado delimitar o escopo e aplicabilidade dessas especificações de conhecimento inferencial reusáveis no desenvolvimento de KBSs. Tal interesse deve-se a importância da definição do comportamento inferencial requerido por aplicações baseadas em conhecimento, em contraste com as dificuldades naturais de compreender e expressar o raciocínio humano em computador. Desta forma, o presente trabalho focaliza a noção de PSMs, os quais são atualmente considerados componentes inferenciais reusáveis centrais no processo de construção de KBSs.

PSMs fazem parte de diferentes abordagens da KE, tais como: Métodos de Limitação de Papéis (Marcus [MAR 88]), Tarefas Genéricas (Chandrasekaran [CHA 86]), CommonKADS (Schreiber et al. [SCH 94]; Schreiber et al. [SCH 99]); MIKE (Angele et al. [ANG 98]); Protégé-II (Grosso et al. [GRO 99]), entre outras. Além disso, diferentes bibliotecas de PSMs reusáveis têm sido descritas na literatura, assim disponibilizando métodos para diferentes tipos de problemas. Porém, também há bibliotecas para tarefas específicas bastante conhecidas, tais como: a biblioteca de PSMs para diagnóstico (Benjamins et al. [BEN 94]; Benjamins et al. [BEN 95]), a biblioteca de PSMs para projeto parametrizado (Motta et al. [MOT 98]), etc. Conforme Gómez-Pérez et al. [GOM 99], a pesquisa atual em PSMs tem abordado diferentes áreas, tais como: identificação de PSMs específicos de tarefas, organização de PSMs em bibliotecas e formalismos de descrição de PSMs, por exemplo.

Segundo Benjamins et al. [BEN 98], um PSM permite descrever os passos de inferência que devem ser executados para alcançar os objetivos de uma tarefa, definir uma ou mais estruturas de controle sobre esses passos, assim demonstrando como efetivamente alcançar tais objetivos pelo uso de conhecimento de domínio, além de definir os papéis que este conhecimento de domínio desempenha em cada passo de inferência detalhado. Logo, usando PSMs é possível descrever o processo de raciocínio de um KBS em uma maneira independente de implementação e domínio, abstraindo aspectos que muitas vezes obscurecem a análise e o reuso do raciocínio especificado.

Conforme Gómez-Pérez et al. [GOM 99], PSMs têm um papel importante para a KE, pois eles podem: 1) descrever uma forma eficiente de alcançar os objetivos de uma tarefa pela aplicação de conhecimento de domínio; 2) guiar o processo de aquisição de conhecimento de domínio e 3) tornar o processo de desenvolvimento de KBSs mais simplificado, pois eles podem ser reutilizados. Em contraste com outras formas de representação de conhecimento mais tradicionais, PSMs podem expressar o

conhecimento inferencial de forma mais abstrata e estruturada, assim tornando explícito o conhecimento de controle muitas vezes obscurecido em grandes bases de conhecimento. Porém, PSMs não substituem formalismos mais tradicionais de especificação de conhecimento, pois formalismos de representação como regras de produção, frames ou lógica, entre outros, podem ser usados para modelar o conhecimento requerido para viabilizar os passos de inferência especificados em PSMs.

Para que PSMs sejam reusados em aplicações baseadas em conhecimento, é necessário encontrar o PSM correto para a tarefa requerida, avaliar a sua aplicabilidade para a situação e, então, adaptá-lo para os requisitos do problema sendo tratado. Para isso, um PSM estabelece uma noção de competência para solucionar uma tarefa, onde tal competência é expressa em termos do quê um PSM pode realizar, como ele pode realizar e, por fim, o quê é necessário para realizar. Não sendo encontrado um PSM usável no problema que se quer resolver, um novo PSM pode ser desenvolvido do zero. Para realizar tal processo de desenvolvimento, existem diferentes propostas metodológicas de desenvolvimento de PSMs, mas ainda não há uma metodologia amplamente aceita, tal como um padrão de fato.

Mesmo que o reuso de PSMs seja recomendável, o desenvolvimento de PSMs a partir do zero é uma abordagem bastante pesquisada atualmente, pois podem ser desenvolvidos novos métodos para bibliotecas de PSMs conhecidas, como também ainda existem muitos tipos de problemas que podem ser descritos por novas bibliotecas de PSMs. Entre outras propostas metodológicas de desenvolvimento de PSMs que poderiam ser citadas: fundamentado em uma especificação da competência (Wielinga et al. [WIE 98]); desenvolvimento estruturado baseado em um sistema de eixos (Fensel et al. [FEN 2000]) e fundamentado em um tipo específico de problema de configuração (Teije et al. [TEI 96]). Contudo, estas metodologias ainda requerem uma validação prática dos passos de desenvolvimento propostos segundo requisitos de aplicações reais.

Em resumo, PSMs podem ser usados para descrever o processo de raciocínio requerido por problemas baseados em conhecimento e, dessa forma, analisar e projetar mecanismos de inferência para KBSs. Contudo, ainda existem poucas descrições de experiências práticas de uso e reuso de PSMs em aplicações reais de KBSs e, principalmente, em domínios de aplicação que envolvem grandes volumes de dados e conhecimento, tal como é o domínio de problemas do projeto *PetroGrapher*, onde este trabalho está inserido.

O objetivo desse trabalho é caracterizar e analisar os PSMs que compõem um esquema de raciocínio para tarefas de interpretação de rochas. Este esquema é centralmente definido pelos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos, os quais são usáveis no domínio do projeto *PetroGrapher*, e possivelmente reusáveis em outros domínios. O esquema de raciocínio identifica alguns dos principais requisitos conceituais necessários para a solução de problemas baseados em conhecimento, em especial para o domínio de problemas de interpretação de rochas. A partir dele, o trabalho propõe um mecanismo de inferência para interpretação de rochas, descrito como uma arquitetura de raciocínio para interpretação de rochas. Tal arquitetura é o projeto de um mecanismo de inferência envolvendo dados e conhecimento, assim como requerido para desenvolver um

módulo inferencial simbólico para uma aplicação baseada em conhecimento - sistema *PetroGrapher* (Silva [SIL 97]; Silva [SIL 99]).

Os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos são modelos conceituais de *uma análise da perícia de petrógrafos especialistas* (Abel et al. [ABE 98]; Abel [ABE 99]; Abel [ABE 2001]) na solução de tarefas de descrição e interpretação de rochas sedimentares clásticas. A descrição de passos de especificação de raciocínios humanos é importante em diferentes domínios de aplicação, pois normalmente poucos especialistas mantêm o conhecimento necessário para resolver problemas baseados em conhecimento.

Logo, esse trabalho é parte de um projeto que envolve a análise cognitiva de um especialista, mais a aquisição e modelagem de conhecimento no domínio da Petrografia, os quais estão sendo desenvolvidos no âmbito do projeto *PetroGrapher*. Tal projeto vem sendo desenvolvido com o objetivo prático de construir uma aplicação de Banco de Dados Inteligentes (Parsaye et al. [PAR 89]; Bertino et al. [BER 98]) para apoio a tarefa de descrição e interpretação de rochas-reservatório de petróleo.

O PSM Interpretação de Rochas é uma especificação abstrata da perícia desenvolvida na Petrografia Sedimentar. O PSM Interpretação de Ambientes Diagenéticos, por sua vez, é uma especificação mais operacional descrita em função das possibilidades de aquisição e representação de conhecimento experimentadas no projeto *PetroGrapher*. Ambos os PSMs permitem compreender e caracterizar os limites de aplicabilidade de um mecanismo de inferência, como também permitem identificar requisitos práticos reais de desenvolvimento de KBSs.

A Petrografia Sedimentar é o domínio de aplicação deste trabalho. Tal domínio é representativo de problemas pouco estruturados e intensivos em conhecimento, onde KBSs têm sido aplicados durante os últimos anos. Dentre outras dificuldades encontradas neste domínio, a organização e utilização de grande volume de dados e conhecimento têm se destacado, pois cada vez mais é necessário preservar e disponibilizar o conhecimento de solução de problemas normalmente mantido por poucos especialistas em petrografia. Desta forma, tal domínio de aplicação requer sistemas inteligentes que auxiliem na descrição e interpretação de rochas, os quais podem ser importantes no processo de *gestão de conhecimento* (Wiig [WII 97]) de empresas de exploração de petróleo, tal como são essenciais em outros domínios de aplicação.

Os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos são usados como modelos de análise do raciocínio desenvolvido para interpretar rochas. As tarefas de interpretação de rochas são fundamentais para caracterizar quantitativa e qualitativamente o potencial econômico de reservatórios de óleo, pois elas podem delimitar a produtividade esperada para reservatórios, assim reduzindo custos de exploração normalmente muito altos. A partir desses modelos, o presente trabalho busca descrever o raciocínio de interpretação de rochas em um modelo abstrato e independente de implementação, assim permitindo compreender e explicar a forma que o conhecimento de domínio é usado para resolver problemas neste domínio de aplicação.

De forma geral, não foram encontrados PSMs descritos para *tarefas de interpretação* na literatura, assim como são definidos e apresentados para outros *tipos de problemas* (Breuker [BRE 94]). Logo, bibliotecas de PSMs para interpretação poderiam ser usadas para construir KBSs em diferentes domínios, pois estas tarefas são elementos centrais de muitas aplicações importantes (Doyle et al. [DOY 96]), além de viabilizarem o processo de negócio de diferentes tipos de empresas (como exemplificado por Cochran et al. [COC 97]). Todavia, não foi encontrado um estudo mais abrangente dos requisitos de conhecimento de tarefas de interpretação, tal como modelos de conhecimento genéricos descrevendo os principais conceitos comuns a processos de interpretação baseados em modelos, tal como poderia ser visto o esquema de raciocínio para interpretação de rochas.

Além disso, poucas aplicações industriais de PSMs têm sido descritas na literatura, não permitindo consolidar de forma prática essa abordagem de análise e representação de conhecimento inferencial. Mesmo assim, algumas aplicações mais conhecidas podem ser citadas (Gómez-Pérez et al. [GOM 99]: uso de uma biblioteca de PSMs de diagnóstico para construir um KBS destinado para o diagnóstico de processos de produção química, típicos em indústrias farmacêuticas; uso de PSMs para desenvolver um KBS para gerenciamento de tráfego rodoviário nas cidades de Madri e Barcelona (Espanha); uso de PSMs de escalonamento para construir um KBS para escalonamento de trabalho em processos de produção na IBM/Japão; uso de PSMs de projeto parametrizados no projeto de automóveis, como também no projeto de peças mecânicas (Motta et al. [MOT 98]); etc. Neste intuito, esse trabalho apresenta um estudo de caso de descrição de processos de raciocínio de interpretação de rochas usados em uma aplicação de Bancos de Dados Inteligentes destinada a auxiliar a análise de rochas-reservatórios de petróleo.

O texto desse trabalho é principalmente descrito em função de objetivos, ações e conhecimento, tal como proposto por Newell [NEW 82]. Para apresentar uma validação prática dos PSMs propostos, também é apresentada uma discussão simbólica de um mecanismo de inferência desenvolvido. Além disso, o texto não faz uma distinção significativa entre *informação* e *conhecimento*. Conforme Schreiber et al. [SCH 99], para a KE atual, conhecimento pode ser simplificadaamente considerado como informação complexa, geralmente afirmando algo sobre outra informação. Quando o conhecimento é modelado, a distinção entre conhecimento e informação torna-se mais difusa.

Entre outros termos usados, o termo *conceito* é usado com uma conotação simplificada, assim como proposto por diferentes abordagem de especificação de conhecimento baseadas na noção de ontologias (Gruber [GRU 93]); os termos *método* e *PSM* são usados com um mesmo significado; o termo *domínio de problemas* é usado com um sentido de expressar o conjunto de problemas comuns a um determinado domínio de conhecimento, tais como, por exemplo: o domínio de problemas de diagnóstico, o domínio de problemas de projeto, entre outros.

O trabalho está organizado da seguinte forma: o segundo capítulo apresenta uma revisão da KE atual segundo a perspectiva de PSMs. Tal revisão caracteriza esta noção de representação de conhecimento inferencial reusável, descreve diferentes abordagens da KE que empregam esta noção e define os componentes básicos de uma arquitetura de descrição de PSMs genérica. Em seguida, recentes propostas

metodológicas de desenvolvimento de PSMs são apresentadas e, por fim, os temas da KE que envolvem o estabelecimento de suposições e a caracterização de bibliotecas de PSMs são apresentados.

O terceiro capítulo apresenta a análise de um estudo de caso de desenvolvimento de PSMs. Tal estudo foi realizado para especificar os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos. Antes de descrever os passos de desenvolvimento destes PSMs, uma síntese do processo de construção de KBSs a partir de PSMs reusáveis é apresentada, bem como é descrito um resumo das principais abstrações do domínio de problemas de interpretação de rochas que são empregadas. A partir disso, os passos de desenvolvimento dos PSMs propostos são sintetizados e uma análise desse processo de especificação é apresentada.

O quarto capítulo descreve o esquema de raciocínio para interpretação de rochas. Tal esquema é centralmente definido em função dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos. Estes PSMs são detalhados em termos de suas competências, especificações operacionais e requisitos/suposições. Como o PSM Interpretação de Ambientes Diagenéticos é fundamentado em modelos de conhecimento de domínio construídos no âmbito do projeto *PetroGrapher*, tais modelos também são apresentados e analisados. Por fim, uma classificação simplificada de ambos os PSMs é realizada.

O quinto capítulo descreve os componentes simbólicos do mecanismo de inferência para interpretação de rochas. Em seguida, os componentes da arquitetura do sistema *PetroGrapher* são descritos brevemente. Tal sistema é ampliado pelo mecanismo de inferência proposto. Este mecanismo é projetado em uma arquitetura de inferência para interpretação de rochas, onde são definidos componentes de manipulação de conhecimento e dados. Para fundamentar estes componentes, são apresentados os mapeamentos desenvolvidos de modelos de conhecimento para modelos de dados, assim permitindo descrever o funcionamento de um algoritmo simbólico de inferência implementado.

O último capítulo apresenta algumas das conclusões alcançadas pelo desenvolvimento desse trabalho. Mesmo assim, pode ser adiantado que se espera que o esquema de raciocínio para interpretação de rochas seja usado para guiar novos passos de refinamento dos modelos de conhecimento do projeto *PetroGrapher*, como também venha a orientar o desenvolvimento de aplicações baseadas em conhecimento em outros domínios de aplicação. Desta forma, esse trabalho descreve um estudo de caso prático de especificação de um esquema de raciocínio para tarefas intensivas em conhecimento, as quais são comuns em aplicações reais de KBSs.

2 Uma visão da engenharia de conhecimento

Nos últimos anos, a KE tem buscado métodos e ferramentas adequados para construir KBSs em um modo sistemático e controlável (Studer et al. [STU 2000]). A pesquisa nessa área tem resultado em grandes avanços com respeito a estruturação de modelos de conhecimento, bem como a construção e reuso desses modelos. Mais recentemente, modelos de conhecimento de KBSs estão sendo bastante discutidos, pois o conhecimento está sendo identificado com um dos principais fatores de produção nas instituições, além de trabalho e capital.

Sistemas convencionais e KBSs podem ser simplificadaamente distintos, pois KBSs explicitamente representam grande quantidade de conhecimento usado para resolver uma tarefa desejada. Considerar explicitamente o *conhecimento sobre solução de problemas* é uma das principais diferenças entre o desenvolvimento de KBSs e outros tipos de software. Segundo Studer et al. [STU 2000], KBSs podem ser compreendidos como sistemas de computador que lidam com problemas complexos fazendo uso de conhecimento. Para isso, tal conhecimento pode ser adquirido de humanos ou automaticamente derivado com técnicas de raciocínio.

Segundo Studer et al. [STU 99], o objetivo da KE é similar ao objetivo da Engenharia de Software: construir sistemas de uma maneira sistemática e controlável. Isto requer a análise do processo de construção, manutenção e uso de métodos e ferramentas adequados para o desenvolvimento de KBSs. O desenvolvimento de KBSs por muito tempo foi tratado como um *processo de transferência*, onde o conhecimento humano era transferido para uma base de conhecimento. Esse conhecimento era, e ainda tem sido, implementado em alguma linguagem de representação, tal como, por exemplo: regras, frames, lisp, etc consideradas compreensíveis tanto por humanos, quanto por computadores. Todavia, alguns problemas da abordagem de transferência podem ser citados (Clancey [CLA 87]):

- diferentes tipos de conhecimento são representados de maneira uniforme;
- alguns tipos de conhecimento não são representados;
- nível de detalhe é muito alto para um modelo conceitual;
- a descrição do problema em si é constantemente misturada com aspectos de implementação.

Além disso, a aquisição de conhecimento de fontes de conhecimentos previamente existentes, como proposto pelo processo de transferência, não permite modelar de forma adequada a importância do *conhecimento tácito* (Nonaka et al. [NON 97]) para as capacidades de solução de problemas de especialistas humanos. Para tentar modelar algumas destas capacidades, a recente visão da KE tem sido compreendida como um *processo de modelagem*.

Para a KE atual, o conhecimento é modelado de forma independente de aspectos de implementação, permitindo identificar e modelar explicitamente diferentes tipos de conhecimento. Dessa forma, um engenheiro de conhecimento

desenvolve um KBS especificando e detalhando um conjunto de modelos que buscam realizar as capacidades de solução de problemas comparáveis as de um especialista humano. Portanto, o processo de aquisição de conhecimento, bem como o desenvolvimento de KBSs como um todo, pode ser sintetizado como um processo de construção de modelos.

Studer et al. [STU 99] apresenta uma caracterização mais precisa da KE como um processo de modelagem:

- assim como todo modelo, um conjunto de modelos de conhecimento desenvolvidos é somente uma aproximação da realidade;
- processo de modelagem é um processo cíclico, onde novas observações sobre a realidade podem levar a refinamentos ou modificações dos modelos construídos. Em contrapartida, os modelos podem orientar o processo de aquisição de conhecimento a ser desenvolvido;
- processo de modelagem é dependente de interpretações, muitas vezes subjetivas, do engenheiro de conhecimento. Dessa forma, tal processo está sujeito a falhas, requerendo uma avaliação dos modelos com respeito a realidade.

A focalização da KE com um processo de modelagem foi principalmente estruturada a partir do conceito de *nível do conhecimento (Knowledge Level - KL)* (Newell [NEW 82]). Segundo esse conceito, a descrição do processo de solução de problemas deve ser especificada em termos de *objetivos* a serem alcançados, *ações* necessárias para alcançar esses objetivos e *conhecimento* necessário para executar essas ações. Nesse nível de abstração, segundo Stefik [STE 95], a discussão envolve a análise de *qual* conhecimento é necessário para executar uma tarefa e *como* o conhecimento pode ser usado. Logo, é proposto um nível mais abstrato de descrição de sistemas, abstraindo características típicas do acesso e manipulação de padrões de símbolos, descritas em um *nível simbólico*.

A noção de conhecimento de Newell é similar a noção de tipos abstratos de dados: especificar *o que* é requerido e o comportamento desejado a partir de um conjunto de operações, e não as estruturas que podem ser usadas para realizar esse comportamento. O conceito de nível do conhecimento pode ser compreendido como um nível mais abstrato de descrição de problemas, tal como é visto o *nível conceitual* em bancos de dados.

2.1 As noções de ontologia e método de solução de problemas

Conforme Studer et al. [STU 2000], os maiores avanços recentes da KE compreendida como um processo de modelagem são as noções de *ontologias e métodos de solução de problemas*. Tais noções tem como objetivo o desenvolvimento de modelos de conhecimento e raciocínio reusáveis entre diferentes domínios e tarefas. Tal objetivo segue os princípios da Engenharia de Software, onde modelos são projetados como componentes de desenvolvimento reusáveis, os quais podem facilitar a construção de sistemas de informações clássicos. De maneira análoga,

espera-se que a composição de ontologias e PSMs reusáveis também orientem o desenvolvimento de KBSs, uma vez que então podem ser construídos a partir de componentes de software reusáveis.

Segundo Gruber [GRU 93], uma ontologia pode ser compreendida como "*uma especificação explícita formal de uma conceitualização compartilhada*". Uma especificação explícita significa que os tipos de conceitos (por exemplo: entidades, propriedades, relações, funções, axiomas, etc) e as restrições em seu uso são explicitamente definidos. Formal significa que a especificação é compreensível tanto por humanos, quanto por máquinas. Uma conceitualização faz referência a um modelo que abstrai algum fenômeno da realidade pela identificação dos conceitos relevantes daquele fenômeno. Por fim, o termo compartilhada reflete a noção que uma ontologia captura conhecimento que é de comum acordo de uma comunidade, o qual não é particular de um indivíduo, mas é aceito por um grupo. Segundo Chandrasekaran et al. [CHA 99], uma ontologia pode ser simplificada como "*um vocabulário de representação, tipicamente especializado para algum domínio ou assunto de interesse*". Mais precisamente, não é o vocabulário que é tratado como uma ontologia, mas as conceitualizações que os termos no vocabulário pretendem capturar.

Esse trabalho não pretende realizar uma análise mais objetiva dos temas de pesquisa que envolvem ontologias. Mesmo assim, uma caracterização bastante precisa e completa dessa área pode ser encontrada nos trabalhos de Chandrasekaran et al. [CHA 99], Uschold et al. [USC 96] e Fernández-López [FER 99], entre outros.

Tradicionalmente, KBSs têm utilizado mecanismos de raciocínio genéricos para derivar nova informação, os quais fazem uso de conhecimento representado como, por exemplo: regras de produção, frames, entre outros formalismos. Além disso, mecanismos de inferência como *encadeamento para frente* e *encadeamento para trás*, entre outros, são encarregados de tratar a dinâmica envolvida no processo de derivar nova informação. Entretanto, pode ser observado que especialistas costumam usar conhecimento sobre a dinâmica do processo de solução de problemas para extrair o comportamento de solução desejado. Por exemplo, especialistas muitas vezes implicitamente codificam algum conhecimento de controle pela ordenação de regras de produção, e estas regras associadas a mecanismos de inferência genéricos derivam o comportamento de solução desejado.

Segundo Fensel et al. [FEN 98b], esse conhecimento de controle, que muitas vezes é expresso implicitamente, é fundamental para viabilizar a solução de problemas na prática. Tornar o conhecimento de controle explícito e considerá-lo como parte importante de todo o conhecimento envolvido em um KBS é o princípio racional que envolve PSMs.

Como descrito por Benjamins et al. [BEN 98], PSMs descrevem o processo de raciocínio de um KBS de uma maneira independente de implementação e domínio. Nessa mesma linha, segundo Fensel et al. [FEN 2000], PSMs descrevem os componentes de raciocínio de KBSs como padrões de comportamento que podem ser reutilizados entre aplicações. Portanto, um PSM é um modelo que expressa, de modo abstrato e independente de implementação, a forma que uma inferência deve acontecer como um padrão de raciocínio reusável entre domínios.

Conforme Benjamins et al. [BEN 98], PSMs podem ser usados para vários propósitos. Dentre eles, alguns podem ser destacados:

- construção de KBSs - um PSM pode ser reutilizado na construção de um KBS, assim facilitando o desenvolvimento de mecanismos de solução de problemas;
- especificação de KBSs - um PSM pode descrever um processo de raciocínio que alcança os objetivos de uma tarefa, assim orientando o projeto de um KBS;
- modelagem cognitiva - um PSM pode descrever um modelo cognitivo humano de solução de problemas. Todavia, ainda não há uma compreensão muito precisa sobre a extensão em que um PSM possa ser usado para produzir explicações adequadas de processos de raciocínio humanos.

Uma caracterização da noção de PSMs em função de exemplos e relações com outras formas de especificação conhecidas é apresentada na seção seguinte desse trabalho, assim permitindo compreender melhor o foco principal de pesquisa abordado.

2.2 Uma caracterização de métodos de solução de problemas

O detalhamento de um PSM envolve descobrir um comportamento de solução de problema comum a uma classe de problemas e abstrair esse comportamento em um padrão de inferência genérico reusável em diferentes domínios. Logo, PSMs especificam mecanismos de inferência particulares (típicos de tarefas) de forma generalizada, assim permitindo a descrição e controle mais direta do processo de raciocínio em análise.

O primeiro PSM proposto na literatura foi o *PSM Classificação Heurística* (Clancey [CLA 85]). Este PSM foi desenvolvido a partir da engenharia reversa de sistemas especialistas desenvolvidos para diferentes aplicações e domínios, os quais utilizavam diferentes formalismos de representação, tais como: regras, frames, lógica, etc. Segundo Clancey, o processo de solução de problemas desses sistemas poderia ser sintetizado em um processo de solução comum, e este poderia ser descrito de forma abstrata.

A Figura 2.1 apresenta a descrição do processo de solução de problemas do PSM Classificação Heurística. Tal descrição está especificada como uma *estrutura de inferência*.

Em resumo, uma *estrutura de inferência* modela todos os caminhos possíveis de solução de um problema, mas não indica uma estratégia de solução particular. Como pode ser percebido, o PSM Classificação Heurística é descrito por 4 *papéis de conhecimento* e 3 *passos de inferência*:

- um passo de abstração de dados (abstrair) que abstrai valores concretos como, por exemplo, "temperatura = 40 °C" para valores abstratos de dados como "febre alta";
- um casamento de padrões heurístico (confrontar heurístico - *heuristic match*) que usa a descrição abstrata dos dados para heurísticamente estabelecer possíveis classes de hipóteses, por exemplo: "classes de doenças";
- um passo de especificação (especificar) que encontra soluções por discriminação, por exemplo: "gripe", a partir de classes de hipóteses.

O conhecimento sobre "diagnóstico médico" apresentado como exemplo pode estar em diferentes passos do processo de solução de problemas. Nesse caso, em determinando momento o conceito "gripe" pode ser considerado como *soluções*, mas em outro momento do processo de raciocínio esse conceito pode ser considerado como *dados*, onde *soluções* poderia ser um tratamento possível, tal como: "vacina para gripe".

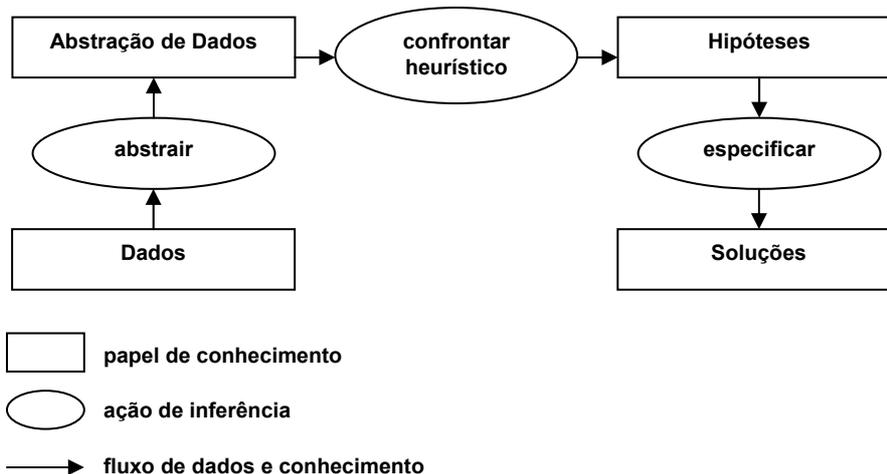


FIGURA 2.1 - Estrutura de inferência do PSM Classificação Heurística

Cada um dos passos de inferência do PSM Classificação Heurística requerem tipos de conhecimento específicos, os quais são considerados como recursos, por exemplo. Para o PSM Classificação Heurística, um passo de abstração de dados só pode ser executado se está disponível *conhecimento hierárquico sobre dados*. Um passo de casamento de padrões heurístico só pode ser executado se houver *conhecimento heurístico* que relacione abstrações de dados com hipóteses. Além disso, um passo de especificação de soluções só pode ser realizado se está disponível *conhecimento hierárquico sobre classes de soluções*.

Outro exemplo bastante conhecido que pode ser apresentado é o *PSM Diagnóstico Sistemático* (Gardner et al. [GAR 98]). Esse PSM "*determina a causa e a localização de um problema pelo uso de hipóteses e testes*" (Figura 2.2).

A estrutura de inferência do PSM Diagnóstico Sistemático apresenta 9 papéis de conhecimento: *reclamação, modelo do sistema, modelo do subsistema, hipóteses,*

testes apropriados, valores de dados, normas de testes, diferencial e resultado. O mesmo apresenta 5 ações de inferência distintas: *selecionar, especificar, decompor, comparar* e *confrontar*. O método apresentado na Figura 2.2 foi descrito com instâncias de conhecimento de domínio, as quais foram mapeadas para papéis de conhecimento definidos. Estas instâncias podem ser usadas para exemplificar o funcionamento desse padrão de raciocínio em um "diagnóstico de falhas mecânicas em automóveis".

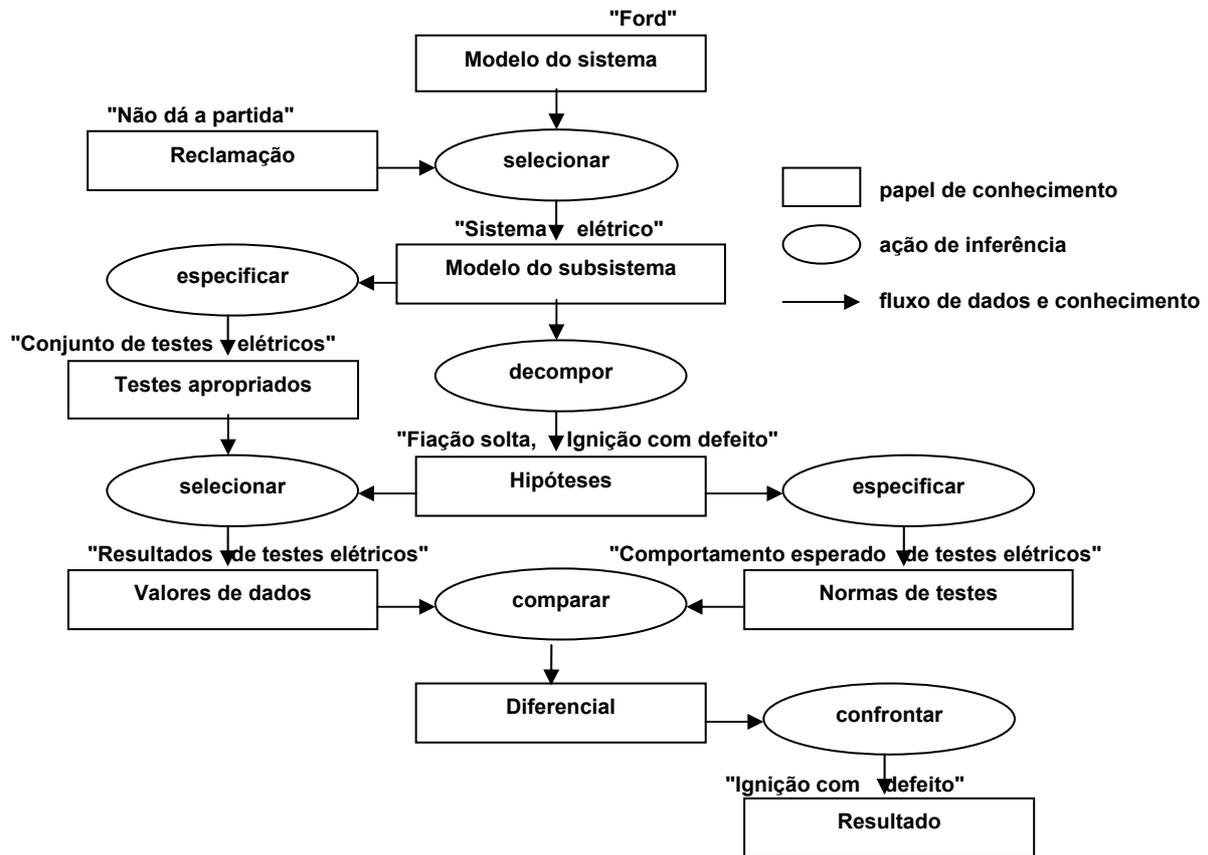


FIGURA 2.2 - PSM Diagnóstico Sistemático:
um exemplo para diagnóstico de falhas mecânicas em automóveis

Em princípio, é tomado como *modelo de sistema* um automóvel “Ford”. Uma *reclamação*, tal como: “não dá a partida”, também é usada para disparar o processo de raciocínio. Esse conhecimento de domínio associado ao *modelo de sistema* e a *reclamação* então é usado para selecionar um *modelo de subsistema*, por exemplo, um “sistema elétrico”. O “sistema elétrico” é então especificado em *testes apropriados* a serem realizados, tais como, um “conjunto de testes elétricos”. Em paralelo, o conhecimento de domínio associado ao *modelo do subsistema* é decomposto em *hipóteses* a serem analisadas, tais como: “fiação solta e ignição com defeito”. Em seguida, o conhecimento sobre *testes apropriados* é selecionado pelas *hipóteses* decompostas, resultando em *valores de dados*, tais como, “resultados de testes elétricos”. Em paralelo, as *hipóteses* são especificadas em *normas de testes*, tais como, um “comportamento esperado de testes elétricos”. Os “resultados de testes elétricos” e o “comportamento esperado de testes elétricos” são então comparados produzindo um *diferencial*. Este conhecimento diferencial é confrontado com

conhecimento sobre *resultados*, derivando como resultado para o processo de raciocínio que a “ignição está com defeito”.

Usando a mesma estrutura de inferência da Figura 2.2, podemos aplicar o mesmo padrão de raciocínio para um "problema de diagnóstico médico" (assim descrito na Figura 2.3).

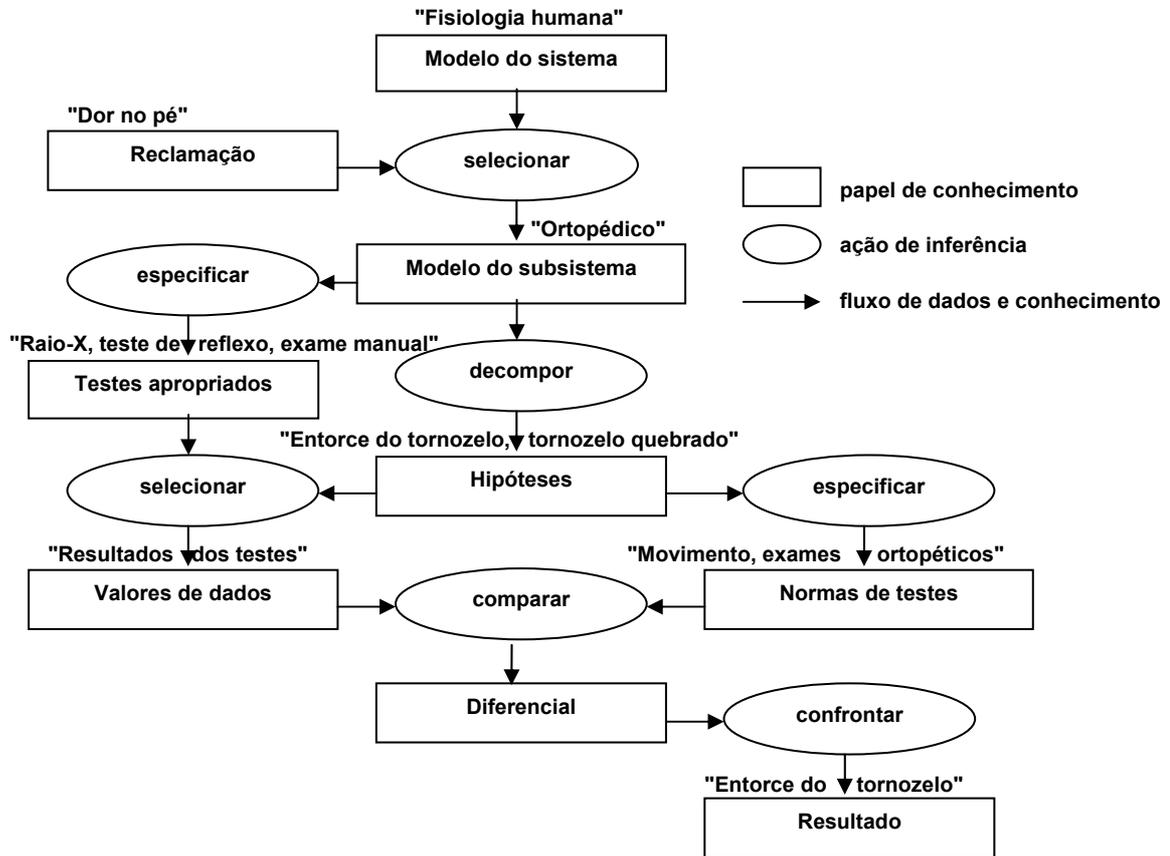


FIGURA 2.3 - PSM Diagnóstico Sistemático: um exemplo para diagnóstico médico

Em princípio, uma "dor no pé" é tomada como *reclamação*. Um *modelo do sistema*, tal como: um modelo da “fisiologia humana”, também é usado para dar início ao processo de raciocínio. O conhecimento associado ao *modelo de sistema* e a *reclamação* selecionam um *modelo de subsistema*, por exemplo, um subsistema “ortopédico”. Este subsistema é decomposto em *hipóteses*, tais como: “entorce do tornozelo e tornozelo quebrado”. Em paralelo, são especificados *testes apropriados*, tais como: “raio-X, testes de reflexo e exames manuais”. As *hipóteses* estabelecidas são especificadas em *normas de testes*, tais como: normas de “movimentos e exames ortopédicos”. Em paralelo, as *hipóteses* são usadas para selecionar *valores de dados* dos *testes apropriados*, produzindo “resultados dos testes”. Em seguida, os “resultados dos testes” e o conhecimento utilizado como *normas de testes* são comparados produzindo um comportamento *diferencial*. Este *diferencial* é então confrontado com *resultados* que podem ser alcançados, assim validando uma das hipóteses, tal como: “entorce de tornozelo”, por exemplo.

Como pode ser notado, um mesmo PSM pode ser aplicado em problemas diferentes, mesmo em domínios de aplicação distintos: "diagnóstico de falhas

mecânicas em automóveis" (Figura 2.2) e "diagnóstico médico" (Figura 2.3). Dessa forma, esse método não é compreendido como a descrição de um raciocínio particular de uma aplicação, mas como um padrão de raciocínio que pode ser usado em diferentes domínios de aplicações, bem como reutilizado em diferentes KBSs.

Conforme Schreiber et al. [SCH 99], a modelagem de PSMs pode ser complementada pela modelagem de dados, pois um PSM requer conhecimento e dados de domínio para alcançar os objetivos de uma tarefa. O conhecimento de domínio requerido pode ser modelado utilizando um modelo de dados convencional, tal como, por exemplo, um modelo de objetos. Este modelo então associado a descrição do PSM poderá descrever uma estrutura típica para o conhecimento de domínio requerido pelo método, assim viabilizando a execução dos passos de inferência especificados no PSM. Um exemplo disso pode ser analisado nas Figuras 2.4 e 2.5.

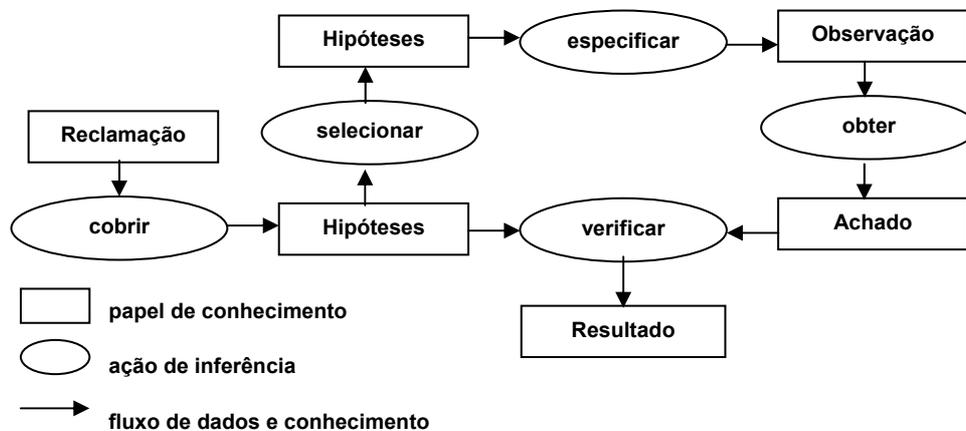


FIGURA 2.4 - Um PSM de Diagnóstico

A Figura 2.4 apresenta a estrutura de inferência de um PSM de Diagnóstico. A partir dessa estrutura, pode ser derivado um esquema de domínio típico representado em algum modelo de dados. Um desses esquemas típicos é apresentado na Figura 2.5 em um formalismo simples, tal como um modelo de objetos.

A Figura 2.5 apresenta um esquema de domínio requerido para viabilizar a solução de problemas de "diagnósticos de falhas". Esse modelo assume que um sistema sendo diagnosticado pode ser caracterizado em termos de "características observáveis" e "não-observáveis". "Características não-observáveis" são aquelas que representam estados internos do sistema. Nesse caso, "falhas no sistema" são definidas como subtipos de "estados internos do sistema", expressando que nem todo "estado interno" pode ser uma "falha".

O modelo causal do sistema usado pelo passo de inferência *cobrir* pode ser representado como uma "dependência causal", modelada possivelmente como uma regra. Esta regra descreveria "estados do sistema" que poderiam causar uma "característica do sistema". Tal característica pode ser outro estado ou um valor observado. Portanto, pode ser percebido que um PSM pode ser apresentado com um modelo de dados representando um esquema típico para o conhecimento de domínio requerido.

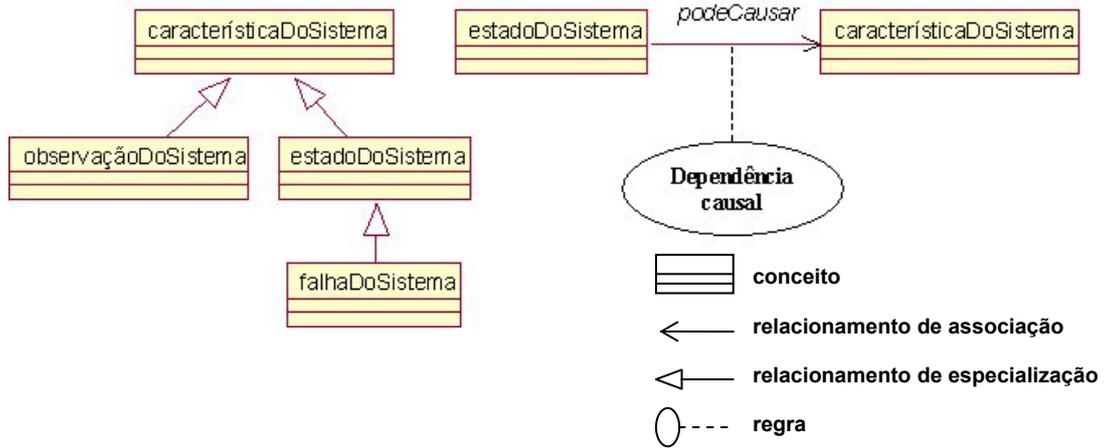


FIGURA 2.5 - Esquema de domínio típico de um PSM de Diagnóstico

A modelagem de conhecimento envolvida na noção de PSMs, por sua vez, pode ser considerada bastante diferente de abordagens funcionais convencionais de modelagem, onde objetos do domínio são diretamente associados com funções identificadas (Schreiber et al. [SCH 99]). Para exemplificar esta distinção, uma parte de um "PSM Diagnóstico" (Figura 2.4) pode ser analisada em uma aplicação de "diagnóstico de falhas mecânicas em automóveis". A Figura 2.7 apresenta um Diagrama de Fluxo de Dados (DFD) modelado a partir da representação apresentada na Figura 2.6.

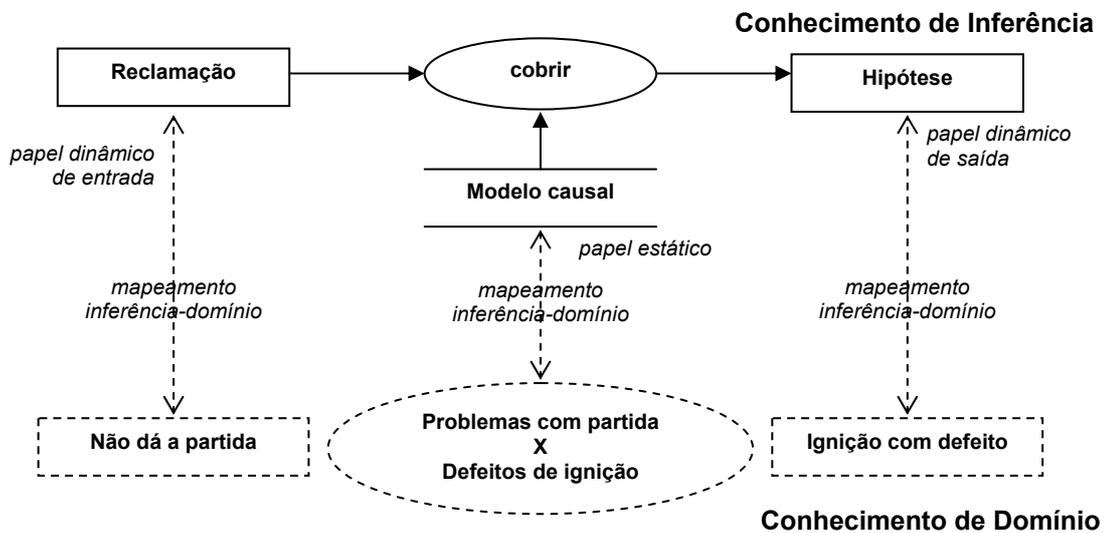


FIGURA 2.6 - A relação entre conhecimento de domínio e conhecimento de inferência

Uma inferência simples, tal como *cobrir*, utiliza uma *reclamação*, tomada como uma área de armazenamento dinâmica de entrada (papel de entrada), e uma *hipótese*, tomada como uma área de armazenamento dinâmica de saída (papel de saída). Também usa um *modelo causal* (papel estático) que disponibiliza o conhecimento estático de domínio requerido pelo passo de inferência *cobrir*. Em modelagem de conhecimento, estes papéis de conhecimento abstraem aspectos de

domínio. Dessa forma, para utilizar este método de raciocínio devem ser demonstrados os devidos mapeamentos dos conceitos de domínio modelados para os papéis de conhecimento especificados no PSM. No caso da Figura 2.6, "não dá a partida" pode ser mapeado para *reclamação*, conhecimento causal relacionando "problemas com partida com defeitos de ignição" pode ser mapeado para *modelo causal* e "ignição com defeito" pode ser mapeado para *hipótese*.



FIGURA 2.7 - Uma representação em Diagrama de Fluxo de Dados

Na Figura 2.7, os papéis de conhecimento desapareceram e foram substituídos por conceitos de domínio específicos ("não dá a partida", "ignição com defeito" e "problemas com partida *versus* defeitos de ignição"). Mesmo sendo mais simplificada, essa representação em DFD torna o reuso do conhecimento de inferência especificado mais difícil. A noção de papéis de conhecimento de PSMs viabilizam algum reuso mais explícito do conhecimento de inferência, pois eles são normalmente especificados em uma terminologia distinta de aplicação, abstraindo explicitamente o comportamento de solução de problemas de um domínio de conhecimento específico.

Por fim, a noção de PSMs está presente em diversas abordagens de KE, embora apresentando princípios de representação algumas vezes diferentes. Uma breve descrição de algumas dessas abordagens será apresentada na seção seguinte desse trabalho.

2.3 Abordagens de engenharia de conhecimento

De forma geral, as metodologias de Engenharia de Software estão voltada para os requisitos de usuários, os quais são modelados em função de dados, informações, processos e procedimentos. A KE, por sua vez, focaliza os requisitos de perícia de especialistas e de usuários, os quais são sistematicamente especificados de forma a demonstrar a aplicação dos dados, informações, processos e procedimentos modelados. Além disso, a KE envolve conceitos e estratégias de solução de problemas, os quais muitas vezes são compostos por dados heurísticos e subjetivos eliciados. Tal como é notado, a KE atual focaliza aspectos de aquisição de conhecimento e técnicas de análise como um processo de construção de modelos, assim buscando capturar, estruturar e representar mais fielmente o conhecimento de domínio e de inferência requeridos por KBSs.

Conforme Benjamins et al. [BEN 97], as abordagens metodológicas da KE moderna têm fornecido uma estrutura para o processo de desenvolvimento de KBSs. Tal estrutura é identificada por modelos intermediários de especificação, além da definição de linguagens e aspectos de organização requeridos para definir estes

modelos. Essas abordagens de KE têm sido apresentadas para responder aos requisitos industriais de desenvolvimento sistemático de KBSs, como oposto à prática de prototipação rápida ainda muito difundida, mas um tanto ultrapassada perante os requisitos atuais de desenvolvimento de software.

Segundo Schreiber et al. [SCH 99], a principal diferença entre várias metodologias de KE atuais e outras de Engenharia de Software tradicionais é a proposta explícita de um *modelo de conhecimento*, originalmente denominado de *modelo de perícia (expertise model)*. Esse modelo especifica os requisitos de conhecimento e raciocínio de KBSs, descrevendo os principais tipos de conhecimento requeridos por uma aplicação baseada em conhecimento. Ele é muitas vezes desenvolvido como parte do processo de análise e, desta forma, é descrito no nível do conhecimento.

Nas seções seguintes desse trabalho, serão apresentadas brevemente as seguintes abordagens de KE: Métodos de Limitação de Papéis, Tarefas Genéricas, CommonKADS, MIKE e Protégé-II. Tal apresentação pretende demonstrar que a noção de PSMs é comum a diferentes propostas metodológicas da KE.

2.3.1 Métodos de Limitação de Papéis

Métodos de Limitação de Papéis (*Role-Limiting Methods* - RLM) (Marcus [MAR 88]) foram umas das primeiras tentativas de suportar o desenvolvimento de KBSs explorando a noção de PSMs reusáveis. O enfoque RLM pode ser caracterizado por uma ferramenta (*shell*) de desenvolvimento de KBSs, onde esta ferramenta contém a implementação de um PSM específico. Somente esse PSM pode ser usado na solução de diferentes tipos de tarefas, as quais o PSM implementado está apto a resolver. Além disso, o PSM disponibilizado determina uma representação de conhecimento para papéis de conhecimento que são definidos e fixos, de forma que o engenheiro de conhecimento precisa instanciar os conceitos e relacionamentos que são definidos nos papéis do PSM.

A estrutura pré-definida da base de conhecimento do enfoque RLM pode ser usada para orientar o processo de aquisição de conhecimento, pois os tipos de conhecimento que devem ser fornecidos pelo especialista do domínio são especificados a priori. Porém, em situações reais, é difícil determinar se uma tarefa particular pode ser resolvida usando um RLM. Além disso, o enfoque RLM tem uma estrutura fixa, não fornecendo uma base adequada para casos onde uma tarefa particular é resolvida pela combinação de vários PSMs.

Para tornar RLM mais flexível, o conceito de RLM configurável (*Configurable Role-Limiting Methods* - CRLM) foi adotado, assim explorando a idéia que um PSM complexo pode ser decomposto em várias subtarefas, onde cada uma delas pode ser resolvida por métodos alternativos. Como tarefas podem possuir subtarefas em comum, o enfoque CRLM disponibiliza um conjunto pré-definido de diferentes métodos para resolver diferentes subtarefas. Segundo este enfoque, um PSM pode ser "configurado" pela seleção de um método para cada subtarefa identificada.

O enfoque CRLM disponibiliza meios de configurar a ferramenta CRLM para diferentes tipos de tarefas. Essa ferramenta contém um esquema fixo de tipos de conhecimento, onde este pode ser estendido/modificado para permitir a introdução de um novo método. O esquema fixo de tipos de conhecimento e relacionamentos pré-definidos entre seus componentes é uma restrição importante que distingue CRLM de abordagens de configuração mais flexíveis, tais como, por exemplo, CommonKADS (apresentada em uma seção seguinte desse trabalho). Segundo Studer et al. [STU 98], a introdução dessa flexibilidade no enfoque de RLM remove uma das suas desvantagens, enquanto ainda permite explorar a vantagem de um esquema fixo de tipos de conhecimento, o qual pode ser utilizado para construir ferramentas de aquisição de conhecimento. Em contrapartida, a configuração da ferramenta CRLM pode dificultar o desenvolvedor de KBSs, visto que é necessário conhecer como configurar o sistema da maneira correta.

2.3.2 Tarefas Genéricas

Conforme Chandrasekaran [CHA 86], Tarefas Genéricas (*Generic Tasks - GTs*) são blocos de desenvolvimento de software (tais como componentes) genéricos que podem ser reutilizados para a construção de diferentes KBSs. A idéia dessa abordagem de KE pode ser sintetizada como:

- uma GT é associada com uma descrição genérica de suas entradas e saídas;
- uma GT tem um esquema fixo de tipos de conhecimento especificando a estrutura de conhecimento de domínio necessária para resolver uma tarefa;
- uma GT inclui uma estratégia de solução de problema detalhando passos de inferência, bem como uma seqüência em que esses passos devem ser executados.

Esse enfoque é baseado na hipótese de que a estrutura e representação de conhecimento de domínio é completamente determinada pelo seu uso. Portanto, uma GT contém uma estratégia de solução de problema e uma coleção de estruturas de conhecimento fixas. Como uma GT fixa o tipo de conhecimento que é necessário para resolver uma tarefa, um vocabulário específico de tarefa é disponibilizado, assim podendo ser explorado no processo de aquisição de conhecimento. Além disso, uma ferramenta genérica denominada *arquitetura específica de tarefa* pode implementar um KBS específico como instâncias de tipos de conhecimento pré-definidos, utilizando termos específicos de domínio. Porém, esta abordagem apresenta algumas desvantagens, pois a noção de tarefa se confunde com a do PSM usado para resolver a tarefa, visto que cada GT inclui uma estratégia de solução de problema pré-definida.

Baseado nessa desvantagem, o enfoque denominado *estrutura de tarefa* foi proposto (a Figura 2.8 apresenta um exemplo de uma estrutura de tarefa para diagnóstico). Esta estrutura distingue uma tarefa, que é usada para referir um *tipo de problema* (Breuker [BRE 94]), de um PSM, que é o modo de atender uma tarefa. A estrutura associa uma tarefa com um conjunto de métodos alternativos utilizados para resolvê-la, onde cada método pode ser decomposto em outras subtarefas. A estrutura de decomposição pode ser refinada até um nível onde subtarefas elementares sejam

definidas, e estas podem diretamente ser resolvidas usando o conhecimento disponível.

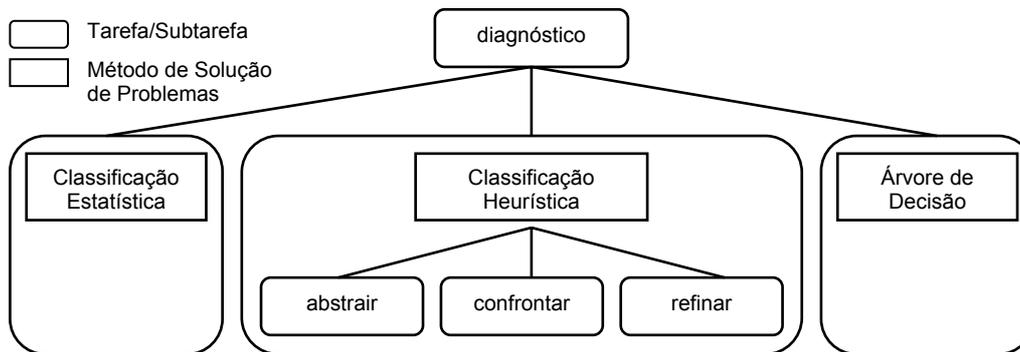


FIGURA 2.8 - Um exemplo de uma estrutura de tarefas para diagnóstico modelada como Tarefas Genéricas

2.3.3 CommonKADS

A metodologia CommonKADS (Schreiber et al. [SCH 94]; Schreiber et al. [SCH 99]) destaca-se por ser a metodologia de KE mais difundida e testada em projetos reais. CommonKADS propõe a construção de um conjunto de modelos, onde cada modelo captura diferentes aspectos do KBS, bem como seu ambiente. Essa metodologia é composta pelos modelos descritos a seguir:

1) *modelo de organização*: permite analisar as principais características de uma organização, no sentido de identificar problemas e oportunidades para KBSs, bem como os impactos do uso explícito de conhecimento. Nesse modelo, podem ser descritas a estrutura da organização e a especificação das funções que são executadas por cada unidade organizacional, entre outros. A partir disso, é possível identificar deficiências em processos de negócios, bem como identificar possíveis melhorias que podem ser obtidas com a introdução de KBSs.

2) *modelo de tarefa*: descreve as hierarquias das principais tarefas executadas por processos de negócio em unidades organizacionais onde o KBS será instalado. Essa análise identifica a estrutura global da tarefa, suas entradas e saídas, pré-condições e critérios de eficiência, bem como os recursos e competências requeridos.

3) *modelo de agente*: especifica as capacidades de cada agente (considerados executores de tarefas) envolvido na execução de uma tarefa. Em geral, agentes podem ser humanos, sistemas ou qualquer outra entidade capaz de executar uma tarefa. Esse modelo principalmente descreve a competência e restrições dos agentes. O modelo de comunicação posterior então relaciona estas descrições na execução de uma tarefa.

4) *modelo de comunicação*: especifica as interações entre os diferentes agentes envolvidos na execução de uma tarefa. Entre outras coisas, os tipos de informação trocados entre os agentes são especificados de modo independente de implementação, em um nível conceitual.

5) *modelo de conhecimento*: é fundamentado em diferentes tipos de conhecimento: *conhecimento de tarefa*, *conhecimento de inferência* e *conhecimento de domínio*. Esses tipos de conhecimento são organizados em diferentes camadas, onde primitivas de representação de conhecimento são disponibilizadas para a modelagem estruturada de conhecimento (um dos principais desafios da KE). O modelo tem uma estrutura similar aos tradicionais modelos de análise de Engenharia de Software, onde uma tarefa é descrita através de uma decomposição hierárquica de funções e processos. Os tipos de dados operados por funções, por sua vez, são descritos através de um esquema que lembra modelos de dados ou modelos de objetos tradicionais.

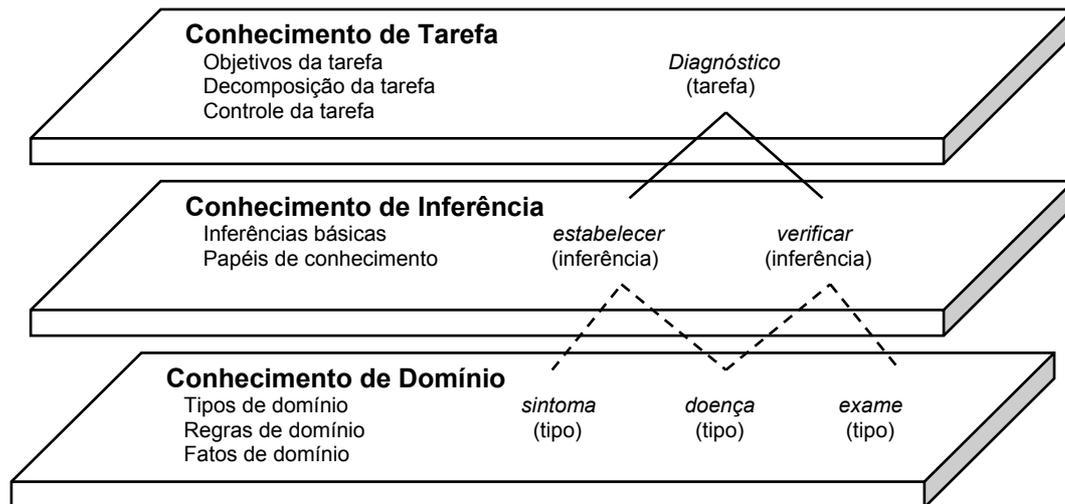


FIGURA 2.9 - As camadas do modelo de conhecimento

O modelo de conhecimento tal como apresentado na Figura 2.9 tem três camadas, onde cada uma expressa uma *categoria de conhecimento*:

- *conhecimento de tarefa* - permite descrever os objetivos de uma aplicação, e como esses objetivos podem ser satisfeitos através de uma decomposição tarefas/subtarefas e, por fim, inferências elementares. O conhecimento de tarefa pode ser considerado similar a altos níveis de decomposição funcional em Engenharia de Software, mas também inclui algum tipo de controle sobre as tarefas envolvidas. Um exemplo típico desse conhecimento é uma "tarefa de diagnóstico", entre outros;
- *conhecimento de inferência* - permite descrever os passos básicos de inferência que serão realizados utilizando o conhecimento de domínio e os papéis desempenhados por esse conhecimento, assim como requeridos por um PSM. Inferências são melhor compreendidas como blocos de desenvolvimento de mecanismos de raciocínio. Para a Engenharia de Software, inferências poderiam ser descritas como níveis mais básicos de decomposições funcional, tal como folhas em uma árvore de funções/subfunções. Exemplos de inferências poderiam ser: *estabelecer* relações entre sintomas e doenças, *verificar* a existência de exames para as doenças estabelecidas, entre outros;

- *conhecimento de domínio* - permite especificar os tipos de informação e o conhecimento da aplicação que podem ser utilizados. Tais tipos poderiam expressar conceitos, propriedades, relações, como também poderiam expressar regras e fatos de um domínio. Por exemplo, o conhecimento de domínio de uma aplicação de diagnóstico médico poderia conter definições relevantes para "doenças", "sintomas" e "exames", bem como relacionamentos entre esses conceitos (Figura 2.9).

Um exemplo desse modelo de conhecimento é apresentado na Figura 2.10.

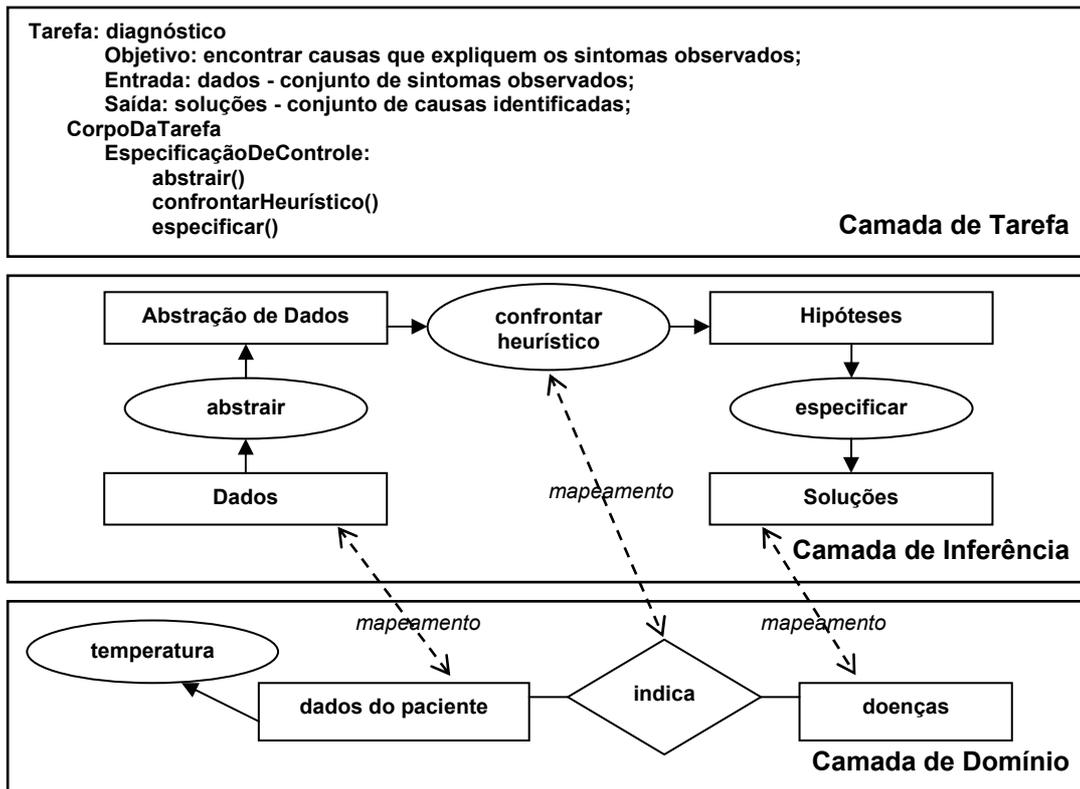


FIGURA 2.10 - Um Exemplo para o Modelo de Conhecimento

Essa divisão em camadas tem como objetivo viabilizar algum tipo de reuso do modelo de conhecimento. Conforme apresentado na Figura 2.10 (notação CML simplificada - Schreiber et al. [SCH 94]), os conceitos de domínio (no exemplo: "dados do paciente", "temperatura" e "doenças") e relacionamentos entre esses conceitos (no exemplo, o relacionamento "indica") especificados na camada de domínio podem ser reutilizados para resolver diferentes tarefas (no exemplo: "tarefa de diagnóstico") pela aplicação de diferentes PSMs detalhados na camada de inferência. Um PSM genérico (no exemplo: Classificação Heurística) também pode ser reusado em diferentes domínios pela definição de outros mapeamentos para a camada de domínio.

6) *modelo de projeto*: os modelos apresentados anteriormente podem ser vistos como especificações de requisitos de KBSs. Em especial, os modelos de conhecimento e comunicação capturam os requisitos funcionais do KBS. Baseado nesses requisitos, o modelo de projeto apresenta uma especificação técnica para o sistema, descrita em termos de arquitetura, plataforma de implementação e outros

mecanismos computacionais necessários para implementar a funcionalidade requerida. A metodologia CommonKADS propõe que a estrutura do modelo de projeto deve refletir a estrutura do modelo de conhecimento tanto quanto possível.

2.3.4 MIKE

O enfoque Engenharia de Conhecimento Incremental e Baseada em Modelo (*Model-Based and Incremental Knowledge Engineering - MIKE*) (Angele et al. [ANG 98]) fornece um método de desenvolvimento de KBSs com o objetivo de integrar técnicas de especificação semi-formal, formal e prototipação em um estrutura de KE. Conforme Studer et al. [STU 99], a integração da prototipação e suporte para um processo de desenvolvimento incremental e reversível em um estrutura baseada em modelos é a principal distinção entre MIKE e CommonKADS, pois:

- MIKE utiliza o *modelo de conhecimento* de CommonKADS como um padrão de modelo. Esta abordagem permite um processo de transição que parte de uma representação semi-formal: o *modelo de estrutura*, para uma representação formal: o *modelo KARL*, e posteriormente para uma representação de implementação denominada *modelo de projeto*;
- o modelo KARL é executável, assim permitindo validar o modelo de conhecimento por meio de prototipação, bem como permitindo integrar o especialista no processo de desenvolvimento.

O processo de desenvolvimento em MIKE pode ser sintetizado pelas seguintes atividades: *eliciação, interpretação, formalização/operacionalização, projeto e implementação*:

- a eliciação envolve a aquisição de conhecimento, basicamente realizada com técnicas de *entrevistas estruturadas*. Como resultado dessa primeira atividade, o conhecimento é expresso em linguagem natural, assim denominado *de protocolos de conhecimento*;
- a interpretação representa os *protocolos de conhecimento* em uma especificação semi-formal do *modelo de conhecimento* denominada de *modelo de estrutura*;
- a formalização/operacionalização é aplicada sobre o *modelo de estrutura*, resultando em um *modelo de conhecimento formal*, o *modelo KARL*. Este modelo formalizado captura os requisitos funcionais para o KBS final;
- o projeto envolve detalhar os requisitos funcionais especificados segundo aspectos computacionais, tais como, por exemplo: eficiência e facilidade de manutenção, entre outros;
- a implementação, por fim, implementa o KBS em um ambiente de software e hardware.

O resultado das fases de desenvolvimento de MIKE são refinamentos do *modelo de conhecimento* e, dessa forma, do modelo de PSM. De maneira geral, o conhecimento no modelo de estrutura é relacionado ao conhecimento correspondente nos protocolos de conhecimento via ligações explícitas. Conceitos e ações de inferência são relacionados aos nodos de protocolos inicialmente descritos em linguagem natural. O modelo de projeto refina o modelo KARL pelo refinamento das inferências nos algoritmos especificados e pela introdução de estruturas de dados adicionais. Estas partes do modelo de projeto são relacionadas com inferências correspondentes do modelo KARL e assim também são ligadas aos protocolos de conhecimento. Esse conjunto de relacionamentos permitem preservar a estrutura do *modelo de conhecimento* durante a fase de projeto do KBS.

O processo inteiro é executado em um ciclo de desenvolvimento, onde cada ciclo produz um protótipo do KBS, o qual já pode ser avaliado. Os resultados das avaliações podem então ser usados nos próximos ciclos para corrigir, modificar ou estender o protótipo corrente. Dessa forma, o enfoque MIKE pode ser considerado restrito à modelagem de um KBS sob desenvolvimento. Para capturar um KBS em um ambiente mais abrangente, tal como um ambiente de negócios, o enfoque MIKE tem sido estendido com novos modelos voltados para a definição de visões de negócio. Com isso, o conjunto de modelos pode vir a expressar uma transição desde um modelo de negócios, até um modelo de processos de solução de problemas.

2.3.5 Protégé-II

O principal objetivo do enfoque Protégé-II (Grosso et al. [GRO 99]) é suportar o reuso de ontologias e PSMs. Protégé-II está baseado em uma *estrutura de decomposição de tarefa-métodos*, onde uma tarefa é decomposta em subtarefas pela aplicação dos métodos. Esta estrutura de decomposição é refinada até um nível no qual métodos primitivos são usados para resolver as subtarefas. As entradas e saídas de um método são especificadas em uma *ontologia de método* (Chandrasekaran et al. [CHA 99]), onde esta ontologia fornece a definição dos conceitos e relações relevantes usados na especificação de um processo de raciocínio destinado a resolver uma tarefa.

Uma *ontologia de método* corresponde a uma terminologia genérica como descrita nos papéis de conhecimento de um PSM. Protégé-II também utiliza *ontologias de domínio*, onde algum conhecimento de senso comum de um domínio pode ser definido. Para o enfoque Protégé-II, PSMs e ontologias são compreendidos como componentes reusáveis para construir diferentes KBSs.

Devido ao *problema da interação* entre ontologias de domínio e PSMs, neste caso, tomados como ontologias de método, o enfoque Protégé-II propõe a noção de *ontologias de aplicação* (Heijst et al. [HEI 97]). De forma simplificada, uma ontologia de aplicação contém o conhecimento necessário para modelar uma aplicação particular. Estas ontologias estendem ontologias de domínio com conceitos e relacionamentos específicos de PSMs, assim viabilizando o mapeamento de uma ontologia de domínio para uma *ontologia de método*. A Figura 2.11 apresenta o relacionamento entre ontologias e PSMs do enfoque Protégé.

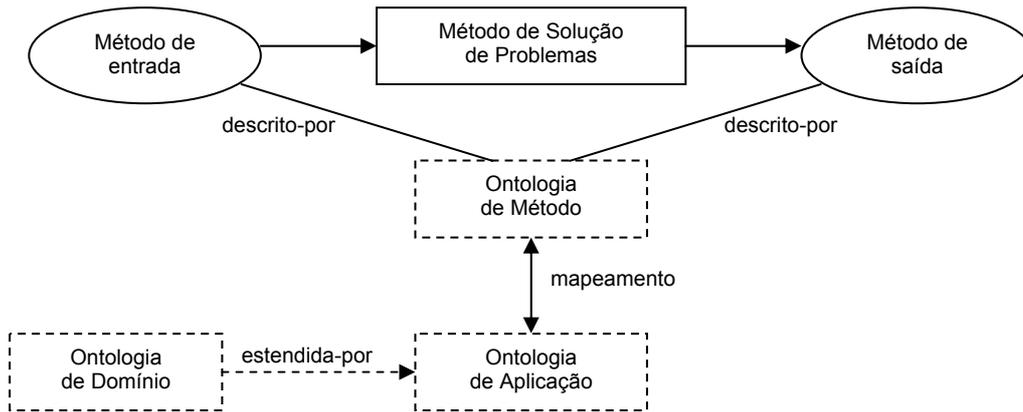


FIGURA 2.11 - Ontologias e PSMs em Protégé-II

Como citado por Grosso et al. [GRO 99], Protégé-II também tem se destacado por gerar ferramentas de aquisição de conhecimento personalizadas a partir de ontologias. Esse processo tem início com a especificação dos conceitos correspondentes a uma ontologia. A partir desses conceitos, Protégé-II gera automaticamente como saída uma ferramenta de aquisição de conhecimento, assim permitindo que especialistas de domínios de problemas sendo abordados acrescentem instâncias de conceitos modelados.

Recentemente, o enfoque Protégé-II está sendo associado a um ambiente WWW no âmbito do projeto IBROW3 (Armengol et al. [ARM 98]), assim permitindo que PSMs e ontologias especificados sejam acessíveis e reusáveis a partir de diferentes ambientes de software. O objetivo dessa associação é permitir a construção de KBSs em um ambiente internet, assim como estão sendo desenvolvidas diversas aplicações de software recentes.

2.4 Arquitetura de métodos de solução de problemas

A noção de PSMs é comum a diversas abordagens de KE, mas cada uma estabelece uma forma de especificação adequada aos seus requisitos conceituais de modelagem. Todavia, há consenso entre diferentes formas de especificação de PSMs que eles sejam compostos de 3 partes relacionadas (Figura 2.12). Tais partes descrevem *o quê* um PSM pode realizar, *como* ele pode realizar e, por fim, *o quê é necessário* para realizar.

As partes da arquitetura de PSMs genérica são referenciadas respectivamente como: *competência*, *especificação operacional* e *requisitos/suposições* (Benjamins et al. [BEN 98]). Simplificadamente, a especificação operacional descreve um modo de realizar a competência do método, desde que os requisitos/suposições em termos de conhecimento sejam satisfeitos. A Figura 2.12 apresenta um descrição das partes dessa arquitetura.

Cada uma das partes da FIGURA 2.12 é descrita a seguir.

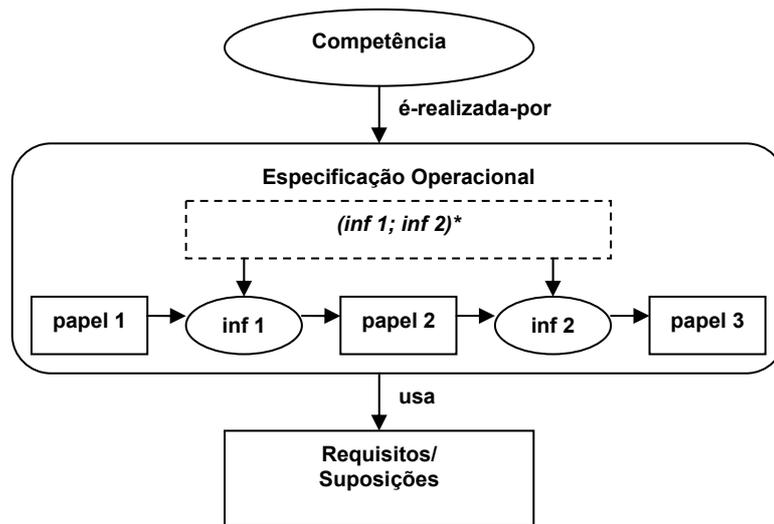


FIGURA 2.12 - A arquitetura de PSMs

2.4.1 Competência

Um PSM expressa uma competência para uma tarefa de raciocínio, ou seja, descreve *o quê* um PSM pode realizar de forma independente da realização em si. A competência de um PSM é descrita declarativamente em função de entradas e saídas do método. Outras propostas também estabelecem a descrição da competência de um PSM como uma extensão da descrição composicional do método, definindo os objetivos que podem ser alcançados independente de como estes objetivos serão alcançados.

Entretanto, ainda não existe um padrão de fato de como uma competência deva ser descrita, mas existem diferentes abordagens de descrição típicas de diferentes metodologias de KE. Em uma proposta bastante recente, uma linguagem de descrição de PSMs denominada UPML (*Unified Problem-Solving Method Development Language* - Fensel et al. [FEN 99]) estabelece que uma competência seja descrita de forma declarativa em função de papéis de entrada, papéis de saída, pré-condições, subtarefas e pós-condições. Mas tal proposta ainda requer uma validação mais adequada.

A competência de um PSM especifica a funcionalidade atual de um KBS (visto que o conhecimento de domínio atende os requisitos de inferência do PSM usado pelo KBS). Conforme Fensel et al. [FEN 99], a distinção entre a competência de um PSM e a competência de uma tarefa que deve ser resolvida pelo KBS pode ser sintetizada pelo seguinte:

- a competência de um PSM pode usar uma terminologia diferente da terminologia de uma tarefa;
- a funcionalidade de um PSM nem sempre é forte suficiente para resolver completamente um problema (ou tarefa);

- PSMs introduzem suposições sobre conhecimento de domínio para alcançar seus objetivos, enquanto a competência de uma tarefa não considera suposições sobre conhecimento disponível.

De forma geral, estabelecer a competência de um PSM requer uma definição de controle, pois somente assim é definida uma ordem de execução para os passos de inferência de um PSM. Além disso, a competência de um método está intimamente relacionada a solução de um problema de raciocínio uma vez disponibilizado o conhecimento de domínio requerido, pois somente dessa forma um método é capaz de alcançar seus objetivos.

Descrição de competência bastante simplificadas podem ser citadas. Por exemplo, para o PSM Classificação Heurística: "estabelecer hipóteses e encontrar uma solução usando conhecimento heurístico" e para PSM Diagnóstico Sistemático: "determinar a causa e a localização de um problema pelo uso de hipóteses e testes", ambos apresentados em Gardner et al. [GAR 98].

2.4.2 Especificação operacional

A especificação operacional de um PSM descreve *como* realizar uma tarefa de raciocínio em termos de *papéis de conhecimento, ações de inferência e especificação de controle*:

1) papéis de conhecimento

Expressam os papéis desempenhados pelo conhecimento de domínio em cada passo de inferência do PSM. Estes papéis são normalmente descritos em uma terminologia genérica, abstraindo aspectos de domínio. Segundo Gómez-Pérez et al. [GOM 99], podem ser distintos dois tipos de papéis de conhecimento:

- *papéis estáticos* - descrevem o conhecimento de domínio requerido pelo PSM;
- *papéis dinâmicos* - áreas de armazenamento que atuam como entrada e saída de passos de inferência.

A distinção explícita entre papéis estáticos e dinâmicos muitas vezes não é representada e, desta forma, simplificando a descrição do processo de raciocínio.

2) ações de inferência

Descrevem passos de raciocínio primitivos que juntamente realizam a competência do método. Por exemplo, o PSM Classificação Heurística (Figura 2.1) apresenta 3 ações de inferência: *abstrair*, *confrontar heurístico* e *especificar*. Cada uma das ações de inferência requer tipos de conhecimento específicos, os quais são considerados requisitos do método.

As ações de inferência mais comuns são padronizadas, assim definindo uma semântica precisa para a representação. Uma dessas padronizações mais conhecidas é a proposta pela metodologia CommonKADS (Schreiber et al. [SCH 99]).

3) especificação de controle

Descreve a ordem de execução das ações de inferência de forma que, uma vez sendo disponibilizado o conhecimento de domínio requerido pelo método, os objetivos de uma tarefa sejam eficientemente alcançados. Um PSM pode ter uma ou mais especificações de controle, onde elas podem ser expressas inicialmente (se forem conhecidas), ou podem ser oportunamente determinadas.

Uma especificação de controle é normalmente detalhada em termos de papéis de conhecimento e ações de inferência especificados. A partir deles, um fluxo de execução é detalhado com uma linguagem simplificada, por exemplo, contendo laços e desvios (tal como uma linguagem de *script*). Todavia, uma especificação de controle pode ser descrita em diferentes níveis de formalismo, desde uma representação informal, até uma representação formal, tal como a lógica. Tais especificações também podem ser descritas em um formato executável, assim sendo empregada como módulo de inferência de um KBS.

Entre outros formalismos de descrição de controle que poderiam ser citados, a linguagem CML (*Conceptual Modeling Language* - Schreiber et al. [SCH 94]) de CommonKADS é bastante conhecida. Mais recentemente, a linguagem UPML (Fensel et al. [FEN 99]) propõe um formalismo lógico para especificação de controle, o qual pode ser usado para operacionalização do método. De maneira mais geral, Fensel et al. [FEN 94] apresenta uma comparação de linguagens destinadas a operacionalizar e formalizar modelo de conhecimento CommonKADS.

2.4.3 Requisitos / suposições

Requisitos/suposições descrevem *o quê é necessário* pelo método em termos de conhecimento de domínio. Requisitos podem indicar, entre outras coisas, o conhecimento de domínio necessário para suportar os passos de inferência. Suposições, por sua vez, são necessárias para caracterizar a competência exata do PSM em termos das tarefas que podem ser resolvidas, além do conhecimento de domínio requerido para realizá-las.

A caracterização da funcionalidade de um KBS por meio de suposições é uma questão muito importante, pois mesmo em uma tarefa simples como a seleção do maior elemento dentre uma lista de elementos é requerido um relação de preferência como conhecimento de domínio. Assim, suposições são usadas para definir os requisitos de tal relação de preferência, por exemplo: transitividade, simetria, etc. O papel das suposições para a KE será melhor descrito na seção 2.6.

2.5 Desenvolvimento de métodos de solução de problemas

Mesmo com vários trabalhos recentes abordando PSMs, poucos têm focalizado o processo de modelagem envolvido no desenvolvimento desses métodos. Dentre os trabalhos analisados que apresentam propostas metodológicas de desenvolvimento, poucos foram testados e validados em domínios de problemas reais, os quais poderiam consolidar essas propostas de especificação de conhecimento inferencial ainda um tanto acadêmicas. Além disso, as metodologias de KE mais difundidas somente consideram PSMs completos (prontos), os quais seriam reusáveis em diferentes domínios e problemas. Portanto, tais metodologias de KE apresentam poucas indicações sobre os passos de desenvolvimento de PSMs a partir do zero.

Segundo Gómez-Pérez et al. [GOM 99], duas abordagens de desenvolvimento de PSMs consideradas tarefas criativas, com pouco suporte metodológico testado e validado, podem ser citadas:

- abordagem cognitiva - desenvolvidos a partir da análise do comportamento de solução de problemas de especialistas humanos. Isto tem sido tratado em uma área relacionada à Inteligência Artificial denominada Psicologia Cognitiva;
- engenharia reversa - desenvolvidos a partir da engenharia reversa de KBSs existentes. Esta foi a abordagem utilizada para desenvolver o PSM Classificação Heurística (Clancey [CLA 85]).

Porém, outras abordagens de desenvolvimento de PSMs, que não se enquadram nas duas anteriormente citadas, seguem alguns princípios no processo de desenvolvimento de PSMs:

- centralizado na tarefa - segundo esse princípio, PSMs são desenvolvidos a partir da tarefa a ser resolvida, onde esse processo envolve a seleção e refinamento/especialização de uma estratégia de solução de problemas;
- baseado em suposições - uma especificação mais precisa de métodos é derivada de requisitos informais pela introdução de suposições sobre o problema. Essa especificação é então refinada em uma especificação funcional de PSM por meio de suposições sobre estratégia de solução de problemas e teoria de domínio disponíveis.

A partir desses princípios, as seguintes propostas metodológicas podem ser citadas:

- desenvolvimento fundamentado em uma especificação da competência (Wielinga et al. [WIE 98]) - é um enfoque de desenvolvimento semi-automático, onde uma competência para o PSM é especificada em relação a uma tarefa e, passo-a-passo, esta competência é especializada em um PSM operacional resultante;

- desenvolvimento estruturado fundamentado em um sistema de eixos (Fensel et al. [FEN 2000]) - tais eixos são denominados: *estratégias de solução de problemas*, *suposições sobre conhecimento de domínio e compromissos de problemas (tarefas)*. Essa abordagem envolve o estabelecimento de diferentes *suposições* e *compromissos* sobre cada um dos eixos, assim especializando gradativamente uma especificação de referência do PSM para um novo PSM;
- desenvolvimento fundamentado em um tipo específico de problema de configuração (Teije et al. [TEI 96]) - processo fundamentado nos componentes básicos de especificação de PSMs. A idéia aqui é que esses componentes detalham parâmetros que devem ser configurados. Nessa abordagem, um *método de projeto* é usado como esquema de modelagem para gradativamente estabelecer, criticar e modificar os valores para parâmetros de um novo PSM em desenvolvimento, até que seja alcançado um PSM adequado.

De forma geral, não existe uma metodologia de desenvolvimento de PSMs consolidada e considerada como um padrão de fato. Portanto, as metodologias que serão analisadas nas seções seguintes desse trabalho ainda são propostas que requerem uma validação mais ampla perante o desenvolvimento de aplicações reais de KBSs.

2.5.1 Enfoque baseado na teoria da competência para a construção de métodos de solução de problemas

O enfoque baseado na teoria de competência para a construção de PSMs (Wielinga et al. [WIE 98]) é uma abordagem semi-automática de desenvolvimento. Para essa proposta, uma suposição central para PSMs é que estes podem ser formulados como uma descrição abstrata de um processo de raciocínio especificado de forma independente de domínio de aplicação. Então, para saber se um PSM é candidato a resolver um problema há a necessidade de responder duas questões básicas:

- o PSM pode resolver o problema da maneira requerida?
- o PSM resolve o problema de uma maneira eficiente?

A resposta da primeira questão já apresenta alguns indícios do que pode ser considerado uma *noção de competência de PSMs*. A resposta da segunda, por sua vez, está relacionada a aspectos de eficiência. Portanto, o processo fundamentado na teoria de competência explora passos de construção de PSMs que permitem estabelecer uma relação entre competência e eficiência. De maneira simplificada, um PSM pode ter uma competência para uma classe de problemas que ele é capaz de resolver, assim eficientemente alcançando um objetivo pelo uso de conhecimento.

A noção de competência de um PSM pode ser definida como uma descrição genérica da habilidade que um PSM requer para resolver uma certa classe de problemas. A utilidade conceitual e prática de um PSM pode ser demonstrada pela

competência dele em relação a um problema de aplicação. Para que isso seja realizado, essa proposta metodológica afirma que é necessário descrever como o processo de construção de PSMs é guiado pelo refinamento conceitual e operacionalização de uma teoria de competência inicial, onde esta é associada a uma especificação abstrata de objetivos.

O processo de construção de um PSM usualmente tem início com uma declaração informal de um objetivo a ser alcançado. A especificação desse objetivo inicial deve descrever as estruturas de informação que a solução deve satisfazer, tais como: tipos de conhecimento a serem usados, critérios de confiabilidade das entradas, número de soluções permitidas ou requeridas, etc. Tal declaração é considerada uma *teoria de competência inicial*, onde são definidos compromissos ontológicos e critérios de solução do método. Em seguida, a teoria de competência inicial é refinada com a introdução de novos vocabulário e suposições, os quais são descritos em função do PSM e de uma teoria de domínio definida. O processo de refinamento seguinte transforma esta *teoria de competência refinada* em um *formato operacional*, onde este é especificado como um conjunto de passos de raciocínio básicos e uma estrutura de controle.

Conforme Wielinga et al. [WIE 98], encontrar um PSM adequado desde uma declaração informal de problema pode ser compreendido como um processo de *refinamento sucessivo*, onde cada passo introduz suposições sobre refinamentos específicos. Esse processo de desenvolvimento pode ser simplificada visualizado na Figura 2.13.

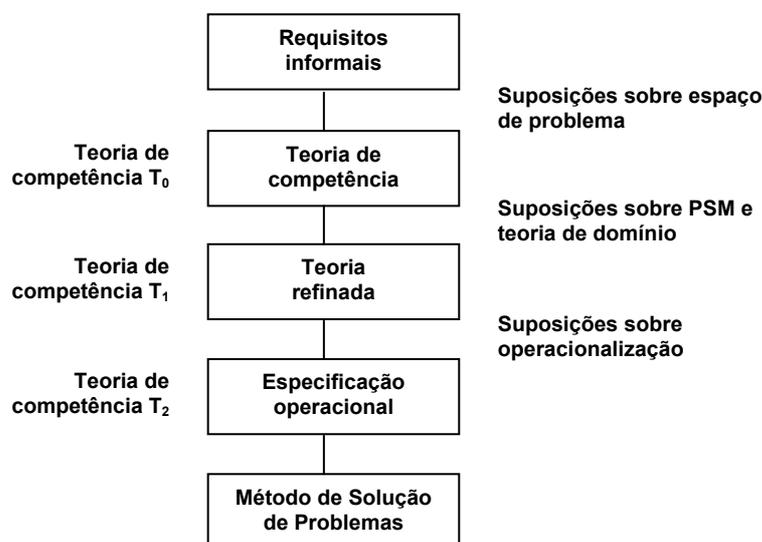


FIGURA 2.13 - O enfoque baseado na teoria de competência para construção de PSM

Esse processo de construção pode ser descrito por três passos:

1) *especificar o espaço de problemas e os requisitos da solução*

O primeiro passo de construção de um PSM é a especificação de um *espaço de problemas*. Este espaço é definido como: 1) um conjunto de problemas, 2) um conjunto de soluções e 3) uma relação entre problemas e soluções. Além desses

componentes, um espaço de problemas requer uma teoria de domínio adicional, assim viabilizando a modelagem de *problemas intensivos em conhecimento* pelo uso de conhecimento de aplicação. Tais problemas podem ser facilmente identificados, pois suas soluções requerem um processo de raciocínio envolvendo passos de inferência não-triviais e frequentemente não-determinísticos que usam conhecimento de domínio.

Baseado na definição do espaço de problemas, o processo de especificação do PSM passa a detalhar os componentes e compromissos de domínio desse espaço, resultando em uma primeira definição da *teoria de competência* (teoria de competência T_0). Entretanto, esta declaração informal inicial ainda não é precisa suficiente para representar um PSM, sendo necessário especificar algumas características, tais como: i) qual é a estrutura de informação do problema, ii) qual é a estrutura de informação da solução, iii) qual é a natureza do conhecimento de domínio e iv) quais são as suposições sobre conhecimento de domínio e tarefa?

2) refinamento conceitual da teoria de competência inicial

A teoria de competência T_0 é ainda uma especificação um tanto abstrata do que o PSM deve alcançar e, dessa forma, não é considerada adequada para operacionalização. Nesse caso, requisitos de solução especificados (tais como, por exemplo: predicados de solução) ainda podem ser refinados de forma que passos de raciocínio individuais sejam identificados. Por exemplo, refinando propriedades do método inicialmente especificadas e definindo propriedades auxiliares que aproximem mais a descrição de um PSM de um formato operacional de especificação. Esta atividade é baseada em um *paradigma de solução de problema e suposições sobre o conhecimento* requerido, onde são introduzidas tarefas intermediárias, restrições sobre o vocabulário de domínio (segundo suposições sobre a teoria de domínio disponível) e, por fim, refinamentos sobre o modo que o objetivo da tarefa é alcançado pelo método.

O refinamento resultante na *teoria de competência refinada* T_1 deve, portanto, manter a teoria inicial especificada em T_0 . Esse processo principalmente demonstra a transformação de uma teoria geral (geralmente computacionalmente intratável) para uma forma mais específica da teoria, assim reduzindo a complexidade computacional do método pela introdução de restrições sobre a utilização de expressões de domínios particulares. A forma típica de uma teoria de competência T_1 é uma conjunção de predicados de solução, tais como: descrições de expressões lógicas que podem ser compreendidas como passos de inferência refinados.

3) operacionalização da teoria de competência refinada

O último passo desse processo de desenvolvimento de PSMs trata da transformação da teoria refinada T_1 em uma especificação operacional. Nesse caso, "operacional" envolve especificar um procedimento (ou estratégia) para resolver um problema na prática (isto é, descrever estruturas de inferência e regimes de controle para estas, de modo a tornar o método suficientemente detalhado e prático para atuar como base do projeto e implementação de KBSs). Isso requer que sejam detalhados passos de raciocínio (inferências) e como eles são ordenados para resolver o problema

(estrutura de controle). O objetivo final da operacionalização é definir uma estratégia possível que satisfaça os requisitos de aplicação.

Dentre outras estratégias possíveis para operacionalizar uma competência T_1 , podem ser citadas as seguintes:

- usar uma especificação operacional simples como *gerar&testar*. Segundo esta estratégia de operacionalização, uma solução é gerada de maneira aleatória. Em seguida, esta solução é analisada segundo os requisitos do problema. Caso esta solução atenda todos os requisitos do problema, a solução final foi encontrada. Caso contrário, uma nova solução é gerada, e o processo é repetido novamente;
- usar a estratégia *gerar&testar* para gerar um conjunto de soluções candidatas. Em seguida, eliminar as soluções candidatas inconsistentes com requisitos de domínio especificados inicialmente. Para isso, uma inferência adicional para *selecionar* requisitos de um conjunto de requisitos pré-estabelecidos é requerida. Em seguida, a inferência *testar* não produziria uma solução, mas somente derivaria um valor verdadeiro indicando se a solução candidata atende ou não os requisitos e teoria de domínio selecionados;
- usar *gerar&testar*, mas gerando soluções candidatas a partir de especificações pré-determinadas. Especialistas frequentemente usam métodos mais refinados para gerar o conjunto de soluções candidato inicial. Sendo assim, o passo *gerar* toma como entrada um requisito de problema e gera somente aqueles candidatos que atendem este requisito. No passo *testar*, a estratégia é idêntica a estratégia anterior, onde somente valores lógicos são derivados. Em resumo, a estrutura de inferência é modificada por: um requisito de problema usado como entrada na inferência *gerar*; o conhecimento usado pelo passo *gerar* é modificado para um predicado como *tem-característica*, selecionando apenas candidatos com características desejáveis.

Por fim, esse método de construção sistemático de PSMs é definido como um refinamento sucessivo de uma teoria de competência requerida por um PSM, resultando em um modelo operacional do método. A partir disso, essa proposta metodológica permite declarar os compromissos e suposições que são estabelecidos por um PSM, o qual pode ser usado para resolver uma classe particular de tarefas.

2.5.2 Desenvolvimento estruturado fundamentado em um sistema de eixos

O processo de desenvolvimento estruturado de PSMs proposto por Fensel et al. [FEN 2000] é descrito como uma introdução de *suposições e compromissos* ao longo de um espaço tridimensional de especificação de PSMs, assim especializando gradativamente uma especificação de referência de PSM. Tal espaço é definido em termos de 1) uma *estratégia de solução de problemas* 2) *compromissos de tarefa (problema)* e 3) *suposições sobre conhecimento de domínio* (Figura 2.14). Esse

método de desenvolvimento é caracterizado como um processo de navegação no espaço tridimensional definido pelos eixos de especificação de PSMs, onde movimentos através deste espaço são explicitamente representados por meio de *adaptadores* (Fensel et al. [FEN 97a]).

Segundo Fensel et al. [FEN 2000], o uso de adaptadores torna possível a especificação de tarefas, métodos de solução de problemas e conhecimento de domínio de forma independente entre si, possibilitando algum reuso destas descrições. Para essa proposta metodológica, o uso de adaptadores no desenvolvimento de PSMs tem as seguintes características:

- ser um meio de refinamento gradativo de PSMs, o qual permite estabelecer um relacionamento com circunstâncias de tarefa e domínio;
- identificar uma tipologia de adaptadores baseada no espaço tridimensional de especificação do método;
- caracterizar adaptadores de forma que eles próprios possam ser reutilizados.

O processo de desenvolvimento aqui descrito, portanto, parte de padrões genéricos a partir dos quais métodos alternativos podem ser derivados por meio de adaptação. Além disso, o processo apresenta uma proposta de modelagem para facilitar os passos de adaptação necessários, os quais são especificados em uma linguagem de descrição e desenvolvimento de PSMs denominada UPML (*Unified Problem-Solving Method Development Language* - Fensel et al. [FEN 99]).

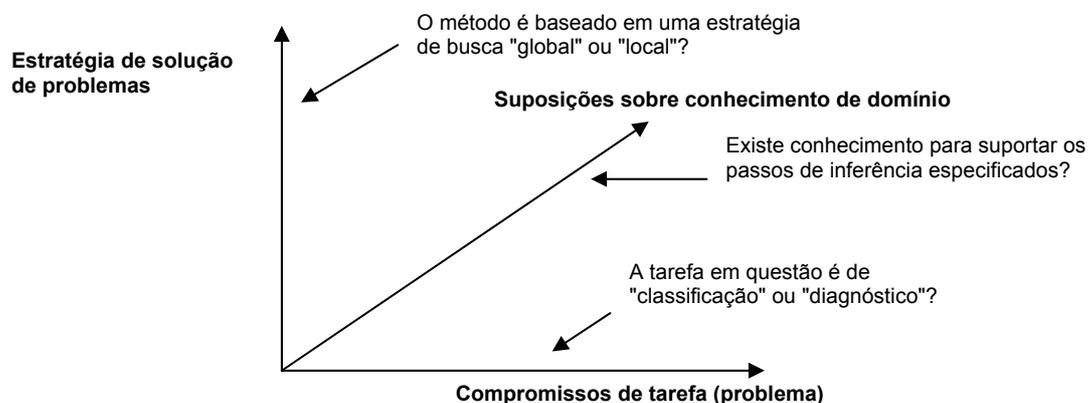


FIGURA 2.14 - As três dimensões da descrição e desenvolvimento de PSMs

- 1) *estratégia de solução de problemas* - em vez de descrever um algoritmo particular, é modelado um tipo de solução de problemas, o qual descreve uma classe de algoritmos. Esta estratégia de solução de problemas estabelece estruturas de dados básicas, fornece uma decomposição inicial de tarefa-subtarefas e uma especificação de controle genérica, onde tais declarações devem ser compartilhadas por todos os PSMs que podem ser resumidos por esta estratégia. Estratégias de solução de problemas que poderiam ser exemplificadas são as seguintes: "gerar&testar"

(*generate&test*) e "busca localizada" (*local search*), as quais são bastante simplificadas;

- 2) *compromissos de tarefa (problema)* - são compromissos ontológicos específicos para o tipo de problema que é resolvido pelo PSM. Estes compromissos são expressos pela adaptação de um PSM em relação a uma *ontologia de tarefa* (Mizoguchi et al. [MIZ 95]) particular. Por exemplo, uma ontologia de tarefa de "projeto" fornece definições para termos, tais como: "modelo de projeto", "componente", "restrição" e "constante", os quais podem ser usados pelo PSM. Tais compromissos ontológicos introduzidos pela tarefa também podem refinar a competência do PSM, bem como a estrutura e a transição dos estados computacionais que o método pode expressar. Por exemplo, um "método de busca genérico" pode ser transformado em um método especializado para "diagnóstico" ou "projeto" pela introdução de compromissos ontológicos de tarefas desses tipos. Cabe salientar que o refinamento de um PSM para uma tarefa pode fornecer suporte mais explícito para a aquisição de conhecimento e o desenvolvimento de aplicações (método mais usável) que um PSM independente de tarefa;
- 3) *suposições sobre conhecimento de domínio* - são requeridas para instanciar um PSM para uma aplicação particular. Estas suposições especificam os tipos e propriedades das estruturas de conhecimento que devem ser fornecidas por um modelo de domínio, além daquelas requeridas para atender os objetivos e restrições da tarefa. Por exemplo, quando é resolvido um "problema de projeto" usando o método "propor&revisar", um domínio precisa disponibilizar o conhecimento requerido para expressar "procedimentos" e "consertos", além do conhecimento de tarefa necessário para modelar o "problema de projeto particular", tais como, por exemplo, "partes" e "restrições" do projeto.

Os movimentos no espaço de especificação de PSMs podem ser descritos em termos de sua orientação, conforme um *processo de adaptação por refinamentos*:

i) movimentos em uma dimensão - um movimento pode ocorrer em um dos três eixos apenas, refinando uma estratégia de solução de problemas, ou uma suposição de domínio ou um compromisso de tarefa:

- *refinar um paradigma de solução de problemas* - como cada inferência normalmente estabelece algum parâmetro que pode ser refinado, o refinamento do paradigma de solução de problemas pode ser realizado pelo refinamento de inferências. Porém, esse caso somente é caracterizado se o refinamento das inferências não introduz novos requisitos e suposições sobre conhecimento de domínio. Um exemplo desse caso que pode ser citado é o refinamento de um método de "busca de soluções" para um método de "busca de soluções locais";
- *refinar conhecimento de domínio* - o conhecimento de domínio é refinado de forma independente de seu uso, ou seja, independentemente do esquema de solução de problemas e definição de tarefa. Esta situação não ocorre no

processo de desenvolvimento de PSMs sendo descrito, pois esse processo é orientado para resultar em uma especificação de método e não de conhecimento de domínio. No entanto, podem ocorrer casos que uma ontologia genérica seja refinada para um certo domínio, tal como, por exemplo, uma ontologia descrevendo "motores de automóveis" refinada para "motores de automóveis Ford";

- *refinar o objetivo de tarefa* - tal refinamento caracteriza um movimento no eixo de compromissos de tarefa. Todavia, esses refinamentos de tarefas normalmente introduzem novos requisitos sobre conhecimento de domínio. Por exemplo, uma "tarefa de otimização" requer uma "função de preferência"; uma "tarefa de projeto" requer "restrições", uma "função de custo"; etc. Logo, tais refinamentos afetam as dimensões de domínio e tarefa. Porém, é possível ter refinamentos que afetam apenas a dimensão de tarefas. Por exemplo, refinando uma tarefa para buscar "todas as soluções para um problema", em vez de apenas "uma solução", ou uma "tarefa de projeto" pode ser refinada para uma "tarefa de projeto parametrizada".

ii) movimentos em duas dimensões - um movimento pode ocorrer no plano especificado por quaisquer duas dimensões, refinando uma estratégia de solução de problemas pela introdução de suposições de domínio, ou adaptando uma estratégia de solução de problemas para um tipo de problema, ou refinando um tipo de problema em termos de suposições de domínio:

- *refinando uma estratégia de solução de problemas pela introdução de suposições de domínio* - especializar alguns passos de inferência, ao mesmo tempo que tipos adicionais de conhecimento são requeridos pelo método resultante. Por exemplo, especializar um passo de "gerarSolução" para "gerarMelhorSolução" segundo critérios de domínio especificados. Dessa forma, essa adaptação também denota um movimento ao longo do eixo de suposições de domínio, caracterizando um movimento em duas dimensões;
- *adaptando uma estratégias de solução de problemas para um tipo de problema* - esse movimento é caracterizado por adaptações de esquemas de solução de problemas genéricos para diferentes tarefas. Por exemplo, adaptando um método como "propor&revisar" para resolver "problemas de projeto". Porém, tais adaptações não podem introduzir novas suposições sobre conhecimento de domínio;
- *refinando um tipo de problema em termos de suposições de domínio* - a definição dos objetivos da tarefa pode ser refinada ao mesmo tempo que são introduzidos novos requisitos de conhecimento de domínio. Por exemplo, uma solução para um "problema de otimização" pode ser refinada em termos de um "critério de otimização" de domínio; uma solução para um "problema de projeto" pode ser refinada para ser "válida" e "completa", assim como requerido para esses problemas.

iii) movimentos em três dimensões - um processo de adaptação normalmente relaciona métodos com tarefas, bem como diminui a lacuna entre eles pela introdução de suposições sobre conhecimento de domínio, assim afetando todas as três dimensões do espaço de especificação do método. Como esse tipo de movimento é bastante complexo, ele deve ser decomposto em termos de movimentos em duas dimensões, tal como afirmado por Fensel et al. [FEN 2000]. Em princípio, pode haver um movimento em duas dimensões, segundo as direções da estratégia de solução de problemas e compromissos de tarefa. Por exemplo, refinando a competência do método pela introdução de compromissos específicos de tarefa. Em seguida, pode haver um movimento nas direções da estratégia de solução de problemas e suposições sobre conhecimento de domínio. Por exemplo, adicionando suposições sobre conhecimento de domínio no sentido de refinar a competência do PSM. O método resultante então poderia ser considerado como similar ao método descrito como um movimento em três dimensões.

Normalmente, quando se desenvolve um KBS, é necessário delegar partes da solução de problemas para o conhecimento de domínio, como também enfraquecer a especificação de uma tarefa. Nesse contexto, movimentos em três dimensões são frequentemente inevitáveis, mas podem ser divididos em movimentos em duas dimensões. Conforme Fensel et al. [FEN 2000], dividir movimentos em três dimensões em uma seqüência de movimentos bidimensionais pode reduzir a complexidade do processo de desenvolvimento de PSMs. No entanto, a divisão de movimentos em três dimensões em movimentos em duas dimensões geralmente está associada a aspectos como eficiência, ou tratabilidade de métodos de inferência complexos. Nesse caso, muitos movimentos divididos podem levar a métodos ineficientes ou intratáveis, pois eles podem não considerar explicitamente o conhecimento de domínio requerido.

Segundo Fensel et al. [FEN 2000], a proposta metodológica de desenvolvimento estruturado de PSMs apresenta alternativas para resolver o problema associado a usabilidade *versus* reusabilidade de PSMs, pois apresenta suporte para:

- especificar esquemas algorítmicos independentes de tarefas, assim descrevendo a estratégia de solução de problemas de um método;
- detalhar esses esquemas de acordo com a tarefa e circunstâncias específicas de domínio. Para isso, os passos de desenvolvimento e adaptação do PSM são formalmente modelados, permitindo o reuso de diferentes níveis de refinamento;
- modelar o processo de adaptação. Essa característica é fundamental para possibilitar o reuso em diferentes níveis de refinamento, pois inclusive o refinamento pode ser reutilizado.

Esse processo de desenvolvimento tenta apresentar soluções para três problemas relacionados a pesquisa em PSMs: 1) quais são os blocos de desenvolvimento de PSMs e como eles se interrelacionam?, 2) como modelar o processo de desenvolvimento de PSMs? e 3) como desenvolver e organizar bibliotecas gerenciáveis e compreensíveis de PSMs? Assim como o processo de desenvolvimento foi apresentado, ele claramente unifica diferentes enfoques de

desenvolvimento anteriormente propostos, mas sem ainda apresentar uma validação em aplicações reais.

2.5.3 Construção de métodos de solução de problemas como projetos baseados em parâmetros

Como proposto por Teije et al. [TEI 96], a construção de PSMs como projetos parametrizados (*parametric design*) pode ser compreendida como um processo de desenvolvimento automático de PSMs. Nessa abordagem, PSMs são representados em um *esquema de inferência*, o qual permite capturar a funcionalidade de uma classe de métodos (problemas).

Os componentes do esquema de inferência (tal como os passos de inferência propostos) são considerados como parâmetros a serem configurados, modelando diferentes membros de uma classe de métodos. Estes membros correspondem a especificação de diferentes definições para os parâmetros detalhados no esquema. A representação de uma classe inteira de PSMs em um esquema genérico permite visualizar o processo de construção de PSMs como um *processo de configuração*, o qual é típico de *tarefas de projeto parametrizadas* (Motta et al. [MOT 98]).

O esquema de representação que serve como base para o processo de construção automático de PSMs está relacionado com a noção geral de PSMs apresentada por Benjamins et al. [BEN 96b]. Segundo essa noção, dado um objetivo a ser resolvido, um PSM contém uma especificação funcional que pode ser alcançada e uma especificação operacional que implementa a funcionalidade. Para alcançar os objetivos desejados, um PSM deve fazer suposições sobre o conhecimento de domínio requerido. A Figura 2.15 apresenta os componentes dessa noção de PSMs.

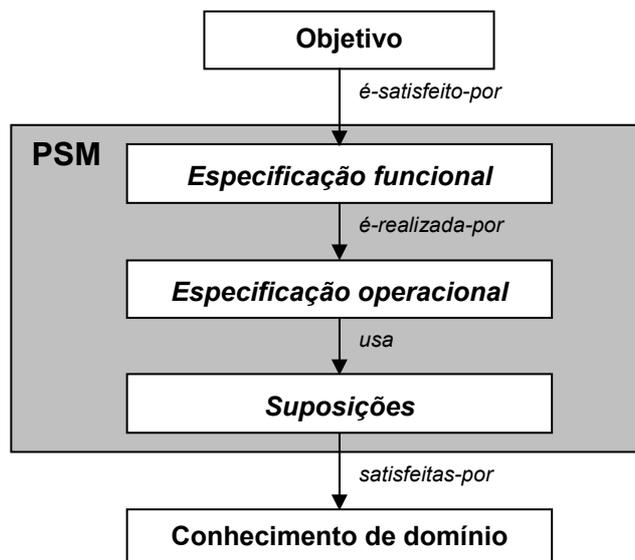


FIGURA 2.15 - Os componentes da especificação de um PSM

Conforme Teije et al. [TEI 96], as principais entradas do processo de construção automático de PSMs são: i) um *problema* que requer a construção de um método. Por exemplo: um "problema de diagnóstico" pode ser descrito por um

"comportamento observado" (dados) e um "modelo de comportamento" (conhecimento); ii) os *objetivos* que o método resultante deve satisfazer. Por exemplo: o método deve computar o número máximo de diagnósticos e iii) as *suposições* que o método deve operar. Por exemplo: suposição sobre *falhas únicas* típicas de tarefas de diagnóstico de falhas.

As principais saídas são: i) a descrição do método construído (ou configurado); ii) as soluções computadas pelo método e iii) as versões possivelmente ajustadas do problema, objetivos e suposições informados inicialmente.

O relacionamento entre entradas e saídas são: i) uma saída deve ser a representação de um PSM; ii) não deve haver conflito em relação às suposições (possivelmente ajustadas); iii) os objetivos (possivelmente ajustados) devem ser satisfeitos e iv) as entradas adaptadas (suposições, objetivos e problema) devem ser intimamente relacionadas com as entradas originais.

O *processo de configuração* é caracterizado como o seguinte: a partir das entradas especificadas, encontrar um PSM cuja especificação funcional atenda os objetivos requeridos segundo as suposições estabelecidas. A configuração anterior a qualquer validação é denominada *configuração estática*, onde se espera encontrar um PSM ótimo dentre PSMs possíveis. A configuração que posteriormente utiliza resultados de passos de validação é denominada *configuração dinâmica*, onde é avaliado o quê deve ser feito caso o PSM não esteja ainda apresentando a solução desejada.

Na linha da configuração estática e dinâmica, são distintos *objetivos estáticos e dinâmicos*. Objetivos estáticos são requisitos que podem ser garantidos usando somente a descrição do método. Por exemplo, um objetivo estático muito comum em PSMs de diagnóstico é a produção de (somente) diagnósticos únicos, entre outros tipos de diagnósticos possíveis. Por sua vez, objetivos dinâmicos são requisitos que podem ser validados somente após a execução do método. Por exemplo, requerer um número máximo de diagnósticos pode ser usado como exemplo de objetivo estático de PSMs de diagnóstico. Entretanto, segundo Teije et al. [TEI 96], a distinção entre objetivos estáticos e dinâmicos não é fixa, pois depende do conhecimento disponível.

O processo de construção automático de PSMs é executado em dois passos. O primeiro passo realiza a configuração de um método segundo os objetivos estáticos especificados. Caso este método não exista, um segundo passo adapta o problema, suposições e objetivos especificados para que o método construído venha a satisfazer os objetivos estáticos. Se este método resultante também satisfizer os objetivos dinâmicos, a especificação de um método adequado foi alcançada. Caso contrário, a adaptação do método é realizada em função dos objetivos dinâmicos. No entanto, quando isto não é possível, o problema, suposições ou objetivos são novamente adaptados e o PSM é configurado para estas novas entradas. O objetivo desse processo é tentar especificar um método capaz de computar a melhor solução possível para um problema, suposições e objetivos desejados. Todavia, o objeto desse processo de construção é o PSM em si. As suposições, objetivos e problema possivelmente ajustados são efeitos adicionais da configuração do método para um dado problema.

Para realizar a configuração de PSMs, a família de métodos *propor-criticar-modificar* (Motta et al. [MOT 98]) é usada. Entre outras características, o processo de configuração desta família de métodos usa resultados de testes intermediários para determinar possíveis ações de configuração, em vez de gerar configurações a partir do zero na ocorrência de uma configuração inadequada. Tal família de PSMs pode ser descrita por quatro passos de inferência:

- *propor* - determinar uma configuração parcial ou completa de um esquema de inferência usado para representar um PSM. Para realizar essa configuração, algumas ações podem ser citadas: 1) uma solução desejada pode ser decomposta, 2) restrições estabelecidas podem ser satisfeitas, 3) definições aleatórias sobre suposições e problema podem ser escolhidas, entre outras;

Para configurar métodos descritos como estruturas de inferência de *projeto parametrizadas*, valores derivados dos requisitos de problema sendo abordado devem ser propostos para alguns dos parâmetros do esquema de inferência. O resultado deste passo é a definição de um método (método possível), onde os objetivos estáticos requeridos pelo método são atendidos tanto quanto possível. Este passo também resulta em objetivos dinâmicos, os quais são parte dos objetivos estáticos inicialmente estabelecidos, mas que ainda não são garantidos pelo método proposto.

- *verificar* - o passo de verificação é dividido em *verificação estática* (ou *verificação de conhecimento*), onde são analisados os objetivos estáticos do método antes de qualquer passo de execução, e *verificação dinâmica* (ou *verificação de simulação*), onde são analisados os objetivos dinâmicos do método a partir do comportamento derivado por simulação;

A *verificação estática* analisa as restrições sobre os componentes do esquema de inferência do método, bem como restrições derivadas de suposições estabelecidas inicialmente. O resultado deste passo é um *método válido*, onde este é um método possível que não tem conflito de suposições ou restrições de componentes do esquema de inferência. De maneira objetiva, um método é *válido* se todas as restrições de componentes são válidas (método componente-válido) e não ocorrerem conflitos entre suposições (método suposição-válido). Caso a verificação estática falhar, um novo passo de *propor* é executado, resultando em um novo método candidato.

Por sua vez, a *verificação dinâmica* testa um método simulado, caso os objetivos dinâmicos não sejam atendidos. Por exemplo, verificando o "tamanho do espaço de soluções" resultante da simulação. Caso a verificação dinâmica venha a obter sucesso, o método é considerado adequado. Nesse caso, o sucesso é indicado por um método válido que atende todos os objetivos dinâmicos, como também pode ser indicado por um método e seus respectivos objetivos violados.

- *criticar* - analisa as falhas ocorridas na fase de configuração, as quais possivelmente levaram a um método não-apropriado. Para isso, o comportamento não-desejado do método é mapeado para componentes do esquema de inferência escolhidos para configuração, onde tais propriedades possivelmente são responsáveis pela falha do processo de configuração. Para realizar essa crítica, geralmente é usado conhecimento

sobre "violação de objetivos" e "reparos possíveis". No entanto, ainda não é tomada nenhuma ação de reparo neste passo, pois um passo de crítica seguinte é necessário porque o processo de verificação pode envolver diversas possibilidades alternativas de configuração. Desta forma, o passo de crítica viabiliza algum nível de controle associado a busca do método adequado, ao invés de gerar e testar métodos arbitrariamente até encontrar um correto;

- *modificar* - o resultado do passo de crítica é usado para encontrar uma ação de reparo apropriada para o método então criticado. Uma ação de reparo pode envolver a modificação de um componente individual do método, ou um método inteiro em direção a uma propriedade desejada, como também pode especializar os componentes do método tornando-os mais compatíveis entre si.

A Figura 2.16 apresenta um cenário desse processo de configuração descrito.

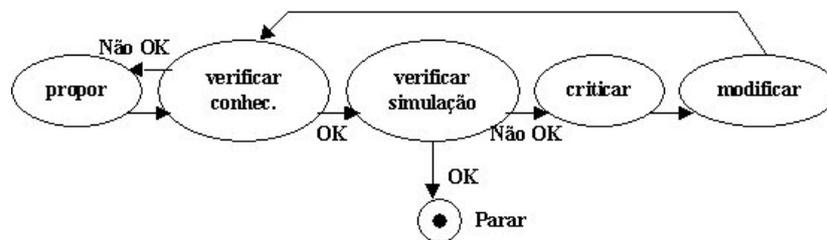


FIGURA 2.16 - Um diagrama de estados para o controle do método de configuração de PSMs

Após receber uma descrição do problema como entrada (objetivos estáticos, problema e suposições), o primeiro passo do processo de configuração é *propor* PSMs candidatos para resolver o problema sendo tratado. Alguns desses candidatos podem ser eliminados imediatamente, conforme o resultado da análise dos objetivos estáticos e suposições especificados. Um PSM candidato é gerado, e este sobrevive ao passo de verificação do conhecimento. Este PSM válido é simulado e são computados resultados possivelmente não desejados, segundo objetivos dinâmicos determinados a priori. Tais soluções descrevendo objetivos violados são então analisadas em um passo de crítica seguinte do método simulado. Os passos *criticar* e *modificar* propõem modificações (possivelmente ações de reparo) nesse PSM simulado, as quais tentam reparar os objetivos violados apresentados. A modificação realizada no PSM é novamente analisada, bem como um novo passo de simulação pode resultar em outros problemas que devam ser ajustados nesse processo de configuração. Sendo assim, outros passos *criticar* e *modificar* verificam as modificações realizadas no candidato corrente. Caso ocorram resultados satisfatórios no passo de verificação da simulação, um PSM adequado então foi alcançado, e o processo de configuração é finalizado.

2.6 Suposições

Conforme Benjamins et al. [BEN 96b], a utilização de PSMs deve ser compreendida através de duas relações. A primeira relação é entre o PSM e o

conhecimento de domínio utilizado. A segunda relação é entre o PSM e o objetivo que se deseja alcançar. Antes que um PSM possa ser aplicado, tais relações são consideradas lacunas a serem preenchidas por *suposições* (Figura 2.17), pois descrevendo as suposições que orientam um sistema de raciocínio podem ser determinados os limites de aplicabilidade de um PSM.

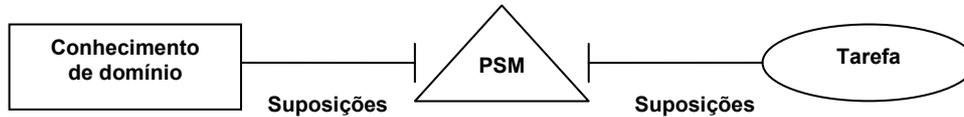


FIGURA 2.17 - As lacunas que isolam um PSM

Conforme Benjamins et al. [BEN 96b], suposições normalmente estabelecem os requisitos sobre o conhecimento de domínio usado pelo PSM, ou restringem o tamanho do problema que é tratado pelo PSM. Descrever uma suposição como um requisito requer esforço de aquisição de conhecimento durante a construção do sistema de raciocínio. Em contrapartida, especificar uma suposição como uma restrição requer um esforço computacional adicional durante o processo de solução do problema, pois o método resultante tem disponível menos conhecimento de domínio, o qual pode restringir o espaço de busca de soluções, entre outras coisas.

Entre outras contribuições para a KE, suposições podem ser usadas para:

- caracterizar a competência de um PSM, a qual é expressa em termos de tarefas que podem ser resolvidas e conhecimento de domínio requerido pelo método;
- viabilizar a possibilidade de solução de problemas, pois elas reduzem a complexidade de computação requerida, como também podem simplificar o problema a ser resolvido pelo sistema;
- assegurar a interação correta do mecanismo de solução de problemas com seu ambiente, pois elas podem caracterizar mais precisamente o KBS.

Conforme Fensel et al. [FEN 98a], um PSM deve descrever não somente a realização de uma funcionalidade, mas uma funcionalidade que considera as restrições do processo de raciocínio e a complexidade da tarefa. Nesse caso, suposições podem restringir a complexidade do problema enfraquecendo a definição da tarefa de modo que a competência do PSM seja suficiente para realizar a tarefa. Em contraste, suposições também podem fortalecer a competência do método assumindo (extra) conhecimento de domínio. As ações (efeitos) destas suposições são complementares, como descrito a seguir:

$$\text{TAREFA}_{\text{objetivo}} - \text{Suposição}_{\text{enfraquecer}} = \text{PSM}_{\text{objetivo}} + \text{Suposição}_{\text{fortalecer}}$$

Segundo a fórmula anterior, o objetivo de uma tarefa menos as suposições que enfraquecem este objetivo é igual ao objetivo do PSM mais suposições que fortalecem este segundo objetivo. Dessa forma, a fórmula seguinte indica que a soma destas

suposições é constante (Δ) para objetivos fixos (exemplificado na Figura 2.18). Isso é denominado *Lei da Conservação das Suposições* (Benjamins et al. [BEN 96b]).

$$\text{TAREFA}_{\text{objetivo}} - \text{PSM}_{\text{objetivo}} = \Delta = \text{Suposição}_{\text{enfraquecer}} + \text{Suposição}_{\text{fortalecer}}$$

Com essa soma de suposições constante, uma suposição de um tipo pode ser substituída por outra de outro tipo, desde que elas tenham a mesma relevância (tal como um mesmo peso). Porém, mesmo que tais suposições sejam compensadas, a direção de sua influência permanece diferente, pois uma é aplicada sobre os objetivos da tarefa e outra é aplicada sobre os objetivos do PSM. Sendo assim, dois tipos de suposições podem ser identificadas:

- *suposições teológicas* - envolvem um "enfraquecimento" do objetivo da tarefa que pode ser alcançado, pois estas suposições traduzem um objetivo original para um objetivo mais fraco e com menor complexidade, ou que pode ser resolvido mais eficientemente em casos práticos de raciocínio. Por exemplo, esse tipo de modificação pode ser exemplificado por um método que passa a não requerer mais uma "solução ótima" como resultado, mas requer apenas uma "solução aceitável";
- *suposições ontológicas* - envolvem um "fortalecimento" da funcionalidade de um PSM, pois elas dizem respeito ao conhecimento de domínio disponível. Tais suposições são feitas para reduzir a complexidade de parte do problema resolvido pelo PSM. Por exemplo, a disponibilidade de bom conhecimento em "problemas de projeto" viabiliza "projetos iniciais melhores" e, sendo assim, passos de revisão desses projetos podem ser simplificados.

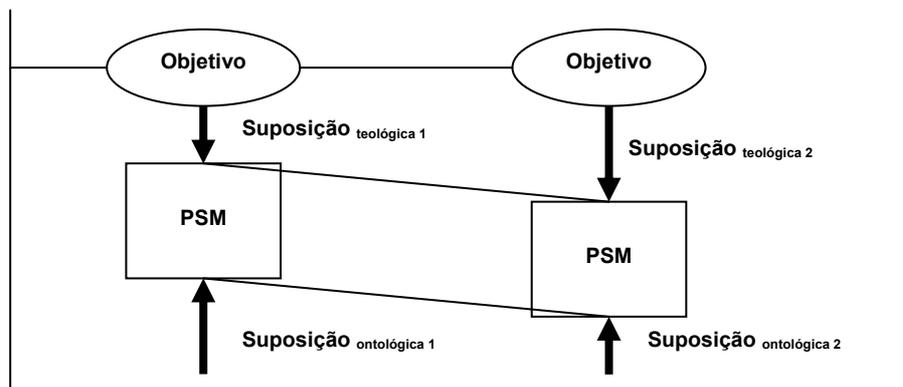


FIGURA 2.18 - Os efeitos das suposições

A Figura 2.18 demonstra o efeito desses tipos de suposições. Segundo esses efeitos, para um objetivo constante, quanto mais suposições ontológicas são feitas, suposições teológicas mais fracas devem ser estabelecidas sobre o objetivo a ser alcançado. Também indica que quanto menos suposições ontológicas, suposições teológicas fortes devem ser estabelecidas. Cabe ressaltar que ambos os tipos de suposições servem para o mesmo propósito: diminuir a lacuna entre o PSM e o

objetivo da tarefa a ser alcançado. Sendo assim, os requisitos em termos de conhecimento de um PSM podem ser enfraquecidos ou fortalecidos de acordo com:

- conhecimento de domínio disponível (suposições ontológicas);
- esforço que é requerido para derivar e modelar conhecimento. Como citado por Fensel et al. [FEN 98a], não existe nada como uma representação livre de suposições;
- as suposições teológicas que podem ser feitas para reduzir o objetivo que deve ser alcançado.

A Lei da Conservação das Suposições pode ser exemplificada com uma suposição típica de "problemas de diagnóstico de falhas": a suposição de "modelos de falhas completo". Como uma suposição ontológica, esta suposição indica que é possível diagnosticar "todas as falhas possíveis", visto que todas estão representadas em um "modelo de falhas". Como uma suposição teológica, esta suposição estabelece que o PSM funcionará para casos onde "apenas ocorrem falhas mencionadas no modelo de falhas", as quais poderiam ser suficientemente atendidas por rotinas de diagnóstico.

Em Benjamins et al. [BEN 96b] e Fensel et al. [FEN 98a] são apresentadas algumas suposições necessárias para relacionar a funcionalidade de um "método de diagnóstico" com a tarefa a ser resolvida (nesse caso, uma tarefa de "diagnóstico de falhas"), como também com o conhecimento de domínio requerido para definir a tarefa. Algumas delas são citadas a seguir:

- *identificar anormalidades* - antes do processo de diagnóstico ser iniciado, supõe-se que seja possível identificar um comportamento anormal associado ao sistema. A partir desse comportamento (ou falhas), podem ser descritas explicações para as anormalidades identificadas;
- *conhecimento de comparação (threshold)* - supõe-se que exista conhecimento para decidir se uma anormalidade (discrepância) de comportamento do sistema é relevante para o processo de solução de problemas;
- *tipos de entradas confiáveis* - essas suposições garantem que se acontecer um comportamento diferente do comportamento desejado, então o componente do sistema tem uma falha ou defeito. Dessa forma, as entradas do método de raciocínio são consideradas confiáveis;
- *nenhuma falha na estrutura* - supõe-se que as interações entre os componentes do sistema sejam correta e completamente modeladas;
- *independência de hipóteses* - uma hipótese individual pode explicar um conjunto de observações, não havendo a necessidade de considerar outras hipóteses para derivar uma explicação;

- *não-intermitência* - as observações sobre o comportamento do sistema não modificam durante o processo de raciocínio;
- *crença nas observações* - as observações do comportamento do sistema identificadas são confiáveis;
- *ações de reparos não interagem* - o resultado da aplicação de um reparo de falha não cancela algum outro, de modo que eles possam ser aplicados em qualquer ordem.

Fensel et al. [FEN 98b] descreve algumas suposições típicas de PSMs eficientes, as quais normalmente são empregadas para reduzir o esforço computacional requerido para resolver uma tarefa. Tais suposições podem ser resumidas de dois modos:

- *suposições para reduzir a complexidade do pior caso de solução de problemas* - a suposição de "falha única" é o melhor exemplo desse tipo de suposição. Caso esta suposição seja estabelecida, o comportamento incorreto de um componente do sistema pode ser completamente explicado por apenas um componente que apresenta a falha;
- *suposições para reduzir a complexidade do caso médio de solução de problemas* - o conhecimento sobre como os componentes do sistema comportam-se quando estão falhando (denominado *modelo de falhas*) pode ser usado para restringir o conjunto de diagnósticos derivados pelo método. Logo, reduzindo a complexidade de casos médios pode reduzir o esforço necessário para atender a maioria dos casos práticos de solução de problemas.

Em resumo, suposições podem caracterizar precisamente as especificações de tarefa, PSM e conhecimento de domínio, as quais orientam um sistema de raciocínio. Desta forma, a aplicabilidade do KBSs pode ser definida e descrita de forma explícita pelo uso de suposições.

2.7 Bibliotecas de métodos de solução de problemas

Conforme Gómez-Pérez et al. [GOM 99], PSMs representam componentes que demonstram as "melhores formas" de raciocínio disponíveis para a construção de KBSs, assim como *padrões de projeto* (*design patterns* - Gamma et al. [GAM 95]) são para a orientação a objetos. Mesmo assim, muitos PSMs novos têm sido construídos a partir do zero, assim disponibilizando experiências de sucesso de construção de métodos efetivamente usados em KBSs. Desta forma, tais experiências descritas como PSMs têm uma conotação mais simplificada que a conotação normalmente envolvida na descrição de um padrão de projeto.

Existem diversas bibliotecas (catálogos) de PSMs disponíveis na literatura. Dentre elas, algumas podem ser citadas:

- biblioteca de CommonKADS (Schreiber et al. [SCH 99]) - PSMs para tarefas de classificação, diagnóstico, previsão, avaliação (*assessment*), monitoramento, projeto, planejamento, designação (*attachment*), escalonamento e modelagem;
- biblioteca de diagnóstico (Benjamins et al. [BEN 94]; Benjamins et al. [BEN 95]) - 38 PSMs para 14 tarefas de diagnóstico;
- biblioteca de projeto parametrizado (Motta et al. [MOT 98]) - contendo PSMs derivados a partir de variações do método *propor&revisar*;
- biblioteca de planejamento (Barros et al. [BAR 96]) - contendo PSMs destinados a tarefas de planejamento.

Segundo Gómez-Pérez et al. [GOM 99], bibliotecas de PSMs podem ser divididas em diferentes tipos, conforme critérios como:

- *generalidade* - PSMs para tarefas/domínios específicos ou genéricos. PSMs para projeto e diagnóstico podem ser citados como exemplo de PSMs para tarefas específicas, onde esta especificidade de tarefas é basicamente identificada pela terminologia que o PSM é descrito. Por exemplo, a biblioteca de CommonKADS apresenta métodos específicos para tarefas de raciocínio. Em contraste, PSMs como *gerar&testar* são formulados em uma terminologia independente de tarefa;
- *formalidade* - PSMs podem ser especificados informalmente como descrições em texto, semiformalmente como diagramas, formalmente como especificações em lógica ou até com descrições executáveis, onde a especificação operacional do PSM é descrita em uma linguagem executável;
- *granularidade* - segundo esta classificação, PSMs podem ser compostos por componentes complexos ou simplificados. Bibliotecas de PSMs complexos são formadas por métodos que realizam tarefas complexas, enquanto as compostas por componentes simplificados envolvem métodos que realizam pequenas partes de tarefas maiores, os quais são associados para resolver uma tarefa;
- *tamanho* - variam conforme o número de PSMs disponibilizados. Dentre as bibliotecas disponíveis na literatura, a maior é a biblioteca de CommonKADS (Schreiber et al. [SCH 99]), a qual contém métodos para diversos tipos de tarefas. Outra biblioteca que pode ser citada é a biblioteca de diagnóstico (Benjamins et al. [BEN 94]), a qual contém 38 PSMs para realizar 14 tarefas de diagnóstico.

Conforme Studer et al. [STU 98], a organização e posterior indexação dessas bibliotecas ainda é um tema de pesquisa bastante recente. No entanto, algumas formas de organização podem ser citadas:

- *estrutura de decomposição de tarefas-métodos* - esta proposta de organização considera que um tarefa pode ser realizada por métodos alternativos, os quais podem ser compostos de subtarefas primitivas ou compostas. Subtarefas compostas podem então ser realizadas por outros métodos alternativos;
- *funcionalidade* - essa organização indica que PSMs com funcionalidades similaridades devem ser armazenados conjuntamente. Tais funcionalidades são principalmente identificadas por especificações de competência similares;
- *suposições* - envolve uma organização que considera as condições que um PSM pode ser aplicado, as quais são especificadas como suposições. Tais condições são descritas em termos de conhecimento de domínio ou conhecimento de tarefa requeridos pelo método;
- *tipos de problemas* - esta organização é baseada nas dependências que existem entre tipos de problemas, ou nos modos que problemas dependem uns dos outros. Por exemplo, a saída de um problema de "predição" pode ser a entrada de um problema de "monitoramento", o qual pode resultar em entradas para um problema de "diagnóstico".

Em resumo, existe a necessidade de classificar e organizar bibliotecas de PSMs, assim permitindo demonstrar os requisitos de conhecimento de domínio e a competência dos PSMs disponibilizados. A partir disso, é possível analisar um PSM e avaliar sua aplicabilidade para um problema que se deseja resolver. Conforme Schreiber et al. [SCH 99], tais bibliotecas de PSMs podem facilitar o processo de KE, pois a escolha de algum PSM dentre muitos é ainda melhor que não possuir nenhum método para orientar o desenvolvimento de um KBSs.

3 Um estudo de caso de desenvolvimento de métodos de solução de problemas

O processo de KE que permite desenvolver KBSs envolve especificar modelos de conhecimento novos, os quais representam a solução do problema sendo tratado. Tal processo de construção pode ser realizado a partir de modelos de conhecimento genéricos e PSMs reusáveis, desde que esses sejam usáveis no domínio de problemas sendo abordado. Mesmo assim, quanto mais usáveis são os PSMs especificados, menos reusáveis eles tendem a ser. Beys et al. [BEY 96] discutem alternativas para compatibilizar os problemas de *usabilidade X reusabilidade* em PSMs, mas este assunto ainda permanece um tema de pesquisa em diversas áreas da KE, da mesma forma que o desenvolvimento de KBSs a partir de componentes reusáveis tem sido pesquisado (Fensel et al. [FEN 97b]; Benjamins et al. [BEN 97]).

A construção de KBSs, por sua vez, pode ser fundamentada em uma metodologia de KE. Para isso, tal metodologia deve prover algum suporte para os passos de desenvolvimento a serem executados, desde a análise dos requisitos de negócio do(s) problema(s) sendo abordado(s), até os requisitos de projeto e implementação do(s) KBS(s). Todavia, o processo de orientar o desenvolvimento de um KBS segundo uma metodologia de KE por si só pode ser custoso e demorado, visto que muitos problemas baseados em conhecimento são resolvidos a partir de esquemas de conhecimento e dados previamente especificados, bem como a partir de recursos de software implementados. Em especial, este último cenário de desenvolvimento de KBSs é aplicável ao trabalho sendo apresentado, pois o projeto *PetroGrapher* vem sendo desenvolvido a partir de modelos de dados, modelos de conhecimento e recursos de software sendo construídos durante os últimos anos.

Em resumo, uma abordagem baseada em PSMs reusáveis pode ser empregada para desenvolver um novo método de raciocínio. A partir de uma revisão na área de pesquisa de PSMs, como também deste estudo de caso de aplicação de PSMs realizado, foi elaborado um conjunto de passos que podem ser seguidos para empregar PSMs no desenvolvimento de KBSs. Logo, a especificação de um PSM adequado à solução de uma tarefa pode seguir os passos:

- *avaliar bibliotecas de PSMs* - antes de qualquer especificação do raciocínio de um novo problema, é necessário definir o *tipo de problema* (Breuker [BRE 94]) que se deseja tratar. A partir disso, algumas bibliotecas de PSMs existentes podem ser analisadas, buscando uma biblioteca descrita ou adequada para a tarefa requerida. Mesmo com muitos tipos de problemas ainda pouco conhecidos e descritos por bibliotecas de PSMs, os PSMs que compõem bibliotecas descritas na literatura (algumas citadas na seção 2.7) podem ser reutilizados, desde que disponibilizem competências adequadas ao problema requerido;
- *analisar a competência dos PSMs* - as *competências* dos PSMs disponibilizados em bibliotecas de PSMs selecionadas devem ser analisadas em função da competência requerida pelo problema que se deseja resolver. Mesmo que muitos requisitos de solução de problemas baseados em conhecimento reais não sejam atendidos por PSMs

conhecidos, competências simplificadas podem ser encontradas em PSMs disponibilizados. Tais PSMs então podem ser reutilizados na análise e projeto do novo método de raciocínio requerido;

- *reutilizar os PSMs* - PSMs selecionados de bibliotecas escolhidas podem ser usados para orientar o processo de aquisição de conhecimento de domínio, assim permitindo executar os passos de inferência requeridos pelo problema. Dessa forma, esse problema pode ser resolvido a partir do PSM reutilizado associado ao conhecimento de domínio adquirido. Entretanto, alguns passos de adaptação dos PSMs selecionados devem ser realizados;
- *adaptar os PSMs* - aplicações reais normalmente possuem requisitos de solução de problemas particulares, os quais podem requerer a especificação de novos passos de inferência, diferentes fluxos de dados e conhecimento, suposições, etc. Tais adaptações são realizadas no sentido de expressar mais fielmente os requisitos de problema e domínio requeridos pelo novo PSM sendo detalhado. Ao final desse processo de adaptação de PSMs reusáveis, um PSM adequado aos principais requisitos de solução do problema requerido pode ser encontrado.

Todavia, alguns problemas podem ser encontrados durante o desenvolvimento (reutilização) do novo PSM requerido. Durante a execução do passo de avaliação de bibliotecas de PSMs, bibliotecas descritas para o mesmo tipo de problema da tarefa que se deseja resolver podem não ser encontradas. Caso não existam bibliotecas de PSMs para o tipo de problema requerido, uma nova biblioteca de PSMs pode ser desenvolvida.

Aplicações baseadas em conhecimento também podem ser definidas em função de um tipo de problema particular. Este tipo particular representa um passo de especialização de uma família de PSMs, resultando em uma família de PSMs usáveis em um domínio de aplicação. Este é o caso do presente trabalho, pois ele aborda problemas de interpretação de rochas, os quais podem ser solucionados com métodos próprios, mas que podem ser compreendidos como problemas de interpretação genéricos.

Mesmo encontrando um PSM para o mesmo tipo de problema que se deseja resolver, um PSM usável ainda pode não ser encontrado. Um PSM usável é simplificarmente compreendido como um PSM competente para resolver o problema requerido. Caso as competências dos PSMs selecionados em bibliotecas de PSMs reusáveis não sejam suficientes para alcançar as soluções requeridas no domínio de problema sendo tratado, novos PSMs podem ser desenvolvidos por adaptação de PSMs existentes, como também podem ser desenvolvidos do zero. Este também é o caso do trabalho sendo apresentado, pois os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos foram especificados do zero.

Para desenvolver um PSM a partir do zero, é possível empregar uma das propostas metodológicas de desenvolvimento de PSMs descritas na literatura e resumidamente apresentadas na seção 2.5. Porém, tais metodologias ainda não são suficientemente testadas e validadas em problemas reais e, desta forma, descrevem

passos de desenvolvimento pouco viáveis de serem executados de maneira sistemática em aplicações reais.

Uma abordagem que pode então ser empregada é utilizar os passos de desenvolvimento de diferentes metodologias de desenvolvimento de PSMs, assim buscando empregar os princípios de desenvolvimento que melhor podem auxiliar no processo de especificação dos novos PSMs requeridos. Este é o caso do desenvolvimento dos PSMs propostos, pois não é objetivo desse trabalho propor uma nova metodologia de desenvolvimento de PSMs, mas sobretudo busca-se alcançar a especificação de PSMs usáveis no domínio de problemas de interpretação de rochas.

Para facilitar a compreensão dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos, bem como definir os principais conceitos de domínio modelados, uma breve caracterização do domínio de problemas de interpretação de rochas é apresentada a seguir.

3.1 Interpretação de rochas

A Petrografia Sedimentar é a área de conhecimento da Geologia que estuda as rochas sedimentares, quanto à sua importância, descrição, classificação, origem e evolução. A interpretação de rochas no domínio da Petrografia Sedimentar envolve, entre outras atividades, um estudo detalhado da rocha e produção de uma *descrição petrográfica*. Uma descrição petrográfica é construída com dados coletados através de análises de amostras de rocha, os quais são usados para sugerir uma teoria sobre os processos de formação da rocha. Esta teoria genética envolve reconhecer e compreender o produto de um fenômeno geológico específico e, a partir desse produto, formalizar uma explicação plausível em termos de processos geológicos conhecidos.

Segundo Abel [ABE 2001], identificar quais são e como são os mecanismos de raciocínio humanos usados para interpretar rochas é um processo mais complexo do que a modelagem dos conceitos e relações que são utilizados durante um processo de inferência. Entre outros motivos citados, um processo cognitivo de raciocínio, tal como o desenvolvido por especialistas para interpretar rochas, utiliza conhecimento tácito mais do que conhecimento explícito (Nonaka et al. [NON 97]), e estes podem ser analisados e abstraídos sob diferentes perspectivas. Conforme Abel, em um domínio de problemas que aplica inferências orientadas por imagens, em especial o domínio da Petrografia Sedimentar, muitos componentes cognitivos que suportam o raciocínio simplesmente não podem ser reproduzidos e permanecem ainda pouco compreendidos.

Mesmo diante de dificuldades de compreensão, algumas tarefas de raciocínio da Petrografia podem ser reconhecidas, assim permitindo desenvolver um modelo cognitivo de raciocínio para interpretação de rochas. Conforme o modelo de perícia proposto para esse domínio de problemas (Abel et al. [ABE 98]; Abel [ABE 99]), a primeira tarefa notável desenvolvida por especialistas é o *reconhecimento visual*, onde feições de rocha relevantes para análise são identificadas e descritas. Esta tarefa acontece antes de uma tarefa de busca, que seleciona hipóteses e busca feições de rochas que possam comprová-las. A tarefa de busca tem um papel secundário no

processo de interpretação de rochas, mas é fundamental para indicar interpretações que podem ou não ser alcançadas.

Portanto, o modelo cognitivo propõe que a perícia envolvida na descrição e interpretação de rochas sedimentares clásticas (domínio especialmente abordado no sistema *PetroGrapher*) é fundamentada nas tarefas de *reconhecimento de pacotes visuais e secundariamente em estratégias de busca ou métodos analíticos*. Conforme Abel [ABE 99], o raciocínio desenvolvido para as tarefas de reconhecimento e busca pode ser caracterizado como progressivo e regressivo combinados, pois petrógrafos especialistas demonstram conhecer a interpretação de uma rocha, mas apenas secundariamente buscam dados que reforçam esta solução.

O desempenho diferenciado de petrógrafos especialistas pode ser explicado através de *pacotes visuais* conhecidos que selecionam *esquemas de rochas* que contêm interpretações complexas de diferentes tipos de rochas. A partir dos conceitos de esquemas de rochas e pacotes visuais, é possível elaborar um modelo de *como* um geólogo aplica seus mecanismos cognitivos para realizar as tarefas envolvidas na interpretação de amostras de rocha em Petrografia Sedimentar.

Como definido por De Groot [DEG 65] e VanLehn [VAN 89], um *pacote visual (chunk)* é uma "*abstração de um padrão visual associado ao seu significado particular*". Em domínios de problemas que aplicam inferências orientadas por imagens, tal como é o domínio da Petrografia, esses pacotes visuais podem ser compreendidos como índices para estruturas mentais de conhecimento mais elaboradas, as quais permitem desenvolver raciocínio complexos. Essas estruturas mentais, por sua vez, podem ser simplificadaamente descritas como *esquemas* (Mattos [MAT 91]). Nesse trabalho, *esquemas de rochas* são abstrações dos conceitos de domínio comuns de um conjunto de casos de descrição e interpretação de rochas, mantendo uma interpretação mais completa de ocorrências geológicas. Esquemas de rochas têm o papel de orientar o processo de raciocínio desenvolvidos por especialistas nesse domínio, preenchendo as lacunas onde dados coletados forem imperfeitos ou incompletos e, dessa forma, oferecendo o suporte para diferentes tarefas de interpretação.

Conforme o petrógrafo especialista, a interpretação de amostras de rochas sedimentares clásticas pode ser sintetizada pelas tarefas de *classificação composicional, caracterização de proveniência, processos/ambientes diagenéticos, processos/ambientes deposicionais e seqüência diagenética*. Dentre estas, esse trabalho focaliza a *tarefa de interpretação de processos/ambientes diagenéticos*, pois é a tarefa mais estudada durante o desenvolvimento do projeto *PetroGrapher*.

O objetivo da tarefa de interpretação de processos/ambientes diagenéticos é identificar os ambientes pelos quais a rocha passou durante sua evolução, podendo envolver ambientes próximos à superfície, sob influência de condições continentais (sob climas diversos) e marinhas, e ambientes de soterramento raso ou profundo sob as sucessões das bacias sedimentares, e de exposição superficial de rochas antes soterradas (De Ros [DER 98]; Anjos et al. [ANJ 2000]). Tais informações servem para principalmente indicar a formação de novos minerais e conseqüente consolidação da rocha sendo analisada.

Nas seções seguintes desse trabalho, serão apresentados alguns passos de desenvolvimento que foram seguidos para especificar os PSMs no domínio de problemas de interpretação de rochas. Tais passos demonstram o desenvolvimento de PSMs em um domínio de conhecimento complexo e intensivo em conhecimento, assim caracterizando alguns requisitos reais de desenvolvimento de PSMs.

3.2 Os passos de desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos

Mesmo diante de várias abordagens de KE (algumas descritas na seção 2.3), o desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos não empregou formalmente uma abordagem de KE em especial. Mesmo assim, os princípios de modelagem de conhecimento propostos na metodologia CommonKADS (Schreiber et al. [SCH 94]; Schreiber et al. [SCH 99]) foram utilizados para analisar e descrever os aspectos funcionais dos PSMs propostos nesse trabalho. A escolha da metodologia CommonKADS, em detrimento de outras abordagens de KE, deu-se pelo fato de que esta é a metodologia mais difundida e consolidada em aplicações reais, assim facilitando a compreensão desse trabalho, bem como possível reutilização dos PSMs desenvolvidos.

De maneira geral, as abordagens de KE analisadas focalizam a reutilização de bibliotecas de PSMs, mas não descrevem passos que podem ser seguidos para desenvolver novos PSMs a partir do zero. Logo, algumas propostas metodológicas de desenvolvimento de PSMs foram analisadas na seção 2.5. Estas descrevem passos de modelagem que podem ser seguidos, mas tais passos não são considerados como metodologias amplamente aceitas e validadas.

Para desenvolver os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos, os passos das propostas metodológicas de desenvolvimento de PSMs que melhor poderiam caracterizar o processo de especificação dos PSMs propostos foram utilizados. Portanto, os passos de desenvolvimento dos PSMs propostos são descritos em termos das metodologias de desenvolvimento de PSMs analisadas, mas não podem ser descritos em termos de uma das propostas em particular.

De forma geral, o desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos passou pelos seguintes passos:

- aquisição de conhecimento
- detalhamento do problema
- reutilização de bibliotecas de PSMs
- definição de um esquema de inferência básico
- adaptação/refinamento

- validação das especificações
- abstração dos PSMs

Mais uma vez, o estabelecimento de uma nova proposta metodológica de desenvolvimento de PSMs não é um dos objetivos desse trabalho. Além disso, não havia ferramentas de modelagem disponíveis para apoiar os passos de desenvolvimento de PSMs das metodologias analisadas, assim dificultando a construção dos PSMs em uma forma mais sistemática e controlável, tal como almejado por recentes visões da KE. Mesmo assim, os passos de desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos podem ser usados para orientar a construção de ferramentas de especificação de PSMs mais adequadas aos requisitos reais de desenvolvimento testados.

3.2.1 Aquisição de conhecimento

Após entrevistas realizadas com especialistas em Petrografia, nota-se claramente a utilização intensiva de grande quantidade de conhecimento e dados geológicos para alcançar melhores interpretações de rochas. Tal capacidade de solução de problemas poderia ou não ser modelada nos PSMs propostos, assim resultando em métodos mais completos e competentes, ou mais simplificados e com competências reduzidas. Mesmo assim, o uso efetivo de conhecimento e dados de domínio seria requerido por qualquer esquema de raciocínio que representasse um desempenho próximo ao desempenho real de solução de problemas de um especialista em Petrografia.

Para adquirir conhecimento e dados do domínio da Petrografia Sedimentar, as principais fontes de aquisição de conhecimento utilizadas foram:

- relatórios técnicos e artigos sobre o projeto *PetroGrapher*, os quais abordavam o desenvolvimento de uma aplicação de bancos de dados inteligentes no domínio da Petrografia;
- a tese de doutorado de Abel [ABE 2001] descrevendo a aquisição e representação de conhecimento em domínios naturais, em especial no domínio da Petrografia Sedimentar;
- entrevistas com especialistas em descrição e interpretação de rochas, em especial com o Dr. Luiz Fernando De Ros - Instituto de Geociências/UFRGS e petrógrafos do CENPES/Petrobrás;
- entrevistas com geólogos novatos (alunos do Instituto de Geociências/UFRGS).

O estudo realizado sobre esse domínio de aplicação permitiu adquirir e modelar grande quantidade de conhecimento de domínio e estratégias de controle efetivamente usadas por petrógrafos especialistas. Tal conhecimento adquirido foi usado para desenvolver e caracterizar o *esquema de raciocínio para interpretação de*

rochas proposto no capítulo seguinte desse trabalho, bem como exemplificar e validar os PSMs *Interpretação de Rochas e Interpretação de Ambientes Diagenéticos*.

3.2.2 Detalhamento do problema

O passo de *detalhamento do problema* pode ser sintetizado como *caracterização do problema e decomposição de tarefas*:

- *caracterização do problema*

Segundo as propostas de desenvolvimento de PSMs analisadas (seção 2.5), um dos primeiros passos de desenvolvimento que deve ser realizado é a caracterização do *tipo de problemas* a ser tratado. Tal análise demonstrou que o comportamento de solução de problemas desenvolvido por petrógrafos especialistas para interpretar rochas é melhor caracterizado como um *problema particular de interpretação*. Contudo, a decomposição desse comportamento poderia ser compreendida como um seqüência de tipos de problemas distintos, onde a resultado de um deles poderia ser usado como entrada em outro.

Conforme descrito por Harmon et al. [HAR 90], problemas de interpretação são aqueles que realizam análises sistemáticas e completas de dados para determinar seus significados, buscando encontrar interpretações consistentes e corretas para os dados analisados. Tal compreensão pode sintetizar a tarefa de interpretação de ambientes diagenéticos, a qual foi especialmente analisada para fundamentar a especificação dos PSMs propostos. Os PSMs *Interpretação de Rochas e Interpretação de Ambientes Diagenéticos*, portanto, podem ser caracterizados como *representativos de tarefas de interpretação particulares*, pois ambos são o resultado da análise de um processo de solução de problemas de uma tarefa bastante específica da geologia.

- *decomposição de tarefas*

O enfoque de desenvolvimento de PSMs baseado na decomposição de tarefas em subtarefas foi simplificada durante o detalhamento do problema sendo tratado. No entanto, as tarefas *de classificação composicional, caracterização de proveniência, processos/ambientes diagenéticos, processos/ambientes deposicionais e seqüência diagenética* foram identificadas. Dentre estas, a tarefa de *processos/ambientes diagenéticos* foi escolhida para análise e, desta forma, fundamenta os PSMs especificados.

Entre outros motivos, essa tarefa foi escolhida porque havia previamente uma compreensão de muitos dos passos de raciocínio básicos que poderiam ser representados como um único PSM, assim não descrevendo explicitamente uma seqüência de decomposição de tarefas. Essa opção de modelagem também pode ser notada em bibliotecas de PSMs que cobrem *tarefas complexas*, tais como as bibliotecas de CommonKADS (Schreiber et al. [SCH 99]) e diagnóstico (Benjamins et al. [BEN 94]; Benjamins et al. [BEN 95]), por exemplo.

3.2.3 Reutilização de bibliotecas de PSMs

O passo de *reutilização de bibliotecas de PSMs* pode ser sintetizado como *análise das bibliotecas disponíveis e análise da competência dos PSMs disponíveis*:

- *análise das bibliotecas disponíveis*

Não foram encontradas bibliotecas de PSMs de interpretação, as quais poderiam disponibilizar métodos claramente reusáveis no domínio de problemas de interpretação de rochas. Contudo, as bibliotecas de PSMs de CommonKADS (Schreiber et al. [SCH 94]; Schreiber et al. [SCH 99]), diagnóstico (Benjamins et al. [BEN 94]) e planejamento (Barros et al. [BAR 96]) foram avaliadas para modelar os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos, pois eram as bibliotecas melhor descritas que foram encontradas.

O PSM Avaliação (*Assessment*) descrito na biblioteca de CommonKADS foi considerado o PSM analisado que melhor descreve operacionalmente o comportamento de solução de problemas desenvolvido nesse domínio. Tal método consiste em tomar um "caso" e uma "descrição do sistema" como entrada, e resultar em uma "decisão" como saída. Aplicado a uma tarefa de interpretação de rochas, esse PSM poderia usar um caso de descrição de amostra de rocha e uma descrição de normas de interpretação de rochas aplicáveis como entradas, assim realizando uma ampla análise sob esses dados e conhecimento informados a fim de alcançar uma interpretação. Como proposto por Abel [ABE 2001], o PSM Avaliação é utilizado como um esquema básico para explicar o processo de interpretação de rochas.

- *análise da competência dos PSMs disponíveis*

Em uma análise mais precisa dos PSMs das bibliotecas disponíveis, em especial sobre o PSM Avaliação, foi verificado que não era disponibilizada uma competência adequada aos requisitos de tarefa e domínio a serem modelados, tal como requerida pelo projeto *PetroGrapher*. Mesmo assim, não foi realizada uma análise metodológica das competências dos PSMs encontrados em relação às tarefas de interpretação de rochas, pois nem mesmo havia indicações na literatura consultada a respeito de estratégias adequadas à realização desta análise. Logo, somente características relativas a terminologia de descrição de papéis de conhecimento e fluxo de dados e conhecimento do PSM Avaliação foram reutilizados na especificação dos PSMs consolidados como Interpretação de Rochas e Interpretação de Ambientes Diagenéticos.

3.2.4 Definição de um esquema de inferência básico

O passo de *definição de um esquema de inferência básico* pode ser sintetizado como *definição do espaço de problemas, estabelecimento da competência, definição de uma estratégia de solução de problemas e definição de um PSM mais operacional*:

- *definição do espaço de problemas*

O *espaço de problemas* do *esquema de raciocínio para interpretação de rochas* (descrito no capítulo 4) foi inicialmente bem delimitado, pois somente dessa forma os PSMs desenvolvidos poderiam produzir explicações plausíveis sobre os processos de formação de amostras de rochas sedimentares clásticas, assim como requerido no domínio da Petrografia. Resultar *explicações plausíveis* certamente foi um dos primeiros *requisitos/objetivos* identificados durante o processo de desenvolvimento, tal como uma *noção básica de competência* (Wielinga et al. [WIE 98]).

- *estabelecimento da competência*

Um dos primeiros passos que foram realizados foi o *estabelecimento da competência* do PSM Interpretação de Rochas em relação ao comportamento de solução de problemas desenvolvido pelo especialista em Petrografia, assim como para descrever *uma teoria de competência* de processos de interpretação de rochas, tal como proposto por Wielinga et al. [WIE 98]. Essa relação de competência não foi, e continua não sendo, facilmente estabelecida, pois a competência de processos de raciocínio humanos é de difícil compreensão e formalização.

Mesmo diante dessas dificuldades, o PSM Interpretação de Rochas descreve um comportamento de solução de problemas em um nível abstrato, o qual permite compreender como funciona o raciocínio desenvolvido por petrógrafos especialistas. Abel [ABE 2001] também confirma esse processo, pois afirma que a automatização por computador de mecanismos cognitivos para realizar a descrição e interpretação de rochas pode não corresponder exatamente aos modelos mentais do especialista, mas corresponde a adaptação desses modelos às limitações computacionais.

- *definição de uma estratégia de solução de problemas*

A *estratégia de solução de problemas* mais genérica alcançada e especificada pode ser compreendida como *seleção múltipla de padrões de domínio*: pacotes visuais que selecionam esquemas de rochas, os quais permitem selecionar pacotes visuais mais relevantes para uma análise localizada de soluções. Desde o início da modelagem dos PSMs propostos, tal estratégia foi intimamente relacionada com a tarefa de interpretação de rochas. Esse relacionamento aplicado inicialmente poderia ser caracterizado como *um movimento bidimensional de especificação desenvolvido no sentido dos eixos de compromissos de tarefa e estratégia de solução de problemas*, assim como proposto por Fensel et al. [FEN 2000].

- *definição de um PSM mais operacional*

Uma vez tendo especificado um PSM Interpretação de Rochas aceitável segundo requisitos do domínio de problemas sendo tratado, tal especificação foi usada como *um esquema de inferência básico* para o desenvolvimento do PSM Interpretação de Ambientes Diagenéticos, pois o primeiro não representava um modelo adequado para fins de operacionalização. Todavia, o próprio PSM Interpretação de Rochas é fundamentado em um modelo de raciocínio bastante elaborado, pois é elaborado em função de estruturas de conhecimento muito particulares. As estruturas de pacotes visuais e esquemas de rocha poderiam ser

citadas como exemplos disso, pois elas já detalham a *natureza do conhecimento de domínio* (Wielinga et al. [WIE 98]) sendo tratado nesse processo de raciocínio.

O PSM Interpretação de Rochas, portanto, pode ser considerado uma abstração de um processo de raciocínio significativamente representativo do domínio de problemas de interpretação de rochas, mas também pode ser compreendido como uma especificação ainda um tanto intratável computacionalmente. Além disso, esta especificação de método é muito pouco semelhante a *esquemas genéricos e simplificados*, tais como: *gerar&testar* e *propor&revisar*, os quais são usados para orientar o desenvolvimento de bibliotecas de PSMs, tal como proposto por Fensel et al. [FEN 2000], e também demonstrado na biblioteca de PSMs para projeto parametrizado (Motta et al. [MOT 98]). Contudo, os passos de desenvolvimento dos PSMs não empregaram esses esquemas inferenciais simplificados para derivar a especificação de nenhum dos PSMs propostos.

3.2.5 Adaptação / refinamento

De maneira geral, o passo de adaptação/refinamento pode ser dividido em passos de especificação mais gerais e dois passos de especificação mais específicos, descritos como *estabelecimento de suposições e detalhamento operacional*. Os passos de especificação mais gerais, por sua vez, são divididos em *especificação usando movimentos, especificação de inferências, especificação do PSM Interpretação de Ambientes Diagenéticos e especificação de PSMs alternativos*:

- *especificação usando movimentos*

As adaptações/refinamentos realizados para desenvolver os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos podem ser compreendidos como *movimentos bidimensionais de especificação*, assim como propostos por Fensel et al. [FEN 2000]. Tais movimentos desenvolvidos possuem uma característica em comum: envolvem os eixos de *compromissos de tarefas*, ou *estratégias de solução de problemas*, sempre associados ao eixo de *suposições sobre conhecimento de domínio*.

Por exemplo, um movimento nos eixos de compromissos de tarefas e suposições sobre conhecimento de domínio realizado poderia ser explicado como sendo o refinamento de uma *tarefa de interpretação* para uma *tarefa de interpretação de rochas*, assim supondo a disponibilidade de conhecimento de domínio sobre "tarefas de interpretação de amostras de rochas". Outro movimento bastante claro envolvendo os eixos de compromissos de tarefas e suposições sobre conhecimento de domínio foi realizado na adaptação do PSM Interpretação de Rochas para o PSM Interpretação de Ambientes Diagenéticos, pois este último efetivamente considera o conhecimento de domínio adquirido e representado no âmbito do projeto *PetroGrapher*.

- *especificação de inferências*

As inferências básicas definidas em especificações intermediárias de ambos os PSMs também foram *adaptadas/refinadas* segundo *requisitos de conhecimento de*

domínio. Um exemplo que pode ser citado é o refinamento da inferência de "composição de soluções" do PSM Interpretação de Rochas. Este passo de inferência foi especializado para apresentar somente "composições possíveis de soluções" no PSM Interpretação de Rochas, e devido às dificuldades de aquisição de conhecimento foi posteriormente simplificado para apenas apresentar uma "combinação de soluções" no PSM Interpretação de Ambientes Diagenéticos.

- *especificação do PSM Interpretação de Ambientes Diagenéticos*

O processo de adaptação/refinamento mais explicitamente realizado durante o desenvolvimento foi a derivação do PSM Interpretação de Rochas para o PSM Interpretação de Ambientes Diagenéticos. Tal processo poderia ser comparado a *um processo de adaptação/refinamento* de um modelo de análise para um modelo que considera requisitos de conhecimento mais fortes, tais como: limitações de aquisição de conhecimento, representação de conhecimento, derivação de soluções possíveis, confiabilidade das entradas, etc. Tais passos foram realizados no sentido de detalhar *uma teoria de competência mais operacional* (tal como proposto por Wielinga et al. [WIE 98]), representada mais precisamente no PSM Interpretação de Ambientes Diagenéticos.

- *especificação de PSMs alternativos*

Durante o processo de adaptação/refinamento do PSM Interpretação de Rochas para o PSM Interpretação de Ambientes Diagenéticos, vários *PSMs alternativos* foram gerados. O detalhamento desses PSMs alternativos poderia ser caracterizado como *um processo de configuração* (Teije et al. [TEI 96]), ou como *um processo de especificação de uma teoria de competência refinada* (Wielinga et al. [WIE 98]). Porém, os PSMs intermediários alcançados perfeitamente poderiam descrever esquemas de inferência usáveis na Petrografia, pois eles representavam especificações conceituais que demonstravam requisitos de competência requeridos para resolver o problema, mas competências ainda não adequadas aos requisitos de inferência do projeto *PetroGrapher*.

Dentre os PSMs alternativos adaptados do PSM Interpretação de Rochas, apenas um foi selecionado e descrito como o PSM Interpretação de Ambientes Diagenéticos. Tal PSM foi usado como modelo de análise mais refinado, o qual orientou o desenvolvimento de um algoritmo de inferência para o sistema *PetroGrapher* (descrito na seção 5.5). Mesmo assim, a descrição de tais *PSMs variantes* (Fensel et al. [FEN 2000]), ou possivelmente *ajustados* (Teije et al. [TEI 96]), poderia ser relevante para diferentes domínios de aplicação (ou tarefas no domínio da Petrografia) que viessem a reutilizar tais descrições mais operacionais.

3.2.5.1 Estabelecimento de suposições

O passo de *estabelecimento de suposições* pode ser descrito como *caracterização da competência, descrição de suposições básicas e descrição de suposições mais abrangentes*:

- *caracterização da competência*

Existe um consenso atualmente que o processo de desenvolvimento de PSMs seja *dirigido por suposições*. Tal afirmação tem grande relevância para a caracterização dos PSMs construídos, pois diversas suposições foram estabelecidas durante o passo de adaptação/refinamento dos modelos. Entre outras suposições estabelecidas, as suposições sobre existência de *associações entre pacotes e esquemas e suporte para composições originais mais comuns* (descritas respectivamente nas seções 4.1.4 e 4.3.4) podem exemplificar o processo de caracterização da competência dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos, respectivamente.

- *descrição de suposições básicas*

Somente a partir da definição de *uma estrutura de inferência básica* para representar um processo de raciocínio de interpretação de rochas desenvolvido pelo especialista em Petrografia, algumas suposições simplificadas foram identificadas. Uma das primeiras suposições estabelecidas que exemplificam esse fato foi a descrição do PSM Interpretação de Rochas em função de esquemas de rochas e pacotes visuais. Estes são os principais conceitos (ou se supõe que sejam) que suportam a perícia neste domínio de problemas.

Segundo Abel et al. [ABE 98], especialistas em Petrografia efetivamente utilizam estruturas de conhecimento bastante elaboradas, as quais permitem desenvolver raciocínios mais complexos. Todavia, pacotes visuais e esquemas de rochas são interpretações simplificadas destas estruturas de conhecimento mentais, assim permitindo compreender e explicar o comportamento de solução de problemas desenvolvido por humanos.

- *descrição de suposições mais abrangentes*

Muitas das suposições que foram estabelecidas dizem respeito ao detalhamento do esquema de raciocínio para interpretação de rochas (descrito no capítulo 4) como um todo, caracterizando o comportamento de solução de problemas derivável pelo conjunto dos modelos especificados. Assim como pôde ser notado, a descrição formal das suposições, tal como proposto por Teije et al. [TEI 96], pode dificultar muito o processo de validação das especificações de PSMs intermediários junto a especialistas, além de tornar o processo de detalhamento dos PSMs demorado e cansativo.

Mesmo assim, a descrição de suposições para tarefas de diagnóstico em um formato informal, tal como aquelas apresentadas por Fensel et al. [FEN 98a] e Benjamins et al. [BEN 96b], auxiliaram na identificação das suposições estabelecidas para os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos. Tal como essas suposições sobre diagnóstico, as suposições que foram estabelecidas nesse trabalho (seções 4.1.4 e 4.3.4) possivelmente podem ser reutilizadas para orientar o estabelecimento de novas suposições em outras tarefas de raciocínio em domínios similares a este, bem como em outros domínios de aplicação.

3.2.5.2 Detalhamento operacional

Está explícito em diferentes abordagens de desenvolvimento de PSMs (seção 2.5) que eles sejam utilizados tanto como *um meio de formalização de processos de raciocínio eficientes*, como também *um meio de operacionalização* destes. Como requerido pelo projeto *PetroGrapher*, um dos critérios que nortearam o desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos foi alcançar uma *estratégia de solução de problemas* para expressar *aspectos operacionais de inferência* possíveis de serem empregados no domínio da Petrografia. No entanto, especificações de PSMs eficientes não foram consideradas logo em princípio, pois somente quando *uma estratégia operacional preliminar* estava completamente especificada (tal como atendendo a *objetivos estáticos* - Teije et al. [TEI 96]), então aspectos de eficiência requeridos nesse domínio de problemas foram explicitamente analisados.

Segundo Fensel et al. [FEN 96], é necessário considerar que especialistas também desempenham tarefas de raciocínio sob limitações reais, onde grande parte do conhecimento perito está associado com algum tipo de raciocínio eficiente considerando essas limitações. O PSM Interpretação de Rochas, neste caso, pode representar mais fielmente *uma estratégia de solução de problemas* descrita a partir de uma análise do comportamento de solução de problemas do especialista em Petrografia. Portanto, tal modelo já considera *como* alcançar soluções melhores de maneira mais eficiente, assim descrevendo quando efetivamente utilizar o conhecimento de domínio para resolver problemas no domínio da Petrografia Sedimentar.

Todavia, a descrição dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos em um formalismo executável, o qual poderia ser utilizado para implementar e testar algoritmos de inferência derivados automaticamente a partir da especificação dos PSMs, não foi realizada. Para declarar as *especificações de controle* dos PSMs modelados, uma linguagem contendo estruturas de controle simplificadas foi empregada. Desta forma, optou-se por uma especificação semiformal (tipo *script*), não descrevendo as especificações operacionais dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos em um formalismo muito preciso, normalmente não muito claro para leigos em computação. Isso facilitou a validação dessas especificações operacionais junto ao especialista em Petrografia, mas não permitiu realizar sobre elas nenhum processo de *configuração dinâmica* dos PSMs especificados, assim como proposto por Teije et al. [TEI 96].

3.2.6 Validação das especificações

Uma abordagem de compreensão, como também de validação, bastante prática empregada durante o processo de desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos foi instanciar as especificações intermediárias deles com exemplos reais de interpretações de rochas adquiridos com o petrógrafo especialista. Tais exemplos permitiam demonstrar o funcionamento da inferência em uma maneira simples, tal como se fosse expressa como um DFD.

Conforme proposto por Wielinga et al. [WIE 98], *requisitos de correção e confiabilidade* dos PSMs representados foram explicitamente analisados, mas foram objetivamente considerados ao final de passos de adaptação/refinamento resultantes no PSM Interpretação de Ambientes Diagenéticos. Tais requisitos foram principalmente derivados a partir de entrevistas com o especialista em Petrografia, bem como de *uma teoria de domínio refinada* detalhada como modelos de conhecimento de domínio especificados (apresentados na seção 4.2). Tais requisitos de informação do problema foram então descritos principalmente como suposições, assim permitindo refinar e validar sucessivamente a competência de ambos os PSMs.

3.2.7 Abstração dos PSMs

O passo de *abstração dos PSMs* pode ser descrito como *abstração de aspectos de implementação e abstração de terminologia de descrição*:

- *abstração de aspectos de implementação*

As especificações dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos resultantes abstraem quaisquer aspectos de implementação, mas já consideram alguns *tipos de conhecimento de domínio requeridos*, assim como proposto por diferentes abordagens de especificação de PSMs. O PSM Interpretação de Ambientes Diagenéticos é um método mais voltado para o conhecimento de domínio adquirido e representado, pois ele foi desenvolvido considerando mais objetivamente as possibilidades de aquisição de conhecimento experimentadas.

O PSM Interpretação de Ambientes Diagenéticos, portanto, modela mais precisamente o conhecimento de controle requerido por projetos de mecanismos de inferência típicos de KBSs. O PSM Interpretação de Rochas, por sua vez, é uma especificação de raciocínio mais abstrata, tal como um esquema de raciocínio representativo de *uma família de métodos* de interpretação de rochas, assim como proposto por Teije et al. [TEI 96].

- *abstração de terminologia de descrição*

Durante o desenvolvimento, a terminologia usada para descrever os diferentes PSMs intermediários especificados esteve muito relacionada ao domínio de problemas de interpretação de rochas, facilitando *um refinamento e validação sucessivos* de modelos intermediários construídos junto ao especialista em Petrografia. Percebe-se que essa terminologia usada para descrevê-los ainda está muito associada ao domínio da Petrografia, de forma que os PSMs especificados representam modelos mais práticos e usáveis no domínio de problemas do projeto *PetroGrapher*.

Todavia, alguns passos de *abstração e simplificação dos termos* usados para descrever os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos foram realizados. Tais passos de adaptação tornaram as *estruturas de inferência* dos PSMs propostos mais simplificadas, de forma que as especificações possam vir a expressar mais fielmente PSMs reusáveis.

3.3 Uma análise do desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos

Em um primeiro momento, o desenvolvimento dos PSMs para o domínio de problemas de interpretação de rochas veio a proporcionar uma melhor compreensão e posterior refinamento do modelo de perícia em interpretação de rochas proposto originalmente em Abel et al. [ABE 98] e Abel [ABE 99]. Uma síntese desses passos de refinamento pode ser encontrada em Abel [ABE 2001].

Como projetos preliminares de mecanismos de inferência para interpretação de rochas tinham sido previamente realizados no projeto *PetroGrapher* (Silva [SIL 97]; Silva [SIL 99]), muitos dos passos básicos de inferência já podiam ser previstos, tal como para detalhar *uma teoria operacional de competência* (Wielinga et al. [WIE 98]). Dessa forma, a especificação do PSM Interpretação de Ambientes Diagenéticos, por exemplo, foi desenvolvida também como um processo de *engenharia reversa de mecanismos de inferência prototípicos* construídos antes e durante o desenvolvimento dos PSMs propostos.

O processo de adaptação/refinamento sucessivo realizado sobre os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos permitiu, entre outras coisas, detalhar os principais fluxos de inferência, com também o conhecimento e dados de domínio efetivamente utilizados no raciocínio. Conforme Fensel et al. [FEN 96], esse conhecimento de controle, muitas vezes representado implicitamente, é fundamental para viabilizar a solução de problemas na prática.

Entretanto, tais adaptações/refinamentos não foram realizadas no sentido de especificar um mecanismo de raciocínio eficiente, tal como PSMs muitas vezes são compreendidos. O objetivo principal do processo de modelagem, portanto, foi *demonstrar um esquema de raciocínio corretamente expresso e compreendido, tanto pelo modelador, quanto por especialistas em Petrografia*. Sendo assim, os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos são esquemas de análise corretos e aceitáveis no domínio de problemas do projeto *PetroGrapher*.

A maior dificuldade encontrada durante o processo de desenvolvimento dos PSMs foi a compreensão do domínio de problemas da Petrografia Sedimentar. Tal domínio de problemas complexo, o qual demandou muito esforço para adquirir alguma fluência, ainda é notável nos exemplos reais apresentados nesse trabalho. Dessa forma, o processo de especificação dos PSMs desenvolvidos foi demorado e complexo, pois quanto mais os modelos foram detalhados, mais e melhor compreensão sobre o processo de raciocínio desenvolvido nesse domínio era adquirida, e mais passos de adaptação/refinamento então foram realizados. Conforme Studer et al. [STU 99], tal fato é característico em processos de modelagem de conhecimento, tal como é compreendida a KE atual, pois "*novas observações sobre a realidade podem levar a refinamentos ou modificações dos modelos construídos*".

Entretanto, os passos de adaptação/refinamento e validação desenvolvidos poderiam ter sido mais simplificados, desde que objetivos claros para cada passo

intermediário tivessem sido previamente estabelecidos. Assim como pode ser percebido, tais objetivos deveriam ser típicos do domínio de problemas sendo tratado, assim evitando o processo de desfazer adaptações/refinamentos intermediários realizados. Mas para sempre considerar objetivos típicos do domínio de problemas sendo tratado, o especialista deveria estar constantemente integrado ao processo de modelagem, e isso demandava tempo escassamente conquistado para entrevistas.

De forma geral, o desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos poderia ser caracterizado como um processo de refinamento gradativo de modelos de conhecimento inferenciais no sentido de tarefas de raciocínio e circunstâncias específicas de domínio. Também pode ser notado que esse processo de desenvolvimento foi guiado por suposições, as quais permitiram caracterizar mais precisamente a competência dos PSMs especificados, tal como detalhando os tipos de soluções possíveis de serem derivadas por eles. Contudo, a identificação e a descrição das suposições típicas do domínio de problemas de interpretação de rochas não foi nada simplificada, demandando inúmeras entrevistas com o especialista em Petrografia. Além disso, tais suposições não são, e continuam não sendo, facilmente descritas em termos de conceitos simplificados de domínio, assim como proposto pelas metodologias de desenvolvimento de PSMs analisadas.

O desenvolvimento desse estudo de caso de especificação de PSMs permitiu concluir que a proposta descrita por Wielinga et al. [WIE 98] apresenta o esquema conceitual de desenvolvimento de PSMs mais aplicável aos requisitos de análise dos métodos de raciocínio do especialista. No entanto, os passos de modelagem propostos são simplificados e, desta forma, não fornecem o suporte para a construção de PSMs de forma sistemática.

As propostas de Fensel et al. [FEN 2000] e Teije et al. [TEI 96] propõem passos de modelagem mais básicos, os quais expressam mais fielmente o desenvolvimento de PSMs a partir de modelos de conhecimento e dados detalhados. Além disso, Teije et al. [TEI 96] estabelece passos de validação das especificações, os quais são essenciais para alcançar descrições simplificadas e representativas de métodos de raciocínios humanos. Porém, tais passos são realizados em função de descrições formais de PSMs, dificultando a validação dos modelos junto a especialistas. Por fim, Fensel et al. [FEN 2000] estabelece um esquema de desenvolvimento que integra com sucesso os princípios de especificação e refinamento das outras propostas metodológicas analisadas, assim disponibilizando uma seqüência de passos que poderia descrever este estudo de caso. Todavia, alguns dos princípios desta proposta podem não ser facilmente visíveis no processo de especificação de PSMs derivados de raciocínios humanos, mas que certamente são aplicáveis à representação de processos de inferência típicos de KBSs.

O desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos, no entanto, pode ser melhor caracterizado como um estudo de caso prático de especificação de um esquema de raciocínio para uma tarefa de raciocínio real, assim como é apresentado o trabalho de Speel et al. [SPE 96]. A descrição desse estudo pode ser útil para transformar uma proposta de especificação de métodos de raciocínio ainda um tanto acadêmica, em uma alternativa conceitual e prática que efetivamente venha a auxiliar a caracterização de requisitos de modelagem e construção de mecanismos de inferência de KBSs.

4 Um esquema de raciocínio para interpretação de rochas

O *esquema de raciocínio para interpretação de rochas* proposto neste capítulo é baseado em uma análise do comportamento de solução de problemas em descrição e interpretação de rochas em Petrografia Sedimentar. Essa compreensão é expressa em dois PSMs inéditos: o PSM Interpretação de Rochas e o PSM Interpretação de Ambientes Diagenéticos. Além destes métodos, este esquema de raciocínio é composto por modelos de conhecimento de domínio especificados. Tais modelos indicam técnicas de aquisição e representação de conhecimento que podem ser usadas para desenvolver esquemas de raciocínio em diferentes domínios de aplicação, assim como foram empregadas no domínio da Petrografia Sedimentar.

A caracterização dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos e dos modelos de conhecimento de domínio especificados torna explícito o processo de análise e descrição do conhecimento inferencial usado para desenvolver um KBS em um domínio de aplicação considerado complexo. Tal processo descreve as principais abstrações realizadas na representação de um esquema de raciocínio para uma aplicação real, assim detalhando modelos que possam orientar futuros passos de pesquisa no projeto *PetroGrapher*.

O *esquema de raciocínio para interpretação de rochas* foi desenvolvido a partir do estudo de um especialista em Petrografia Sedimentar, complementado pela análise cognitiva de um grupo de geólogos que incluiu especialistas, intermediários e novatos nesse domínio. Uma caracterização do domínio, os resultados da análise cognitiva, a descrição dos métodos de aquisição de conhecimento e um detalhamento preliminar dos modelos de conhecimento especificados podem ser obtidos em Abel et al. [ABE 95], Abel et al. [ABE 98] e Abel et al. [ABE 99].

4.1 O PSM Interpretação de Rochas

Dentre outras tarefas de raciocínio desenvolvidas por petrógrafos especialistas quando realizam uma descrição petrográfica, a tarefa escolhida para análise nesse trabalho envolve a sugestão dos processos e ambientes que formaram uma rocha. Logo, os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos são especialmente suportados pela análise desta tarefa. Todavia, esses PSMs também podem ser considerados como especificações de uma análise cognitiva fundamentada em domínios de problemas orientados por imagens, assim como proposto por Abel [ABE 2001]. Desta forma, espera-se que os métodos propostos sejam reusáveis na solução de problemas nestes domínios de aplicação mais genéricos.

A Figura 4.1 apresenta a *estrutura de inferência* do PSM Interpretação de Rochas. Essa estrutura representa os caminhos possíveis do processo de inferência usando *papéis de conhecimento e passos de inferência*, tal como proposto como formalismo de descrição de PSMs. As setas indicam o *fluxo de dados e conhecimento* do PSM Interpretação de Rochas, mas não indicam a direção do raciocínio na estrutura como um todo ou quantas vezes cada um dos caminhos será percorrido.

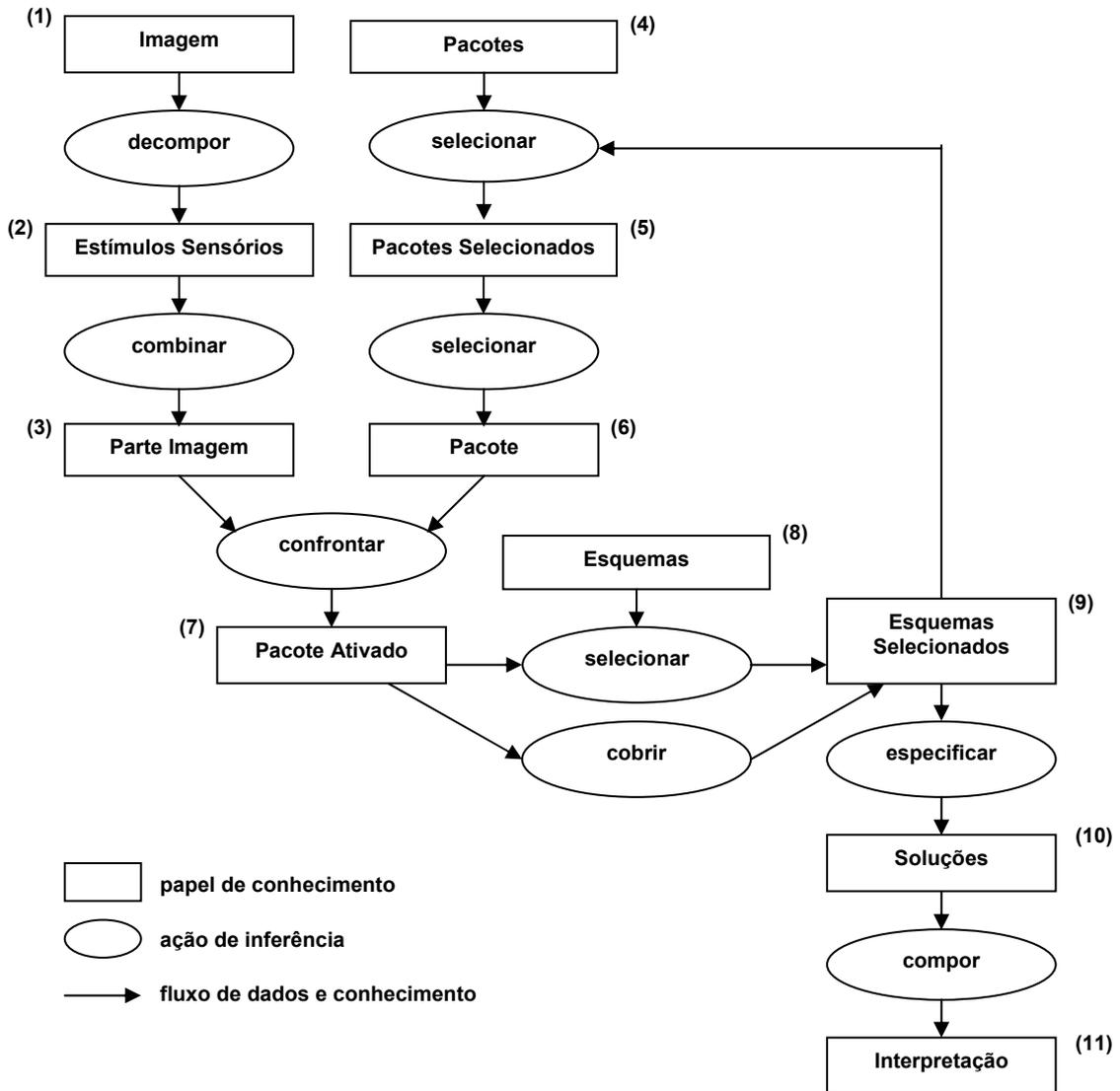


FIGURA 4.1 - Estrutura de inferência do PSM Interpretação de Rochas

A estrutura de raciocínio descrita na Figura 4.1 pode ser explicada por um raciocínio progressivo e um regressivo combinados. Desta forma, o fluxo de inferência detalhado pode ser percorrido progressivamente (iniciando pela análise de uma imagem) ou regressivamente (iniciando por esquemas conhecidos), da mesma forma que o conhecimento de domínio requerido por cada ação de inferência pode ser analisado progressivamente ou regressivamente, assim permitindo executar cada um dos passos de inferência detalhados no PSM Interpretação de Rochas.

Caso seja percorrido progressivamente, o processo de interpretação de rochas modelado pode ser explicado da seguinte forma: a primeira parte do raciocínio tem início com o exame visual de uma amostra de rocha (1), sendo capturado um conjunto de estímulos sensoriais (2). Tais estímulos são combinados de diferentes formas resultando na seleção de uma parte da imagem (3). Paralelamente, o petrógrafo seleciona do seu conjunto de pacotes visuais conhecidos (4) aqueles (5) que podem estar associados ao tipo particular de rocha representada em um esquema de rochas

(9). Caso não haja esquemas de rochas selecionados (9), os pacotes visuais (4) mais significativos para todos os esquemas de rochas conhecidos (8) são selecionados. Desse subconjunto (5) de pacotes selecionados, um dos pacotes visuais é selecionado (6) (normalmente pelo seu índice de relevância) e confrontado com a parte da imagem reconhecida na rocha (3). A imagem (1) será sucessivamente comparada a cada um dos pacotes visuais do subconjunto (5) até que um ou mais pacotes sejam similares (ativados) (7) ao aspecto visto na rocha (3).

Um pacote ativado (7), por sua vez, dispara a segunda parte do processo de raciocínio modelado, selecionando dos esquemas de rochas (8) um ou mais esquemas associados ao pacote ativado (9). Estes esquemas selecionados (9) provocarão a seleção de mais subconjuntos de pacotes (5), os quais serão analisados em um processo localizado de busca que tem como objetivo validar um esquema (9) considerado, neste caso especial, como hipótese.

O raciocínio se repetirá até que toda a rocha seja examinada visualmente e todos os esquemas de rochas (8) tenham sido avaliados. Os esquemas (9) que forem cobertos (validados) pelo maior número de pacotes visuais ativados (7), ou que tenham pacotes ativados que por si só sejam suficientemente relevantes para indicar uma solução, serão empregados para especificar as soluções (10). Como uma rocha pode passar por diferentes processos durante a sua formação, mais de uma solução pode ser encontrada. Tais soluções são, por sua vez, combinadas na interpretação (11), desde que não sejam conflitantes.

Para melhor caracterizar este processo de raciocínio, ele será exemplificado a seguir por meio de conceitos de domínio típicos da Petrografia Sedimentar, os quais são verbalizados pelo especialista quando ele realiza a interpretação uma amostra de rocha.

Em princípio, uma "amostra de rocha" é uma imagem (1) e pacotes visuais conhecidos, tais como, por exemplo: "oxidação, dissolução, caulinição, calcita, ...", são um conjunto de pacotes (4). Em seguida, a "amostra de rocha" é decomposta em um conjunto de "feições de rocha" descritas como um conjunto de estímulos sensoriais (2). Estas "feições" são então combinadas resultando em "aspectos de rocha relevantes", vistos como parte da imagem (3). Em paralelo, um subconjunto de pacotes visuais (5), por exemplo: "caulinição e dissolução", é selecionado. Esta seleção pode ser aleatória, como também podem ser selecionados pacotes visuais mais indicativos de diferentes esquemas de rochas (9). Deste subconjunto de pacotes (5), um pacote visual (6) é selecionado, por exemplo, "caulinição".

Os padrões indicados pelos "aspectos de rocha relevantes" e pelo pacote visual "caulinição" são confrontados. Caso estes padrões sejam similares, o pacote "caulinição" é um pacote ativado (7). Este pacote seleciona de um conjunto de esquemas de rochas (8), por exemplo: "arenitos", um esquema (9) relacionado a ele, por exemplo: "arenito originado por telodiagênese". Estando em um processo de busca de pacotes conhecidos que confirmem um esquema já selecionado, um pacote ativado é usado para reforçar a certeza sobre um esquema (9) já selecionado, nesse caso, "caulinição" reforçaria a certeza sobre um esquema "arenito originado por telodiagênese". Porém, em uma visão progressiva do raciocínio, o "arenito originado por telodiagênese" é usado para selecionar um novo subconjunto de pacotes (5) de

todos os pacotes conhecidos (4). Os pacotes deste subconjunto serão sucessivamente buscados na rocha, assim tentando confirmar um esquema (9) considerado, neste caso especial, como hipótese.

Em determinado momento do processo de interpretação de rochas, o esquema (9) "Arenito originado por telodiagênese" pode ser suficientemente validado por pacotes ativados para sugerir uma interpretação parcial. Ocorrendo isso, uma solução (10), tal como, por exemplo: "Telodiagênese sob condições meteóricas" pode ser especificada a partir do esquema (9) sendo considerado. Como uma rocha pode passar por diferentes processos durante sua formação, mais de uma solução (10) pode ser sugerida. Sendo assim, estas interpretações parciais são compostas em uma interpretação final (11), tal como, por exemplo: "Telodiagênese sob condições meteóricas" e "Eodiagênese meteórica continental sob clima úmido", desde que estas soluções parciais não sejam conflitantes.

Em resumo, o PSM Interpretação de Rochas descreve um processo de raciocínio real especificado em função de papéis desempenhados por conceitos abstratos, os quais podem ser identificados em diferentes domínios de problemas e aplicações. De forma mais genérica, diferentes problemas de interpretação possivelmente podem ser descritos em função de imagens, pacotes e esquemas, os quais compõem um modelo de análise do raciocínio. Tal modelo pode orientar o processo de aquisição de conhecimento de domínio requerido para construir um KBS, assim como é empregado para explicar uma análise do raciocínio desenvolvido por especialistas.

4.1.1 A competência do PSM Interpretação de Rochas

Como proposto por diferentes abordagens de descrição de PSMs (seção 2.4), a competência do método de interpretação de rochas é descrita de maneira simplificada, apenas apresentando as principais entradas e saídas do PSM. Portanto, *o PSM Interpretação de Rochas sugere uma interpretação de rocha a partir de uma imagem de amostra de rocha analisada segundo associações de pacotes visuais e esquemas de rochas conhecidos*. Esta competência ainda não considera como este PSM realiza a tarefa de interpretação de rochas. No entanto, pode ser percebido que ela é descrita em função de componentes abstratos: *imagem de amostra de rocha, pacotes visuais e esquemas de rocha*. Portanto, demonstrar esse processo de raciocínio em função destes conceitos de domínio requeridos é uma das primeiras suposições estabelecidas sobre o processo de interpretação de rochas especificado.

4.1.2 Uma especificação operacional para o PSM Interpretação de Rochas

A especificação operacional do processo de interpretação de rochas é descrita em um formalismo operacional simplificado. O PSM Interpretação de Rochas é composto pelos seguintes papéis de conhecimento: *imagem, estímulos sensoriais, parte imagem, pacotes, pacotes selecionados, pacote, pacote ativado, esquemas, esquemas selecionados, soluções e interpretação*.

O PSM Interpretação de Rochas é composto pelas seguintes ações de inferência distintas:

- *decompor* - decompõe aspectos visuais individuais de uma imagem sendo analisada, assim formando uma coerência em termos de estímulos sensoriais;
- *combinar* - combina estímulos sensoriais de diferentes formas resultando em diferentes partes da imagem sendo analisada;
- *selecionar* - seleciona um subconjunto de padrões visuais conhecidos e, deste subconjunto, seleciona um pacote em particular (normalmente pelo seu índice de relevância). Esta inferência elementar também é usada para selecionar de um conjunto de esquemas algum esquema relacionado a um pacote reconhecido em uma parte da imagem sendo analisada;
- *confrontar* - determina se uma combinação de estímulos visuais mapeada como parte de uma imagem é um padrão similar a um pacote conhecido;
- *cobrir* - verifica se um esquema pode ser reforçado (explicado) por um pacote reconhecido em uma imagem sendo analisada. Em caso afirmativo, o fator de confiança sobre um esquema já selecionado é incrementado;
- *especificar* - especifica uma sugestão de interpretação que pode ser descrita por um esquema validado por pacotes encontrados na imagem sendo analisada;
- *compor* - processo de arranjar interpretações intermediárias alcançadas em uma completa coerência, formando uma interpretação final sugerida.

Para complementar a especificação operacional do PSM Interpretação de Rochas, uma especificação de controle para este método é apresentada na Figura 4.2.

A descrição da Figura 4.2 é uma das especificações de controle que podem ser derivadas da estrutura de inferência do PSM Interpretação de Rochas. Nesta especificação, externamente é realizada uma análise de uma imagem de amostra de rocha perante os esquemas de rochas conhecidos pelo especialista. Centralmente a esse laço, uma busca sistemática de pacotes visuais conhecidos na imagem da rocha é realizada, de forma a permitir reforçar (*cobrir*) a certeza sobre algum esquema de rocha selecionado para análise.

Tal como foi demonstrado, o processo de interpretação de rochas não tem esquemas de rochas selecionados inicialmente (*esquemasSelecionados := ∅*), pois normalmente os pacotes visuais mais relevantes para diferentes interpretações são inicialmente selecionados pelo especialista. Em seguida, um processo de casamento de padrões entre partes da imagem da rocha e estes pacotes visuais mais diagnósticos é executado (*confrontar*). Esse processo de casamento de padrões é realizado enquanto não for aleatoriamente encontrado um pacote visual relevante para selecionar um esquema de rocha em especial.

```

psm interpretaçãoDeRochas

controlEspecification:
  begin
    esquemasSelecionados := ∅;
    selecionar(pacotes + esquemasSelecionados → pacotesSelecionados);
    repeat
      decompor(imagem → estímulosSensórios);
      combinar(estímulosSensórios → parteImagem);
      selecionar(pacotesSelecionados → pacote);
      confrontar(parteImagem + pacote → pacoteAtivado);
      if ExistePacote(pacoteAtivado)
        then
          begin
            selecionar(esquemas + pacoteAtivado → esquemasSelecionados);
            selecionar(pacotes + esquemasSelecionados → pacotesSelecionados);
            repeat
              decompor(imagem → estímulosSensórios);
              combinar(estímulosSensórios → parteImagem);
              selecionar(pacotesSelecionados → pacote);
              confrontar(parteImagem + pacote → pacoteAtivado);
              if ExistePacote(pacoteAtivado)
                then cobrir(pacoteAtivado → esquemasSelecionados);
            until PacotesAnalisados(pacotesSelecionados);
          end then;
        until ImagemAnalisada(imagem) or EsquemasAnalisados(esquemas);
      especificar(esquemas → soluções);
      compor(soluções → interpretação);
    end controlEspecification
  end psm.

```

FIGURA 4.2 - Especificação de controle do PSM Interpretação de Rochas

Caso ocorra um casamento de padrões relevante para análise, um predicado denominado *ExistePacote* retorna o valor *verdadeiro*. Dessa forma, o pacote visual encontrado na amostra permite selecionar um esquema de rocha (*esquemasSelecionados*) que então seleciona um subconjunto de pacotes visuais (*pacotesSelecionados*) relacionados a ele. O predicado *PacotesAnalisados* é usado para indicar uma análise sobre todos os pacotes visuais selecionados a partir do esquema de rocha considerado como hipótese. *PacotesAnalisados* retorna *verdadeiro* quando esses pacotes visuais forem completamente buscados na amostra de rocha.

Como foi afirmado, o processo de raciocínio desenvolvido é repetido até que toda as partes de uma imagem de uma amostra de rocha (*parteImagem*) tenham sido avaliadas pelo especialista, no caso dessa especificação de controle: o predicado *ImagemAnalisada* retornando o valor *verdadeiro*, ou todos os esquemas de rochas conhecidos (*esquemas*) pelo especialista tenham sido analisados, neste caso: o predicado *EsquemasAnalisados* retornando o valor *verdadeiro*. Ao final da especificação de controle modelada, diferentes soluções podem ser especificadas a partir de esquemas de rochas indicados e validados por pacotes visuais encontrados na amostra analisada. Se houver conflito entre as soluções encontradas, tais conflitos são

resolvidos durante a composição das soluções em uma completa coerência resultante na interpretação.

A especificação de controle apresentada reforça a idéia que o raciocínio desenvolvido para interpretar rochas pode ser explicado por casamentos de padrões consecutivos, os quais são dirigidos por esquemas de rochas e pacotes visuais conhecidos. O resultado final desse processo de raciocínio é uma interpretação derivada de esquemas de rochas, os quais foram devidamente alcançados e reforçados por pacotes visuais encontrados pelo especialista na amostra de rocha analisada.

4.1.3 Um esquema de domínio típico do PSM Interpretação de Rochas

O esquema de domínio típico apresentado na Figura 4.3 descreve os principais tipos e propriedades do conhecimento de domínio requerido pelo PSM Interpretação de Rochas. Esta descrição pode ajudar na compreensão do método proposto, como também pode orientar a aquisição de conhecimento de domínio em outro domínio de aplicação que venha a reutilizar este PSM. O modelo descrito na Figura 4.3 é apresentado em um formalismo simples, tal como um modelo de objetos.

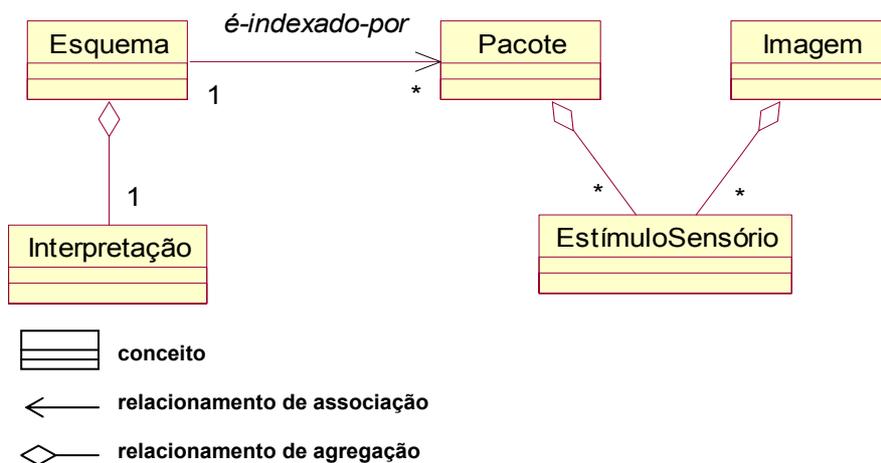


FIGURA 4.3 - Esquema de domínio típico do PSM Interpretação de Rochas

O esquema de domínio do PSM Interpretação de Rochas indica que o conhecimento de domínio requerido pelo método deve abordar diferentes conceitos, tais como: *esquema*, *interpretação*, *pacote*, *imagem*, *estímulo sensório*. Para suportar o processo de raciocínio especificado, um esquema deve ser composto por uma interpretação, além de ser *indexado-por* múltiplos pacotes. Tais associações entre esquemas e pacotes são requeridas pelos passos de *múltipla seleção* especificados no PSM Interpretação de Rochas. Por sua vez, esquemas compostos por interpretações são requeridos pelo passo de especificação de soluções e composição de interpretações, pois esses passos resultam nas interpretações componentes de cada esquema. Por fim, pacotes e imagens devem ser formados por estímulos sensórios similares, os quais são combinados e recombinaados no sentido de formar diferentes pacotes e imagens a serem confrontados.

4.1.4 Alguns requisitos/suposições do PSM Interpretação de Rochas

As suposições estabelecidas foram descritas no *nível do conhecimento* (Newell [NEW 82]), de forma a descrever os principais tipos de conhecimento requeridos pelo PSM Interpretação de Rochas, bem como caracterizar a funcionalidade do método em relação a um objetivo a ser alcançado. Como pode ser percebido, o PSM é descrito em função de componentes conceituais identificados no modelo de perícia proposto por Abel [ABE 2001]. Desta forma, sua funcionalidade é bastante fortalecida, pois é considerado que o conhecimento de domínio requerido pelo método está disponível, assim expressando um PSM mais competente para atender as tarefas de raciocínio do domínio de problemas da Petrografia Sedimentar.

A Figura 4.4 apresenta os principais tipos de conhecimento requeridos pelo PSM Interpretação de Rochas.

O PSM Interpretação de Rochas é capaz de alcançar os objetivos almejados desde que seja disponibilizado conhecimento de domínio requerido para viabilizar os passos de inferência especificados. Desta forma, tais passos foram usados para identificar as suposições deste método. A Figura 4.4 apresenta algumas indicações sobre esses tipos de conhecimento de domínio requeridos, os quais são apresentados mais detalhadamente a seguir:

(1) *Seleção/abstração de estímulos sensoriais* - esta suposição viabiliza o reconhecimento visual de diferentes combinações de feições geológicas, as quais são relevantes para a análise de uma imagem de rocha. Pode ser percebido que o especialista decompõe aspectos relevantes de uma imagem de rocha, de forma a filtrar aspectos não importantes para a interpretação. Em seguida, ele combina e recombina estes estímulos sensoriais de diferentes formas, até alcançar uma parte da imagem relevante para análise, assim confrontando-a com aspectos visuais conhecidos. Conforme Abel [ABE 2001], esses passos de reconhecimento visual podem ser explicados por conhecimento sobre seleção/abstração de estímulos sensoriais, o qual varia segundo diferentes ordenações posicionais de feições geológicas visuais em uma amostra de rocha analisada em microscópio.

(2) *Associações de estímulos sensoriais em termos de objeto-atributo-valor* - esta suposição indica que uma porção de imagem de rocha e um pacote visual conhecido podem ser confrontados, desde que sejam formados por componentes decompostos em granularidades "semelhantes". Para realizar esse passo de inferência em computadores, o PSM requer algum conhecimento que viabilize a visualização de feições de rocha em termos de *objetos-atributos-valores*, os quais podem ser confrontados com pacotes visuais conhecidos. Tal casamento de padrões é realizado por especialistas em Petrografia até encontrar um casamento relevante para análise, resultando em um pacote visual ativado. A ocorrência de um casamento aceitável então indica o reconhecimento de aspectos de rocha conhecidos, os quais podem selecionar um esquema de rocha com interpretações a serem analisadas.

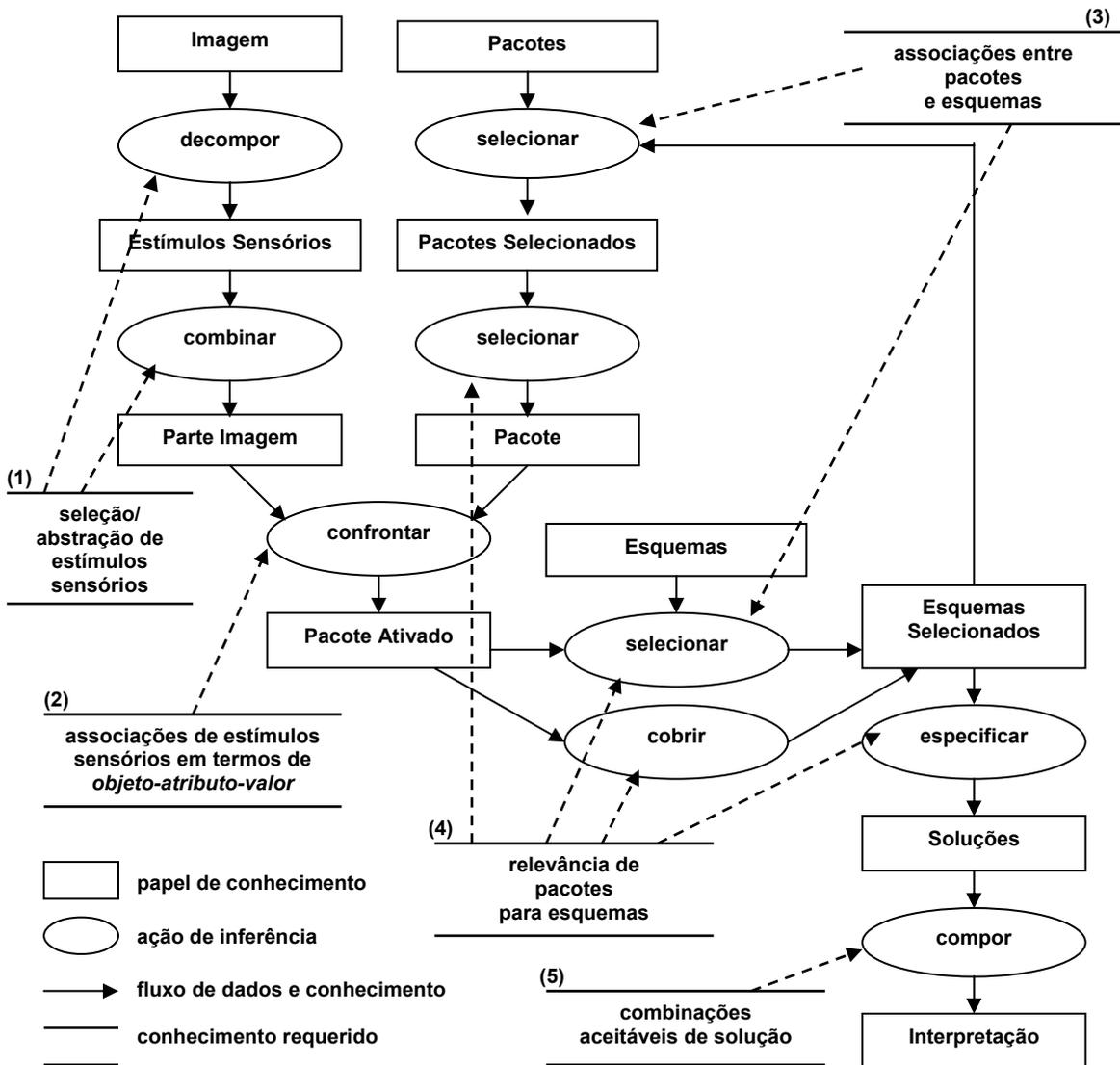


FIGURA 4.4 - Conhecimento de domínio requerido pelo passo de inferência do PSM Interpretação de Rochas

(3) *Associações entre pacotes e esquemas* - segundo Abel [ABE 2001], o processo de raciocínio de interpretação de rochas desenvolvido por especialistas pode ser explicado por associações entre pacotes visuais e esquemas de rochas, ou seja, em função destes tipos de conhecimento. Durante o raciocínio, diferentes pacotes visuais encontrados em uma amostra selecionam esquemas de rochas particulares a serem analisados, e esses esquemas de rochas selecionam outros pacotes visuais conhecidos relacionados a eles. Estes pacotes conhecidos são então buscados na amostra de rocha, assim tentando validar um esquema selecionado considerado como hipótese. Nesse processo de múltipla seleção, percebe-se que quanto mais e melhores forem as associações entre esquemas e pacotes, mais complexos serão os processos de raciocínio desenvolvidos pelo especialista. Conforme Abel [ABE 2001], o raciocínio desenvolvido para interpretar rochas sedimentares clásticas é tanto mais eficaz quanto mais indexados forem pacotes e esquemas. Tal suposição também simplifica fortemente a complexidade envolvida no processo de raciocínio de interpretação de rochas.

Como proposto por Abel [ABE 99], a estratégia de raciocínio desenvolvida por especialistas em Petrografia Sedimentar é baseada em *processos de casamentos de padrões visuais*, e apenas secundariamente em processos de busca de soluções. A partir disso, pode ser afirmado que o PSM Interpretação de Rochas seja competente para resolver tarefas de raciocínio onde o casamento de padrões visuais é predominante. Entretanto, processos de busca de soluções podem ser identificáveis, e são tanto mais notados quanto menores os níveis de perícia do especialista que realiza a interpretação.

Expressando um processo dirigido por esquemas de rochas e pacotes visuais conhecidos, o PSM Interpretação de Rochas pode ser sintetizado e centralmente explicado como um *processo de raciocínio progressivo de um passo*: de pacotes visuais diretamente para esquemas de rochas, e vice-versa. Neste caso, o PSM proposto não expressa um encadeamento de hipóteses, tal como representado em uma rede de inferência (Tanimoto [TAM 95]). Todavia, o PSM Interpretação de Rochas modela passos únicos de solução de pacotes para interpretações, os quais são suportados por múltiplas seleções e um casamento de padrões central.

O processo de interpretação de rochas modelado é fundamentado em esquemas mantidos pelo especialista. Além disso, existem fortes indícios que esses esquemas sejam organizados em hierarquias de classes e subclasses de conceitos com múltiplos relacionamentos, os quais envolveriam diferentes categorias de rochas, processos diagenéticos, feições geológicas, constituintes minerais e outros aspectos geológicos. Como salientado por Abel [ABE 98], uma solução para esses problemas de interpretação é tanto melhor e mais rapidamente pode ser alcançada quanto mais detalhados e interrelacionados forem os esquemas do especialista.

Todavia, as diferentes hierarquias de rochas que poderiam ser modeladas como esquemas foram simplificadas em apenas uma classe de rocha no projeto *PetroGrapher*: a classe das rochas sedimentares clásticas. Mesmo sem requerer conhecimento hierárquico sobre esquemas de rochas, o PSM Interpretação de Rochas pode representar casos práticos de interpretação descritos por especialistas. Como proposto por Fensel et al. [FEN 98a], tal suposição simplifica os objetivos da tarefa de interpretação de rochas, assim viabilizando sua modelagem e solução computacional.

Caso o projeto *PetroGrapher* venha a ser ampliado para outras classes de rochas (tal como rochas carbonáticas), o PSM Interpretação de Rochas pode ser adaptado para esquemas de rochas organizados segundo diferentes níveis hierárquicos. Tais adaptações poderiam ser expressas por refinamentos na estrutura dos passos de *seleção de esquemas de rocha*, bem como *especificação de soluções*.

(4) *Relevância de pacotes para esquemas* - o conhecimento sobre relevância de pacotes visuais para esquemas de rocha é empregado por especialistas da Petrografia em diversos pontos do processo de raciocínio modelado. Este processo de raciocínio supõe que tais índices de relevância de pacotes visuais para esquemas de rocha sejam efetivamente usados pelo especialista, assim estabelecendo um foco para o raciocínio, bem como aumentando sua eficiência. Tal indicação é usada de diferentes formas: 1) para diminuir o espaço de busca de soluções pela seleção de pacotes mais diagnósticos a serem analisados, ou pela simples seleção de pacotes

mais relevantes para diferentes esquemas, 2) para selecionar um esquema relevante para análise e reforçar a certeza sobre esquemas selecionados sendo analisados e 3) para fundamentar a confiabilidade de uma solução especificada a partir de um esquema validado por pacotes encontrados na amostra de rocha.

Conforme a análise da perícia em interpretação de rochas (Abel [ABE 2001]), os índices de relevância para cada pacote visual indicativo de uma interpretação de rocha são variáveis. Em muitos casos, a significância de cada pacote visual é atribuída pelo especialista quando ele realiza a primeira análise da amostra de rocha. Além disso, tal índice pode aumentar ou diminuir durante o raciocínio.

Como o tratamento de variações sobre índices de relevância de pacotes visuais é bastante complexo em termos de aquisição de conhecimento e representação computacional, as possibilidades de variação não foram explicitamente modeladas no PSM Interpretação de Rochas proposto. No entanto, o conhecimento sobre índices de relevância variáveis é requerido pelos passos de seleção de pacotes conhecidos, seleção e reforço de esquemas e especificação de soluções, os quais foram expressos no PSM Interpretação de Rochas.

(5) *Combinações aceitáveis de solução* - como uma rocha pode passar por diferentes processos durante a sua formação, o PSM Interpretação de Rochas supõe que esteja disponível o conhecimento de domínio requerido para avaliar a composição de diferentes soluções intermediárias muito comuns em uma interpretação de rocha final. Como citado por especialistas em Petrografia, freqüentemente é necessário compor diferentes interpretações de rochas individuais em uma completa coerência. Tal coerência pode ser alcançada pelo uso de conhecimento geral de domínio sobre processos geológicos, o qual pode eliminar dúvidas e complementar a falta de informação durante uma interpretação, bem como resolver possíveis conflitos na apresentação de um resultado final.

Tal suposição, portanto, refina a competência do PSM Interpretação de Rochas no sentido de alcançar *múltiplas soluções parciais*, sem haver conflito entre elas. A possibilidade de múltiplas soluções foi modelada no método como um passo de especificação de soluções seguido de uma composição destas soluções em uma interpretação, assim definindo *uma estrutura de informação* necessária para derivar uma solução, tal como proposto por Wielinga et al. [WIE 98].

Conforme especialistas em Petrografia entrevistados, interpretações de rochas confiáveis normalmente são fundamentadas na quantidade e relevância de instâncias de pacotes visuais encontradas em uma amostra de rocha. Quanto maior número e mais relevantes forem os pacotes visuais encontrados (pacotes ativados), mais confiável é uma interpretação particular derivável a partir destes pacotes. No PSM especificado, pacotes ativados reforçam gradativamente a confiança sobre esquemas selecionados. A partir de esquemas então validados, uma interpretação confiável pode ser especificada.

Em síntese, as suposições descritas para o PSM Interpretação de Rochas permitiram: 1) apresentar uma explicação para uma análise do raciocínio desenvolvido por um especialista; 2) considerar o raciocínio analisado em termos de modelos abstratos; 3) simplificar a complexidade da modelagem, sem inviabilizar a

solução do problema; 4) identificar possibilidades de refinamento do PSM especificado; 5) caracterizar mais precisamente a competência do PSM; 6) identificar aspectos de eficiência associados ao uso de conhecimento de domínio e 7) caracterizar as soluções deriváveis pelo método.

4.2 Um modelo de conhecimento de domínio

O modelo de conhecimento de domínio para a descrição e interpretação de rochas desenvolvido empregou níveis de abstração distintos. Em princípio, o conhecimento do domínio foi coletado como *casos* (Kolodner [KOL 93]). Segundo Abel et al. [ABE 95], casos nesse domínio de problemas descrevem ocorrências geológicas particulares e suas possíveis interpretações, representando o conhecimento que o petrógrafo pode recuperar após algum tempo que uma experiência particular de descrição e interpretação tenha ocorrido.

Dentre outras coisas, a abordagem de casos demonstra uma estrutura básica de organização que abstrai o conhecimento declarativo do domínio, além de permitir identificar os principais conceitos de domínio utilizados por especialistas. A Figura 4.5 apresenta as partes de um caso de amostra de rocha, assim descrevendo os principais aspectos geológicos relacionados ao processo de descrição e interpretação de rochas.

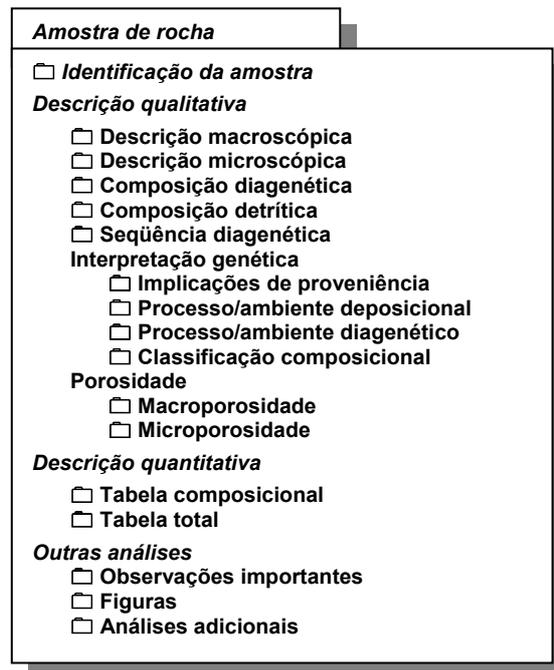


FIGURA 4.5 - Partes de um caso de descrição de amostra de rocha

A Figura 4.5 indica que um caso de descrição de amostra de rocha é composta por uma *identificação da amostra*, uma *descrição qualitativa*, uma *descrição quantitativa* e *outras análises*. Cada caso é composto por múltiplos atributos simbólicos, muitas vezes multivalorados, onde o especialista utiliza diferentes conjuntos de conceitos de domínio entre descrições. Além disso, um caso é uma

descrição normalmente incompleta, pois o petrógrafo só descreve o que é relevante para atender os objetivos da descrição particular sendo realizada.

A abordagem de casos é fundamental para caracterizar o esquema de raciocínio para interpretação de rochas sendo proposto. Todavia, essa abordagem foi apenas utilizada para modelar o domínio, mas não foi empregada para derivar uma especificação inferencial para este problema. Neste caso, a abordagem de Raciocínio Baseado em Casos (Kolodner [KOL 93]) não foi empregada porque não estavam disponíveis descrições representativas de casos de problemas resolvidos, as quais são requeridas para orientar um mecanismo de raciocínio baseado em algoritmos de similaridades. Logo, uma abordagem baseada em modelos mais tradicional foi empregada para representar o conhecimento sobre soluções de problemas de interpretação de rochas.

Uma descrição mais abrangente da aplicação da técnica de casos nesse domínio de problemas está fora do escopo desse trabalho, pois foi desenvolvido em passos anteriores do projeto *PetroGrapher*. Mesmo assim, Abel et al. [ABE 95] apresenta uma avaliação da técnica de Raciocínio Baseado em Casos em problemas da Petrografia Sedimentar.

Em um passo seguinte de especificação do modelo de conhecimento de domínio desenvolvido, as principais partes de um caso de amostra de rocha foram representadas como um modelo de frames (Fikes et al. [FIK 85]). O formalismo de frames foi escolhido por ser uma representação de conhecimento de alto-nível, onde é possível expressar e detalhar as principais divisões semânticas desse domínio de aplicação.

Cada frame modelado é composto por feições geológicas típicas, as quais foram agrupadas segundo a forma de descrição de amostras usada pelo especialista (ou seja, onde os valores são obtidos em um mesmo momento de descrição). Os principais frames modelados são apresentados na Figura 4.6 usando um formalismo simples, tal como um modelo de objetos.

O modelo de conhecimento de domínio representado como frames é composto principalmente pelo *frame amostra* e suas partes identificáveis em um caso de descrição de amostra de rocha. Como uma amostra é composta por diferentes tipos de descrições, estes tipos foram representados como conceitos de domínio expressos como frames, tais como, por exemplo: *frame identificação*, *frame composição-diagenética*, *frame macroporosidade*, etc.

Os frames modelados foram detalhados sucessivamente durante o transcorrer do projeto de pesquisa *PetroGrapher*, e também durante o desenvolvimento desse trabalho, assim buscando atender os requisitos de descrição e interpretação de rochas requeridos por especialistas em Petrografia. O modelo de domínio resultante pode ser considerado bastante refinado, pois foi cuidadosamente orientado para permitir a descrição de amostras de rocha em diferentes níveis de granularidade, desde descrições bastante precisas até descrições mais sucintas e abrangentes. Para isso, grande quantidade de conhecimento de domínio e tarefa foi exaustivamente adquirida e especificada em uma nomenclatura de feições geológicas considerada padronizada pelo especialista em Petrografia. Contudo, conforme diferentes especialistas em

Petrografia, usando os conceitos de domínio explicitamente detalhados no modelo de frames é possível descrever uma amostra de rocha sedimentar clástica correta e completamente.

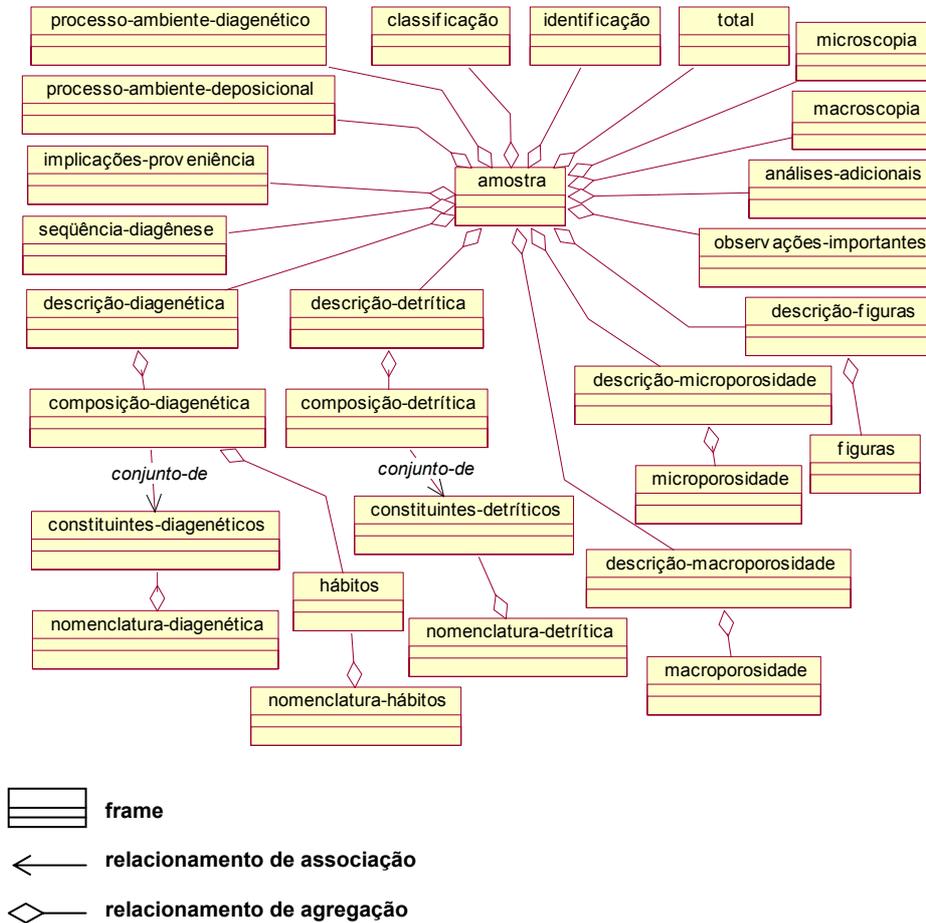


FIGURA 4.6 - Modelo de frames do domínio

O completo detalhamento dos frames especificados pode ser analisado no Anexo 1. Mesmo assim, um exemplo desse refinamento é apresentado na Tabela 4.1.

O *frame composição-diagenética* é composto pelos atributos: *é-um*, *parte-de* e *conjunto-de*, os quais indicam o relacionamento deste frame com outros frames modelados. Os demais atributos descrevem aspectos geológicos típicos de uma descrição de composição diagenética de rocha, bem como seus respectivos domínios de valores permitidos (principais restrições de integridade).

Os domínios de valores especificados para os atributos do *frame composição-diagenética* são descritos em função de cadeias de caracteres, listas de valores simbólicos pré-estabelecidos, valores numéricos com intervalos de valores permitidos, etc. Para o esquema de interpretação de rochas proposto, tais restrições especificadas no modelo de frames detalham principalmente uma terminologia padronizada de feições geológicas adquirida junto a especialistas em Petrografia. Esta terminologia é usada como modelo básico para uma ferramenta de aquisição de instâncias de conceitos de domínio (sistema *PetroGrapher*), as quais são informadas por um

petrógrafo quando este descreve uma amostra de rocha. Assim como realizado no projeto Protégé-II (Grosso et al. [GRO 99]), o conjunto desses conceitos de domínio poderia ser considerado como uma *ontologia de domínio*, a qual poderia ser empregada para a construção automática de ferramentas de aquisição de conhecimento.

TABELA 4.1 - Um exemplo do modelo de frames: frame composição-diagenética

Frame composição-diagenética	
é-um	Objeto
parte-de	frame descrição-diagenética
conjunto-de	frame constituintes-diagenéticos
IDAmostra	string(20), um-de <identificação.IDAmostra>
nomeMineral	string(80), um-de [constituenteDiagenético]
conjuntoConstituinte	string(40), um-de [Silica, Feldspato, Argilas infiltradas, Argilas de pseudomatrix, Argilas autigênicas, Zeolitas, Carbonatos, Sulfatos, Sulfetos, ...]
hábito	string(20), um-de <hábitos.nomeHábito>
quantidade	real, Intervalo [0.0 - 100.00]
quantidadeNominal	string(20), um-de [abundante, comum, raro, traço]
localização	string(40), um-de [recobrimento de poro contínuo intergranular, recobrimento de poro descontínuo intergranular, preenchimento de poro intergranular, intergranular discreto, ...]
modificador	string(40), um-de [dissolvido, zonado, fraturado, recristalizado]
relaçãoParagenética	string(120), um-de [Cobrindo <um-de [constituenteDiagenético]>, Cobrindo <um-de [constituenteDetritico]>, Coberto por <um-de [constituenteDiagenético]>, Substituindo grão de <um-de [constituenteDetritico]>, Substituindo matriz de <um-de [constituenteDetritico]>, Substituindo <um-de [constituenteDiagenético]>, Substituído por <um-de [constituenteDiagenético]>, Alternado com <um-de [constituenteDiagenético]>, Envolvendo <um-de [constituenteDiagenético]>, ...]
relaçãoParagenética conjuntoConstituinte	string(40), um-de [Silica, Feldspato, Argilas infiltradas, Argilas de pseudomatrix, Argilas autigênicas, Zeolitas, Carbonatos, Sulfatos, Sulfetos, ...]
descrição	string(256)

Na maioria dos casos, os valores possíveis para conceitos de domínio representados no modelo de frames foram modelados como listas pré-definidas. O número de valores permitido para cada um desses conceitos é especificado como *facet*, tais como, por exemplo: *um-de*: indicando que o conceito pode conter apenas um valor dentre os valores especificados e *lista-de*: indicando que o atributo aceita uma lista de valores dentre os valores detalhados. Tal como foi especificado, o conhecimento inferencial de domínio modelado nesse trabalho como grafos de conhecimento (Leão [LEA 88]) não considera valores que não foram especificamente detalhados no modelo de frames.

4.2.1 Um modelo de conhecimento inferencial de domínio

O conhecimento inferencial de domínio usado para efetivamente sugerir interpretações de amostras de rocha foi adquirido e modelado como grafos. A idéia de adquirir e representar conhecimento utilizando um formalismo de representação

gráfico é bastante intuitiva, pois várias metodologias de aquisição e representação de conhecimento utilizam esta abordagem. Dentre elas, as metodologias mais difundidas são as *redes semânticas* e *árvores de decisão* (Turban [TUR 92]).

O formalismo de representação de grafos empregado nesse trabalho é baseado em uma metodologia de aquisição de conhecimento proposta por Leão [LEA 88]. Simplificadamente, esta metodologia envolve uma tentativa de melhor compreender os processos cognitivos envolvidos na construção de um diagnóstico médico. Para isso, tal metodologia utiliza a técnica de representação de grafos de conhecimento para adquirir, junto a especialistas, regras heurísticas utilizadas na prática diagnóstica médica.

Nesse trabalho, grafos de conhecimento representam relações de implicação entre conceitos de domínio. Tais implicações são descritas em função de pacotes visuais (considerados como evidências) para uma interpretação (considerada como solução). Pacotes visuais, assim como previstos no PSM Interpretação de Rochas, foram rotulados como um nome e modelados como *expressões de domínio*. Tais expressões são descritas em função de conceitos simplificados de domínio especificados no modelo de frames desenvolvido. Interpretações, por sua vez, são representadas com um nome, a qual indica um significado particular nesse domínio de problemas. Desta forma, os grafos modelados em função de expressões de domínio e interpretações desempenham o papel de relações causais requeridas pelos passos de raciocínio especificados, assim descrevendo as informações requeridas pelo raciocínio de forma estruturada e simplificada. Deve ser ressaltado que essa estratégia de modelagem usada é genérica, podendo ser aplicada na especificação de KBSs em diferentes domínios de aplicação.

Os grafos modelados também realizam um importante ajuste de granularidade ante os dados descritos por usuários e o conhecimento requerido pelos passos de inferência. Tais ajustes são expressos por associações lógicas de conceitos simplificados de domínio, os quais relacionados indicam significados particulares. Esses significados são modelados como expressões de domínio, as quais orientam o processo de inferência proposto. Portanto, *o processo de raciocínio modelado é desenvolvido em uma granularidade diferente do processo de descrição de uma amostra*. Para viabilizar esta abordagem inferencial, o modelo de conhecimento de domínio desenvolvido foi refinado sucessivamente até detalhar explicitamente todas as feições petrográficas usadas para derivar uma interpretação de rocha.

As expressões de domínio modeladas (Anexo 2) representam combinações lógicas de conceitos de domínio do modelo de frames construído, desta forma representando os conceitos considerados essenciais e mais diagnósticos para o raciocínio. Cada expressão de domínio especificada tem um índice de relevância (ou significância) associado, e este foi representado como um peso. O peso de uma expressão de domínio expressa o quanto ela pode influenciar na confirmação de uma interpretação rotulada como conclusão em um grafo, como também usada como identificador do grafo.

De forma mais objetiva, os grafos modelados são formados pelos seguintes componentes:

- na raiz de cada grafo, é indicada uma interpretação e um limiar de conclusão desta solução. Durante o processo de inferência, o limiar é comparado com o somatório dos pesos das expressões de domínio encontradas em uma descrição de amostra e indicativas de uma mesma interpretação;
- em um primeiro nível de nodos de cada grafo, as diferentes expressões de domínio modeladas e seus respectivos pesos em relação a uma interpretação são representados;
- em níveis intermediários de nodos, cada expressão de domínio é descrita por combinações lógicas de conceitos-atributos-valores de domínio descritos no modelo de frames desenvolvido.

A idéia de representar pesos associados às evidências faz parte da proposta original da técnica de grafos de conhecimento (Leão [LEO 88]). No entanto, a representação de limiares de conclusão amplia levemente a proposta original, viabilizando especificar um índice de confiabilidade de solução requerido pelo processo de raciocínio desenvolvido nesse domínio de problemas. Como descrito por Abel [ABE 2001], os limiares de conclusão e pesos de expressões de domínio foram propostos pelo especialista a partir de múltiplas simulações de inferência realizadas em entrevistas de aquisição de conhecimento.

Os limiares identificados para todas as interpretações expressas nos grafos de conhecimento modelados atualmente tem valor igual a 6. Os pesos identificados para as expressões de domínio modeladas, por sua vez, utilizam apenas os valores 1, 3, 5 e 6. Estes valores de limiares e pesos adquiridos e modelados são típicos desta tarefa de interpretação de rochas e, desta forma, provavelmente não sejam usáveis em domínios similares a este. Contudo, a representação de pesos e limiares de conclusão pode ser bastante útil para expressar conhecimento heurístico em diferentes domínios de aplicação. A Tabela 4.2 apresenta o significado de cada um dos valores de pesos modelados.

TABELA 4.2 - Significado dos pesos modelados

Peso	Significado
1	interpretação possivelmente é verdadeira
3	interpretação provavelmente é verdadeira
5	interpretação muito provavelmente é verdadeira
6	interpretação é verdadeira

O conjunto de grafos de conhecimento modelados até o momento expressam o conhecimento de domínio usado para interpretar ambientes diagenéticos a partir de descrições de amostras de rocha sedimentares clásticas. Como já foi citado, um ambiente diagenético é expresso por um nome em um grafo de conhecimento, e a sugestão destes ambientes é suportada por um conjunto de expressões de domínio validadas/encontradas em uma descrição de amostra. Segundo o especialista em Petrografia, as origens de rocha que caracterizam ambientes diagenéticos mais

significativos foram modeladas, bem como as expressões de domínio mais relevantes para indicar esses ambientes.

Para tornar a compreensão desse modelo de conhecimento inferencial mais simplificada, as Figuras 4.7 e 4.8 apresentam um exemplo de partes de um grafo de conhecimento modelado.

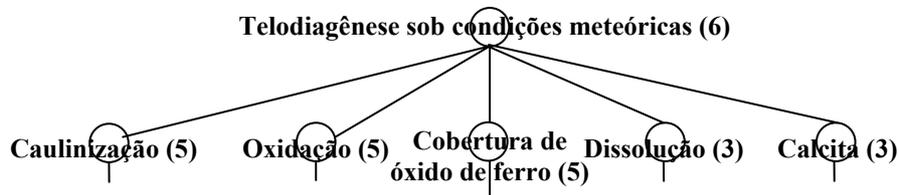


FIGURA 4.7 - Um exemplo de grafo de conhecimento: interpretação e expressões de domínio

O grafo da Figura 4.7 indica a relação entre as expressões de domínio: "Caulinização, Oxidação, Cobertura de óxido de ferro, Dissolução e Calcita", e uma interpretação de ambiente diagenético: "Telodiagênese sob condições meteóricas". Cada expressão individualmente é rotulada com um nome e tem um peso para uma conclusão, tal como, por exemplo: "Caulinização" tem um peso de 5 para sugerir um ambiente tipo "Telodiagênese sob condições meteóricas". Tal conclusão tem um limiar de 6, indicando que esse ambiente pode ser alcançado se a soma das expressões de domínio encontradas na descrição de uma amostra de rocha for igual ou maior ao limiar expresso no grafo. A seguir, a Figura 4.8 apresenta uma das expressões de domínio previstas no grafo da Figura 4.7.

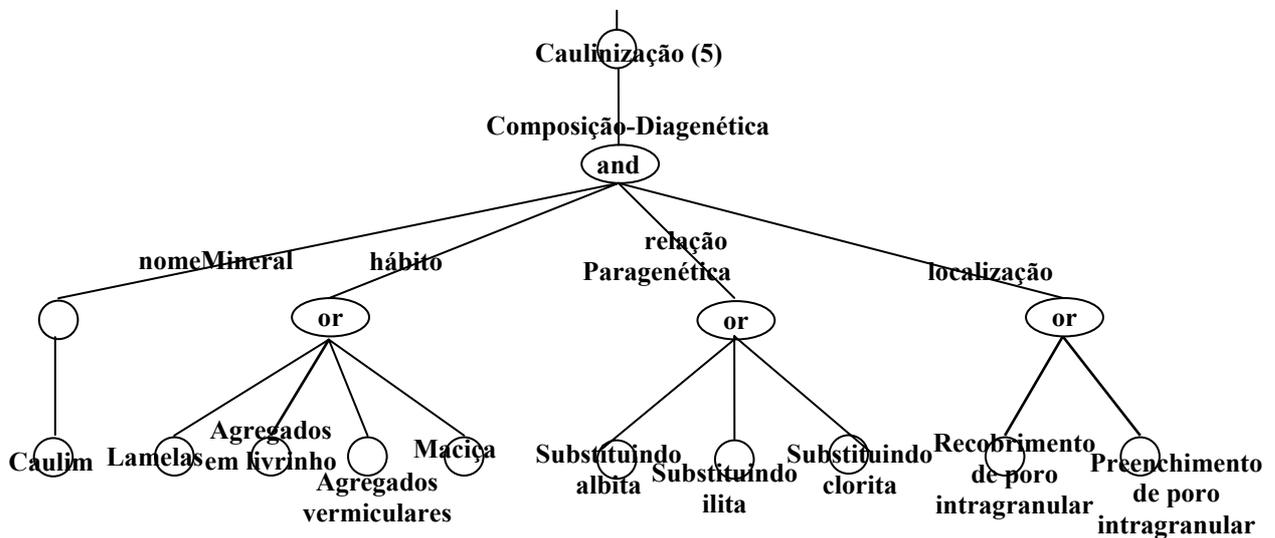


FIGURA 4.8 - Um exemplo de expressão de domínio modelada

O grafo da Figura 4.8 descreve uma expressão de domínio denominada "Caulinização". Tal expressão é descrita em termos de *conceitos-atributos-valores* do modelo de frames desenvolvido (seção 4.2). Por exemplo, "Caulinização" é composta pelo conceito "Composição-Diagenética" (representado como um frame composição-diagenética), seus atributos e respectivos valores esperados: "nomeMineral = Caulim" AND "hábito = (Lamelas OR Agregados em livrinho ...)" AND "relaçãoParagenética

= (Substituindo albita OR Substituindo ilita ...) AND "localização = (Recobrimento de poro intragranular OR Preenchimento de poro intragranular)".

Como pode ser notado no exemplo, uma expressão de domínio é composta por diferentes combinações lógicas envolvendo *conceitos-atributos-valores*. Tais combinações muitas vezes envolvem relações AND entre atributos e OR entre valores destes atributos. Por exemplo, relações OR entre os valores de "hábito", tais como: "Lamelas, Agregados em livrinho, Agregados vermiculares, Maciça".

Nenhuma expressão de domínio adquirida e modelada expressa relações OR entre conceitos de domínio, tal como, por exemplo: uma relação OR entre "composição-diagenética" e "microscopia". A ocorrência desse tipo de relação entre conceitos de domínio é representada como expressões de domínio alternativas, as quais são tratadas individualmente. Entretanto, relações AND entre conceitos de domínio são permitidas. Um exemplo deste caso é apresentado na Figura 4.9.

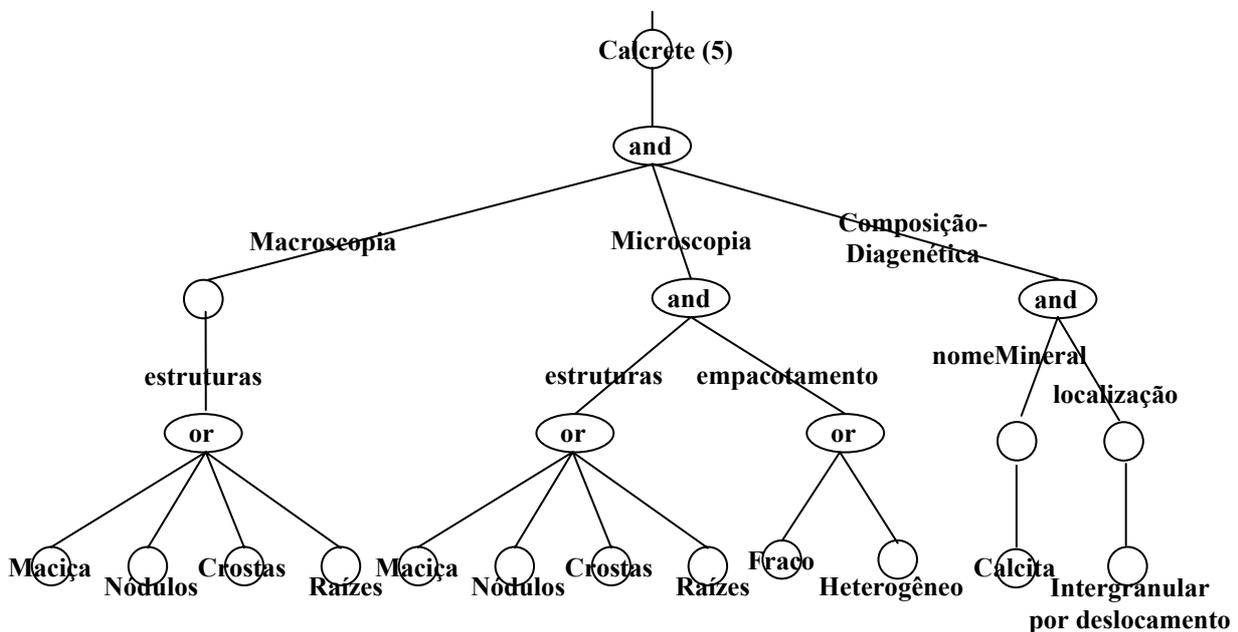


FIGURA 4.9 - Exemplo de expressão de domínio modelada

A expressão "Calcrete" da Figura 4.9 indica "Eodiagênese continental meteórica sob clima seco". Tal expressão envolve diferentes conceitos de domínio modelados, sendo representados por relações AND entre os conceitos "Macroscopia", "Microscopia" e "Composição-Diagenética". Para cada um destes conceitos de domínio, os atributos e valores que devem ser considerados são representados em subgrafos deste grafo de conhecimento.

Nesse trabalho, diferentes tipos de interpretações foram modeladas, indicando ambientes diagenéticos distintos sugeridos por diferentes expressões de domínio modeladas. A Tabela 4.3 apresenta um resumo desse conhecimento inferencial de domínio detalhado.

TABELA 4.3 - Grafos de conhecimento e expressões de domínio modelados

Ambientes Diagenéticos	Limiar	Expressões de Domínio	Pesos
Diagênese de soterramento profundo (Mesodiagênese profunda)	6	Estilolitos	6
		Albita	3
		Ilita	5
		Ilitização	5
		Compactação intensiva em arcabouço rígido	6
		Ilitização de feldspato	3
Eodiagênese Continental Meteórica sob Clima Seco	6	Óxido de ferro/hidróxidos	1
		Argilas infiltradas	3
		Calcrete	5
		Dolomita	3
		Silcrete	6
		Sulfato	6
		Dolocrete	5
Eodiagênese sob Condições Marinhas	6	Glauconita	5
		Bertierina/chamosita	3
		Pirita	3
		Carbonato	3
		Fosfato	5
		Matriz de glauconita	3
		S/Nome	3
Eodiagênese Precoce sob Mistura de Águas Marinhas e Meteóricas	6	Dolomita	5
		Bertierina/chamosita	3
Telodiagênese sob Condições Meteóricas	6	Dissolução	3
		Caulinização	5
		Oxidação	5
		Cobertura de óxido de ferro	5
		Calcita	3
Eodiagênese Continental Meteórica sob Clima Úmido	6	Cimento de caulinita	3
		Caulinização por substituição	5
		Caulinização por deslocamento	6
		Siderita	3
		Óxidos de ferro/hidróxidos (Lateritas/Ferricretes)	1
		Caulinização da matriz	3

A quinta entrada da Tabela 4.3, por exemplo, pode ser compreendida da seguinte forma: um ambiente diagenético do tipo "Telodiagênese sob condições meteóricas" tem limiar de conclusão igual a 6. Este ambiente é sugerido com um peso igual a 3 associado a uma expressão de domínio denominada "Dissolução". Os demais ambientes, limiares, expressões e pesos modelados podem ser compreendidos de forma semelhante. A completa descrição dos grafos de conhecimento modelados é apresentada no Anexo 2.

A simplicidade do formalismo de representação do modelo de grafos de conhecimento significativamente auxiliou no processo de aquisição e validação de conhecimento inferencial junto ao especialista. Além disso, o formalismo de grafos não obscureceu a semântica desse tipo de conhecimento, tal como normalmente seria obscurecida se fosse representada por inúmeras regras de produção. Por fim, tal formalismo de representação permitiu adquirir e representar conhecimento perito

requerido para interpretar ambientes diagenéticos em amostras de rocha sedimentares clásticas. A partir desse modelo de conhecimento de domínio detalhado, o PSM Interpretação de Ambientes Diagenéticos foi especificado.

4.3 O PSM Interpretação de Ambientes Diagenéticos

No sentido de expressar um método de raciocínio que considerasse mais objetivamente os modelos de conhecimento de domínio desenvolvidos, foi especificado o PSM Interpretação de Ambientes Diagenéticos. Este PSM é uma representação mais operacional do processo de raciocínio descrito como PSM Interpretação de Rochas, mesmo sendo ambos modelados a partir da análise da *tarefa de interpretação de processo/ambiente diagenético*. Para detalhar esse processo, a Figura 4.10 apresenta a estrutura de inferência do PSM Interpretação de Ambientes Diagenéticos.

O processo de raciocínio descrito na Figura 4.10 tem início com a descrição dos dados de uma amostra de rocha a ser interpretada. Tal descrição é um caso (1) para o PSM Interpretação de Ambientes Diagenéticos, e este caso é o objeto sob o qual o processo de inferência será realizado. O caso é então abstraído (2) segundo as principais divisões semânticas identificáveis em diferentes tipos de descrições. Ao final dessa fase inicial de inferência, os conceitos de domínio mais diagnósticos previamente especificados também são mapeados para expressões de domínio (4), e estas dirigem a primeira parte do raciocínio.

Na primeira parte do raciocínio, uma expressão de domínio (5) qualquer é selecionada de um conjunto de expressões de domínio conhecido (4), até que todas as expressões tenham sido selecionadas para análise. Essa expressão selecionada é então usada para selecionar conceitos-atributos-valores mais diagnósticos (3) de um caso previamente abstraído (2). A expressão de domínio selecionada (5) também é decomposta em seus conceitos-atributos-valores componentes (6), onde estes serão sucessivamente confrontados com conceitos-atributos-valores selecionados no caso (3). Se os conceitos-atributos-valores selecionados (3) e conceitos-atributos-valores decompostos (6) forem similares, tais objetos são considerados ativados (7), pois foram encontrados na descrição em análise.

Os conceitos-atributos-valores ativados são compostos para assim indicar ou não expressões de domínio ativadas (8). As expressões ativadas são combinadas para formar ou não um conjunto de expressões de domínio encontradas na descrição (9). A partir disso, são especificadas as soluções que podem ser alcançadas pelas expressões de domínio encontradas e pelas normas de interpretação conhecidas.

Na segunda parte do raciocínio, os grafos de conhecimento modelados (tal como hipóteses) são normas (10), e estas são selecionadas repetidamente até que todas sejam analisadas. Para especificar soluções (12), são usados índices de relevância entre expressões de domínio e interpretações, os quais são descritos nas normas (10). O processo de especificação de soluções é repetido para todas as normas (10), onde cada interpretação particular (detalhada em uma norma selecionada) pode ser especificada como uma solução distinta (12), desde que as expressões de domínio encontradas (9) sejam suficientes (ou significativamente relevantes) para resultar na

interpretação. Por fim, como mais de uma interpretação particular pode ser encontrada, estas são combinadas em uma descrição que abrange todas as soluções (12) que foram alcançadas (13) a partir dos dados descritos em um caso (1) de descrição de amostra de rocha.

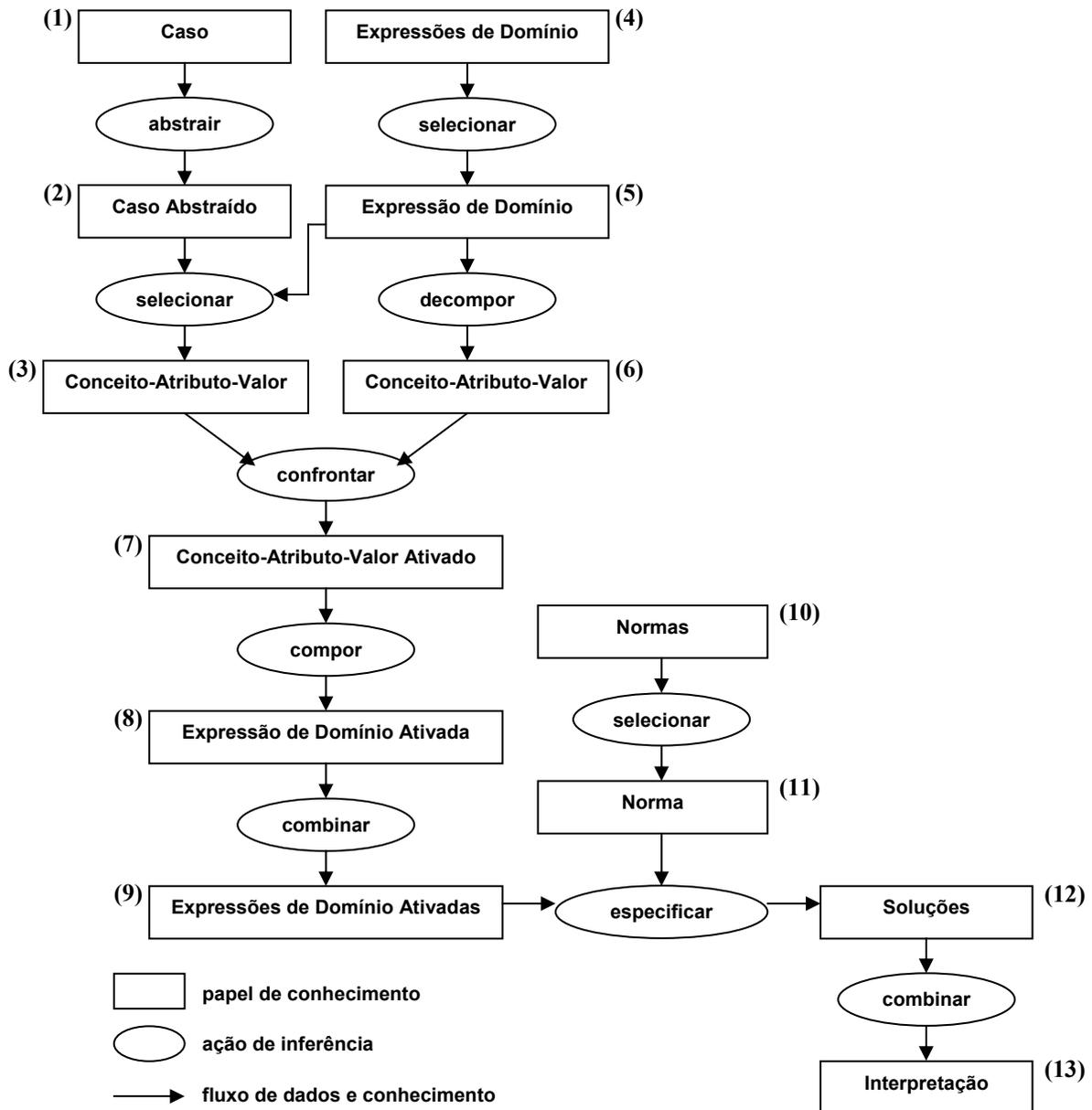


FIGURA 4.10 - PSM Interpretação de Ambientes Diagenéticos

Para caracterizar melhor o processo de inferência, um exemplo do PSM Interpretação de Ambientes Diagenéticos é apresentado usando conceitos do domínio detalhados como grafos de conhecimento e frames (seção 4.2).

Em princípio, uma "descrição de amostra de rocha" é um caso (1), o qual é abstraído (2) por suas principais divisões semânticas, tais como, por exemplo: "identificação, macroscopia, microscopia, ...". Em paralelo, todas as expressões de domínio modeladas, tais como, por exemplo: "calcrete, sulfato, dissolução,

caulinização, ..." são expressões de domínio (4). Uma dessas expressões de domínio (4) é selecionada aleatoriamente. Por exemplo: a expressão "caulinização" é selecionada para análise. Em seguida, esta expressão (5) é usada para selecionar conceitos-atributos-valores (3) descritos na "identificação, macroscopia, microscopia, ..." da amostra. Por exemplo, selecionando feições de rocha descritas para "composiçãoDiagenética.relaçãoParagenética, composiçãoDiagenética.localização, composiçãoDiagenética.nomeMineral, ...". Em paralelo, a expressão de domínio "caulinização" é decomposta em seus conceitos-atributos-valores (8) componentes, tais como, por exemplo:

"composiçãoDiagenética.relaçãoParagenética =
 Substituindo <constituenteDiagenético> AND
 composiçãoDiagenética.localização =
 Recobrimento de poro intragranular AND
 composiçãoDiagenética.nomeMineral = Caulim AND
 (composiçãoDiagenética.conjuntoConstituenteRelaçãoParagenética =
 Albita OR
 composiçãoDiagenética.conjuntoConstituenteRelaçãoParagenética =
 Ilita OR ...) AND ...".

Esses padrões de conceito-atributo-valor (3) e (6) são confrontados sucessivamente até que todos tenham sido analisados individualmente. Caso sejam similares, eles são considerados como ativados (9), assim resultando apenas aqueles padrões similares, por exemplo:

"composiçãoDiagenética.relaçãoParagenética =
 Substituindo <constituenteDiagenético> AND
 composiçãoDiagenética.localização =
 Recobrimento de poro intragranular AND
 composiçãoDiagenética.nomeMineral = Caulim AND
 composiçãoDiagenética.conjuntoConstituenteRelaçãoParagenética =
 Albita AND ..."

Tais conceitos-atributos-valores ativados (7) podem representar ou não uma expressão de domínio ativada (8), tal como, por exemplo: uma expressão "caulinização". Como todas as expressões de domínio (6) serão avaliadas durante o processo de inferência, mais de uma pode ser validada na descrição da amostra, resultando em expressões de domínio ativadas (9), ou seja, reconhecidas, tais como, por exemplo: "caulinização, oxidação, sulfato, ...".

Ao final do processo de análise das expressões de domínio (4), um conjunto de grafos de conhecimento, tais como, por exemplo: "Telodiagênese sob condições meteóricas (limiar: 6) indicada por: dissolução (peso: 3), caulinização (peso: 5), oxidação (peso: 5) ...", "Eodiagênese meteórica continental sob clima úmido (limiar: 6) indicada por: cimento de caulinita (peso: 3), caulinização por substituição (peso: 5), ...", etc, são normas (10) de interpretação. Esses grafos (4) são selecionados aleatoriamente para análise, até que todos tenham sido analisados.

Supondo que tenha sido selecionado o grafo, por exemplo: "Telodiagênese sobre condições meteóricas (limiar: 6) indicada por: dissolução (peso: 3), caulinização

(peso: 5), oxidação (peso: 5) ...". Em uma análise das expressões de domínio ativadas (9), pode ser percebido que "caulinização" e "oxidação" são especificadas no grafo corrente (11), bem como possuem índices de relevância para a interpretação detalhados como: "peso = 5" e "peso = 5", respectivamente. A soma destes pesos ("somaPesos = 10") é comparada com o limiar de conclusão ("limiar = 6") do grafo selecionado (11). Se a soma dos pesos das expressões de domínio ativadas (9) é maior ou igual ao limiar da norma selecionada (11), então uma solução (12) pode ser especificada. Assim sendo, uma solução "Teolodiagênese sobre condições meteóricas" seria especificada a partir do grafo (11) e das expressões de domínio "caulinização" e "oxidação" (9). Por fim, caso mais de um solução diferente seja alcançada pela análise de outros grafos (10) e expressões de domínio ativadas (9), estas soluções (12) serão combinadas em uma interpretação final (13), tal como, por exemplo: "Teolodiagênese sobre condições meteóricas, Eodiagênese meteórica continental sob clima úmido".

Em resumo, o PSM Interpretação de Ambientes Diagenéticos descreve um processo de inferência baseado em modelos de conhecimento inferencial de domínio especificados. Dessa forma, este PSM possivelmente venha a descrever o raciocínio desenvolvido em diferentes domínios de problemas, desde que a inferência seja suportada por conhecimento de domínio representado como casos, expressões de domínio, conceitos de domínio e normas de interpretação.

4.3.1 A competência do PSM Interpretação de Ambientes Diagenéticos

A descrição de competência do PSM Interpretação de Ambientes Diagenéticos é definida de forma bastante simplificada, assim como foi apresentada a competência do PSM Interpretação de Rochas. Logo, *o PSM Interpretação de Ambientes Diagenéticos sugere uma interpretação de ambiente diagenético a partir de um caso de descrição de amostra de rocha analisado segundo associações de expressões de domínio e normas de interpretação de ambientes diagenéticos*. De maneira mais objetiva, as entradas do PSM Interpretação de Rochas: *imagem de amostra de rocha, pacotes visuais e esquemas de rocha*, são previstas pelo PSM Interpretação de Ambientes Diagenéticos como, respectivamente: *caso de descrição de amostra de rocha, expressões de domínio e normas de interpretação de ambientes diagenéticos*. Além disso, o PSM Interpretação de Rochas resulta em interpretações de rocha quaisquer, mas o PSM aqui sendo apresentado é especializado para resultar somente interpretações de ambientes diagenéticos.

4.3.2 Uma especificação operacional para o PSM Interpretação de Ambientes Diagenéticos

O PSM Interpretação de Ambientes Diagenéticos é composto pelos seguintes papéis de conhecimento distintos: *caso, caso abstraído, conceito-atributo-valor, normas, norma, expressões de domínio, expressão de domínio, conceito-atributo-valor ativado, expressão de domínio ativada, expressões de domínio ativadas*,

soluções e interpretação. Esta especificação operacional também é descrita em função das seguintes ações de inferência:

- *abstrair* - uma amostra de rocha é um caso a ser analisado. Esta amostra é abstraída em seus aspectos relevantes para assim expressar um caso indicado no modelo;
- *selecionar* - os conceitos de domínio de um caso abstraído são selecionados por uma expressão de domínio a ser verificada neste caso. Tal expressão de domínio também é selecionada aleatoriamente de um conjunto de expressões de domínio conhecido e modelado. Além desses passos de seleção citados, um conjunto de associações entre expressões de domínio e interpretações (conjunto de normas) é selecionado aleatoriamente resultando em uma das normas componentes;
- *decompor* - uma expressão de domínio selecionada é decomposta em todos as suas combinações lógicas de conceitos-atributos-valores componentes;
- *confrontar* - conceitos-atributos-valores selecionados do caso abstraído são confrontados com outros conceitos-atributos-valores decompostos das expressões de domínio modeladas. O resultado desse casamento de padrões são os conceitos-atributos-valores considerados como similares (ativados);
- *compor* - conceitos-atributos-valores encontrados no caso são compostos formando ou não uma expressão de domínio encontrada neste mesmo caso;
- *combinar* - expressões de domínio encontradas no caso são combinadas em um conjunto de expressões de domínio ativadas que indica uma mesma interpretação particular. Este mesmo passo de inferência é usado em outro momento no PSM, pois interpretações individuais validadas no caso (então consideradas como soluções) são também combinadas para expressar uma interpretação final;
- *especificar* - um modelo de normas descrevendo índices de relevância de expressões de domínio para interpretações é usado para especificar as interpretações que podem ser alcançadas, sendo então consideradas como soluções.

A Figura 4.11 apresenta a especificação de controle do PSM Interpretação de Ambientes Diagenéticos. Tal especificação é descrita em função de papéis de conhecimento, ações de inferência e estruturas de controle básicas.

A especificação de controle do PSM Interpretação de Ambientes Diagenéticos contém três predicados:

- *ExpressõesDomínioAnalisadas* - retorna *verdadeiro* quando todas as expressões de domínio modeladas forem analisadas individualmente. Para realizar essa verificação, esse predicado toma como base o conjunto de expressões de domínio sendo analisado (*expressõesDomínio*), onde cada

expressão depois de analisada individualmente é retirada deste conjunto. Tal predicado pode ser comparado ao predicado *PacotesAnalisados*, o qual foi estabelecido na especificação de controle do PSM Interpretação de Rochas;

Da maneira como essa especificação de controle foi modelada, o predicado *ExpressõesDomínioAnalisadas* também pode sintetizar grande parte da idéia do predicado *ImagemAnalisada* do PSM Interpretação de Rochas, o qual é empregado para indicar uma análise completa (ou muitas vezes suficiente, segundo critérios do especialista) da imagem da rocha. Porém, a especificação de controle do PSM Interpretação de Ambientes Diagenéticos não é explicitamente modelada em função da descrição da amostra, mesmo que as feições de rocha mais diagnósticas sejam individualmente analisadas. Logo, o predicado *ImagemAnalisada* é refinado na especificação de controle sendo apresentada, assim atendendo a restrições de aquisição de conhecimento experimentadas nesse trabalho.

```

psm interpretaçãoDeAmbientesDiagenéticos

controlEspecification:
  begin
    abstrair(caso → casoAbstraído);
    repeat
      selecionar(expressõesDomínio → expressãoDomínio);
      repeat
        selecionar(casoAbstraído + expressãoDomínio → conceitoAtributoValor_1);
        decompor(expressãoDomínio → conceitoAtributoValor_2);
        confrontar(conceitoAtributoValor_1 +
                    conceitoAtributoValor_2 → conceitoAtributoValorAtivado);
        compor(conceitoAtributoValorAtivado → expressãoDomínioAtivada);
      until ExpressãoDomínioAnalisada(expressãoDomínio);
      expressõesDomínio := expressõesDomínio - expressãoDomínio;
      combinar(expressãoDomínioAtivada → expressõesDomínioAtivadas);
    until ExpressõesDomínioAnalisadas(expressõesDomínio);
    repeat
      selecionar(normas → norma);
      especificar(expressõesDomínioAtivadas + norma → soluções);
      normas := normas - norma;
    until NormasAnalisadas(normas);
    combinar(soluções → interpretação);
  end
end psm.

```

FIGURA 4.11 - Especificação de controle do PSM Interpretação de Ambientes Diagenéticos

- *ExpressãoDomínioAnalisada* - retorna *verdadeiro* quando uma expressão de domínio em particular tenha sido analisada por completo. Uma expressão de domínio é analisada pelos seus conceitos-atributos-valores componentes, os quais são sucessivamente decompostos a partir da expressão de domínio selecionada;

Segundo a especificação de controle do PSM Interpretação de Ambientes Diagenéticos detalhada, enquanto uma expressão de domínio não for completamente decomposta, o predicado *ExpressãoDomínioAnalisada* não permite que a análise de conceitos individuais de rocha termine, assim impondo que todos os conceitos sejam verificados. No entanto, poderia ser especificado um novo predicado similar ao *ExpressãoDomínioAnalisada*, onde este novo predicado verificaria quando uma expressão de domínio poderia ser otimizada composta a partir de conceitos de domínio encontrados (*conceitoAtributoValorAtivado*). Desta forma, no instante que a expressão pudesse ser composta (*compor*) com uma devida relevância descrita por critérios do especialista, o processo de casamento de padrões de conceitos de domínios individuais poderia ser finalizado de forma adiantada, assim como é tardiamente concluído quando todos os conceitos de domínio componentes de uma expressão de domínio são analisados.

O predicado *ExpressãoDomínioAnalisada* não tem similar na especificação de controle do PSM Interpretação de Rochas, pois aquela descrição modela um raciocínio desenvolvido em uma granularidade de análise mais alta que a granularidade de conceitos-atributos-valores individuais descrita no PSM Interpretação de Ambientes Diagenéticos. Todavia, esse predicado pode modelar uma idéia de análise similar (ou mais reduzida e simplificada) à idéia modelada no predicado *ExistePacote* da especificação de controle do PSM Interpretação de Rochas.

- *NormasAnalisadas* - retorna *verdadeiro* quando todas as normas modeladas forem analisadas. Cada norma é selecionada aleatoriamente e verificada individualmente, mas expressões de domínio mais diagnósticas encontradas poderiam ser usadas como índices para selecionar apenas as normas mais relevantes para a interpretação atual.

O predicado *NormasAnalisadas* poderia ser comparado ao predicado *EsquemasAnalisados*, estabelecido na especificação de controle do PSM Interpretação de Rochas. Todavia, o predicado *EsquemasAnalisados* tende a expressar uma condição mais heurística de término do processo de raciocínio, pois o predicado *NormasAnalisadas* é especificado apenas em função de normas que são individualmente extraídas de um conjunto de todas as normas conhecidas.

Os passos de inferência abstrair e combinar da especificação de controle do PSM Interpretação de Ambientes Diagenéticos são executados uma única vez durante o processo de inferência. O primeiro deles realiza uma abstração de um caso de descrição de amostra de rocha. Esses dados abstraídos servem como base para análise no processo de inferência especificado. O segundo passo executado uma única vez combina soluções individuais (neste caso, interpretações individuais) alcançadas. Uma análise de combinações possíveis de solução não foi especificada nesse modelo de raciocínio, tal como foi especificada como uma composição de soluções no PSM Interpretação de Rochas.

De maneira simplificada, a especificação de controle da Figura 4.11 expressa uma busca de expressões de domínio conhecidas seguida de uma especificação de soluções a partir de normas de interpretação. Em princípio, todas as expressões de domínio modeladas (*expressõesDomínio*) são analisadas, e as expressões que puderam

ser encontradas na descrição da rocha (*casoAbstraido*) são combinadas em um conjunto de expressões de domínio validadas (*expressõesDomínioAtivadas*).

Para analisar as expressões de domínio modeladas, cada expressão de domínio é selecionada e verificada em um laço central de casamento de padrões. Esse laço seleciona os conceitos-atributos-valores do caso (objeto da análise), os quais são relacionados com a expressão de domínio sendo analisada. A partir disso, decompõe sucessivamente esta expressão de domínio em seus conceitos-atributos-valores e, por fim, realiza um casamento de padrões entre eles. Os conceitos-atributos-valores similares (*conceitoAtributoValorAtivado*) podem ser ou não compostos novamente em uma expressão de domínio (*expressãoDomínioAtivada*), desde que expressem as características mínimas que representam a expressão de domínio selecionada. Ao final desse processo inicial, todas as expressões encontradas na descrição da amostra de rocha são compostas em um conjunto de expressões de domínio (*expressõesDomínioAtivadas*), o qual é usado para especificar soluções.

Para especificar as soluções que puderam ser alcançadas, o processo de inferência analisa todas as normas de interpretação modeladas (grafos de conhecimento mapeados para *normas*). Para isso, cada uma das normas é selecionada individualmente (*norma*), e então são verificados os índices de relevância das expressões de domínio ativadas (índices descritos na norma selecionada). Tais índices verificados em relação a limiares de interpretação podem resultar ou não em uma solução. Como mais de uma solução pode ser alcançada (*soluções*), elas são apenas combinadas em uma interpretação final. Segundo o conhecimento de domínio adquirido, não foi possível modelar uma composição final de soluções resultante em uma completa coerência de interpretação, tal como previsto no PSM Interpretação de Rochas.

4.3.3 Um esquema de domínio típico do PSM Interpretação de Ambientes Diagenéticos

O esquema de domínio do PSM Interpretação de Ambientes Diagenéticos é basicamente fundamentado pelos modelos de domínio desenvolvidos: modelo de casos, modelo de frames e modelo de grafos de conhecimento (seção 4.2). Porém, esse esquema de conhecimento é apresentado de forma mais simplificada, descrevendo apenas os principais tipos de conhecimento requeridos pelo método de inferência especificado. O modelo é apresentado em um formalismo simples, tal como um modelo de objetos descrito na Figura 4.12.

O esquema de domínio típico apresentado na Figura 4.12 indica que o conhecimento de domínio requerido pelo PSM Interpretação de Ambientes Diagenéticos deve abordar conceitos, tais como: *normas*, *interpretações*, *expressões de domínio*, *casos*, *casos abstraídos* e *conceitos-atributos-valores*. Cada interpretação deve ser parte-de uma norma modelada, assim como expressões de domínio modeladas também devem ser *parte-de* uma norma. Tal como foi denotado nesse modelo, uma norma é composta por expressões de domínio que indicam, com um certo índice de relevância, uma interpretação. Esta interpretação pode ser alcançada segundo um limiar de conclusão avaliado em relação aos índices de relevância de expressões de domínio encontradas.

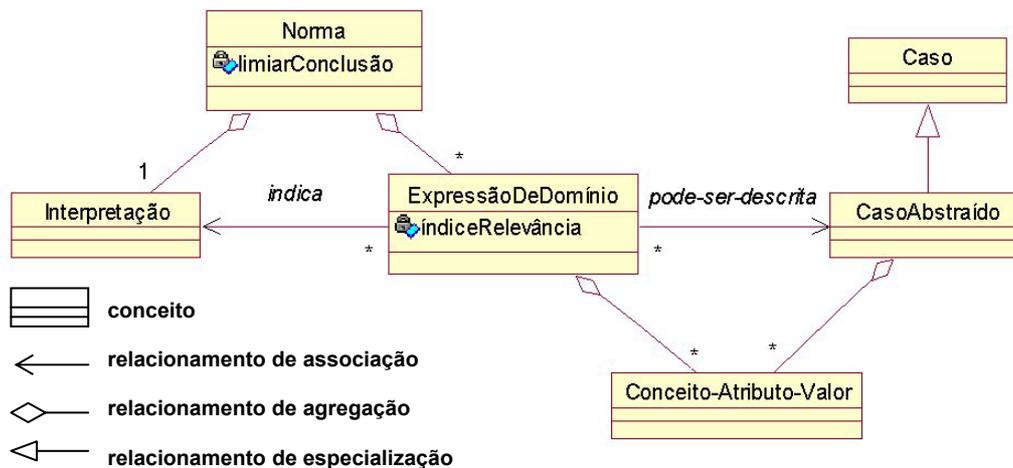


FIGURA 4.12 - Esquema de domínio típico do PSM Interpretação de Ambientes Diagenéticos

O modelo também demonstra que uma ou várias expressões de domínio modeladas podem ser descritas em um caso abstraído. A abstração realizada sobre o caso deve ser tal que o caso abstraído seja descrito em termos de conceitos-atributos-valores, assim como também são descritas as expressões de domínio modeladas. Tais conceitos-atributos-valores são os tipos de conhecimento de granularidade mais fina, sendo que expressões de domínio e casos abstraídos devem ser formados por agregações destes.

4.3.4 Alguns requisitos/suposições do PSM Interpretação de Ambientes Diagenéticos

De maneira geral, as suposições estabelecidas para o PSM Interpretação de Ambientes Diagenéticos servem para caracterizar a utilidade prática deste método de raciocínio proposto. A maioria das suposições identificadas foram derivadas dos modelos de domínio construídos (modelos de frames e grafos de conhecimento - seção 4.2), pois toda representação contém suposições, mesmo que sejam simplificadas (Fensel et al. [FEN 2000]). Em especial, a caracterização precisa do PSM Interpretação de Ambientes Diagenéticos está relacionada com o poder de expressão das técnicas de frames e grafos de conhecimento, pois tais técnicas foram empregadas para efetivamente adquirir e modelar o conhecimento inferencial de domínio requerido pelo método proposto.

Dentre as suposições estabelecidas para o PSM Interpretação de Ambientes Diagenéticos, muitas delas enfraquecem a funcionalidade do PSM Interpretação de Rochas, pois os objetivos originais deste método foram reduzidos no sentido de viabilizar a modelagem e solução do problema considerado. Dessa forma, o PSM Interpretação de Ambientes Diagenéticos considera um contexto de aplicação bem definido e delimitado, o qual permite expressar casos práticos de solução de problemas descritos por especialistas em Petrografia.

As principais suposições estabelecidas podem ser citadas:

(1) *Suporte para composições originais mais comuns* - o PSM Interpretação de Ambientes Diagenéticos foi modelado no sentido de apresentar soluções para as variações mais comuns de composições originais de rocha, as quais foram explicitamente representadas nos modelos de conhecimento de domínio desenvolvidos. Segundo o especialista em Petrografia, as feições geológicas representadas no modelo de frames especificado podem expressar as composições que representam 90% dos casos de interpretação de rochas sedimentares clásticas.

Por conseguinte, rochas com composições exóticas (variações muito diferentes da composição original de uma rocha), resultantes de processos de formação incomuns, podem não ter todas suas feições geológicas descritas em termos de conceitos de domínio especificados no modelo de conhecimento construído. Como o processo de inferência é desenvolvido em função de combinações de conceitos de domínio representados nesse modelo, o PSM Interpretação de Ambientes Diagenéticos pode não alcançar soluções para variações exóticas da composição original de uma rocha.

Interpretações resultantes de composições exóticas são consideradas os piores casos de solução que o PSM Interpretação de Ambientes Diagenéticos pode encontrar. Nestes casos, o PSM não alcança uma solução mesmo para uma amostra correta e completamente descrita. Logo, o modelo de conhecimento inferencial de domínio é completo para a descrição de composições originais mais comuns, onde se supõe que o método alcance uma solução para esses casos práticos de interpretação. Tal suposição foi estabelecida por vários motivos, mas principalmente para caracterizar a competência exata do método de solução desenvolvido. Como pode ser notado, o objetivo a ser alcançado pelo método foi simplificado (tal como uma *suposição teológica* - Fensel et al. [FEN 98a]), assim viabilizando a solução desse problema em meio computacional.

(2) *Dados conhecidos e constantes* - esta suposição indica que todos os dados requeridos pelo PSM Interpretação de Ambientes Diagenéticos são informados antes do início do raciocínio, além de não serem modificados durante o processo de inferência. Os dados, aqui referidos, são as feições geológicas que compõem as expressões de domínio dos grafos de conhecimento modelados (Tabela 4.3). O estabelecimento desta suposição permite que não haja interação do mecanismo de solução com o usuário durante a inferência, assim caracterizando a forma de interação do mecanismo de solução proposto com o seu ambiente. Esta suposição é similar a suposição de *não-intermitência* (Benjamins et al. [BEN 96b]).

(3) *Aspectos mais diagnósticos descritos correta e completamente* - o PSM Interpretação de Ambientes Diagenéticos supõe que os conceitos de domínio mais diagnósticos sejam correta e completamente descritos pelo petrógrafo em um caso. Conforme o especialista em Petrografia, o modelo de conhecimento de domínio desenvolvido (modelo de frames - seção 4.2) permite a descrição de feições de rocha em diferentes níveis de granularidade e extensão. Sendo assim, uma amostra de rocha pode ser descrita com alto grau de precisão, mas o reconhecimento e classificação das feições de uma amostra ainda é totalmente dependente do petrógrafo que analisa a rocha. Sendo assim, quando mais fluência em Petrografia o usuário possuir, mais

feições ele poderá reconhecer e descrever corretamente usando o modelo de conhecimento disponibilizado.

Para que o PSM alcance o comportamento esperado e também contornar esse problema, supõe-se que apenas as feições mais diagnósticas devam ser descritas correta e completamente, pois somente elas são efetivamente analisadas durante a inferência. Tais feições mais diagnósticas são aqueles conceitos que compõem as expressões de domínio modeladas. Desta forma, caso os valores para estes conceitos de domínio sejam correta e completamente reconhecidos e descritos pelo petrógrafo, o PSM modelado nunca resultará uma interpretação incorreta a partir desses dados. Tal suposição é similar a suposição envolvendo *tipos de entradas confiáveis* descrita para *tarefas de diagnóstico* (Fensel et al. [FEN 98a]).

(4) *Confiabilidade simplificada de interpretação* - como afirmado pelo especialista em Petrografia, a presença de feições diagnósticas, mais o número e a abundância e/ou intensidade desses aspectos definem uma interpretação como confiável. Além disso, uma interpretação confiável demanda a ocorrência de mais de um parâmetro diagnóstico na amostra, preferencialmente em substancial abundância ou intensidade. Todavia, o modelo de conhecimento de domínio detalhado não oferece o suporte necessário para que esta medida de confiabilidade do especialista seja representada.

Para representar esse critério, o modelo de conhecimento de domínio desenvolvido deveria permitir a descrição do número de instâncias de conceitos de domínio que são encontradas em uma amostra, pois esses conceitos formam as expressões de domínio modeladas, as quais são efetivamente consideradas durante o processo de inferência. Por exemplo, no grafo exemplificado na Figura 4.8, o modelo de conhecimento de domínio desenvolvido deveria prever explicitamente a descrição de várias instâncias de "nomeMineral = Caulim" com "hábito = Lamelas" ..., os quais são conceitos da expressão de domínio "caulinização". Desta forma, várias instâncias de uma mesma expressão de domínio poderiam ser validadas, assim expressando o critério de confiabilidade de interpretação do especialista.

A medida de confiabilidade adotada (logo, simplificada) pelo PSM Interpretação de Ambientes Diagenéticos está fundamentada na suposição que uma interpretação somente pode ser considerada confiável se for encontrada na descrição da amostra pelo menos uma expressão de domínio significativamente relevante (com um peso de 6) para aquela interpretação. Para tornar esse critério mais flexível para casos práticos de interpretação, uma solução também é considerada confiável pelo PSM se pelo menos sejam encontradas duas ou mais expressões de domínio, de modo que suas relevâncias somadas resultem em um valor maior ou igual a um limiar de conclusão de interpretação, o qual é expresso em cada um dos grafos de conhecimento modelados. Por conseguinte, o critério de confiabilidade adotado indica quando um conjunto de expressões de domínio ativadas é relevante ou não para especificar uma solução. Deste modo, uma competência simplificada para o PSM Interpretação de Ambientes Diagenéticos foi caracterizada, mas também foi reduzido o esforço de modelagem necessário para desenvolver um mecanismo de inferência para esse domínio de problemas, pois o processo de aquisição de conhecimento pôde ser mais simplificado.

Outra medida real de confiabilidade que poderia ser suportada a partir do modelo de conhecimento de domínio atual pode ser obtida com o cruzamento simultâneo de análises de várias amostras de rocha. Todavia, tais amostras necessariamente deveriam ser oriundas de uma mesma localização, pois um ambiente diagenético é dominante em grandes extensões geográficas. Logo, a comprovação de um mesmo grafo (por exemplo, "Telodiagênese sob condições meteóricas" - Figura 4.7) em várias descrições de amostras oriundas em uma mesma localização (tal como uma mesma "unidade sedimentar") poderia indicar com grande confiabilidade uma interpretação derivada a partir do PSM Interpretação de Ambientes Diagenéticos. Contudo, essa funcionalidade ainda não é modelada no método proposto, bem como não há conhecimento inferencial de domínio adquirido que suporte o seu emprego. Este é um refinamento que pode ser acrescentado ao método atual.

(5) *Expressões de domínio auto-exclusivas* - caso as entradas do PSM Interpretação de Ambientes Diagenéticos estejam corretas (aspectos mais diagnósticos correta e completamente descritos), o método não poderá derivar uma interpretação onde todas as expressões de domínio de um mesmo grafo modelado sejam encontradas em uma amostra. Por exemplo, ativando simultaneamente em uma mesma amostra as expressões "dissolução", "caulinização", "oxidação", "cobertura de óxido de ferro" e "calcita", as quais pertencem ao mesmo grafo (Figura 4.7).

Conforme o especialista em Petrografia, as expressões de domínio modeladas, mesmo as pertencentes a um mesmo grafo, representam classes composicionais de rocha distintas. Desta forma, supõe-se que elas sejam variações auto-exclusivas da composição original de uma rocha sedimentar clástica quando submetidas a um ambiente diagenético particular.

A suposição relativa a *expressões de domínio auto-exclusivas* é similar a *suposição de falha única de tarefas de diagnóstico* (Benjamins et al. [BEN 95]). No entanto, esta suposição pode ser usada como um critério de confiabilidade a ser verificado sobre a solução derivada pelo PSM Interpretação de Ambientes Diagenéticos, assim permitindo verificar se as feições mais diagnósticas da amostra sendo analisada não foram corretamente descritas.

(6) *Poucas expressões de domínio modeladas* - mesmo que a base de conhecimento inferencial de domínio modelada venha a ser acrescentada de novas expressões de domínio, não se espera que haja um número muito grande de novas expressões de domínio modeladas. As expressões de domínio modeladas representam pacotes visuais mantidos por especialistas em Petrografia, os quais são gradativamente refinados. Segundo Abel [ABE 2001], depois de algum tempo, especialistas em Petrografia criam poucos pacotes visuais novos, pois o processo de refinamento (abstração) de antigos é mais evidente e possível de ser verbalizado. Além disso, percebe-se que o processo de raciocínio modelado será tanto mais eficaz quanto melhores forem as expressões de domínio modeladas, e não quanto mais expressões de domínio forem modeladas. Tal suposição foi estabelecida para caracterizar a interação do PSM com sua base de conhecimento, viabilizando prever alguma eficiência sobre a estratégia de solução de problemas proposta pelo PSM Interpretação de Ambientes Diagenéticos.

(7) *Uma interpretação pode não ser alcançada* - conforme o especialista em Petrografia, quando não é encontrado nenhum casamento de padrões relevante entre aspectos geológicos conhecidos e uma amostra de rocha, normalmente são usadas teorias gerais sobre processos geológicos e processos de busca genéricos para tentar sugerir alguma interpretação para a amostra. Como esse conhecimento de domínio sobre "teorias genéticas" não foi adquirido e especificado, supõe-se que o PSM Interpretação de Ambientes Diagenéticos pode não alcançar uma interpretação.

O estabelecimento desta suposição enfraquece o objetivo do PSM Interpretação de Ambientes Diagenéticos resultante, pois o método somente utiliza as expressões de domínio modeladas para alcançar soluções, tornando a tarefa de aquisição de conhecimento mais simplificada e viável de ser realizada. Mesmo assim, em casos reais de interpretação é bastante comum que uma solução não seja facilmente derivável a partir dos dados disponíveis para análise, os quais são resultantes de descrição de amostra incompletas ou feições de rocha incorretamente identificadas pelo petrógrafo.

(8) *Conhecimento inferencial de domínio completamente analisado* - durante o processo de raciocínio, o mecanismo de inferência avalia completamente o conhecimento inferencial de domínio modelado. Neste caso, todos os grafos de conhecimento modelados são analisados individualmente em relação a uma descrição de amostra de rocha, resultando ou não em soluções. Esta suposição pode ser considerada típica de tarefas de interpretação (Harmon et al. [HAR 90]), pois essas tarefas requerem uma análise completa dos dados para que a interpretação não venha a desconsiderar fatores importantes.

(9) *Índices de relevância fixos* - como demonstrado anteriormente, foram modelados índices de relevância associados a cada expressão de domínio detalhada. Por exemplo, a expressão de domínio "caulinização" (Figura 4.8) tem relevância de 5 para indicar um ambiente "Telodiagênese sob condições meteóricas". Todavia, conforme verificado por Abel [ABE 2001], os índices de relevância de pacotes visuais para interpretações de rochas são variáveis em casos reais de interpretação, mas foram simplificados e representados como valores de pesos fixos.

Para alcançar esses valores fixos, Abel [ABE 2001] conduziu experimentos de interpretação com o especialista em Petrografia usando diferentes valores de pesos, até que valores fixos adequados para representar o comportamento de solução desejado fossem alcançados. Desta forma, uma relevância de uma expressão X não é afetada pela relevância de outra expressão Y, mesmo que sejam indicativas de uma mesma solução particular. Tais índices são efetivamente empregados no passo de especificação de soluções do PSM Interpretação de Ambientes Diagnéticos, pois são usados para verificar a possibilidade de especificação de uma solução a partir de expressões de domínio encontradas em uma amostra. Supor que a relevância de expressões de domínio é fixa principalmente auxiliou a modelagem desse problema, pois seria muito complexo analisar e projetar um método que considerasse relevâncias variáveis.

(10) *Inexistência de encadeamento de soluções* - o PSM Interpretação de Ambientes Diagenéticos não é fundamentado em diferentes níveis de soluções parciais, assim como um encadeamento de hipóteses em uma *rede de inferência*

(Tanimoto [TAN 95]). Nesse PSM, as expressões de domínio são diretamente relacionadas com interpretações de rocha. Isso indica que as expressões de domínio consideradas como evidências são diretamente relacionadas as conclusões, assim como expresso nos grafos de conhecimento detalhados. Tal suposição, portanto, detalha uma estrutura de informação básica para derivar soluções.

A suposição sobre *inexistência de encadeamento de soluções* foi estabelecida para melhor expressar o processo de raciocínio desenvolvido pelo especialista em Petrografia. Além disso, tal suposição pode simplificar o mecanismo de solução de problemas proposto, pois estratégias de explicação associadas a caminhos de inferência curtos são simples de serem desenvolvidas.

(11) *Interpretações individuais não interagem* - os grafos de conhecimento desenvolvidos supõem que os ambientes diagenéticos que atuam sobre uma rocha são as primeiras transformações diagenéticas sofridas pela amostra de rocha, representando a ação de um processo diagenético considerado individualmente sobre a rocha. Por exemplo, uma expressão de domínio "caulinização" indica um ambiente "Telodiagênese sobre condições meteóricas" sobre a composição original de uma amostra de rocha. Porém, nenhuma expressão modelada indica um ambiente sobre a composição de outro ambiente. Utilizando esta abordagem de modelagem, possíveis interações poderiam ser descritas como, por exemplo: uma nova expressão de domínio possivelmente denominada de "Caulinização e Calcrete", assim simultaneamente representado um ambiente "Telodiagênese sob condições meteóricas" e um ambiente "Eodiagênese meteórica continental sob clima úmido".

Mesmo notando que existem interações permitidas e não permitidas (tais como restrições de integridade de domínio, por exemplo) entre interpretações em casos reais descritos por especialistas em Petrografia, estas interações entre processos diagenéticos não foram modeladas como expressões de domínio. Tais relações, muitas vezes temporais, ou de dependência, são complexas e difíceis de serem adquiridas e validadas junto a especialistas em Petrografia. Além disso, soluções alcançadas pelo método, ou possíveis de serem alcançadas, não são consideradas durante o transcorrer do processo de inferência para derivar novas soluções. Mesmo assim, algum conhecimento de domínio poderia ser modelado e usado para avaliar combinações possíveis de soluções particulares alcançadas, as quais seriam compostas em uma interpretação final. Tal conhecimento poderia indicar amostras de rochas descritas incorretamente, tornando o PSM Interpretação de Ambientes Diagenéticos mais tolerante a falhas, mas aumentando a complexidade do problema a ser modelado e resolvido.

Em síntese, as suposições descritas para o PSM Interpretação de Ambientes Diagenéticos permitiram: 1) caracterizar o PSM em relação a casos reais de raciocínio; 2) determinar a interação do PSM em relação ao seu ambiente; 3) descrever o PSM em função de modelos de conhecimento detalhados; 4) indicar critérios de confiabilidade suportados pelo PSM; 5) caracterizar a competência do PSM em relação a competência do especialista; 6) identificar requisitos do PSM; 7) caracterizar o conhecimento requerido pelo PSM; 8) identificar simplificações que podem ser usadas para desenvolver mecanismos de inferência e 9) descrever as adaptações realizadas em relação a modelos de raciocínio mais genéricos.

4.4 Uma classificação para os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos

A construção de bibliotecas de PSMs reusáveis tem sido um tema bastante pesquisado nos últimos anos. Entre outros motivos, tais bibliotecas podem facilitar o desenvolvimento de KBSs, pois PSMs podem orientar diferentes tarefas de aquisição de conhecimento de domínio, bem como descrever estratégias de inferência complexas, os quais ainda são grandes gargalos da KE. Os PSMs desenvolvidos nesse trabalho, portanto, poderiam representar um primeiro passo no desenvolvimento de *uma biblioteca para interpretação de rochas*, ou mesmo serem utilizados como exemplos de PSMs de uma biblioteca de Interpretação.

Os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos apresentam *funcionalidades similares*, mas competências distintas. Tais funcionalidades foram caracterizadas mais precisamente por suposições típicas de domínio, mas que também podem ser possivelmente reusáveis entre domínios. Todavia, as suposições identificadas são um tanto abrangentes, caracterizando muitas vezes o mecanismo de inferência como um todo. Portanto, a organização de uma possível biblioteca de PSMs para a Petrografia poderia ser realizada segundo aspectos de funcionalidade, mas tais funcionalidades deveriam ser descritas com diferentes níveis de competência expressos por meio de suposições.

A partir de entrevistas com o especialista em petrografia, também foi notado que muitas tarefas de raciocínio desse domínio são dependentes umas das outras, pois os resultados de uma são utilizados como entradas em outra. Por exemplo, os resultados de uma tarefa de "classificação composicional" de uma amostra de rochas são usados para distinguir um tipo de "ambiente diagenético" dessa amostra, o qual também pode ser usado para caracterizar uma "seqüência diagenética", sendo que essas tarefas não podem ser resolvidas por PSMs com competências similares. Além disso, apenas o resultado de uma tarefa não é usado como única fonte de informação para outra tarefa, pois outras entradas são requeridas para resolver as tarefas de raciocínio identificadas nesse domínio. Logo, a organização de PSMs para este domínio de problemas pode ser descrita em termos de *tipos de problemas*. Todavia, somente tipos de problemas podem não ser requisitos suficientes para expressar a organização das tarefas de raciocínio em domínios de problemas reais, tal como é o domínio de interpretação de rochas em Petrografia Sedimentar.

5 Um mecanismo de inferência para interpretação de rochas

O esquema de raciocínio para interpretação de rochas proposto nesse trabalho foi usado como um modelo de análise do raciocínio desenvolvido no domínio de problemas da Petrografia Sedimentar, tal como um modelo descrito no *nível do conhecimento*. Para apresentar uma validação prática desse modelo de análise, foi desenvolvido um *mecanismo de inferência para interpretação de ambientes diagenéticos*. Tal mecanismo simbólico é um projeto operacional destinado a fornecer um recurso de inferência para interpretação de rochas requerido pelo sistema *PetroGrapher*.

Em um primeiro momento, a versão atual do sistema *PetroGrapher* foi analisada para que um *módulo de interpretação de rochas* pudesse ser desenvolvido e integrado neste sistema. A partir dessa análise, *uma arquitetura de inferência para interpretação de rochas* foi planejada, de forma a demonstrar componentes simbólicos de manipulação de dados e conhecimento, assim como requeridos para representar um modelo de projeto de um mecanismo de inferência. Em seguida, os modelos de conhecimento de domínio especificados como um modelo de frames e um modelo de grafos de conhecimento (seção 4.2) foram mapeados para modelos de dados. Tais mapeamentos então foram usados para colocar em prática o projeto e desenvolvimento de um algoritmo de inferência simbólico para interpretação de ambientes diagenéticos. Nas seções seguintes desse capítulo, serão apresentados esses passos de validação práticos desenvolvidos, bem como a construção de um recurso de software fundamental para o sistema *PetroGrapher*.

5.1 A arquitetura do sistema *PetroGrapher*

O sistema *PetroGrapher* é uma aplicação de Bancos de Dados Inteligentes (Parsaye et al. [PAR 89]; Bertino et al. [BER 98]) destinada a descrição e interpretação de amostras de rochas sedimentares clásticas como suporte para a avaliação de reservatórios de petróleo. Uma caracterização deste sistema e do seu processo de desenvolvimento, bem como uma análise preliminar desse domínio de aplicação também pode ser encontrada em Silva [SIL 97] e Silva [SIL 99]. O presente trabalho, por sua vez, contemplou o estudo dos métodos de inferência neste domínio e a proposta de um modelo de inferência para o sistema *PetroGrapher*.

A versão atual do sistema *PetroGrapher* foi desenvolvida em uma ferramenta de programação simbólica orientada a objetos (*Allegro Common Lisp for WindowsTM*) e um SGBD relacional (*Sybase SQL AnywhereTM*). Tal como apresentado na Figura 5.1, esse sistema foi desenvolvido como uma arquitetura de três camadas:

- *camada de interface* - suporta a descrição e análise de amostras de rocha sedimentares clásticas. Esse suporte é fundamentado em uma estrutura de entrada de dados de amostras de rocha padronizada por especialistas em Petrografia, e um conjunto de interfaces de consultas e análises típicas desse domínio de aplicação;

- *camada de aplicação* - realiza as funções de gerenciamento de uma aplicação de Banco de Dados Inteligentes, tais como as funções que envolvem o controle do sistema de BD e de KBS, bem como a manipulação uniforme e integrada de dados e conhecimento de domínio utilizados durante o processo de descrição e interpretação de amostras de rocha;
- *camada de banco de dados* - essa camada mantém um repositório de dados de descrições de amostras de rochas. A camada de BD também é usada para armazenar os dados sobre o conhecimento de domínio de aplicação utilizado para orientar a descrição e interpretação dessas amostras.

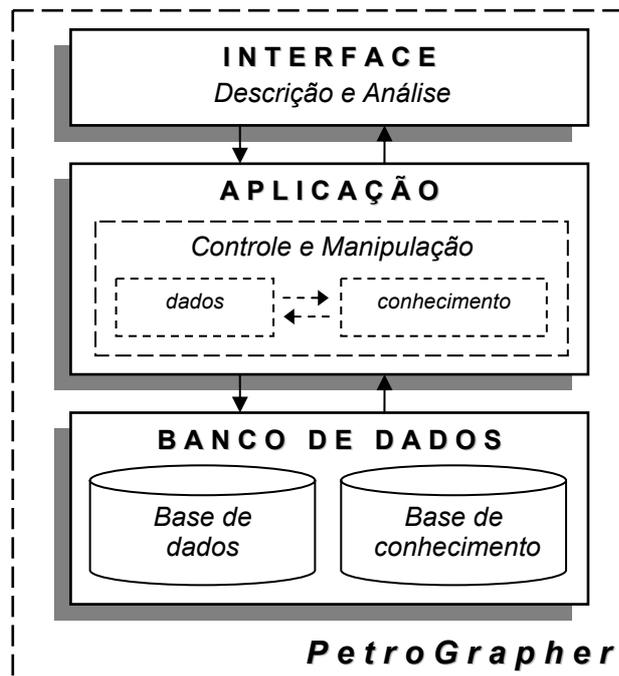


FIGURA 5.1 - Arquitetura do sistema *PetroGrapher*

A interface de aplicação do sistema *PetroGrapher* orienta a descrição de amostras, sugerindo uma ordem de descrição para feições de rocha e disponibilizando uma terminologia geológica bastante flexível para a descrição de diferentes aspectos petrográficos (terminologia detalhada e especificada no modelo de frames desenvolvido). A camada de interface atual não contém recursos disponíveis para auxiliar o petrógrafo no processo de identificação de feições geológicas, mas prevê que esses recursos sejam disponibilizados futuramente em diferentes formatos multimídia, ou mesmo como recursos de aplicação com capacidades inferenciais.

A identificação correta e completa de feições geológicas de rocha sedimentares clásticas pode ser considerado um gargalo para o processo de interpretação de rochas proposto nesse trabalho. Nesse caso, descrições pouco completas, ou muitas vezes incorretas, podem não indicar combinações de feições de rochas diagnósticas (tais como as expressões de domínio modeladas), assim dificultando alcançar uma interpretação de rocha a partir dos dados de amostras descritos.

Em muitos domínios similares ao domínio das rochas sedimentares, o processo de identificação e descrição de feições geológicas foi caracterizado por si só como uma *tarefa de classificação*. Gappa et al. [GAP 98] indicam que a tarefa de *identificação dos constituintes minerais* que compõem amostras de rocha muitas vezes é uma tarefa intensiva em conhecimento, a qual também é alvo de PSMs.

A camada de aplicação, por sua vez, representa simbolicamente como classes os frames de domínio de aplicação modelados, as quais são instanciadas pelo sistema *PetroGrapher* no decorrer do processo de descrição de uma amostra. Tais classes de aplicação possuem métodos que garantem algumas restrições de integridade típicas desse domínio, tais como: domínios de valores válidos para feições de rocha sendo descritas e conferência automática de aspectos geológicos numéricos e simbólicos relacionados, entre outros. Tal como descrito em Silva [SIL 97], a camada de aplicação é responsável por controlar o funcionamento do sistema como um todo, como também preservar o modelo de conhecimento de aplicação detalhado no projeto do sistema *PetroGrapher*.

Além de desempenhar o papel de repositório de dados e conhecimento de domínio de aplicação, a camada de BD representa algumas das principais restrições de integridade do domínio. Tais restrições são verificadas nos dados resultantes de descrições de amostras de rocha, assim mantendo um BD de descrições de amostra de rocha consistente e acessível por outras aplicações desse domínio.

Tal como pode ser percebido, a aplicação alvo do sistema *PetroGrapher* tem funcionalidades de BD requeridas muito fortes, pois descrições de amostras resultam em grande volumes de dados numéricos e simbólicos, os quais podem ser analisados por diferentes componentes inferenciais. No entanto, a versão atual da camada de BD do sistema *PetroGrapher* não tem nenhum papel inferencial automatizado, tal como poderia desempenhar um Banco de Dados Avançado (Beynon-Davis [BED 91]).

A comunicação entre o sistema simbólico é realizada através de uma interface ODBC. Tal interface permite portar o sistema *PetroGrapher* para diferentes SGBDs relacionais, mas não viabiliza o desenvolvimento de componentes de inferência *fortemente integrados* (Beynon-Davis [BED 91]) ao BD. Um requisito que deve então ser notado no mecanismo de inferência é o tempo de resposta resultante, o qual normalmente é incrementado por camadas de comunicação com o BD. A partir dessa breve descrição do sistema *PetroGrapher*, uma *arquitetura de inferência para interpretação de rochas* pode ser apresentada.

5.2 Uma arquitetura de inferência para interpretação de rochas

A *arquitetura de inferência para interpretação de rochas* proposta está fundamentada na interação entre um sistema simbólico e um sistema de BD, assim como também está organizado o sistema *PetroGrapher*. O sistema simbólico desempenha os papéis de controle e monitoração de *um PSM Interpretação de Ambientes Diagenéticos*, assim como previsto nessa arquitetura de inferência. Tal

mecanismo de inferência foi integrado no sistema *PetroGrapher* como um *módulo de interpretação de rochas*. Este módulo é composto por um *mecanismo de interpretação de ambientes diagenéticos*, um *gerador de consultas* e, por fim, uma *memória de trabalho estática* e uma *memória de trabalho dinâmica*. O sistema de BD, por sua vez, desempenha diferentes papéis para o processo de inferência, pois é empregado como:

- repositório de descrições de amostras de rochas entradas por petrógrafos usando o sistema *PetroGrapher*;
- repositório de conhecimento de domínio de aplicação usado para orientar a descrição dessas amostras
- repositório de conhecimento inferencial de domínio usado para guiar um mecanismo de interpretação de ambientes diagenéticos.

A Figura 5.2 apresenta a arquitetura de inferência para interpretação de rochas proposta.

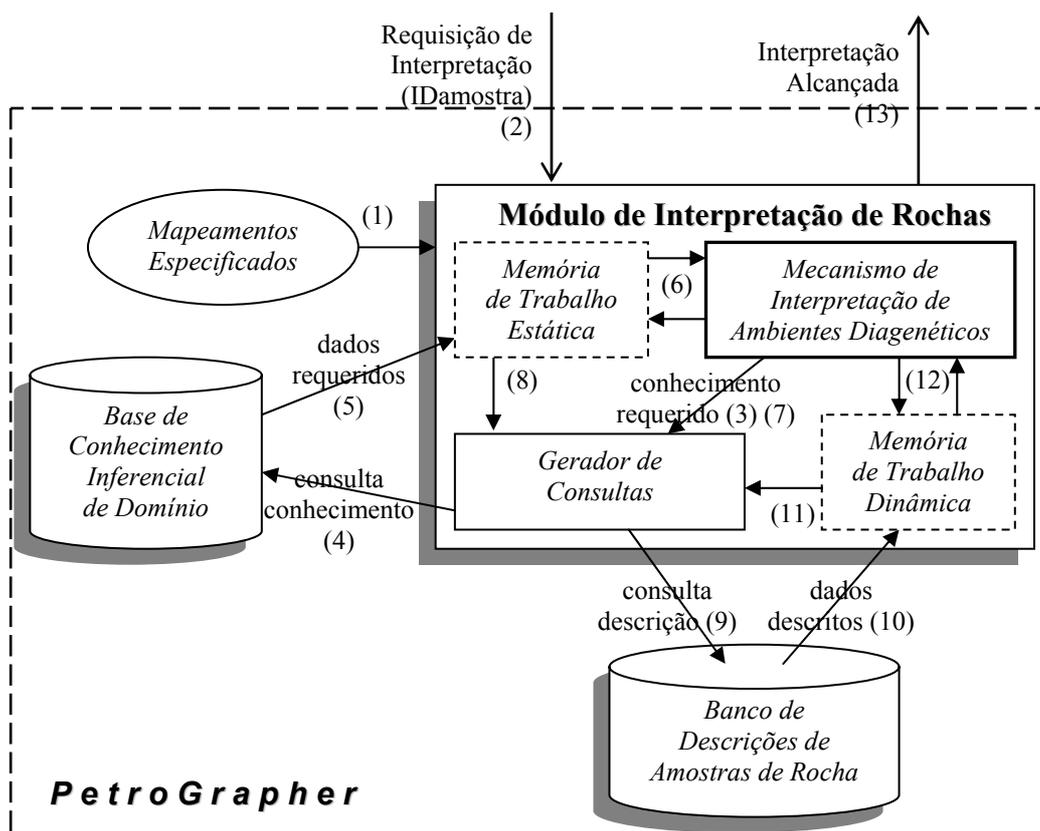


FIGURA 5.2 - Uma arquitetura de inferência para interpretação de rochas

Conforme a *arquitetura de inferência para interpretação de rochas* apresentada na Figura 5.2, o processo de interpretação de ambientes diagenéticos em uma descrição de amostra de rocha pode ser exemplificado:

Antes de iniciar o processo de inferência, os *mapeamentos especificados* entre os modelos de conhecimento e modelos de dados de aplicação desenvolvidos são carregados pelo *módulo de interpretação de rochas* (1). Durante o processo de inferência, esses mapeamentos são conhecidos por todos os componentes do *módulo de interpretação de rochas*, mas são principalmente usados pelo *gerador de consultas*. Esse módulo é responsável pela tradução de requisições de conhecimento para consultas sobre dados que são realizadas sobre a *base de conhecimento inferencial de domínio (consulta conhecimento)* e sobre o *banco de descrições de amostras de rocha (consulta descrição)*.

O processo de inferência tem início com uma *requisição de interpretação* (2) de uma amostra de rocha descrita por um petrógrafo e armazenada no *banco de descrições de amostras de rocha*. A amostra para análise é indicada nesta requisição por uma identificação de amostra (*IDamostra*) única. Tal identificação é usada em todas as consultas realizadas sobre o banco de descrições (*consulta descrição*), as quais analisam a descrição da amostra armazenada em busca de aspectos diagnósticos necessários para alcançar alguma interpretação.

A execução do processo de inferência é controlada pelo *mecanismo de interpretação de ambientes diagenéticos*. Esse mecanismo realiza os passos de inferência do *PSM Interpretação de Ambientes Diagenéticos* (PSM proposto nesse trabalho). Tais passos são executados a partir do conhecimento inferencial de domínio especificado e mantido na *memória de trabalho estática* (6) e pelos dados de uma amostra consultados e mantidos na *memória de trabalho dinâmica* (12).

Para recuperar do BD esse conhecimento e dados requeridos, o mecanismo de interpretação emite requisições de conhecimento (3) (6) (*conhecimento requerido*) para o componente *gerador de consultas*. Em um primeiro momento, essas requisições emitidas são analisadas pelo *gerador de consultas*, assim resultando em consultas (*consulta conhecimento*) (4) que então são executadas para recuperar os dados que compõem o conhecimento inferencial de domínio armazenado na *base de conhecimento inferencial de domínio*. Os dados sobre conhecimento inferencial então retornados (*dados requeridos*) (5) pelas *consultas conhecimento* (4) são organizados (6) pelo mecanismo de interpretação em estruturas de simbólicas de representação mantidas na *memória de trabalho estática* (nesse trabalho, estruturas simbólicas de listas).

Durante o processo de interpretação de ambientes diagenéticos desenvolvido, todos os dados sobre grafos de conhecimento e expressões de domínio modelados são gradativamente recuperados para a *memória de trabalho estática* por consultas (4) emitidas pelo *gerador de consultas*. Tais dados sobre conhecimento são recuperados em termos de requisições de conhecimento (3) e de estruturas simbólicas de armazenamento então mantidas na *memória de trabalho estática* (8). Desta forma, o *mecanismo de interpretação de ambientes diagenéticos* gradativamente carrega para a *memória de trabalho estática* os grafos de conhecimento e as expressões de domínio requeridos pelos passos de inferência implementados para executar o *PSM Interpretação de Ambientes Diagenéticos*.

Para demonstrar o restante do processo de interpretação de uma amostra desenvolvido usando essa arquitetura de inferência, vamos supor que uma expressão

de domínio qualquer recuperada da *base de conhecimento inferencial de domínio* esteja simbolicamente estruturada na *memória de trabalho estática*. Tal expressão é então analisada (6) pelo *mecanismo de interpretação*, e este novo tipo de *conhecimento requerido* (7) é emitido para o *gerador de consultas*. O *gerador de consultas* analisa essa nova requisição, assim gerando e executando uma ou mais consultas (9) (*consulta descrição*) sobre o *banco de descrições de amostras de rocha*. Desta forma, o *mecanismo de interpretação de ambientes diagenéticos* gradativamente verifica os conceitos de domínio que compõem cada uma das expressões de domínio especificadas.

Os *dados descritos* na amostra retornados (10) (ou não, caso não haja dados relativos aos conceitos de domínio buscados) podem ou não validar uma expressão de domínio sendo analisada. Esses *dados* retornados pelas consultas sobre a descrição da amostra (9) são organizados (12) pelo *mecanismo de interpretação de ambientes diagenéticos* em uma estrutura de armazenamento simbólica dinâmica na *memória de trabalho dinâmica* (nesse trabalho, listas de símbolos). Tal estrutura dinâmica é gradativamente acessada pelo *mecanismo de interpretação* (12) durante o processo de inferência, assim verificando se os dados da amostra retornados podem ou não validar a expressão de domínio sendo analisada.

Muitas vezes apenas uma consulta (9) pode não validar uma expressão de domínio por completo. Nesse caso, o *mecanismo de interpretação* requisita (7) novas consultas por *conhecimento requerido* (7). Como o *gerador de consultas* tem acesso aos dados da amostra já retornados (11) em consultas anteriores, ele executa novas consultas sobre a descrição (9), assim completando o processo de casamento de padrões de todos os conceitos de domínio que formam uma expressão de domínio em análise.

O processo de interpretação como um todo é controlado pelo *mecanismo de interpretação de ambientes diagenéticos*. Para completar a inferência, os passos de 3 à 12 são sucessivamente executados, de forma a analisar todo o conhecimento inferencial de domínio modelado perante a descrição de uma amostra de rocha armazenada no *banco de descrições de amostras de rocha*. Caso alguma interpretação de ambiente diagenético seja alcançada a partir dos dados da amostra analisados e então mantidos na *memória de trabalho dinâmica*, essa interpretação resultante é retornada pelo *módulo de interpretação de rochas* (13). Adicionalmente, também são retornados todos os *dados* (12) e *conhecimento* (6) analisados durante o processo de inferência, de forma a demonstrar a amplitude da inferência realizada.

Para viabilizar praticamente a arquitetura de inferência para interpretação de rochas proposta, dois tipos de mapeamentos foram especificados:

- do modelo de frames do domínio especificado (apresentado na seção 4.2) para um modelo de dados usado como esquema para o armazenamento de descrições de amostras de rocha em BD;
- do modelo de conhecimento inferencial de domínio especificado (grafos de conhecimento e expressões de domínio modelados, apresentados na seção 4.2.1) para um modelo de dados usado como esquema para o armazenamento dos dados que compõem esse conhecimento inferencial.

As seções seguintes apresentam e exemplificam os mapeamentos desenvolvidos.

5.3 Um mapeamento do modelo de frames especificado para um modelo de dados

O modelo de frames de aplicação desenvolvido nesse trabalho pode ser compreendido como um modelo conceitual que especifica os conceitos de domínio dessa aplicação, assim como requeridos para a descrição e interpretação de rochas sedimentares clásticas. Esse modelo de conhecimento foi mapeado para um modelo de dados (um modelo Entidade-Relacionamento), assim representando mais explicitamente os aspectos simbólicos de uma descrição de amostra. Tal modelo de dados mapeado resultou em um esquema de BD para o armazenamento de descrições de amostras de rochas, as quais são realizadas com a interface de aplicação do sistema *PetroGrapher*.

De forma geral, as instâncias da base de conhecimento também devem ser consideradas para realizar um mapeamento entre um modelo de conhecimento e um modelo de dados, pois o primeiro muitas vezes não faz uma distinção explícita entre esquema e dado. Logo, o mapeamento entre eles também deve ser realizado em função dos valores que os atributos do modelo de conhecimento podem conter, pois muitas vezes esses valores são descritos por objetos complexos. Mesmo assim, uma organização que represente esses dados complexos deve ser encontrada no modelo de dados resultante.

O mapeamento realizado entre o modelo de conhecimento e o modelo de dados pode ser analisado na Figura 5.3, a qual é apresentada em um formalismo de objetos simplificado.

Entre outros objetivos, o mapeamento especificado busca preservar a semântica do modelo de conhecimento no modelo de dados. No entanto, critérios como normalização, típicos de modelos de dados, também são considerados, assim permitindo o armazenamento e compartilhamento de grande volume de dados de descrições de amostras, ao invés de pequenas bases de conhecimento mantidas em arquivos texto.

A Figura 5.3 apresenta as entidades:

<Frames>_para_<EntidadesPrincipais> e
<AtributosMultivaloradosFrames>_para_<EntidadesSecundárias>

Também apresenta os relacionamentos *outrosRelacionamentos*, os quais são usados para exemplificar o mapeamento dos conceitos de domínio de aplicação especificados em um modelo de frames para um modelo de dados.

Para realizar tal mapeamento, o primeiro passo desenvolvido foi classificar os frames especificados como *principais* e *adicionais*. *Frames principais* foram

identificados como frames do domínio de aplicação especificados contendo múltiplos atributos típicos dessa aplicação, os quais poderiam ser usados para representar atributos de um esquema de dados necessário ao armazenamento de descrições de amostras. *Frames adicionais*, por sua vez, foram identificados como frames do domínio especificados para apenas representar aspectos organizacionais típicos de estruturas de modelagem de conhecimento como frames. Por exemplo, o *frame descrição-diagenética* (Figura 4.6) foi modelado para indicar que vários *frames composição-diagenética* deveriam ser delhadamente descritos em uma mesma descrição de amostra de rocha. Tais frames adicionais não foram mapeados explicitamente para o modelo de dados, mas seus principais aspectos organizacionais foram representados no modelo de dados como relacionamentos entre entidades mapeadas.

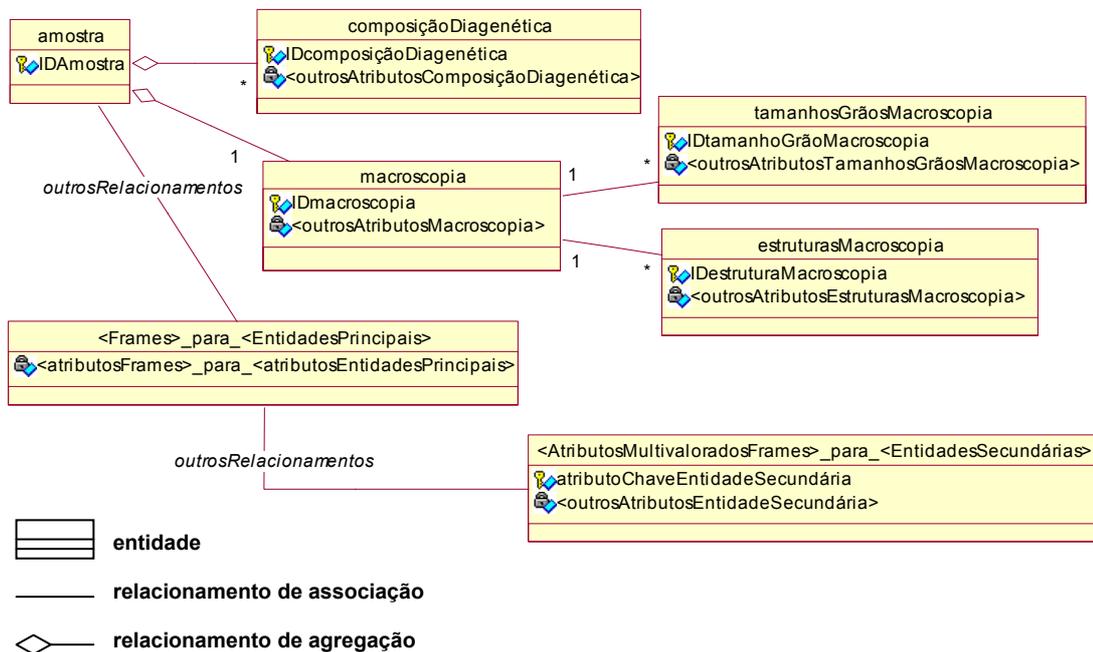


FIGURA 5.3 - Mapeamento do modelo de frames para o modelo de dados

O processo de mapeamento realizado pode ser sintetizado como o seguinte (exemplos vide Figura 5.3):

- 1) cada um dos *frames principais* identificados foram mapeados para *entidades principais* do modelo de dados mapeado;

Por exemplo, o *frame composição-diagenética* foi mapeado para a *entidade composiçãoDiagenética*, o *frame macroscopia* foi mapeado para a *entidade macroscopia*, etc. Além disso, os atributos não-multivalorados e definidos por domínios simplificados dos frames principais foram mapeados para atributos das entidades principais mapeadas (exemplificado como *<atributosFrames>_para_<atributosEntidadesPrincipais>* na Figura 5.3). Frames classificados como adicionais não foram mapeados para o modelo de dados, mas todos os frames classificados como principais foram mapeados de forma similar.

- 2) relacionamentos *parte-de* entre os *frames principais* foram mapeados como relacionamentos de *agregação* entre *entidades principais* mapeadas;

Por exemplo, o *frame composição-diagenética* considerado como *parte-de* um *frame amostra* foi mapeado para uma *entidade composiçãoDiagenética* considerada como uma *agregação* da *entidade amostra*. Outros relacionamentos *parte-de* especificados entre frames principais foram mapeados de forma similar no modelo de dados.

- 3) relacionamentos *lista-de* entre *frames principais* e *frames adicionais* foram mapeados para cardinalidades múltiplas entre *entidades principais* mapeadas;

Por exemplo, um *frame descrição-diagenética* (classificado como adicional) foi especificado como *parte-de* um *frame amostra* no modelo de frames de aplicação desenvolvido. No entanto, esse *frame descrição-diagenética* é uma *lista-de frames composição-diagenética*, o qual foi classificado como frame principal (Figura 4.6). Para esse caso, apenas a *entidade composiçãoDiagenética* foi mapeada, sendo que a cardinalidade do relacionamento desta entidade com a *entidade amostra* também mapeada é múltipla, tal como demonstrado na Figura 5.3. Para viabilizar esse relacionamento múltiplo, um atributo chave adicional (*IDcomposiçãoDiagenética*) para a *entidade composiçãoDiagenética* foi criado. Outros relacionamentos *lista-de* como esses foram mapeados de forma similar no modelo de dados.

- 4) relacionamentos *um-de* entre *frames principais* foram mapeados para cardinalidades únicas entre *entidades principais* mapeadas;

Por exemplo, apenas um *frame microscopia* é *parte-de* um *frame amostra* no modelo de frames desenvolvido. Para esse caso, a cardinalidade da *entidade microscopia* é única em relação à *entidade amostra*, assim como demonstrado na Figura 5.3. Demais relacionamentos desse tipo foram mapeados de forma similar no modelo de dados.

- 5) atributos multivalorados de *frames principais* identificados (atributos com domínios tipo *lista-de*) foram mapeados como *entidades secundárias* associadas a *entidade principal* mapeada do frame principal;

Por exemplo, o *frame macroscopia* foi detalhado com dois atributos multivalorados: *tamanhosGrãos* e *estruturas*, entre outros atributos especificados. Tais atributos multivalorados foram mapeados respectivamente para as *entidades tamanhosGrãosMacroscopia* e *estruturasMacroscopia*. O relacionamento entre estas entidades secundárias foi detalhado de forma a permitir múltiplas entradas de *tamanhosGrãos* e *estruturas*, assim como descrito por *lista-de* valores no *frame macroscopia*. Para representar essa cardinalidade múltipla, um atributo chave adicional foi criado para cada uma das entidades secundárias. Demais atributos multivalorados de outros frames especificados foram mapeados de forma similar no modelo de dados.

- 6) atributos não-multivalorados de *frames principais* (atributos com domínios tipo *um-de*) compostos por valores distintos foram mapeados para mais de um atributo de *entidades mapeadas*;

Por exemplo, o *frame composição-diagenética* possui o atributo *relaçãoParagenética* definido por um domínio de valores distintos, nesse caso: um nome da "relação paragenética" ("cobrindo", "substituindo", etc) e um nome do "constituente diagenético" que está envolvido nessa "relação paragenética". Para esse caso, o atributo originalmente especificado como *relaçãoParagenética* no *frame composição-diagenética* foi mapeado para os atributos *relaçãoParagenética* e *relaçãoParagenéticaNomeMineral* na entidade *composiçãoDiagenética*. Tal estratégia de mapeamento foi usada para todos os atributos não-multivalorados que possuíam domínios similares a esse, assim como exemplificado na Figura 5.3 por *<atributosFrames>_para_<atributosEntidadesPrincipais>* e por *<outrosAtributosEntidadesSecundárias>*.

Os relacionamentos *é-um* especificados no modelo de frames não foram mapeados para o modelo de dados na versão atual desse trabalho. Tais relacionamentos detalhados no modelo de frames representam restrições organizacionais típicas em modelos de frames. Por exemplo, todo frame detalhado no modelo de frames desenvolvido *é-um Objeto*, pois todo frame é "filho-de" um *frame raiz*. Como esses relacionamentos não são típicos do domínio de aplicação e, desta forma, não são utilizados para fins de inferência na versão atual desse projeto, não foi necessário mapeá-los ou mesmo implementá-los no modelo de dados mapeado. Todavia, tais relacionamentos devem ser mapeados em passos futuros do projeto *PetroGrapher*, da mesma forma que BD mais semânticos que o BD relacional também devem ser empregados.

O mapeamento desenvolvido entre os modelo de conhecimento de domínio e o modelo de dados de aplicação poderia ser mais genérico, como também poderia ser descrito explicitamente em uma base de mapeamentos conhecida pelo sistema *PetroGrapher* (tal como demonstrado na Figura 5.2 como *Mapeamentos Especificados*). A vantagem dessa alternativa é tornar o conhecimento de domínio especificado mais independente de aspectos de aplicação, tornando os componentes do sistema *PetroGrapher* mais genéricos e reusáveis. Por sua vez, a desvantagem da representação explícita dos mapeamentos é tornar o sistema resultante mais ineficiente, pois qualquer acesso a BD envolveria efetivamente um consulta aos mapeamentos declarados explicitamente.

O modelo de dados desse domínio de aplicação é apresentado na Figura 5.4 em um formalismo de objetos simplificado.

O modelo de dados apresentado na Figura 5.4 foi empregado como um esquema para um BD de armazenamento de descrições de amostras de rochas. Para isso, cada umas das entidades representadas foi mapeada para uma tabela em um BD relacional, a exceção da *entidade amostra*. Esta entidade possui um único atributo *identificação da amostra - IDamostra*, e tal atributo foi representado em todas as tabelas do BD de armazenamento de amostras, assim explicitamente indicando a amostra de rocha que cada uma das entradas do banco está associada.

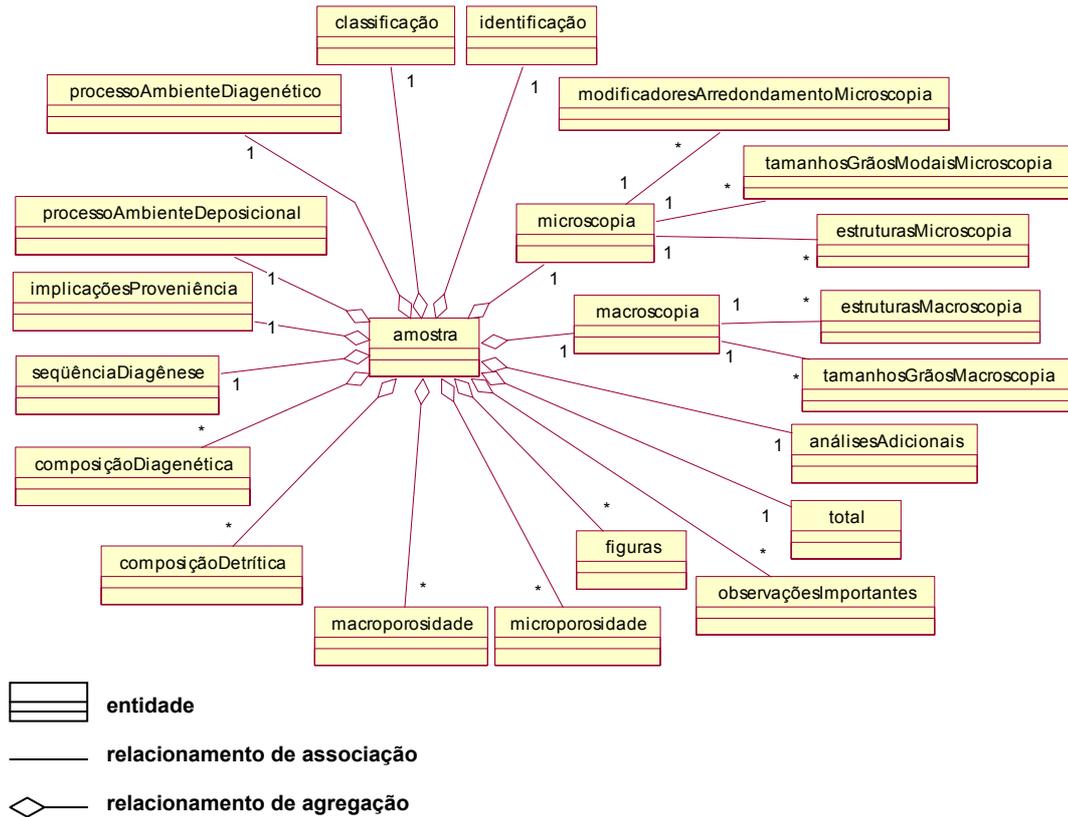


FIGURA 5.4 - Modelo de dados do domínio de aplicação

A indicação explícita da identificação da amostra em todas as entradas do BD de descrição de amostras facilitou o projeto e implementação do *gerador de consultas* previsto na *arquitetura de inferência para interpretação de rochas*. Durante o processo de interpretação, esse *gerador de consultas* formata e executa consultas SQL sobre o BD de descrições de amostras, como foi proposto na arquitetura de inferência para interpretação de rochas. Para realizar a interpretação de uma amostra em especial, tais consultas devem explicitamente ser aplicadas à descrição de uma mesma amostra dentre as muitas descrições armazenadas no *banco de descrições de amostras de rocha*.

5.4 Um mapeamento do modelo de conhecimento inferencial de domínio especificado para um modelo de dados

A estrutura composicional dos grafos de conhecimento e expressões de domínio modelados foi mapeada para um esquema de dados. Esse esquema é usado como um modelo para uma base de armazenamento dos dados que compõem o conhecimento inferencial de domínio de aplicação especificado. Tal base permite armazenar, analisar e alterar a estrutura de informação desse conhecimento, mas não permite que seja representado explicitamente no BD muitas das restrições de integridade conceituais que são expressas com um modelo de conhecimento semanticamente mais expressivo, tal como é o modelo de grafos de conhecimento.

O modelo dessa base de armazenamento de conhecimento inferencial é apresentado a seguir em um formalismo de modelo de objetos simplificado.

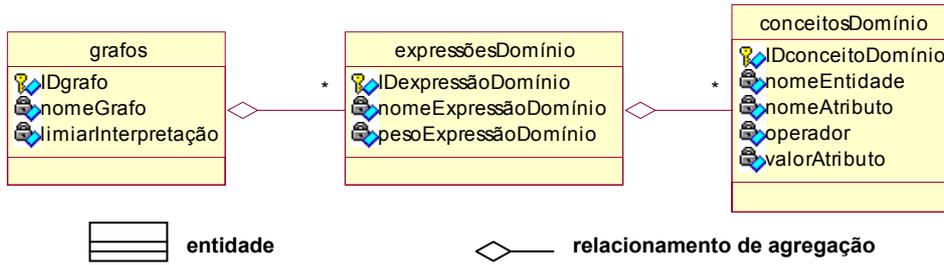


FIGURA 5.5 - Base de conhecimento inferencial de domínio

A Figura 5.5 apresenta o modelo de dados derivado da estrutura de informação que compõe o conhecimento inferencial de domínio modelado como grafos de conhecimento, expressões de domínio e conceitos de domínio. Esses modelos de conhecimento de domínio foram representados como entidades do modelo de dados. Os relacionamentos de agregação modelados entre tais entidades representam os principais relacionamentos entre os diferentes níveis de informação encontrados na estrutura composicional desse conhecimento inferencial de domínio.

A partir das entidades *grafos*, *expressõesDomínio* e *conceitosDomínio* apresentadas na Figura 5.5, foram geradas tabelas em um BD relacional. Essas tabelas são usadas como uma base de armazenamento para os dados inferenciais de domínio requeridos para interpretar ambientes diagenéticos em descrições de amostras de rocha, como previsto na arquitetura de inferência para interpretação de rochas.

Para evitar um mapeamento dinâmico entre frames e entidades durante o processo de inferência, assim tornando mais eficiente o *mecanismo de interpretação de ambientes diagenéticos* atualmente implementado, as entidades de aplicação especificadas no modelo de dados (Figura 5.4) foram usadas para reescrever manualmente todo o conhecimento inferencial de domínio adquirido junto ao especialista em Petrografia. Mais especificamente, as expressões de domínio modeladas foram descritas em função de entidades e atributos do esquema de dados do domínio. Para fazer isso, o processo de reescrita manteve a semântica descrita pelas expressões de domínio modeladas, pois tal semântica pode ser obscurecida pela utilização de um modelo menos expressivo. Dessa forma, toda a base de expressões de domínio de aplicação especificada foi mapeada para uma terminologia simbólica envolvendo entidades e atributos do esquema de dados de aplicação desenvolvido (Figura 5.4).

Sendo assim, na versão atual do *mecanismo de interpretação de ambientes diagenéticos*, as tabelas *expressõesDomínio* e *conceitosDomínio* armazenam a terminologia simbólica desse conhecimento inferencial de aplicação. Por exemplo, em vez de armazenar nessa base de conhecimento o conceito de domínio rotulado como *frame composição-diagenética* no modelo de frames especificado, foi armazenada a entidade rotulada como *composiçãoDiagenética* mapeada desse modelo de frames. Tal estratégia também foi usada para mapear os atributos dos conceitos de domínio componentes das expressões de domínio modeladas, os quais também foram armazenados na tabela *conceitosDomínio*. Logo, os exemplos apresentados na Tabela

5.3 já estão descritos em função da terminologia simbólica do esquema de dados de aplicação.

Em contrapartida, poderiam ser armazenadas nesse BD de conhecimento inferencial os nomes de conceitos e atributos de domínio descritos na terminologia do modelos de frames desenvolvido, assim como foram adquiridos junto aos especialistas em Petrografia. Seguindo essa abordagem, o mapeamento desses conceitos de domínio para entidades do BD deveria ser realizado a cada passo do processo de interpretação de uma amostra, pois o processo de inferência analisa uma descrição de amostra tal como está descrita no esquema de dados de aplicação que é usado para descrever o *banco de descrições de amostras de rocha*.

As tabelas seguintes apresentam exemplos dos grafos de conhecimento, expressões de domínio e conceitos de domínio armazenados na base de conhecimento inferencial de domínio construída.

TABELA 5.1 - Tabela grafos: exemplos de grafos de conhecimento armazenados no BD

IDgrafo	nomeGrafo	limiarInterpretação
05	Telodiagênese sob condições meteóricas	6
02	Eodiagênese continental meteórica sob clima seco	6
...

A Tabela 5.1 apresenta um exemplo de dois grafos de conhecimento modelados e armazenados na *base de conhecimento inferencial de domínio*, os quais são identificados com um identificador único denominado *IDgrafo*. Cada grafo de conhecimento detalhado nesse trabalho é armazenado como uma linha na tabela *grafos*. Por sua vez, a tabela seguinte apresenta um exemplo de algumas das expressões de domínio especificadas e armazenadas no BD, as quais são associadas a estes grafos de conhecimento exemplificados.

TABELA 5.2 - Tabela expressõesDomínio: exemplos de expressões de domínio armazenadas no BD

IDgrafo	IDexpressãoDomínio	nomeExpressãoDomínio	pesoExpressãoDomínio
02	01	Óxido de ferro/hidróxidos	1
02	02	Argilas infiltradas	3
...
05	01	Dissolução	3
05	02	Caulinização	5
...

As expressões de domínio exemplificadas na Tabela 5.2 são explicitamente associadas aos grafos de conhecimento também armazenados no BD, assim como todas as expressões de domínio detalhadas. Esse relacionamento de associação é indicado pela chave estrangeira *IDgrafo* (tabela *expressõesDomínio*), onde cada expressão de domínio armazenada tem como identificador único as colunas (*IDgrafo*, *IDexpressãoDomínio*). O valor de *IDexpressãoDomínio* é único para cada grafo de conhecimento especificado. A seguir, são apresentados exemplos de conceitos-atributos-valores que compõem as expressões exemplificadas na Tabela 5.2.

TABELA 5.3 - Tabela conceitosDomínio: exemplos de conceitos de domínio armazenados no BD

IDgrafo	IDexp. Domínio	IDconc. Domínio	nome Entidade	nome Atributo	oper.	valor Atributo
...
02	01	01	composição Diagenética	nome Mineral	=	Hematita
02	01	02	composição Diagenética	hábito	=	Cutícula
02	01	03	composição Diagenética	hábito	=	Franja
...
02	02	01	composição Diagenética	hábito	=	Cutícula
02	02	02	composição Diagenética	nome Mineral	=	Argilas infiltradas
...
05	01	01	macroporosidade	relação Paragenética	=	Dissolução de <constituente Diagenético>
05	01	02	macroporosidade	localização	=	Intersticial
...
05	02	01	composição Diagenética	relação Paragenética	=	Substituindo <constituente Diagenético>
...

Cada um dos conceitos-atributos-valores componentes das expressões de domínio modeladas são armazenados como linhas da tabela *conceitosDomínio*. Conceitos de domínio relacionados por AND em uma mesma expressão de domínio são identificados pelos valores de (*IDgrafo*, *IDexpressãoDomínio*) comuns. Por exemplo, "composiçãoDiagenética.nomeMineral = 'Argilas Infiltradas' AND composiçãoDiagenética.hábito = 'Cutícula'" são armazenados como (*IDgrafo* = '02', *IDexpressãoDomínio* = '02') em duas entradas na tabela *conceitosDomínio* (Tabela 5.3). Cada conceito de domínio armazenado é unicamente identificado em um grafo e uma expressão de domínio pela coluna *IDconceitoDomínio*, ou seja, a chave da tabela *conceitosDomínio* é (*IDgrafo*, *IDexpressãoDomínio*, *IDconceitoDomínio*).

Relações OR entre valores de atributos (*valorAtributo*) de um conceito de domínio de uma expressão de domínio são armazenadas na tabela *conceitosDomínio* de forma mais implícita. Caso duas linhas da tabela *conceitosDomínio* contenham valores para (*IDgrafo*, *IDexpressãoDomínio*) similares, essas linhas podem estar armazenando o mesmo conceito de domínio com diferentes valores para seus atributos. Para identificar estas entradas, valores similares armazenados para *nomeEntidade* e *nomeAtributo* indicam que existe uma relação OR entre os valores armazenados em *valorAtributo*. Por exemplo, "composiçãoDiagenética.hábito = 'Cutícula' OR composiçãoDiagenética.hábito = 'Franja'" são armazenados como (*IDgrafo* = '02', *IDexpressãoDomínio* = '01') e mesmos *nomeEntidade* = 'composiçãoDiagenética' e *nomeAtributo* = 'hábito', como apresentado na Tabela 5.3.

A forma de representação de relações AND entre conceitos de domínio e OR entre valores de atributos de conceitos de domínio apresentada foi usada para popular a tabela *conceitosDomínio*. As tabelas *expressõesDomínio* e *grafos* também foram populadas durante a fase de desenvolvimento do algoritmo de interpretação de ambientes diagenéticos.

5.5 Um algoritmo de inferência para interpretação de ambientes diagenéticos

O *mecanismo de interpretação de ambientes diagenéticos* foi desenvolvido a partir do modelo de comportamento inferencial especificado no PSM Interpretação de Ambientes Diagenéticos. Todavia, tal PSM não foi empregado para derivar uma especificação executável, como proposto em diferentes formalismos executáveis de descrição de PSMs. Nesse trabalho, o PSM Interpretação de Ambientes Diagenéticos proposto foi empregado como um modelo conceitual para compreender o raciocínio nesse domínio de problemas e, em seguida, projetar e implementar um algoritmo de inferência simbólico para interpretação de ambientes. Esse algoritmo simbólico desenvolvido em Lisp/CLOS implementa funcionalmente os passos primitivos de inferência que são executados entre os diferentes papéis de conhecimento especificados no PSM proposto.

A especificação funcional do algoritmo de inferência desenvolvido não representa explicitamente o conhecimento de controle especificado no PSM Interpretação de Ambientes Diagenéticos, tal como esse tipo de conhecimento poderia ser preservado no modelo de projeto do mecanismo de inferência. Essa opção de desenvolvimento foi tomada para que o mecanismo de inferência construído estivesse mais integrado com o sistema *PetroGrapher*, assim resultando em um módulo de raciocínio mais otimizado para essa aplicação, mas menos reusável entre domínios de aplicações.

Para realizar o processo de interpretação de rochas almejado, o *mecanismo de interpretação de ambientes diagenéticos* desenvolvido recupera os dados que compõem o conhecimento inferencial de domínio armazenados no BD para a *memória de trabalho estática*, assim como demonstrado na arquitetura de inferência proposta. Por exemplo, os dados dos grafos de conhecimento modelados e armazenados no BD são recuperados e organizados como listas de símbolos, tal como apresentado na Figura 5.6.

A lista da Figura 5.6 demonstra como todos os grafos de conhecimento modelados são mantidos na *memória de trabalho estática* durante o processo de inferência. Os dados sobre as expressões de domínio modeladas e recuperadas do BD, por sua vez, são também representados no formato de lista de símbolos, tal como apresentado na Figura 5.7.

A estrutura simbólica das expressões de domínio (Figura 5.7) mantida na *memória de trabalho estática* é analisada sucessivamente pelo *mecanismo de interpretação de ambientes diagenéticos*. Esse mecanismo seleciona cada uma das expressões de domínio simbolicamente especificadas, de forma a requerer esse

conhecimento para o *gerador de consultas*. O *gerador de consultas* então analisa individualmente esse conhecimento requerido, formatando e executando consultas SQL sobre o *banco de descrições de amostras de rochas*.

```
( (nomeGrafo1 limiarInterpretação1
  (nomeExpressãoDomínio11 pesoExpressãoDomínio11
   nomeExpressãoDomínio12 pesoExpressãoDomínio12
   ... nomeExpressãoDomínio1n pesoExpressãoDomínio1n) )
 (nomeGrafo2 limiarInterpretação2
  (nomeExpressãoDomínio21 pesoExpressãoDomínio21
   nomeExpressãoDomínio22 pesoExpressãoDomínio22
   ... nomeExpressãoDomínio2n pesoExpressãoDomínio2n) )
 ... (nomeGrafon limiarInterpretaçãon
  (nomeExpressãoDomínion1 pesoExpressãoDomínion1
   nomeExpressãoDomínion2 pesoExpressãoDomínion2
   ... nomeExpressãoDomínionn pesoExpressãoDomínionn) ) )
```

FIGURA 5.6 - Estrutura simbólica de representação de grafos de conhecimento

```
( (nomeExpressãoDomínio1
  (nomeConceito11 nomeAtributo11 operador11
   (valorAtributo111 valorAtributo112 ... valorAtributo11n) )
  (nomeConceito12 nomeAtributo12 operador12
   (valorAtributo121 valorAtributo122 ... valorAtributo12n) )
  ... (nomeConceito1n nomeAtributo1n operador1n
   (valorAtributo1n1 valorAtributo1n2 ... valorAtributo1nn) ) )
 (nomeExpressãoDomínio2
  (nomeConceito21 nomeAtributo21 operador21
   (valorAtributo211 valorAtributo212 ... valorAtributo21n) )
  (nomeConceito22 nomeAtributo22 operador22
   (valorAtributo221 valorAtributo222 ... valorAtributo22n) )
  ... (nomeConceito2n nomeAtributo2n operador2n
   (valorAtributo2n1 valorAtributo2n2 ... valorAtributo2nn) ) )
 ... (nomeExpressãoDomínion
  (nomeConceiton1 nomeAtributon1 operadorn1
   (valorAtributon11 valorAtributon12 ... valorAtributon1n) )
  (nomeConceiton2 nomeAtributon2 operadorn2
   (valorAtributon21 valorAtributon22 ... valorAtributon2n) )
  ... (nomeConceitonn nomeAtributonn operadornn
   (valorAtributonn1 valorAtributonn2 ... valorAtributonnn) ) ) )
```

FIGURA 5.7 - Estrutura simbólica de representação de expressões de domínio

Para exemplificar esse processo de geração de consultas SQL sobre a descrição de uma amostra sendo analisada, a expressão de domínio "caulinização" será usada. Tal expressão, devidamente mapeada para uma terminologia de entidades e atributos do esquema de dados usado para armazenar descrições de amostras (Figura 5.4), é simbolicamente analisada pelo *gerador de consultas*, assim como descrito na Figura 5.8.

Para cada um dos conceitos-atributos-valores de domínio indicados na estrutura simbólica de representação da Figura 5.8, o *gerador de consultas* formata e executa uma consulta SQL sobre a descrição da amostra. Digamos que uma amostra com "IDamostra = 'Campos 3012'" esteja sendo analisada. Para esta amostra, as consultas SQL exemplificadas na Figura 5.9 seriam automaticamente geradas pelo *gerador de consultas*.

```

("caulinização"
 ("composiçãoDiagenética" "nomeMineral" "=" ("Caulim") )
 ("composiçãoDiagenética" "hábito" "="
 ("Lamelas" "Agregados em livrinho" "Agregados vermiculares" "Maciça") )
 ("composiçãoDiagenética" "relaçãoParagenética" "="
 ("Substituindo <constituenteDiagenético>") )
 ("composiçãoDiagenética" "relaçãoParagenéticaNomeMineral" "="
 ("Albíta" "Ilíta" "Clorita") )
 ("composiçãoDiagenética" "localização" "="
 ("Recobrimento de poro intragranular" "Preenchimento de poro intragranular")))

```

FIGURA 5.8 - Especificação simbólica da expressão de domínio "caulinização"

```

SELECT DISTINCT nomeMineral
FROM composiçãoDiagenética
WHERE IDamostra = 'campos 3012' AND
      (nomeMineral = 'Caulim');

SELECT DISTINCT hábito
FROM composiçãoDiagenética
WHERE IDamostra = 'campos 3012' AND
      (hábito = 'Lamelas' OR
       hábito = 'Agregados em livrinho' OR
       hábito = 'Agregados vermiculares' OR
       hábito = 'Maciça');

<Demais consultas SQL geradas para
os conceitos-atributos-valores da
expressão de domínio "caulinização">

```

FIGURA 5.9 - Consultas SQL geradas a partir da especificação simbólica da expressão de domínio "caulinização"

Como as consultas da Figura 5.9 representam relações AND entre conceitos de domínio da expressão de domínio "caulinização", o *gerador de consultas* foi desenvolvido para executar cada uma das consultas individualmente, desde que elas retornem individualmente algum valor descrito na amostra "campos 3012". Caso uma consulta executada não retorne nenhum valor, ou seja, o conceito de domínio requerido não tenha sido descrito pelo usuário na amostra sendo analisada, o *gerador de consultas* interrompe o processo de casamento de padrões, assim otimizando o processo de validação de uma expressão de domínio então selecionada para análise.

Portanto, o casamento de padrões de conceitos-atributos-valores e implementado como um conjunto de consultas SQL, e o casamento de padrões de expressões de domínio e implementado simbolicamente a partir dos resultados retornados por essas consultas. Neste último caso, por exemplo, se algum valor para todos os atributos "nomeMineral, hábito, relaçãoParagenética, relaçãoParagenéticaNomeMineral e localização" (como exemplificado na Figura 5.9) for retornado como resultado para cada uma das consultas executadas, isso indica que a relação AND entre os conceitos de domínio da expressão de domínio "caulinização" pode ser atendida. Caso essa relação seja atendida, a expressão de domínio foi

encontrada (validada) na descrição da amostra. Desta forma, a expressão de domínio "caulinização" então é considerada ativada no processo de especificação das soluções.

A estratégia simbólica de inferência demonstrada como um processo de análise de estruturas simbólicas de representação, bem como pela geração e execução de consultas SQL sobre um *banco de descrições de amostras de rocha*, é empregada para todas as expressões de domínio modeladas. Isso permite de analisar todos os conceitos-atributos-valores do conhecimento inferencial de domínio representado como expressões de domínio, assim como previsto no PSM Interpretação de Ambientes Diagenéticos.

Uma estruturação simbólica de representação similar à apresentada na Figura 5.7 também é usada para organizar dados de amostras de rochas na *memória de trabalho dinâmica*. Tais dados são resultantes de consultas SQL geradas e realizadas sobre o *banco de descrições de amostras de rocha* durante o processo de validação das expressões de domínio mantidas na *memória de trabalho estática*. A partir dessas expressões de domínio validadas na descrição de amostra, os limiares de conclusão dos grafos de conhecimento mantidos na *memória de trabalho estática* são usados para especificar as soluções que podem ser alcançadas.

Como foi demonstrado no PSM Interpretação de Ambientes Diagenéticos, as expressões validadas (*nomeExpressãoDomínioAtivada*) segundo seus respectivos pesos (*pesoExpressãoDomínioAtivada*) podem indicar diferentes interpretações de ambientes diagenéticos (*nomeGrafoValidado*), desde que respeitem limiares de interpretação (*limiarAlcançado*). Essa estrutura de interpretação é então representada simbolicamente na *memória de trabalho dinâmica*, como apresentado na Figura 5.10.

```
( (nomeGrafoValidado1 limiarAlcançado1
  (nomeExpressãoDomínioAtivada11 pesoExpressãoDomínioAtivada11
  nomeExpressãoDomínioAtivada12 pesoExpressãoDomínioAtivada12
  ... nomeExpressãoDomínioAtivada1n pesoExpressãoDomínioAtivada1n) )
(nomeGrafoValidado2 limiarAlcançado2
  (nomeExpressãoDomínioAtivada21 pesoExpressãoDomínioAtivada21
  nomeExpressãoDomínioAtivada22 pesoExpressãoDomínioAtivada22
  ... nomeExpressãoDomínioAtivada2n pesoExpressãoDomínioAtivada2n) )
... (nomeGrafoValidadon limiarAlcançadon
  (nomeExpressãoDomínioAtivadan1 pesoExpressãoDomínioAtivadan1
  nomeExpressãoDomínioAtivadan2 pesoExpressãoDomínioAtivadan2
  ... nomeExpressãoDomínioAtivadann pesoExpressãoDomínioAtivadann) ) )
```

FIGURA 5.10 - Estrutura simbólica de representação de grafos de conhecimento validados

Na versão atual do algoritmo de inferência desenvolvido, a interpretação final de um ambiente diagenético de uma amostra de rocha analisada é formada pelo nome do(s) ambiente(s) diagenético(s) resultante(s) (*nomeGrafoValidado*) e pelos nomes das expressões de domínio encontradas na descrição da amostra de rocha (*nomeExpressãoDomínioAtivada*). As respectivas feições petrográficas (conceitos-atributos-valores de domínio) descritas na amostra que permitiram validar as expressões de domínio encontradas também são armazenadas e retornadas como parte da interpretação resultante. Esses dados são organizadas na *memória de trabalho*

dinâmica em um formato similar ao apresentado na Figura 5.10 e na Figura 5.7. Adicionalmente, também são apresentados como resultado todos os demais critérios de interpretação testados, assim demonstrando a amplitude de informações analisada pelo algoritmo de interpretação de ambientes diagenéticos, como também apresentando uma explicação simplificada da interpretação que foi alcançada a partir da descrição de rocha analisada.

O algoritmo de inferência para interpretação de ambientes diagenéticos desenvolvido foi testado em um pequeno conjunto de descrições de amostras (19 descrições ao todo). Tais descrições de amostras de rocha somente são realizadas por geólogos com algum nível de especialização em Petrografia. Desta forma, não foi possível obter petrógrafos que pudessem descrever um número significativo de amostras de rochas, os quais poderiam validar com mais precisão o conhecimento inferencial de domínio adquirido e modelado como grafos de conhecimento e expressões de domínio, bem como realizar ajustes no algoritmo de inferência desenvolvido.

O algoritmo de interpretação foi testado em um conjunto de descrições realizadas por um geólogo novato em Petrografia (aluno do curso de geologia - 16 descrições das 19 disponíveis), e em um conjunto de descrições parciais realizadas pelo especialista em Petrografia (3 descrições das 19 disponíveis). Como pôde ser notado, as descrições realizadas pelo geólogo novato disponíveis eram mais abrangentes, mas descreviam apenas feições de rocha mais simplificadas, assim expressando poucas feições diagnósticas. As descrições realizadas pelo especialista em Petrografia, por sua vez, eram menos abrangentes, mas indicavam explicitamente os ambientes diagenéticos modelados.

Além disso, a inferência foi validada contra uma base de casos manualmente descritos pelo especialista. Para as mesmas feições descritas, o sistema é capaz de sugerir os mesmos ambientes. A capacidade do mecanismo de inferência de extrair conclusões é limitada pela capacidade do usuário em reconhecer e descrever as feições de rocha mais diagnósticas da amostra. Muitas dessas feições são de difícil reconhecimento por usuários não-especialistas, o que limita a aplicabilidade do sistema. Mesmo assim, conforme o especialista em Petrografia, qualquer sugestão de interpretação derivada corretamente seria bastante útil para esse domínio de problemas, pois poucos geólogos demonstram conhecer como resolver esse problema mesmo de forma simplificada.

6 Conclusões

O trabalho apresentado focaliza a noção de PSMs, os quais são usados como modelos de processos de raciocínio, possibilitando compreender o raciocínio e implementá-lo em computador. Para isso, foi realizada uma revisão da KE atual segundo o ponto de vista de PSMs e foram descritos os passos de desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos. Tais PSMs orientaram a descrição de um esquema de raciocínio, permitindo desenvolver uma arquitetura de raciocínio em termos de conhecimento e dados. Por fim, essa arquitetura foi usada para construir um algoritmo de inferência simbólico para o sistema *PetroGrapher*.

O presente trabalho propõe e caracteriza dois PSMs inéditos: o PSM Interpretação de Rochas e o PSM Interpretação de Ambientes Diagenéticos. Esses PSMs são os componentes centrais de um esquema de raciocínio para interpretação de rochas usável no domínio da Petrografia Sedimentar, e possivelmente reusável em domínios de problemas similares a este.

O desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos é melhor caracterizado como um estudo de caso prático de especificação de um esquema de raciocínio para uma tarefa de raciocínio real. Para isso, tal domínio de problemas complexo e intensivo em conhecimento foi analisado em profundidade, permitindo refinar os modelos de conhecimento e dados de domínio especificados antes e durante o desenvolvimento desse trabalho. Entre outras coisas, este estudo de caso pode ser útil para transformar a proposta de especificação de conhecimento inferencial sintetizada na noção de PSMs em uma alternativa conceitual e prática que caracterize requisitos reais de problemas abordados por KBSs. Desta forma, o trabalho descreve um processo de especificação, refinamento e mapeamento de modelos de conhecimento e dados, tal como é compreendido o desenvolvimento de mecanismos de inferência de KBSs segundo abordagens recentes da KE.

O trabalho focaliza a caracterização de dois PSMs corretamente expressos e conseqüentemente compreendidos, tanto pelo modelador, quanto por um especialista em Petrografia. As descrições de competências, especificações operacionais e requisitos/suposições dos métodos estabelecem competências mais precisamente especificadas, em detrimento de PSMs mais simplificados e com competências fracamente delimitadas. A especificação detalhada desses PSMs foi requerida no projeto *PetroGrapher*, pois somente assim seria expresso um esquema de raciocínio para interpretação de rochas que pudesse alcançar um desempenho de solução de problemas próximo ao desempenho do petrógrafo especialista entrevistado. Contudo, tal desempenho ainda deve ser testado em um conjunto mais abrangente de casos reais de interpretação de rochas, mas já pode ser considerado expresso conceitualmente nos PSMs demonstrados.

Os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos são descritos como modelos de análise de um mesmo processo de raciocínio. Tais PSMs são usados para compreender o raciocínio desenvolvido e projetar uma arquitetura de inferência para interpretação de rochas. Essa arquitetura então é usada como um projeto envolvendo conhecimento e dados para uma tarefa abordada no

projeto *PetroGrapher*, assim estruturando o desenvolvimento de um módulo inferencial para o sistema *PetroGrapher*.

Todavia, a arquitetura de raciocínio para interpretação de rochas proposta é considerada como uma validação prática dos PSMs propostos, como também é um módulo de software para uma tarefa importante no processo de caracterização de reservatórios de óleo em Petrografia Sedimentar. Portanto, a descrição de uma arquitetura é compreendida como uma forma de validação prática de modelos conceituais de raciocínio, permitindo explicitamente considerar requisitos simbólicos de conhecimento e dados previamente existentes, os quais são abstraídos na especificação de PSMs.

Os PSMs propostos são descritos em níveis de abstração diferentes, mas ambos são descritos no nível de conhecimento. O PSM Interpretação de Rochas modela mais explicitamente os resultados de uma análise cognitiva de processos de raciocínio desenvolvidos por especialistas em domínios orientados por imagens, assim como proposto por Abel [ABE 2001]. Esse PSM é caracterizado em função pacotes visuais e esquemas de rocha, as quais podem ser usados para analisar o raciocínio desenvolvido nestes domínios de problemas. O PSM Interpretação de Ambientes Diagenéticos, por sua vez, descreve um processo de inferência mais operacional. Tal processo é especificado em função de modelos de conhecimento inferencial de domínio, descrevendo mais fielmente uma forma de analisar e representar o conhecimento de solução de problemas requerido por mecanismos de raciocínio. Portanto, o PSM Interpretação de Ambientes Diagenéticos poderia ser mais explicitamente usado para orientar o projeto de um mecanismo de inferência para um KBSs, assim como foi usado para orientar o desenvolvimento de um módulo de inferência para o sistema *PetroGrapher*.

O estudo de caso de desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos pode ser descrito como um processo de adaptação/refinamento gradativo de especificações inferenciais. No desenvolvimento do PSM Interpretação de Rochas, os passos de desenvolvimento foram orientados no sentido de uma análise do raciocínio desenvolvido por petrógrafos especialistas. No desenvolvimento do PSM Interpretação de Ambientes Diagenéticos, por sua vez, o processo de especificação foi orientado no sentido das possibilidades de aquisição e representação de conhecimento de domínio experimentadas no projeto *PetroGrapher*. Embora considerando premissas de desenvolvimento diferentes, tal processo de desenvolvimento proporcionou uma melhor compreensão e refinamento do modelo de perícia de petrógrafos especialistas.

Os passos de especificação dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos foram caracterizados em função de propostas metodológicas de desenvolvimento de PSMs conhecidas, não sendo representativos de uma nova proposta metodológica. Mesmo assim, tais passos descrevem requisitos reais de desenvolvimento de PSMs, os quais podem ser usados para construir ferramentas de desenvolvimento para as metodologias descritas por Wielinga et al. [WIE 98], Fensel et al. [FEN 2000] e Teije et al. [TEI 96], entre outras. Dentre os princípios de desenvolvimento de PSMs experimentados, as validações dos modelos intermediários dos métodos junto ao especialista e também com exemplos reais adquiridos mostraram ser fundamentais para orientar o processo de modelagem. Logo,

alguma validação sistemática é requerida em propostas práticas de desenvolvimento de PSMs, assim facilitando a compreensão das possíveis adaptações requeridas sobre as especificações dos métodos.

O processo de desenvolvimento dos PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos, bem como o refinamento desenvolvido no esquema de raciocínio de interpretação de rochas como um todo, foi guiado por suposições. Tal estratégia permitiu delimitar a competência dos PSMs especificados, bem como a amplitude e aplicabilidade do esquema de raciocínio proposto. Contudo, a identificação de suposições típicas do domínio de problemas sendo abordado demandou grande esforço de aquisição de conhecimento.

As suposições estabelecidas não foram, e continuam não sendo, descritas formalmente em termos de tarefas e conceitos de domínios simplificados, assim como proposto pela metodologias de desenvolvimento de PSMs analisadas. A maioria das suposições estabelecidas dizem respeito ao esquema de raciocínio como um todo, caracterizando o comportamento de solução de problemas derivado pelo conjunto de modelos detalhados. Mesmo assim, as suposições estabelecidas podem orientar futuros passos de aquisição de conhecimento a serem desenvolvidos no projeto *PetroGrapher*, pois muitas delas expressam simplificações atualmente modeladas. Logo, tais simplificações podem ser minimizadas por novos passos de aquisição de conhecimento de domínio, tornando o esquema de raciocínio para interpretação de rochas mais completo e efetivo.

Conforme o petrógrafo especialista, os PSMs Interpretação de Rochas e Interpretação de Ambientes Diagenéticos podem ser usados para explicar a solução de tarefas de interpretação comuns no domínio de problemas da Petrografia Sedimentar. Portanto, há grande possibilidade que os PSMs especificados sejam reutilizados em tarefas similares a tarefa de interpretação de ambientes diagenéticos, tais como, por exemplo: a tarefa de "interpretação de ambientes detríticos" e a tarefa de "caracterização de proveniência", ambas também desenvolvidas no domínio do projeto *PetroGrapher*. Desta forma, tais PSMs podem ser usados como modelos de análise dessas novas tarefas de raciocínio, facilitando a compreensão do raciocínio desenvolvido e orientado a aquisição de conhecimento de domínio para viabilizar os passos de inferências dos PSMs propostos.

Outras tarefas de interpretação genéricas, como a interpretação de "imagens de raios-X" ou "ecografia", também possivelmente podem empregar os PSMs propostos como uma estrutura de raciocínio básica para adaptação. Além disso, conforme o especialista em Petrografia, tarefas em domínio de problemas similares a este, tal como o domínio das rochas carbonáticas, são prováveis candidatas a reutilizar os PSMs propostos, pois ambos podem representar tarefas típicas da Petrografia. De forma mais ampla, problemas de interpretação são potenciais aplicações de KBSs, mas ainda permanecem pouco formalizados e compreendidos. Logo, esse trabalho possivelmente pode vir a facilitar a compreensão e a especificação de modelos de interpretação genéricos reusáveis, os quais são requeridos para facilitar o desenvolvimento de KBSs em diferentes domínios de aplicação.

Anexo 1

Modelo de Frames da Aplicação *PetroGrapher*

Frame Identification	
Is-a	Object
Part-of	Frame Sample
Sample-ID	string(20), value-unique
Thin-ID	string(20)
Unit	string(60)
Basin	string(40)
Field	string(20)
Well	string(20)
Depth	real, range [0.0 - 9999.99]
Place	string(40)
State	string(2)
Country	string(15)
Use	string(80), list-of [Depositional, Diagenetic, Ecologic, Paleogeographic/paleogeologic, Provenance, Reservoir, Stratigraphic, Other use]
Date	date, [DD/MM/YYYY]
Petrographer	string(20)

Frame Depositional-Environment-Process	
Is-a	Object
Part-of	Frame Sample
Sample-ID	string(20), one-of <Sample-ID>
Depositional-environment	string(256)

Frame Additional-Analysis	
Is-a	Object
Part-of	Frame Sample
Sample-ID	string(20), one-of <Sample-ID>
Analysis	string(80)
Description	string(256)

Frame Provenance-Implications	
Is-a	Object
Part-of	Frame Sample
Sample-ID	string(20), one-of <Sample-ID>
Provenance	string(256)

Frame Diagenetic-Environment-Process	
Is-a	Object
Part-of	Frame Sample
Sample-ID	string(20), one-of <Sample-ID>
Diagenetic-environment	string(256)

Frame Important-Observations	
Is-a	Object
Part-of	Sample
Sample-ID	string(20), one-of <Sample-ID>
Observation	string(256)

Frame Classification	
Is-a	Object
Part-of	Frame Sample
Sample-ID	string(20), one-of <Sample-ID>
GrainSize	string(20), one-of [gravel, very coarse sand, coarse sand, medium sand, fine sand, very fine sand, silt, clay]
Sorting	string(20), one-of [very well sorted, well sorted, sorted, moderately sorted, poorly sorted]
Accessory	string(80), one-of [Detrital-Constituent]
RockName	string(30), one-of [Folk's Classification]

Frame Macroscopic	
Is-a	Object
Part-of	Frame Sample
Sample-ID	string(20), one-of <Sample-ID>
Color	string(20)
Structures	string(90), list-of [parallel bedding, cross bedding, normal-graded bedding, reverse-graded bedding, massive bedding, parallel lamination, climbing lamination, liner lamination, flaser lamination, massive, bioturbation, fluidized, load, folded, fractured, stylolites, nodules, crystalline, crusts, roots, spotted, vesicular, amigdaloidal, flow, bioconstruction, bioacumulation], MAX [3 occurrences]
Scales	string(45), list-of [<number and metric unit (mm, cm)> or <number - number and metric unit (mm, cm)>], range [0 - 100], MAX [3 occurrences], [ex.: 23 cm, 12 - 21 mm]
Grain-sizes	string(40), list-of [boulder, cobble, pebble, granule, gravel, very coarse sand, coarse sand, medium sand, fine sand, very fine sand, silt, clay], MAX[2 occurrences].
Sorting	string(20), one-of [very well sorted, well sorted, sorted, moderately sorted, poorly sorted]
Shape	string(30), one-of [equant, oblate, prolate] or [spheroidal, rod, blade, disc]
Orientation	string(20), one-of [homogeneous, parallel, oriented, imbricated, heterogeneous, chaotic]
Support	string(30), one-of [grain-supported, grain to matrix-supported, grain to cement-supported, matrix-supported, matrix to cement-supported, cement-supported]

Frame Detrital-Composition	
Is-a	Object
Part-of	Frame Detrital-Description
Set-of	Frame Detrital-Constituents
Sample-ID	string(20), one-of <Sample-ID>
MineralName	string(80), one-of [Detrital-Constituent].
ConstituentSet	string(40), one-of [Detrital quartz, Detrital feldspar, Plutonic rock fragments, Volcanic rock fragments, Sedimentary rock fragments, Metamorphic rock fragments, Micas/ chlorite, Heavy minerals, Intrabasinal grains, Detrital matrix, Other detrital constituents]
Location	string(40), one-of [in metamorphic rock fragment, in plutonic rock fragment, in sedimentary rock fragment, in volcanic rock fragment, in intrabasinal fragment, as monomineralic grain]
Amount	real, range [0.0 - 100.00]
NominalAmount	string(10), one-of [abundant, common, rare, trace].
Description	string(256)

Frame Microscopic	
Is-a	Object
Part-of	Frame Sample
Sample-ID	string(20), one-of <Sample-ID>
GrainSize	string(20), one-of [gravel, very coarse sand, coarse sand, medium sand, fine sand, very fine sand, silt, clay]
NumericGrainSize	string(15), one-of [<number and metric unit(mm)> or <number - number and metric unit(mm)>], range [0 - 100], [ex.: 23 mm, 12 - 21 mm]
ModalGrainSizes	string(80), list-of [gravel, very coarse sand, coarse sand, medium sand, fine sand, very fine sand, silt, clay], MAX [4 occurrences]
NumericModalGrainSize	string(15), list-of [<number and metric unit(mm)> or <number - number and metric unit(mm)>], range [0 - 100], MAX [4 occurrences], [ex: 23 mm, 2 - 3 mm]
Structures	string(90), list-of [parallel bedding, cross bedding, normal-graded bedding, reverse-graded bedding, massive bedding, parallel lamination, climbing lamination, liner lamination, flaser lamination, massive, bioturbation, fluidized, load, folded, fractured, stylolites, nodules, crystalline, crusts, roots, spotted, vesicular, amigdaloidal, flow, bioconstruction, bioacumulation], MAX [3 occurrences]
Gravel	real, range [0.0 - 100.00]
Sand	real, range [0.0 - 100.00]
Mud	real, range [0.0 - 100.00]
Sorting	string(20), one-of [very well sorted, well sorted, sorted, moderately sorted, poorly sorted]
SortingNumeric	real, range [0.0 - 100.00]
Roundness	string(20), one-of [well rounded, rounded, sub-rounded, sub-angular, angular]
RoundnessModifiers	string(75), list-of [soft intraclasts, intraclasts, pressure dissolution, corrosion, deformation, replacement, overgrowths], MAX [3 occurrences].
Sphericity	string(10), one-of [high, medium, low]
Orientation	string(20), one-of [parallel, imbricated, homogeneous, heterogeneous, oriented, chaotic]
Packing	string(15), one-of [loose, normal, tight, heterogeneous]
PackingIndex	one-of [<number and percent signal>], range [0 -100], [ex.: 30 %]
PointContacts	string(10), one-of [abundant, common, rare, trace] or [real, range [0.0 - 100.00]]
LongContacts	string(10), one-of [abundant, common, rare, trace] or [real, range [0.0 - 100.00]]
ConcavoConvexContacts	string(10), one-of [abundant, common, rare, trace] or [real, range [0.0 - 100.00]]
SuturedContacts	string(10), one-of [abundant, common, rare, trace] or [real, range [0.0 - 100.00]]
Support	string(30), one-of [grain-supported, grain to matrix-supported, grain to cement-supported, matrix-supported, matrix to cement-supported, cement-supported]

Frame Detrital-Description	
Is-a	Object
Part-of	Frame Sample
DetritalComposition	list-of [instance-of <Frame Detrital-Composition>]

Frame Pictures	
Is-a	Object
Part-of	Frame Pictures-Description
Picture-ID	string(20), value-unique
Sample-ID	string(20), one-of <Sample-ID>
Description	string(256)
Picture	BLOB, bitmap picture

Frame Pictures-Description	
Is-a	Object

Part-of	Frame Sample
Pictures	list-of [instance-of <Frame Pictures>]

Frame Diagenese-Sequence	
Is-a	Object
Part-of	Frame Sample
Sample-ID	string(20), one-of <Sample-ID>
Diagenese-Description	string(512)

Frame Diagenetic-Description	
Is-a	Object
Part-of	Frame Sample
DiageneticComposition	list-of [instance-of <Frame Diagenetic-Composition>]

Frame Diagenetic-Nomenclature	
Is-a	Object
Part-of	Frame Sample
MineralName	list-of [instance-of < Frame Diagenetic-Constituents >]

Frame Diagenetic-Constituents	
Is-a	Object
Part-of	Frame Diagenetic-Nomenclature
MineralName	string(80)
MineralSet	string(40), one-of [Silica, Feldspar, Infiltrated clays, Pseudomatrix clays, Authigenic clays, Zeolites, Carbonates, Sulphates, Sulfides, Iron oxides/ hydroxides, Titanium minerals, Other diagenetic constituents]

Frame Macroporosity	
Is-a	Object
Part-of	Frame Macroporosity-Description
Sample-ID	string(20), one-of <Sample-ID>
Porosity	string(20), one-of [Intergranular Porosity, Intragranular Porosity, Intracrystalline Porosity, Intercrystalline Porosity, Moldic Porosity, Grain fracture Porosity, Rock fracture Porosity, Shrinkage Porosity, Oversized Porosity, Lamellar Porosity]
Amount	real, range [0.0 - 100.00]
NominalAmount	string(20), one-of [abundant, common, rare, trace]
Location	string(30), one-of [Framework, Interstitial, Framework and Interstitial]
ParageneticRelations	string(120), one-of [<i>Primary</i> , <i>Dissolution of</i> : <one-of [Detrital-Constituent]>, <i>Dissolution of</i> : <one-of [Diagenetic-Constituent]>, <i>Within</i> <one-of [Detrital-Constituent]>, <i>Within</i> <one-of [Diagenetic-Constituent]>, <i>Contraction of</i> <one-of [Detrital-Constituent]>, <i>Contraction of</i> <one-of [Diagenetic-Constituent]>, <i>Fracturing of</i> <one-of [Detrital-Constituent]>, <i>Fracturing of</i> <one-of [Diagenetic-Constituent]>, <i>Decompaction</i>]
ParageneticRelation ConstituentSet	string(40), one-of [Silica, Feldspar, Infiltrated clays, Pseudomatrix clays, Authigenic clays, Zeolites, Carbonates, Sulphates, Sulfides, Iron oxides/hydroxides, Titanium minerals, Other diagenetic constituents, Detrital quartz, Detrital feldspar, Plutonic rock fragments, Volcanic rock fragments, Sedimentary rock fragments, Metamorphic rock fragments, Micas/chlorite, Heavy minerals, Intrabasinal grains, Detrital matrix, Other detrital constituents]
Modifier	string(40), one-of [Reduced by compaction, Reduced by cementation, Enlarged by dissolution]
Description	string(256)

Frame Diagenetic-Composition	
Is-a	Object
Part-of	Frame Diagenetic-Description
Set-of	Frame Diagenetic-Constituents
Sample-ID	string(20), one-of <Sample-ID>
MineralName	string(80), one-of [Diagenetic-Constituent]
ConstituentSet	string(40), one-of [Silica, Feldspar, Infiltrated clays, Pseudomatrix clays, Authigenic clays, Zeolites, Carbonates, Sulphates, Sulfides, Iron oxides/ hydroxides, Titanium minerals, Other diagenetic constituents]
Habit	string(20), one-of [instance-of <Frame Habits>]
Amount	real, range [0.0 - 100.00]
NominalAmount	string(20), one-of [abundant, common, rare, trace]
Location	string(40), one-of [intergranular continuous pore-lining, intergranular discontinuous pore-lining, intergranular pore-filling, intergranular discrete, intergranular displacive, intragranular replacive, intragranular pore-lining, intragranular pore-filling, intragranular discrete crystals, intragranular displacive, moldic pore-lining, moldic pore-filling, oversized pore-lining, oversized pore-filling, grain fracture-filling, grain fracture-lining, rock fracture-filling, rock fracture-lining, concretions/nodules, massive beds/lenses]
Modifier	string(40), one-of [dissolved, zoned, fractured, recrystallized]
ParageneticRelation	string(120), one-of [<i>Covering</i> <one-of [Diagenetic-Constituent]>, <i>Covering</i> <one-of [Detrital-Constituent]>, <i>Covered by</i> <one-of [Diagenetic-Constituent]>, <i>Replacing grain of</i> <one-of [Detrital-Constituent]>, <i>Replacing matrix of</i> <one-of [Detrital-Constituent]>, <i>Replacing</i> <one-of [Diagenetic-Constituent]>, <i>Replaced by</i> <one-of [Diagenetic-Constituent]>, <i>Alternated with</i> <one-of [Diagenetic-Constituent]>, <i>Engulfing</i> <one-of [Diagenetic-Constituent]>, <i>Engulfing</i> <one-of [Detrital-Constituent]>, <i>Engulfed by</i> <one-of [Diagenetic-Constituent]>, <i>Intergrown with</i> <one-of [Diagenetic-Constituent]>, <i>Overgrowing</i> <one-of [Diagenetic-Constituent]>, <i>Overgrowing</i> <one-of [Detrital-Constituent]>, <i>Overgrown by</i> <one-of [Diagenetic-Constituent]>, <i>Expanding</i> <one-of [Detrital-Constituent]>, <i>Compacted from</i> <one-of [Detrital-Constituent]>, <i>Within intergranular primary porosity</i> , <i>Within intergranular porosity after</i> <one-of [Diagenetic-Constituent]>, <i>Within intergranular porosity after detrital matrix</i> , <i>Within intragranular porosity in</i> <one-of [Detrital-Constituent]>, <i>Within intracrystalline porosity in</i> <one-of [Diagenetic-Constituent]>, <i>Within moldic porosity after</i> <one-of [Detrital-Constituent]>, <i>Within moldic porosity after</i> <one-of [Diagenetic-Constituent]>, <i>Within shrinkage porosity of</i> <one-of [Detrital-Constituent]>, <i>Within shrinkage porosity of</i> <one-of [Diagenetic-Constituent]>, <i>Within grain fracture porosity in</i> <one-of [Detrital-Constituent]>, <i>Within rock fracture porosity in</i> <one-of [Detrital-Constituent]>]
ParageneticRelation ConstituentSet	string(40), one-of [Silica, Feldspar, Infiltrated clays, Pseudomatrix clays, Authigenic clays, Zeolites, Carbonates, Sulphates, Sulfides, Iron oxides/ hydroxides, Titanium minerals, Other diagenetic constituents, Detrital quartz, Detrital feldspar, Plutonic rock fragments, Volcanic rock fragments, Sedimentary rock fragments, Metamorphic rock fragments, Micas/chlorite, Heavy minerals, Intrabasinal grains, Detrital matrix, Other detrital constituents]
Description	string(256)

Frame Habits-Nomenclature	
Is-a	Object
Part-of	Frame Sample
MineralName	list-of [instance-of <Frame Habits>]

Frame Habits	
Is-a	Object
Part-of	Frame Diagenetic-Composition
Part-of	Frame Habits-Nomenclature
HabitName	string(20), one-of [Blocky, Booklet, Botryoid, Bridge, Coarsely-crystalline, Coarse mosaic, Coating, Discrete crystal, Fibro-radiated, Fibrous, Fine mosaic, Framboid, Ingrowth, Internal sediment, Lamella, Large rhomb, Massive, Meniscus, Microcrystalline, Ooid, Outgrowth, Overgrowth, Parallel-prismatic, Pelletoid, Peloid, Pigment, Poikilotopic, Prismatic, Prismatic-radiated, Radiated, Rim, Rosette, Sheaf, Small rhomb, Spherulite, Vermicule]

Frame Detrital-Nomenclature	
Is-a	Object
Part-of	Frame Sample
MineralName	list-of [instance-of <Frame Detrital-Constituents>]

Frame Detrital-Constituents	
Is-a	Object
Part-of	Frame Detrital-Nomenclature
MineralName	string(80)
MineralSet	string(40), one-of [Detrital quartz, Detrital feldspar, Plutonic rock fragments, Volcanic rock fragments, Sedimentary rock fragments, Metamorphic rock fragments, Micas/chlorite, Heavy minerals, Intrabasinal grains, Detrital matrix, Other detrital constituents]

Frame Macroporosity-Description	
Is-a	Object
Part-of	Frame Sample
Macroporosity	list-of [instance-of <Frame Macroporosity>]

Frame Microporosity-Description	
Is-a	Object
Part-of	Frame Sample
Microporosity	list-of [instance-of <Frame Microporosity>]

Frame Microporosity	
Is-a	Object
Part-of	Frame Microporosity-Description
Sample-ID	string(20), one-of <Sample-ID>
MineralName	string(80), one-of [<Diagenetic-Constituent>, <Detrital-Constituent>]
ConstituentSet	string(40), one-of [Silica, Feldspar, Infiltrated clays, Pseudomatrix clays, Authigenic clays, Zeolites, Carbonates, Sulphates, Sulfides, Iron oxides/hydroxides, Titanium minerals, Other diagenetic constituents, Detrital quartz, Detrital feldspar, Plutonic rock fragments, Volcanic rock fragments, Sedimentary rock fragments, Metamorphic rock fragments, Micas/chlorite, Heavy minerals, Intrabasinal grains, Detrital matrix, Other detrital constituents]
Amount	real, range [0.0 - 100.00]
NominalAmount	string(10), one-of [abundant, common, rare, trace]

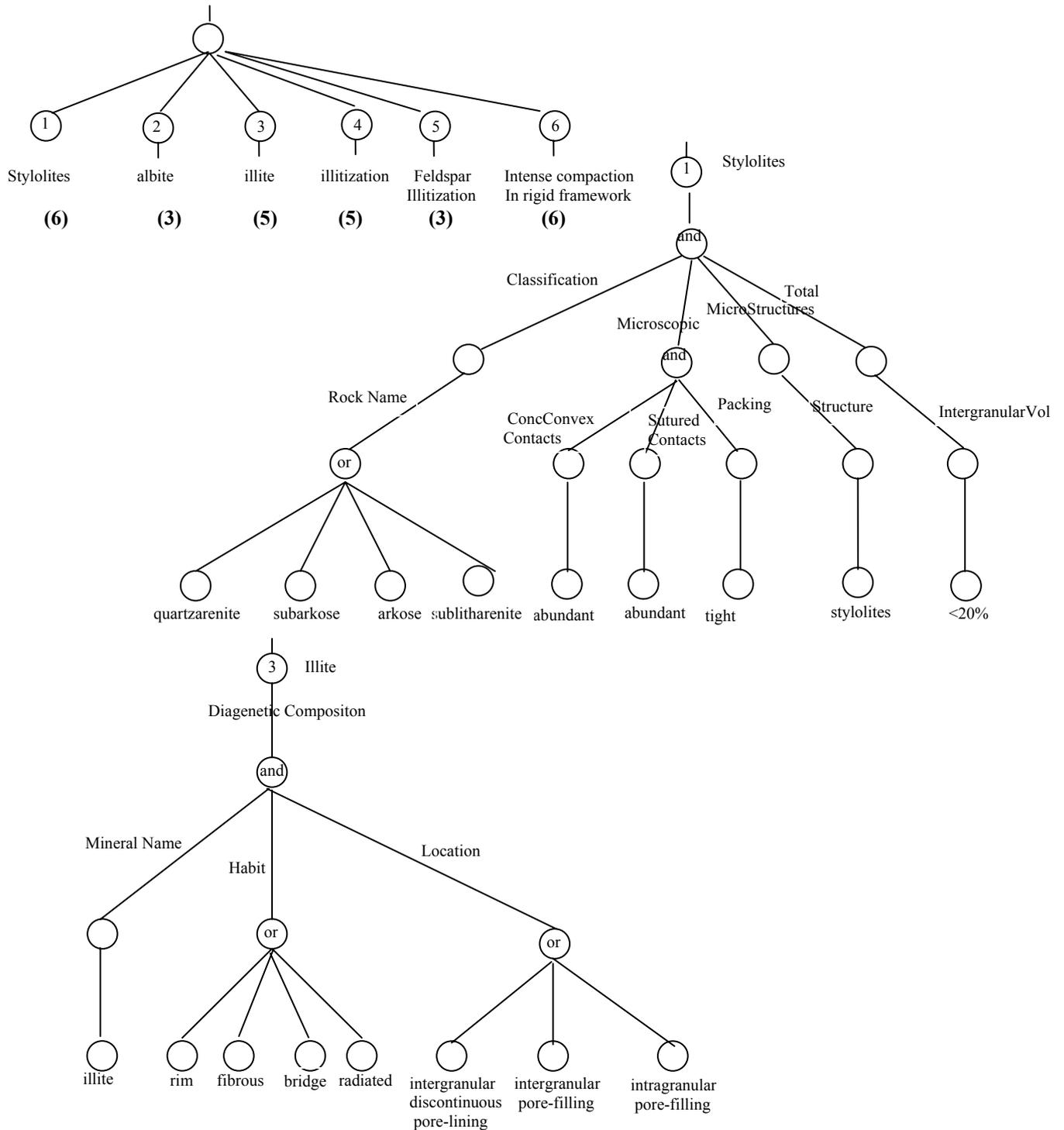
Frame Total		
Is-a	Object	
Part-of	Frame Sample	
Sample-ID	string(20), one-of <Sample-ID>	
Plutonic-RF	real, range [0.0 - 100.00]	
Volcanic-RF	real, range [0.0 - 100.00]	
Metamorphic-RF	real, range [0.0 - 100.00]	
Sedimentary-RF	real, range [0.0 - 100.00]	
Diagenetic	real, range [0.0 - 100.00]	
Intergranular-volume	real, range [0.0 - 100.00]	
Cement-volume	real, range [0.0 - 100.00]	
Matrix-volume	real, range [0.0 - 100.00]	
Grain-volume	real, range [0.0 - 100.00]	
Macroporosity-volume	real, range [0.0 - 100.00]	
Microporosity-volume	real, range [0.0 - 100.00]	
	Volume %	Index %
Folk's classification		
QFLFolk	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
QFolk	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
FFolk	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
LFolk	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
QFLoFolk	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
QoFolk	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
FoFolk	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
LoFolk	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
McBride's classification		
QFLMcBride	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
QMcBride	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
FMcBride	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
LMcBride	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
QFLoMcBride	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
QoMcBride	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
FoMcBride	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
LoMcBride	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Dickinson's classification		
Qt-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Qm-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Qp-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
F-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
P-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
K-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Lt-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
L-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Lv-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Ls-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Qto-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Qmo-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Qpo-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Fo-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Po-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Ko-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Lto-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Lo-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Lvo-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Lso-Dickinson	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]

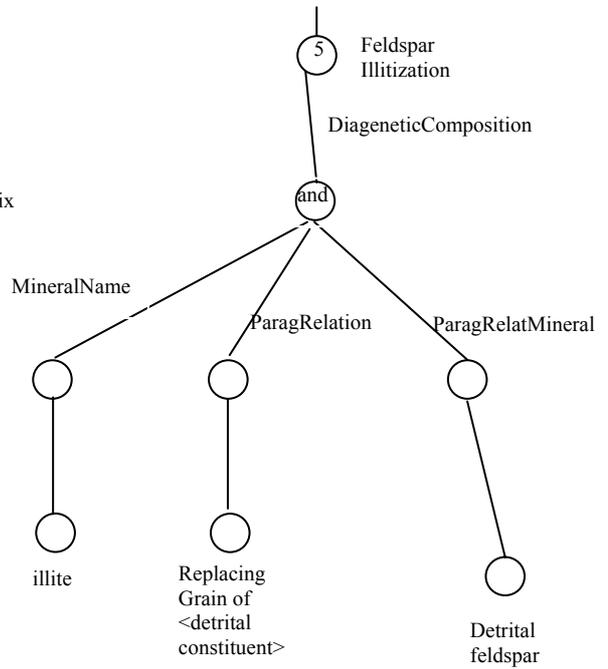
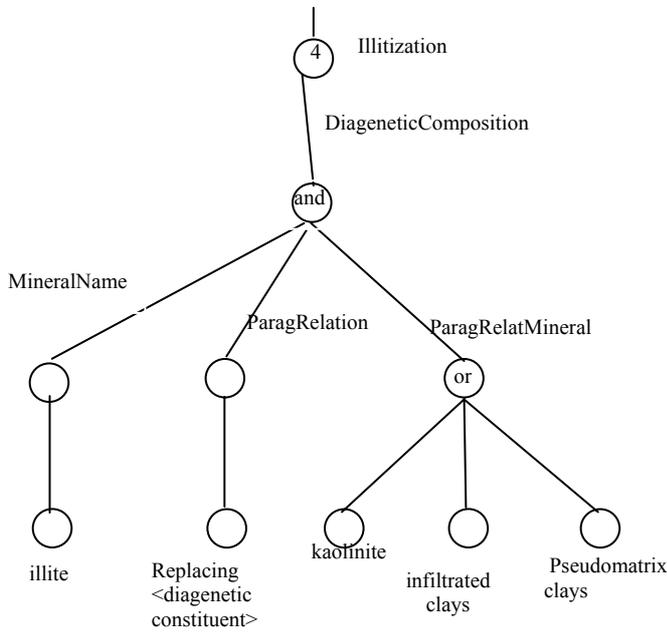
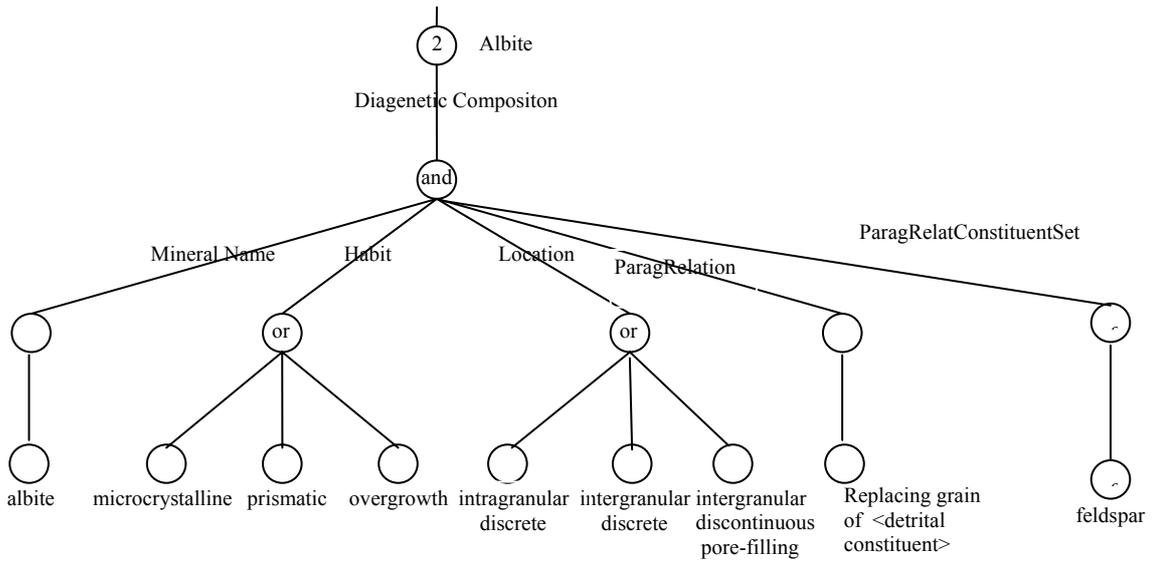
Zuffa's classification		
Total Essential Grains	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Extrabasinal grains	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
NCE	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
CE	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Intrabasinal grains	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
NCI	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
CI	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Total Essential Grains original	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Extrabasinal grains original	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
NCEo	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
CEo	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
Intrabasinal grains original	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
NCIo	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]
CIo	real, range [0.0 - 100.00]	real, range [0.0 - 100.00]

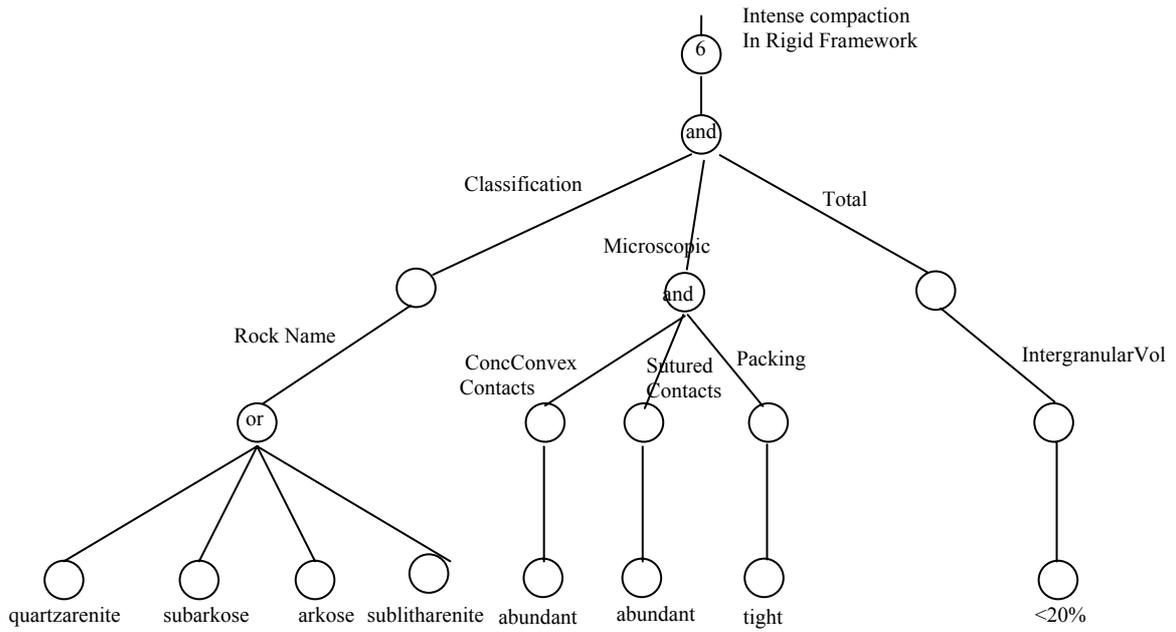
Anexo 2

Modelo de Grafos de Conhecimento da Aplicação *PetroGrapher*

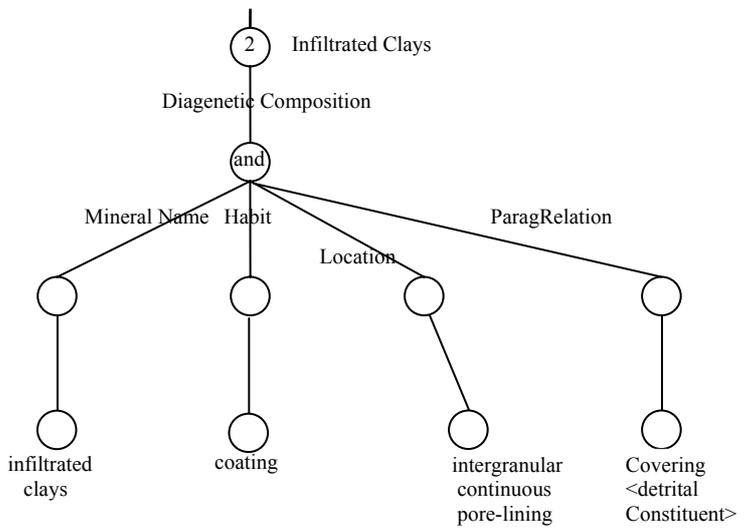
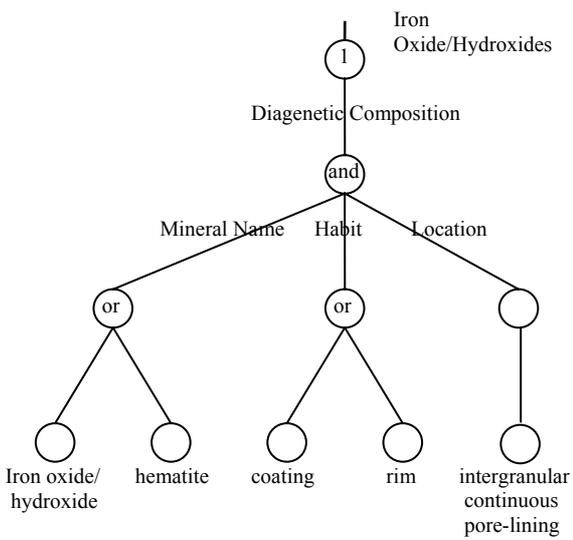
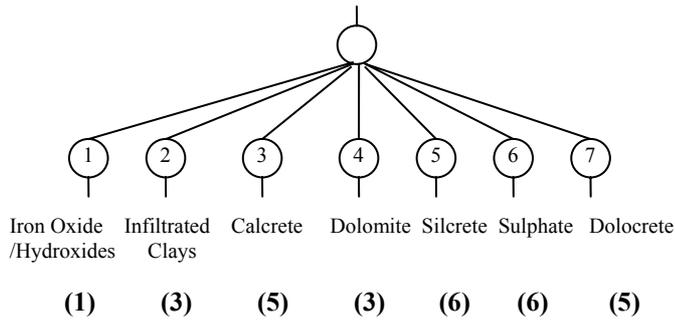
Deep Burial Diagenesis (Deep mesodiagenesis) (6)

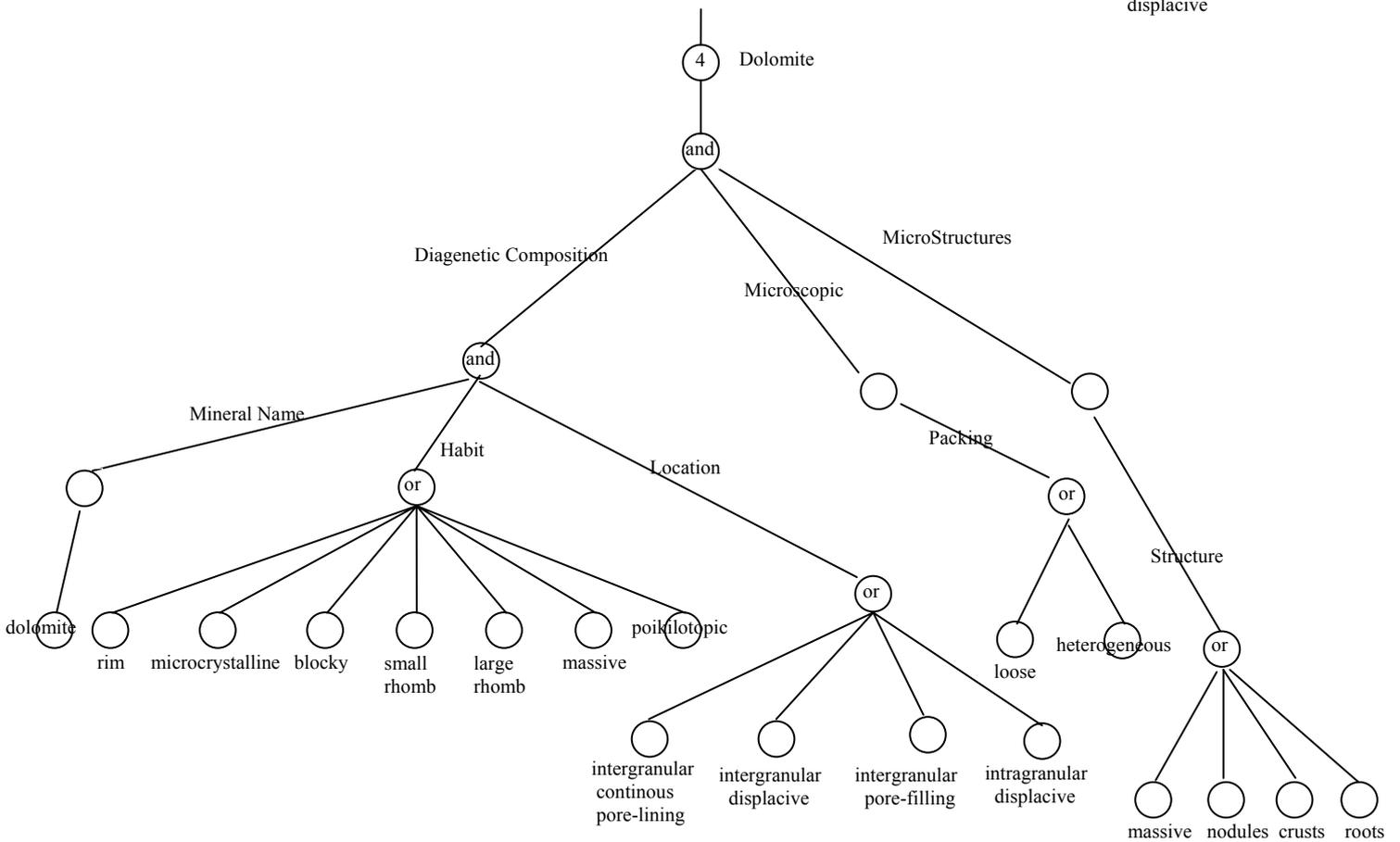
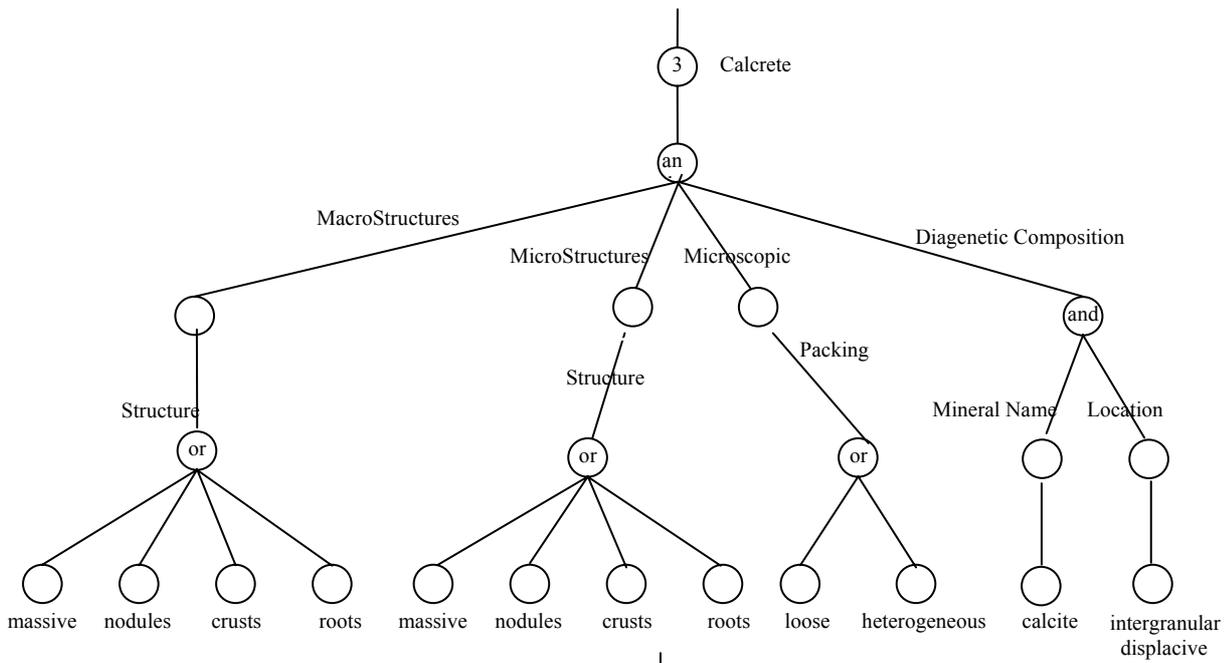


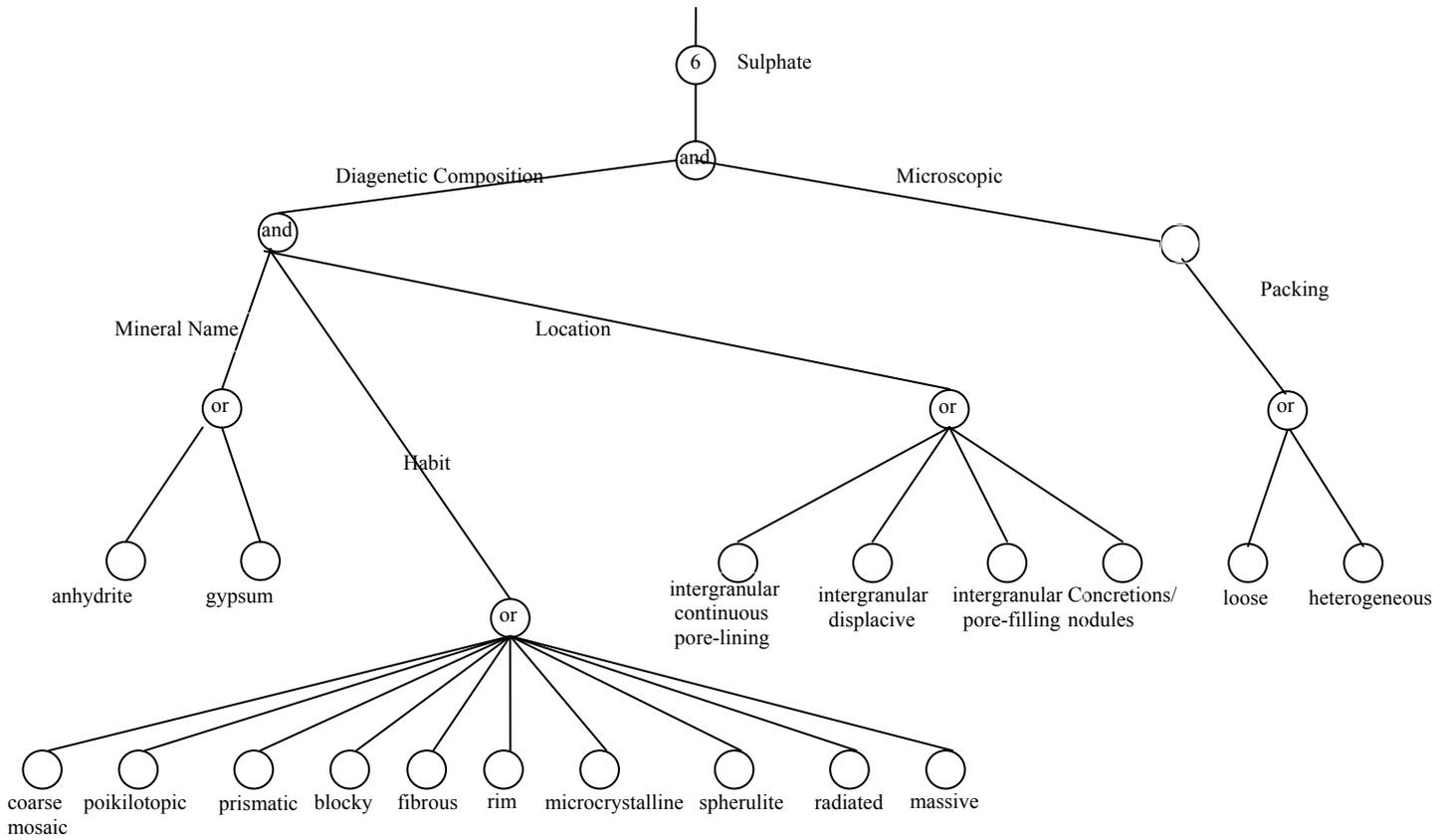
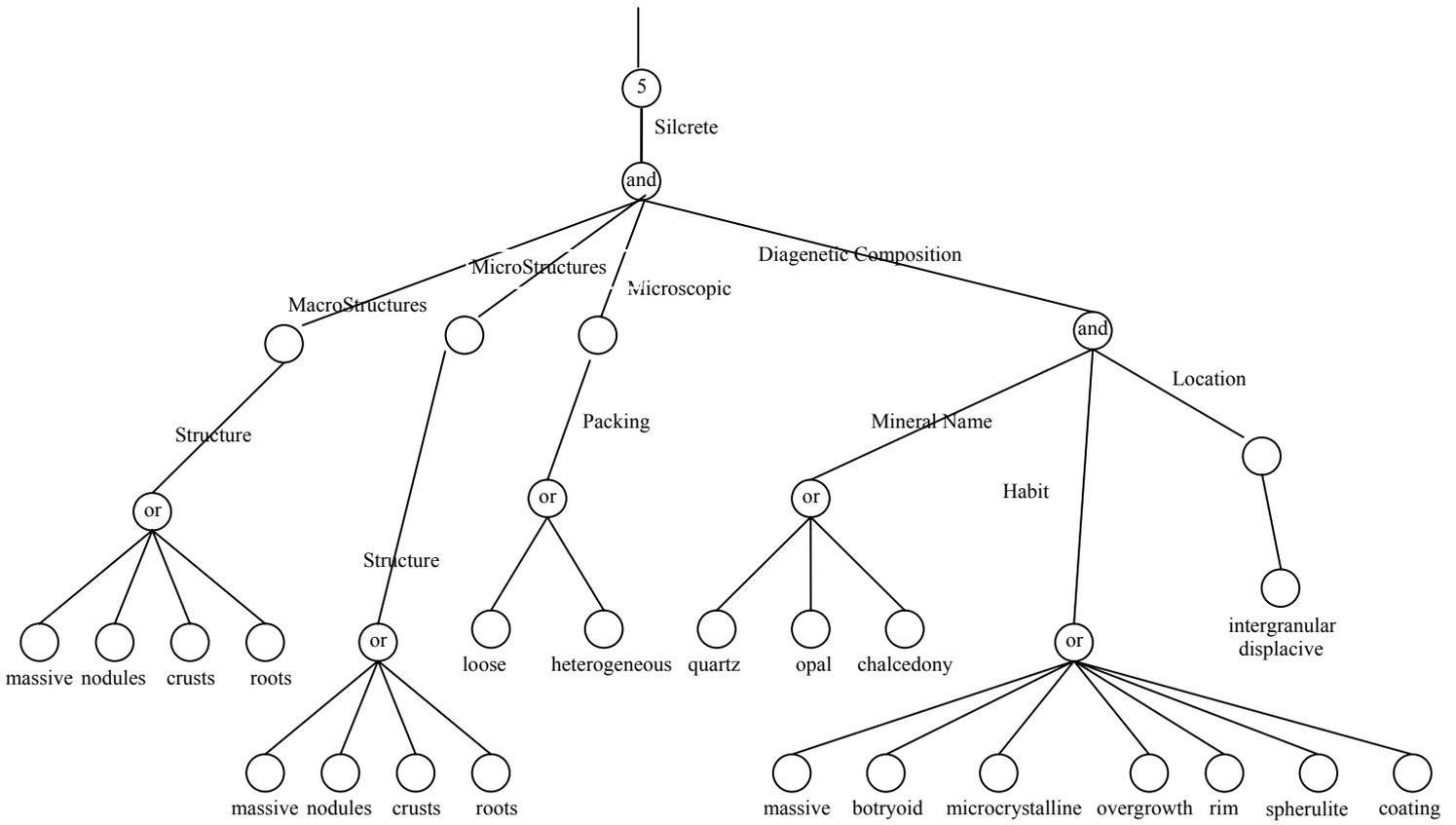


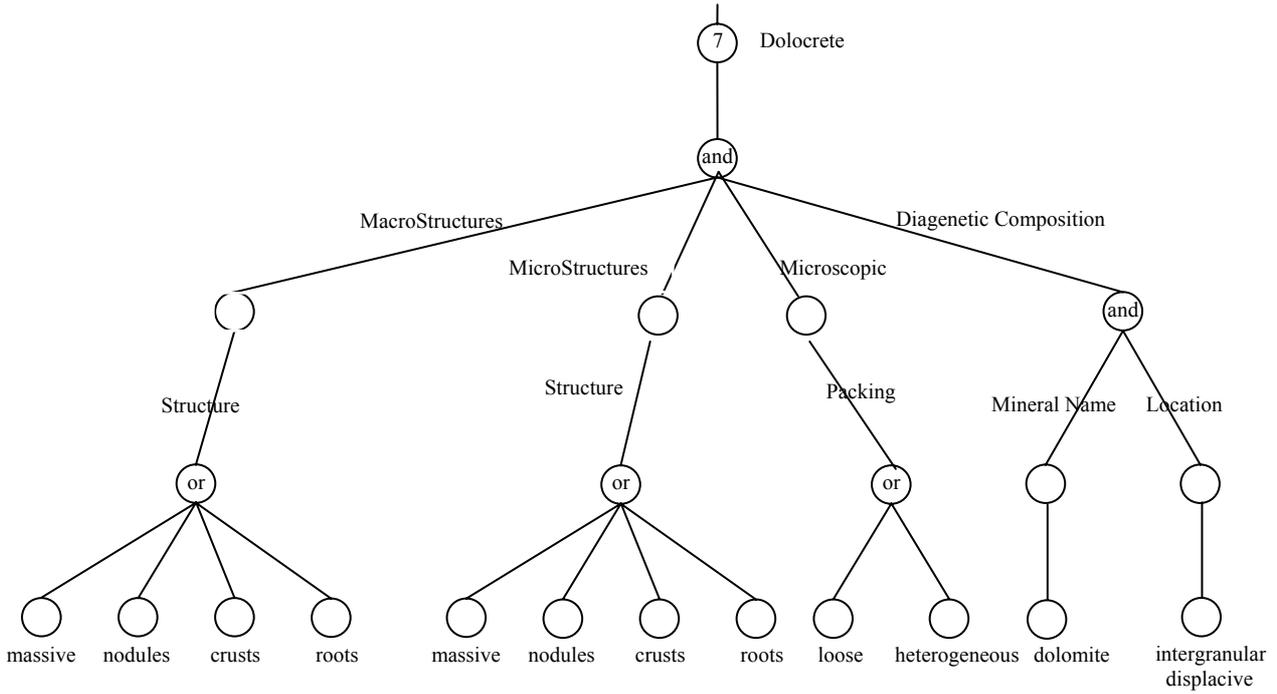


Continental Meteoric Eodiagenesis Under Dry Climate (6)

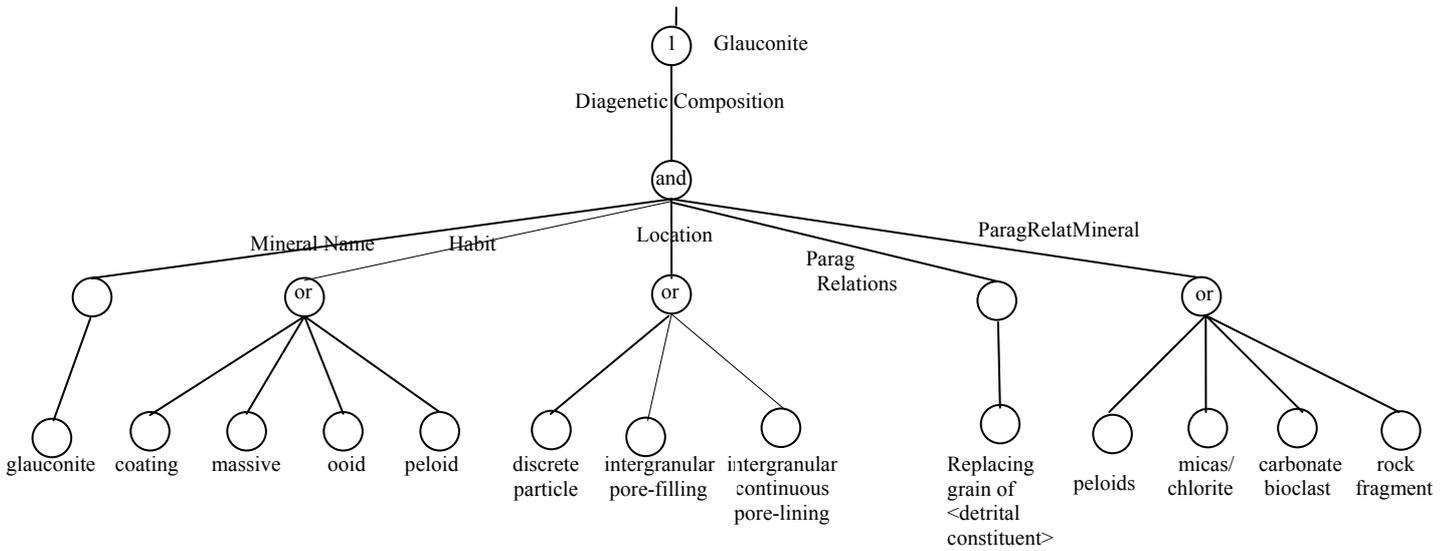
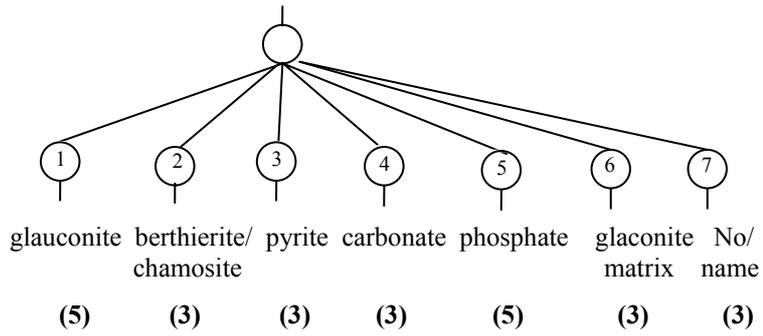


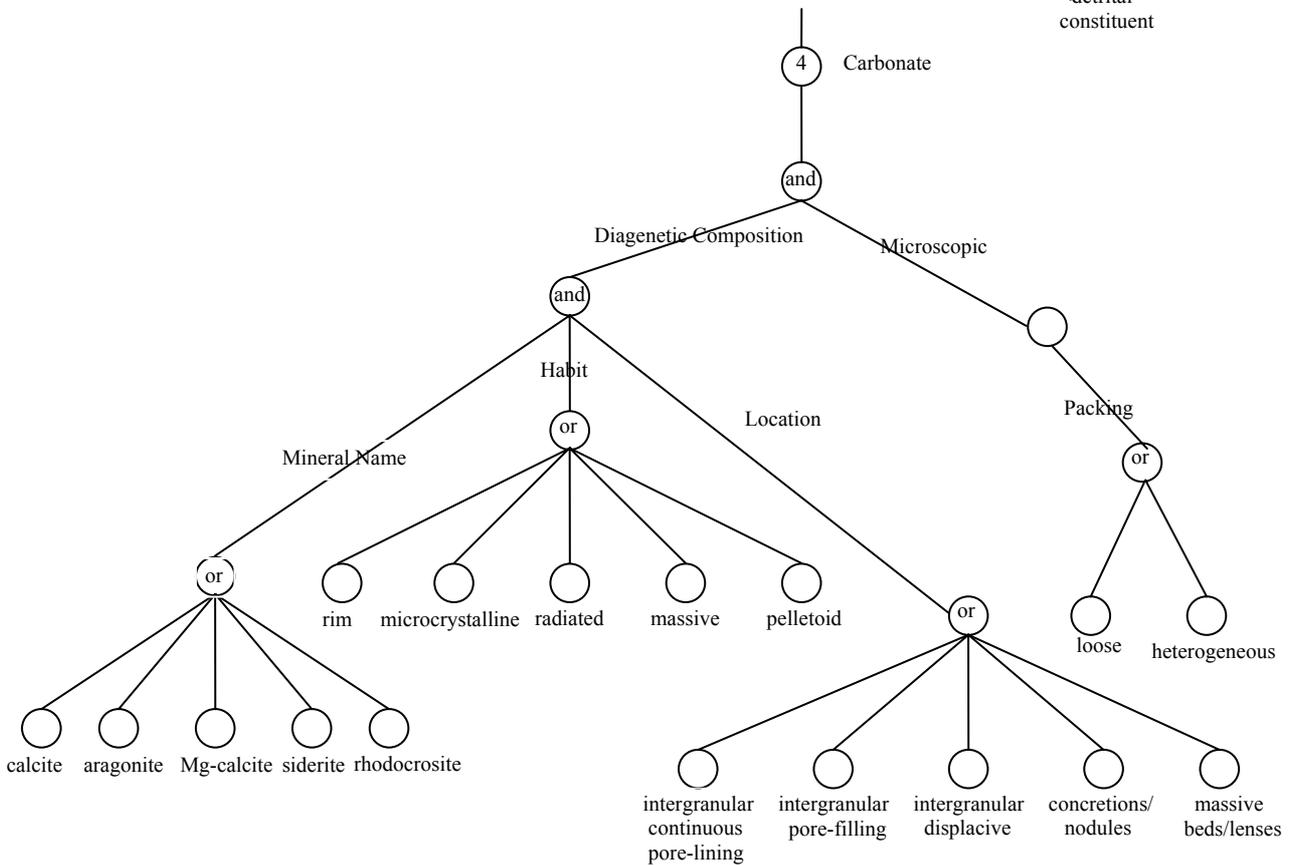
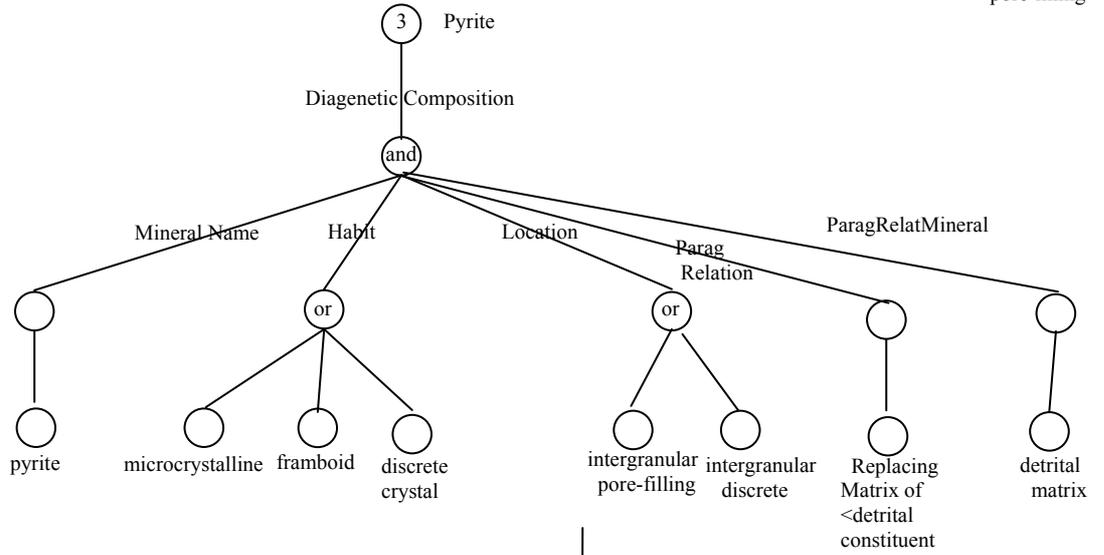
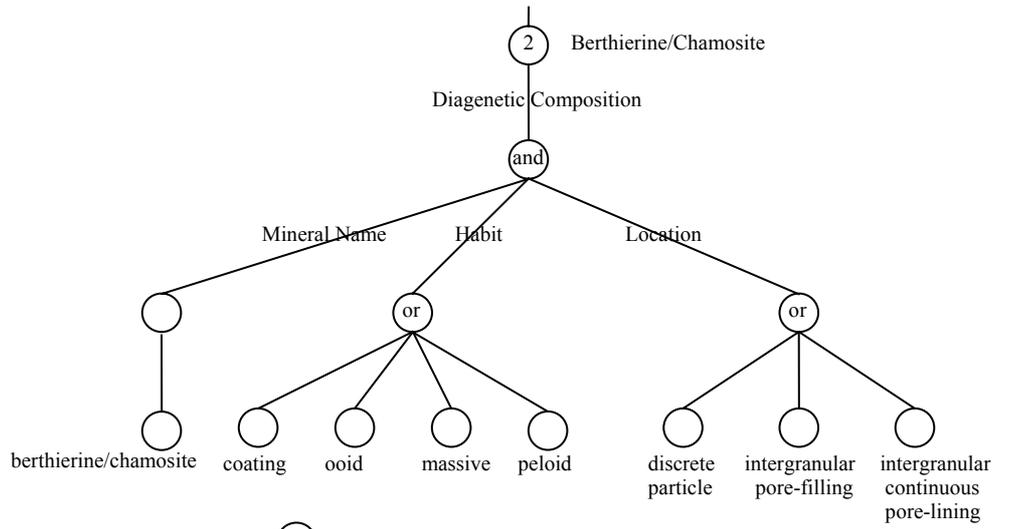


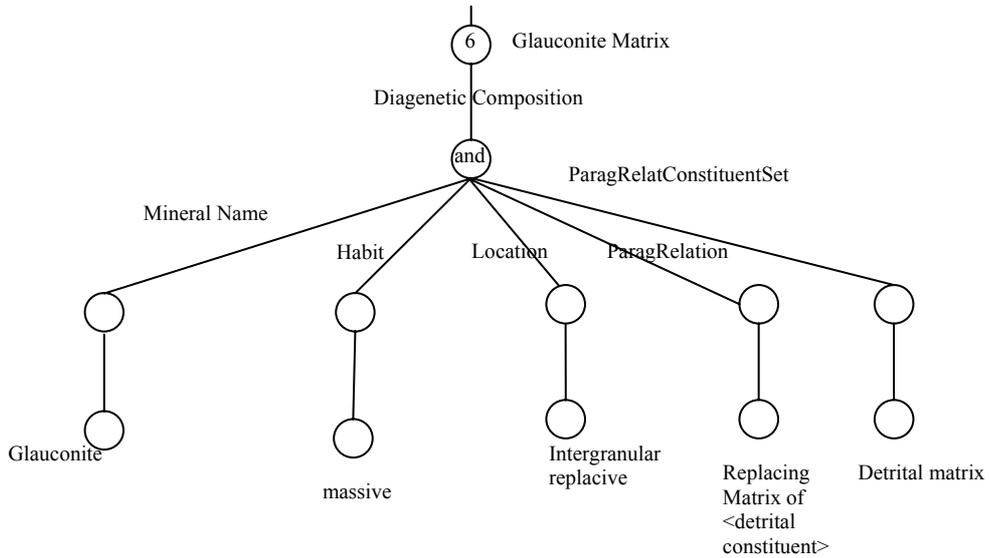
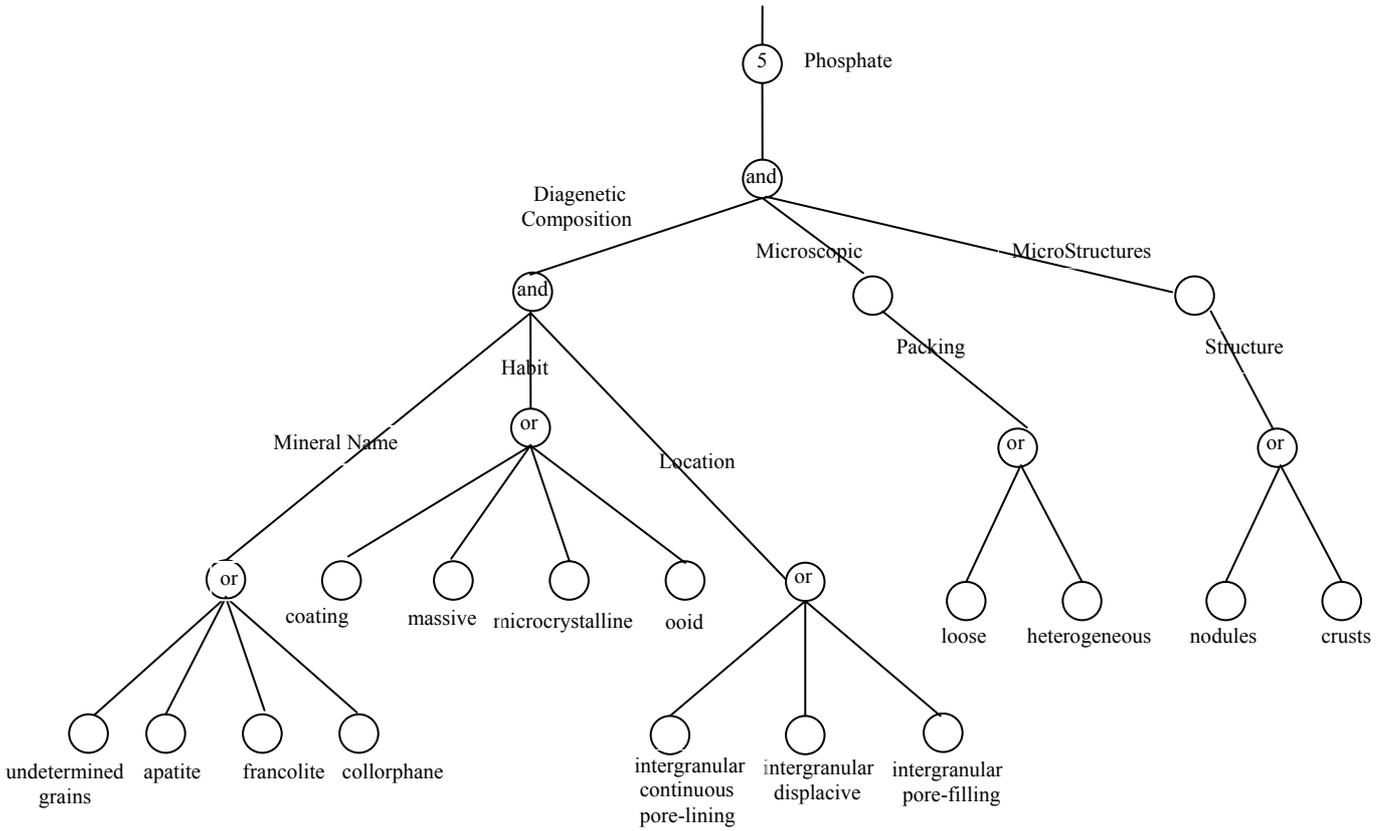


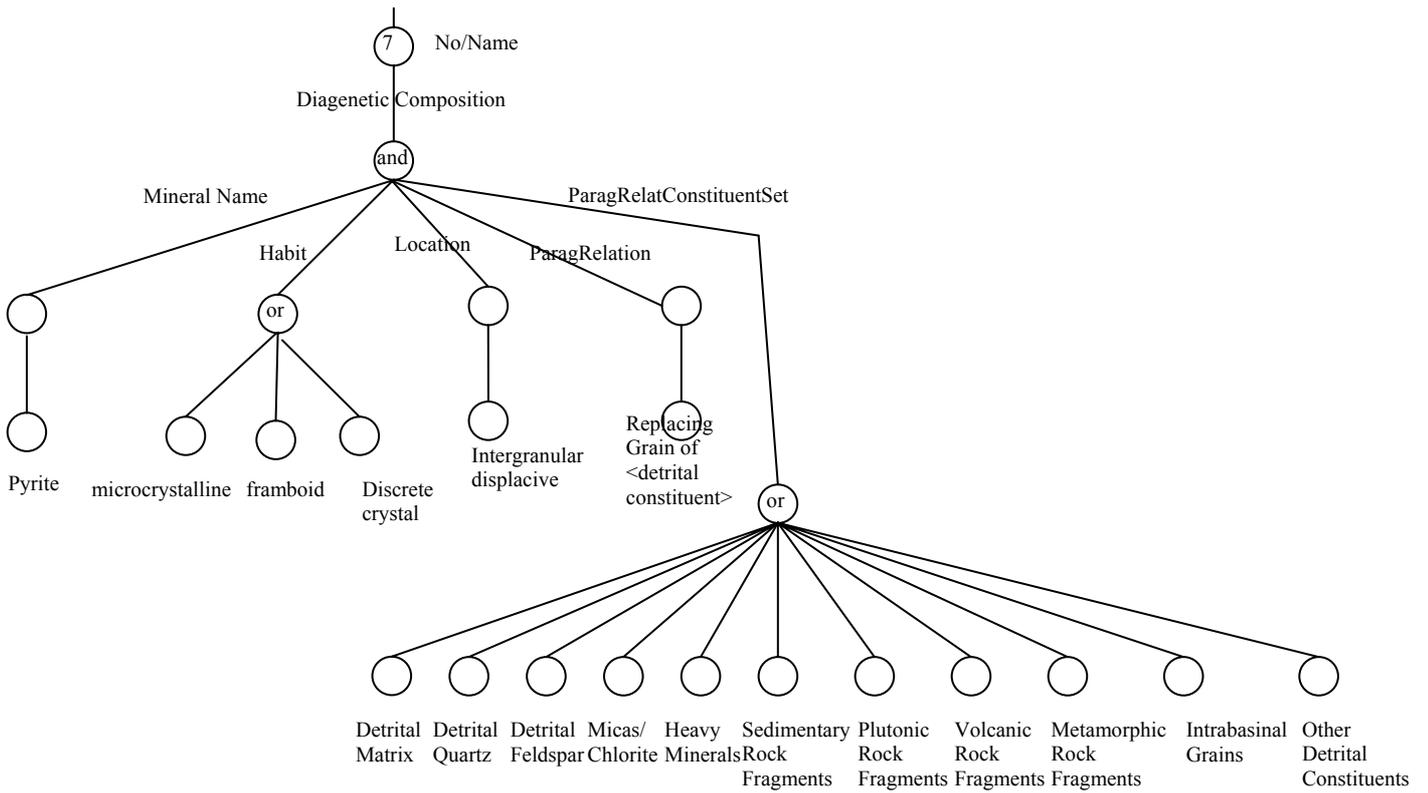


Eodiagenesis Under Marine Conditions (6)

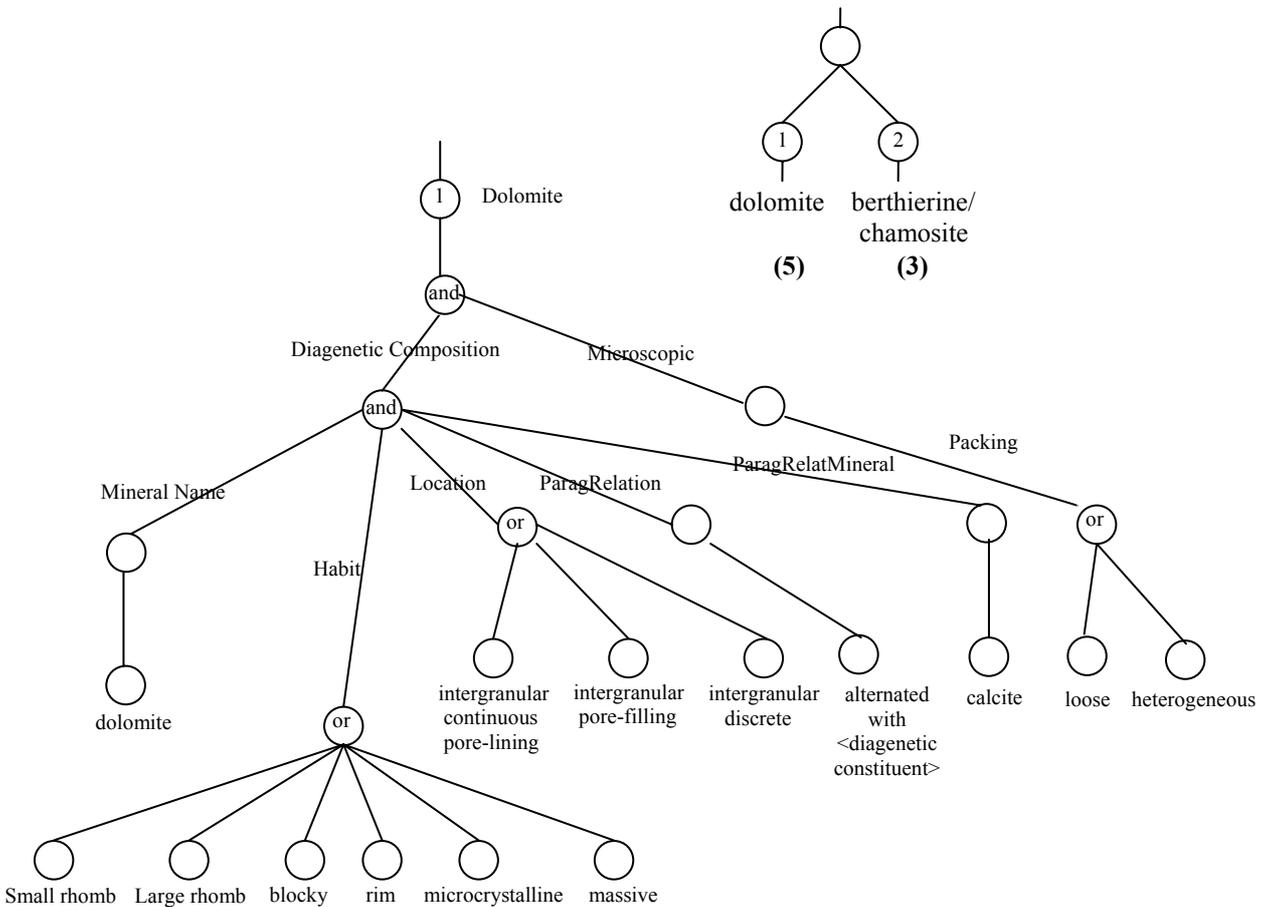


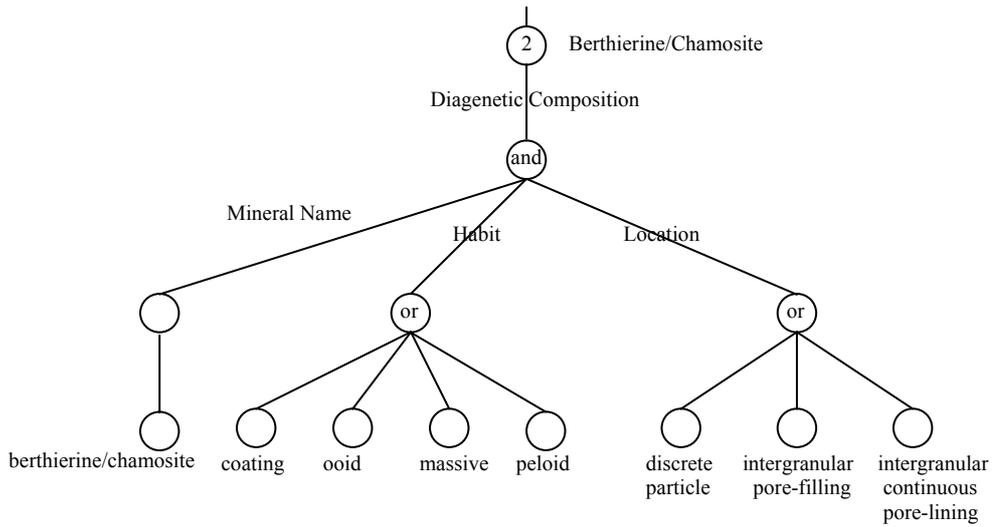




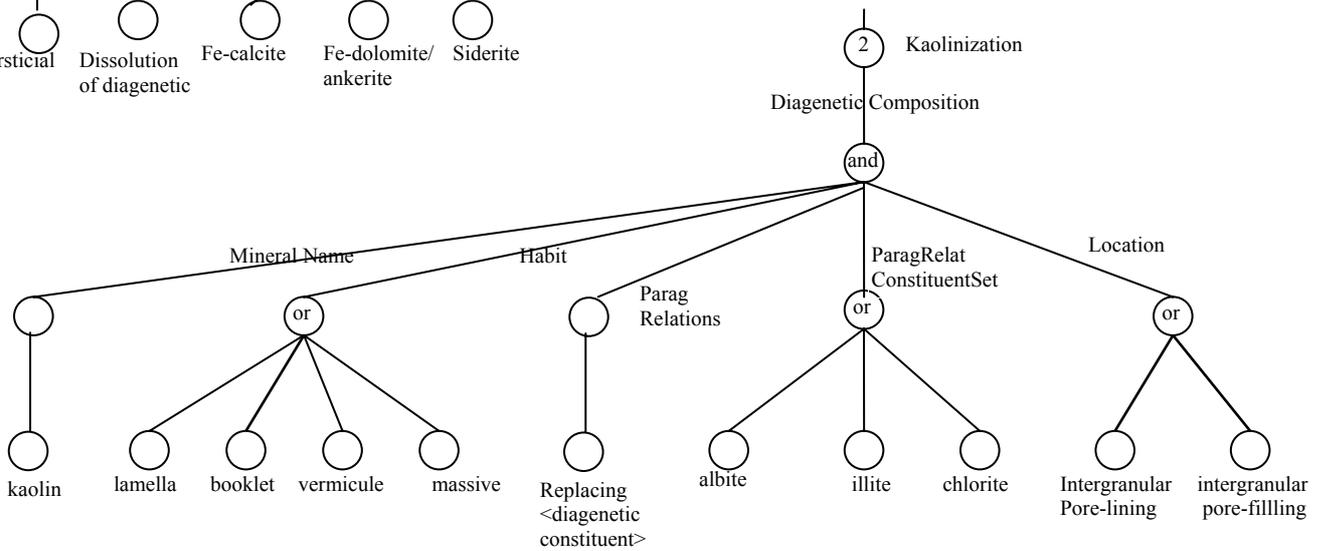
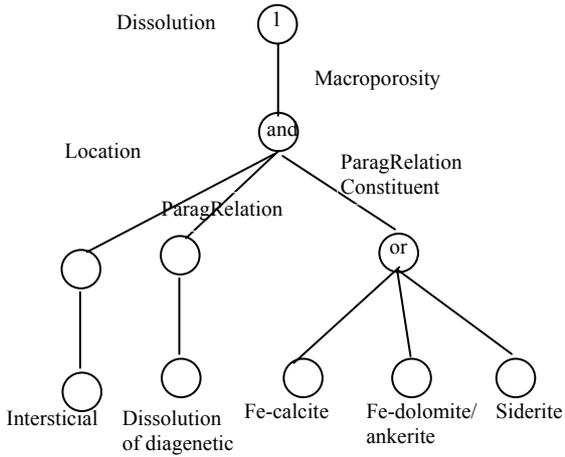
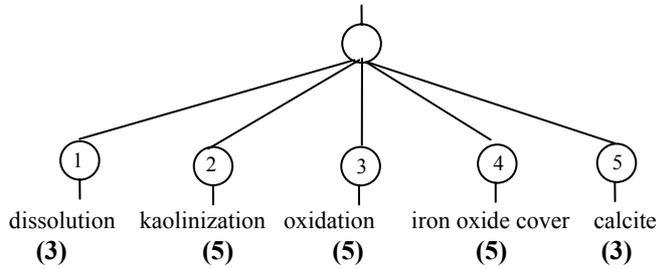


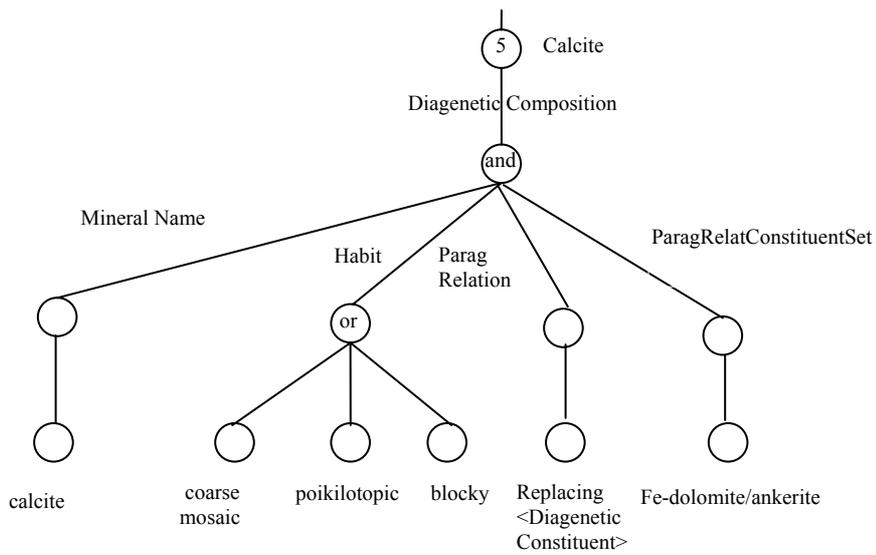
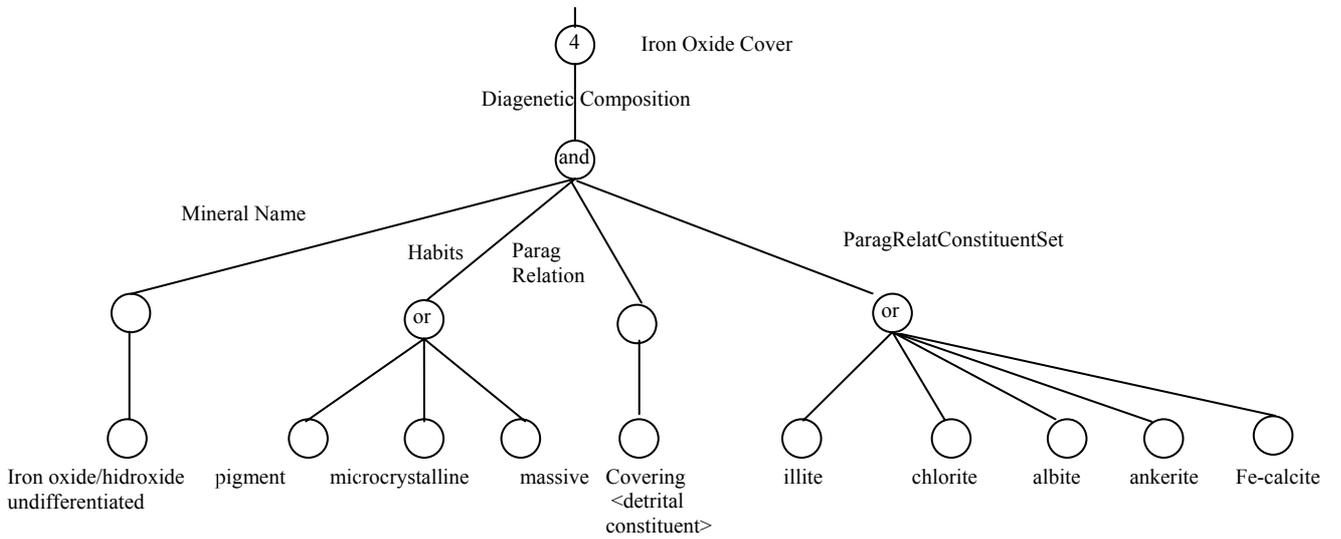
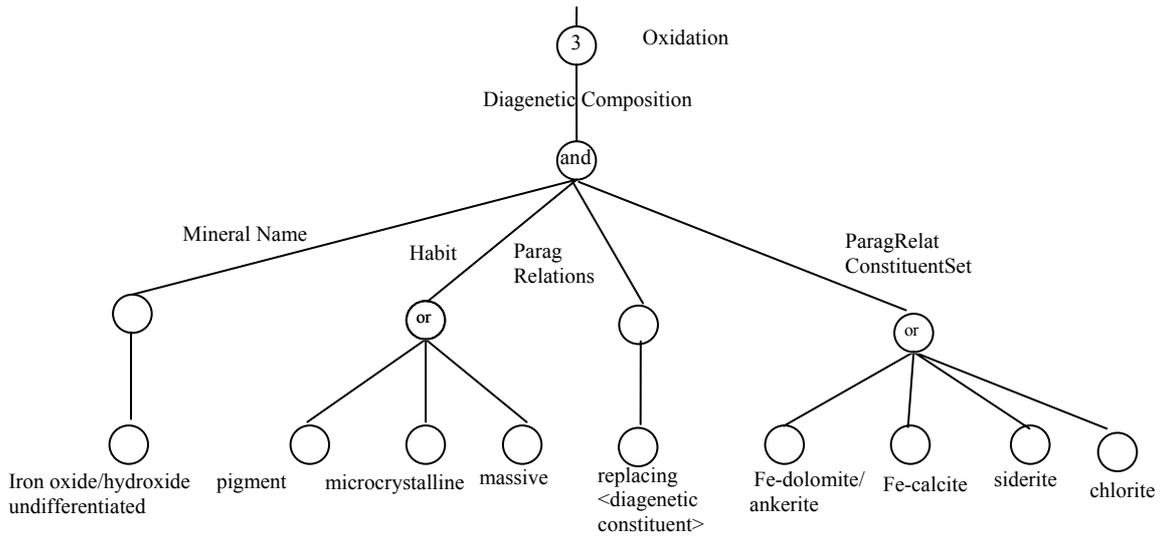
Early Eodiagenesis Under Mixing Between Marine and Meteoric Waters (6)



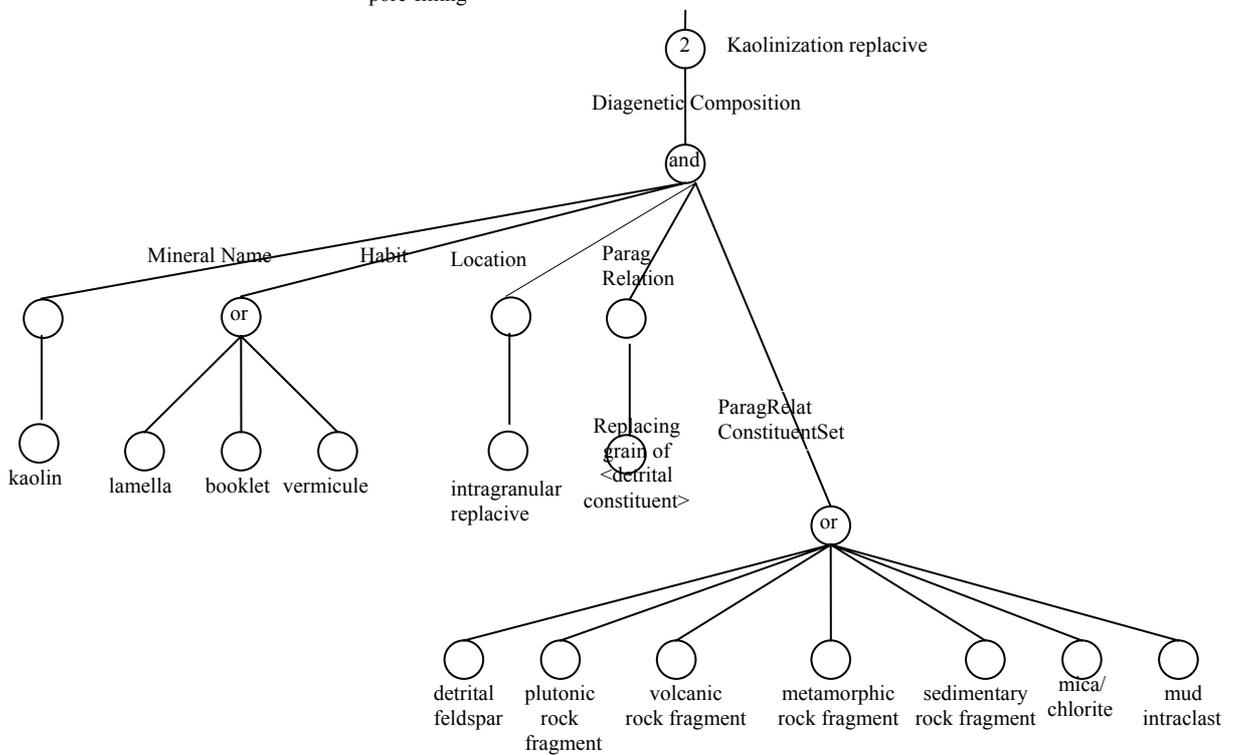
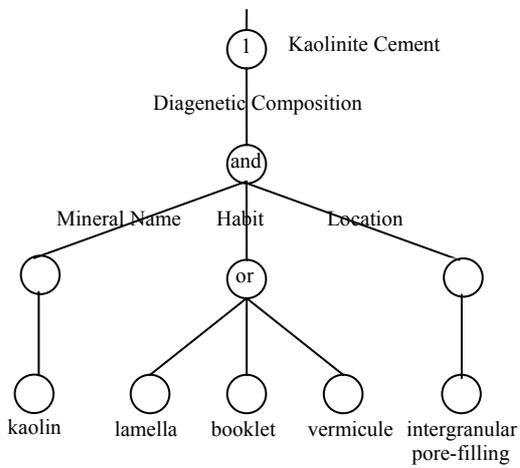
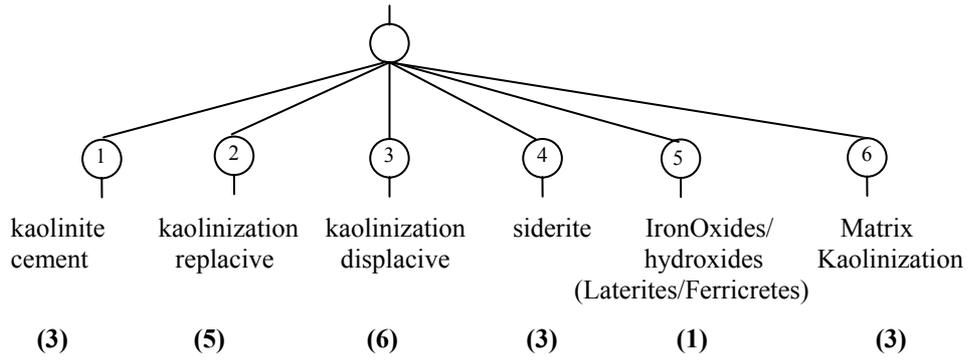


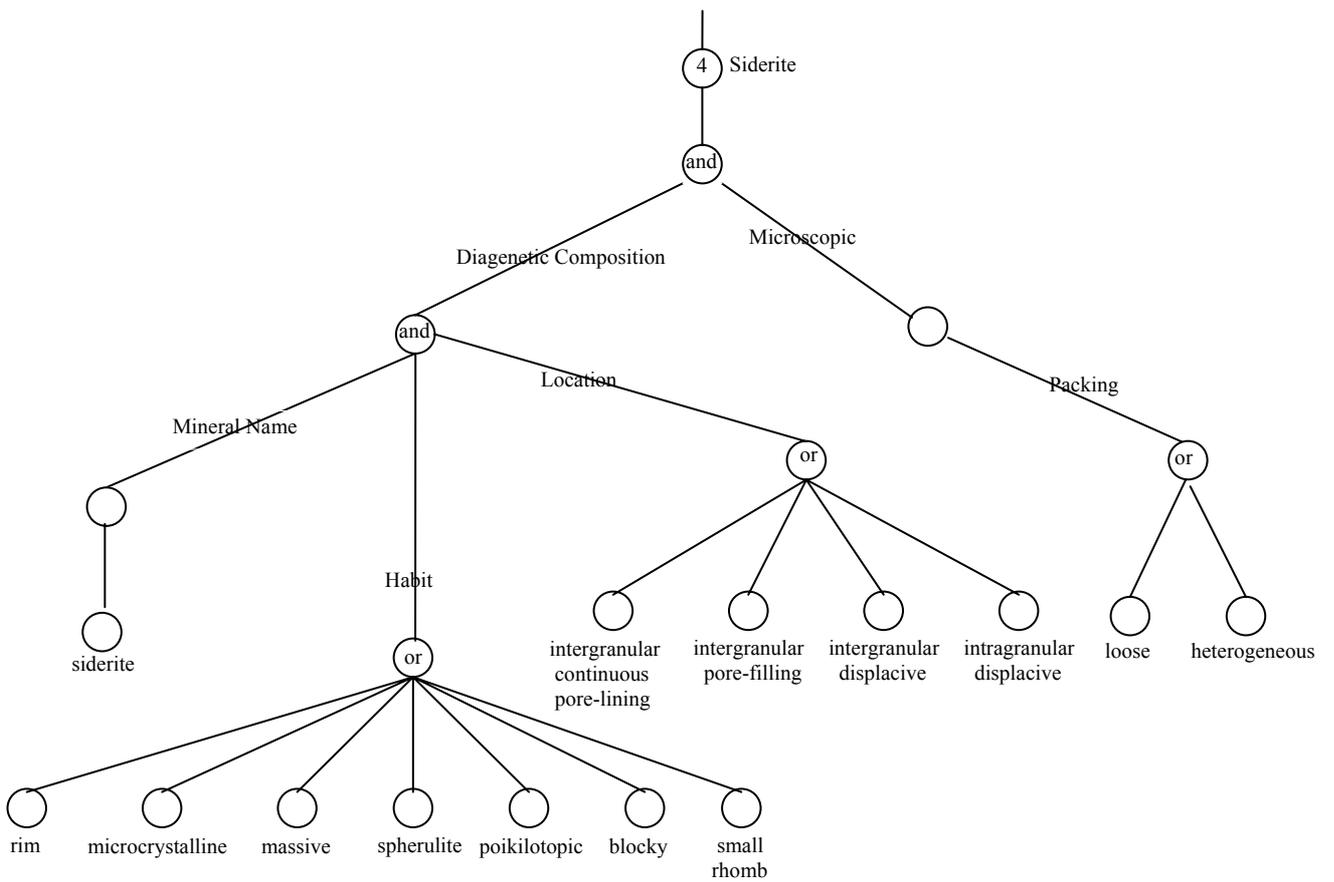
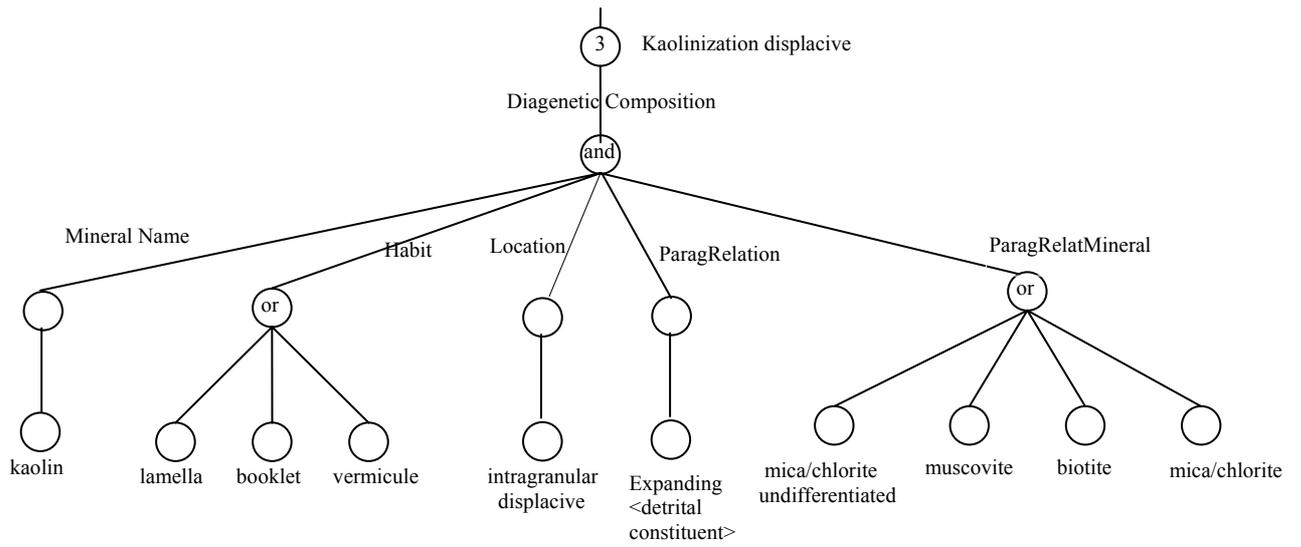
Telodiagenesis Under Meteoric Conditions (6)

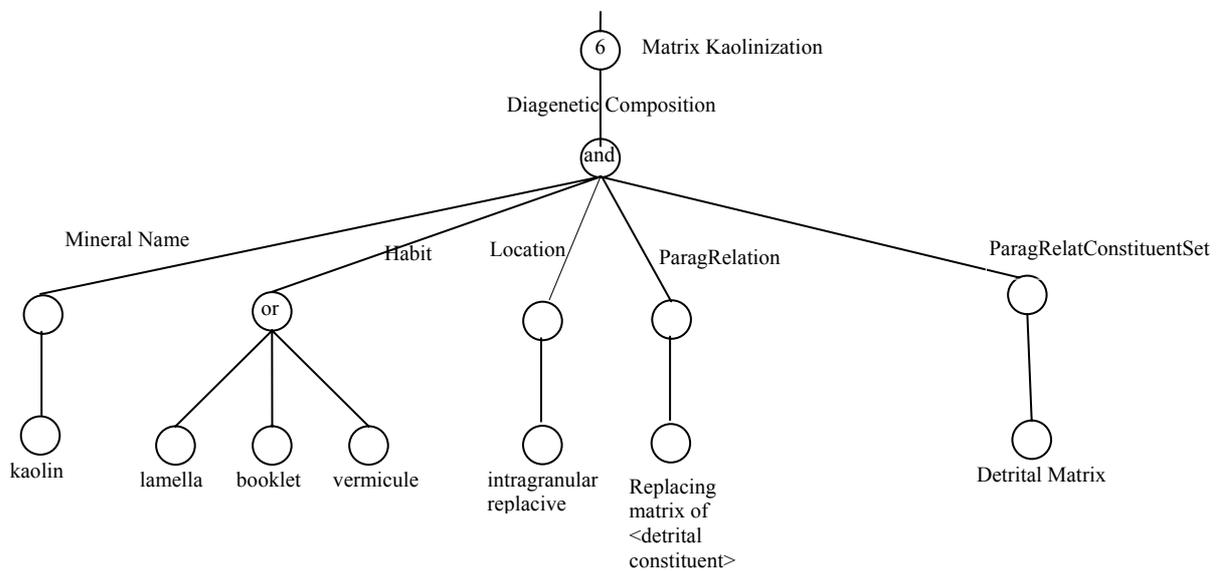
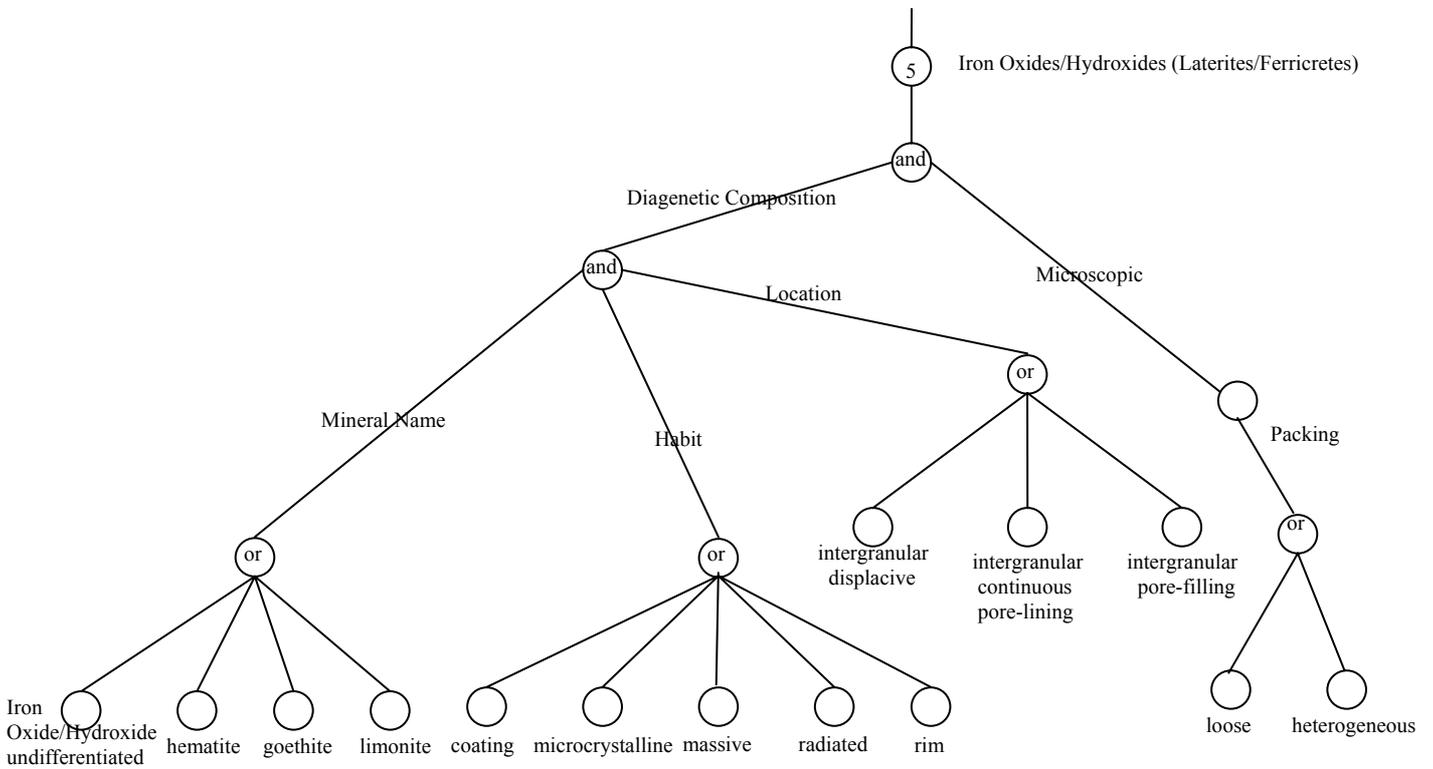




Continental Meteoric Eodiagenesis Under Wet Climate (6)







Bibliografia

- [ABE 95] ABEL, Mara; REATEGUI, Eliseo Berni; CASTILHO, José M.V.; CAMPBELL, John. Evaluating Case-Based Reasoning in a Geological Model. In: INTERNATIONAL CONFERENCE AND WORKSHOP ON DATABASE AND EXPERT SYSTEMS APPLICATIONS, DEXA, 6., 1995, London. **Proceedings...** London: [s.n.], 1995. Disponível em: <<http://www.inf.ufrgs.br/gpesquisa/bdi/marabel.htm>>. Acesso em: jan. 2001.
- [ABE 98] ABEL, M.; CASTILHO, J. M. V.; CAMPBELL, J. Analysis of Expertise for Implementing Geological Expert Systems. In: WORLD CONFERENCE IN EXPERT SYSTEMS, 4., 1998, Mexico City. **Proceedings...** Mexico City: [s.n.], 1998. Disponível em: <<http://www.inf.ufrgs.br/gpesquisa/bdi/marabel.htm>>. Acesso em: jan. 2001.
- [ABE 99] ABEL, Mara. **Fundamentos Cognitivos da Perícia e sua Versão Computacional**. Porto Alegre: PPGC da UFRGS, Abr. 1999. (EQ-33).
- [ABE 2001] ABEL, Mara. **Estudo da Perícia em Petrografia Sedimentar e sua Importância para a Engenharia de Conhecimento**. Porto Alegre: PPGC da UFRGS, 2001. Tese de Doutorado.
- [ANG 98] ANGELE, Jürgen; FENSEL, Dieter; LANDES, D.; STUDER, Rudi. Developing Knowledge-Based Systems with MIKE. **Journal of Automated Software Engineering**, [S.l.], v.5, n.4, p.389-418, Oct. 1998. Disponível em: <<http://www.cs.vu.nl/~dieter/pub.html>>. Acesso em: jan. 2001.
- [ANJ 2000] ANJOS, S. M. C. et al. Depositional and Diagenetic Controls on the Reservoir Quality of Lower Cretaceous Pendência Sandstones, Potiguar Rift Basin, Brazil. **American Association of Petroleum Geologists Bulletin**, [S.l.], v. 84, p.1719-1742, 2000.
- [ARM 98] ARMENGOL, Eva et al. **IBROW3**: State of Art Deliverable. Disponível em: <<http://swi.psy.uva.nl/projects/IBROW3/docs/deliverables/deliverables.html>>. Acesso em: jan. 2001. (ESPRIT Project n.27169, An Intelligent Brokering Service for Knowledge-Component Reuse on the World-Wide Web (IBROW3), Identifier D1.4, ver.3, ver. date 29-05-1998, Status Final).
- [BAR 96] BARROS, L. Nunes de; VALENTE, A.; BENJAMINS, V. R. Modeling Planning Tasks. In: INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE PLANNING SYSTEMS, AIPS, 3., 1996. **Proceedings...** [S.l.:s.n.], 1996. p.11-18. Disponível em: <<http://www.swi.psy.uva.nl/usr/richard/publications/pub-chrono.html>>. Acesso em: jan. 2001.

- [BED 91] BEYNON-DAVIES, Paul. **Expert Database Systems - A Gentle Introduction**. London: McGraw-Hill, 1991.
- [BEN 94] BENJAMINS, V. Richard; JANSWEIJER, Wouter N. H. Toward a Competence Theory of Diagnosis. **IEEE Expert**, [S.l.], v.9, n.5, p.43-52, Oct. 1994.
- [BEN 95] BENJAMINS, V. Richard. Problem-Solving Methods for Diagnosis and their Role in Knowledge Acquisition. **International Journal of Expert Systems: Research & Applications**, [S.l.], v.8, n.2, p.93-120, 1995. Disponível em: <<http://www.swi.psy.uva.nl/usr/richard/publications/pub-chrono.html>>. Acesso em: jan. 2001.
- [BEN 96a] BENJAMINS, Richard; BARROS, Leliane Nunes de; VALENTE, André. Constructing Planners through Problem-Solving Methods. In: WORKSHOP ON KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED SYSTEM, KAW, 10., 1996, Banff, Canada. **Proceedings...** [S.l.]: SRDG Publications, 1996. p.14.1-14.20. Disponível em: <<http://ksi.cpsc.ucalgary.ca/KAW/KAW96/KAW96Proc.html>>. Acesso em: jan. 2001.
- [BEN 96b] BENJAMINS, V. Richard; FENSEL, Dieter; STRAATMAN, Remco. Assumptions of Problem-Solving Methods and Their Role in Knowledge Engineering. In: EUROPEAN CONFERENCE ON ARTIFICIAL INTELLIGENCE, ECAI, 12., 1996, Budapest, Hungary. **Proceedings...** [S.l.]: John Wiley & Sons, 1996. p.408-412. Disponível em: <<http://www.cs.vu.nl/~dieter/pub.html>>. Acesso em: jan. 2001.
- [BEN 97] BENJAMINS, V. Richard et al. Making Knowledge Engineering Technology Work. In: INTERNATIONAL CONFERENCE ON SOFTWARE ENGINEERING & KNOWLEDGE ENGINEERING, SEKE, 9., 1997, Madrid, Spain. **Proceedings...** [S.l.: s.n.], 1997. p.56-61. Disponível em: <<http://www.cs.vu.nl/~dieter/pub.html>>. Acesso em: jan. 2001.
- [BEN 98] BENJAMINS, V. Richard; FENSEL, Dieter. Editorial: Problem-Solving Methods. **International Journal of Human-Computer Studies**, IJHCS, [S.l.], v.49, n.4, p.305-313, Oct. 1998. Disponível em: <<http://www.academicpress.com/ijhcs>>. Acesso em: jan. 2001.
- [BER 98] BERTINO, Elisa; ZARRI, G.P. Intelligent Database Systems - State of Art. In: WORLD CONGRESS ON EXPERT SYSTEMS, 4., 1998, Mexico City, Mexico. **Tutorial...** Mexico City: ITESM, 1998.
- [BEY 96] BEYS, Pascal, BENJAMINS, Richard; HEIJST, Gertjan van. Remedying the Reusability - Usability Tradeoff for Problem-Solving Methods. In: WORKSHOP ON KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED SYSTEM, KAW, 10., 1996, Banff, Canada. **Proceedings...** [S.l.]: SRDG Publications, 1996. p.2.1-2.20. Disponível em:

em: <<http://ksi.cpsc.ucalgary.ca/KAW/KAW96/KAW96Proc.html>>.
Acesso em: jan. 2001.

- [BRE 94] BREUKER, J. Components of Problem Solving and Types of Problems. In: EUROPEAN KNOWLEDGE ACQUISITION WORKSHOP, EKAW, 8., 1994, Berlin, Germany. **Proceedings...** Berlin: Springer-Verlag, 1994. p.118–136. (Lecture Notes in Artificial Intelligence, LNAI, v.867).
- [CHA 86] CHANDRASEKARAN, B. Generic Tasks in Knowledge-Based Reasoning: High-Level Building Blocks for Expert System Design. **IEEE Expert**, [S.l.], v.1, n.3, p.23-30, 1986.
- [CHA 99] CHANDRASEKARAN, B.; JOSEPHSON, John R.; BENJAMINS, V. Richard. What Are Ontologies, and Why Do We Need Them? **IEEE Intelligent Systems and Their Applications**, [S.l.], v.14, n.1, p.20-26, 1999.
- [CLA 85] CLANCEY, W. J. Heuristic Classification. **Artificial Intelligence**, [S.l.], v.27, p.289-350, 1985.
- [CLA 87] CLANCEY, W. J. From Guidon to Neomycin and Heracles in Twenty Short Lessons. In: LAMSWEERDE, A. van (Ed.). **Current Issues in Expert Systems**. [S.l.]: Academic Press, 1987.
- [COC 97] COCHRAN, W. J.; VEDHANAYAGAM, A.; BLAGG, B. O. Knowledge-Based Systems as a Technology Enabler for Business Process Re-Engineering. In: LIEBOWITZ, J.; WILCOX, L. C. (Ed.). **Knowledge Management and Its Integrative Elements**. New York: CRC Press, 1997. p.89-101.
- [DEG 65] DE GROOT, A. D. **Thought and Choice in Chess**. Mouton: The Hague, 1965.
- [DER 98] DE ROS, L. F. Heterogeneous Generation and Evolution of Diagenetic Quartzarenites in the Siluro-Devonian Furnas Formation of the Paraná Basin, Southern Brazil. **Sedimentary Geology**, [S.l.], v.116, p.99-128, 1998.
- [DOY 96] DOYLE, Jon; KOHANE, Isaac; LONG, Willian; SZOLOVITS, Peter. **High-Performance Knowledge Base Support for Monitoring, Analysis, and Interpretation Tasks**. Dec. 1996. Disponível em: <<http://medg.lcs.mit.edu/projects/maita/>>. Acesso em: jan. 2001.
- [FEN 94] FENSEL, Dieter; HARMELEN, Frank van. A Comparison of Languages which Operationalize and Formalise KADS Models of Expertise. **The Knowledge Engineering Review**, [S.l.], v.9, n.2, p.105-146, 1994. Disponível em: <<http://www.cs.vu.nl/~dieter/pub.html>>. Acesso em: jan. 2001.

- [FEN 96] FENSEL, Dieter; STRAATMAN, Remco. Problem-Solving Methods: Making Assumptions for Efficiency Reasons. In: EUROPEAN KNOWLEDGE ACQUISITION WORKSHOP, EKAW, 9., 1996, Nottingham, United Kingdom. **Proceedings...** Berlin: Springer-Verlag, 1996. Disponível em: <<http://www.cs.vu.nl/~dieter/pub.html>>. Acesso em: jan. 2001. (Lecture Notes in Artificial Intelligence, LNAI, v.1076).
- [FEN 97a] FENSEL, Dieter. The Tower-of-Adapters Method for Developing and Reusing Problem-Solving Methods. In: EUROPEAN WORKSHOP ON KNOWLEDGE ACQUISITION, MODELING AND MANAGEMENT, EKAW, 10., 1997, Sant Feliu de Guixols, Catalonia, Spain. **Proceedings...** Berlin: Springer-Verlag, 1997. Disponível em: <<http://www.cs.vu.nl/~dieter/pub.html>>. Acesso em: jan. 2001. (Lecture Notes in Artificial Intelligence, LNAI, v.1319).
- [FEN 97b] FENSEL, Dieter; GROENBOOM, Rix. Specifying Knowledge-Based Systems with Reusable Components. In: INTERNATIONAL CONFERENCE ON SOFTWARE ENGINEERING & KNOWLEDGE ENGINEERING, SEKE, 9., 1997, Madrid, Spain. **Proceedings...** [S.l.: s.n], 1997. p.349-357. Disponível em: <<http://www.cs.vu.nl/~dieter/pub.html>>. Acesso em: jan. 2001.
- [FEN 97c] FENSEL, Dieter; SCHÖNEGGE, Arno. Assumption Hunting as Developing Method for Problem-Solving Methods. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, IJCAI, 15.; WORKSHOP ON PROBLEM-SOLVING METHODS FOR KNOWLEDGE-BASED SYSTEMS, 1997, Nagoya, Japan. **Proceedings...** [S.l.: s.n.], 1997. Disponível em: <<http://www.cs.vu.nl/~dieter/pub.html>>. Acesso em: jan. 2001.
- [FEN 98a] FENSEL, Dieter; BENJAMINS, V. Richard. The Role of Assumptions in Knowledge Engineering. **The International Journal of Intelligent Systems**, IJIS, [S.l.], v.13, n.7, p.715-748, 1998. Disponível em: <<http://www.cs.vu.nl/~dieter/pub.html>>. Acesso em: jan. 2001.
- [FEN 98b] FENSEL, Dieter; STRAATMAN, Remco. The Essence of Problem-Solving Methods: Making Assumptions to Gain Efficiency. **The International Journal of Human Computer Studies**, IJHCS, [S.l.], v.48, n.2, p.181-215, Feb. 1998. Disponível em: <<http://www.academicpress.com/ijhcs>>. Acesso em: jan. 2001.
- [FEN 99] FENSEL, Dieter et al. **The Unified Problem-solving Method Development Language - UPML**. Disponível em: <<http://www.aifb.uni-karlsruhe.de/WBS/dfe/publications99.html>>. Acesso em: jan. 2001. (ESPRIT Project n.27169, An Intelligent Brokering Service for Knowledge-Component Reuse on the World-Wide Web (IBROW3), Deliverable 1.1, chap.1, ver.3, ver. date 01-02-1999, Status Prefinal).

- [FEN 2000] FENSEL, Dieter; MOTTA, Enrico. **Structured Development of Problem-Solving Methods**. Disponível em: <<http://www.cs.vu.nl/~dieter/pub.html>>. Acesso em: jan. 2001.
- [FER 99] FERNÁNDEZ-LÓPEZ, Mariano. Overview of Methodologies for Building Ontologies. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, IJCAI, 16.; WORKSHOP ON ONTOLOGIES AND PROBLEM-SOLVING METHODS, KRR5, 1999, Stockholm, Sweden. **Proceedings...** [S.l.: s.n.], 1999. Disponível em: <<http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-18/>>. Acesso em: jan. 2001.
- [FIK 85] FIKES, R., KEHLER, T. The Role of Frame-Based Representation in Reasoning. **Communications of the ACM**, [S.l.], v.28, n.9, Sept. 1985.
- [GAM 95] GAMMA, E.; HELM, R.; JOHNSON, R.; VLISSIDES, J. **Design Patterns: Elements of Reusable Object-Oriented Software**. Reading : Addison Wesley, 1995.
- [GAP 98] GAPPA, Ute; PUPPE, Frank. A Study of Knowledge Acquisition - Experiences from the SISYPHUS-III Experiment for Rock Classification. In: WORKSHOP ON KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED SYSTEM, KAW, 11., 1998, Banff, Canada. **Proceedings...** [S.l.: s.n.], 1998. Disponível em: <<http://ksi.cpsc.ucalgary.ca/KAW/KAW98/KAW98Proc.html>>. Acesso em: jan. 2001.
- [GAR 98] GARDNER, Karen M. et al. **Cognitive Patterns: Problem-Solving Frameworks for Object Technology**. Cambridge: Cambridge University Press and New York: SIGS Books & Multimedia, 1998.
- [GOM 99] GÓMEZ-PÉREZ, Asunción; BENJAMINS, V. Richard. Overview of Knowledge Sharing and Reuse Components: Ontologies and Problem-Solving Methods. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, IJCAI, 16.; WORKSHOP ON ONTOLOGIES AND PROBLEM-SOLVING METHODS, KRR5, 1999, Stockholm, Sweden. **Proceedings...** [S.l.: s.n.], 1999. Disponível em: <<http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-18/>>. Acesso em: jan. 2001.
- [GOR 99] GROSSO, William E. et al. **Knowledge Modelling at the Millennium - The Design and Evolution of Protégé-2000**, 1999. Disponível em: <<http://www.smi.stanford.edu/projects/protege/papers.html>>. Acesso em: jan. 2001.
- [GRU 93] GRUBER, T. R. A Translation Approach to Portable Ontology Specifications. **Knowledge Acquisition**, [S.l.], v.5, n.2, p.199-221, 1993.

- [HAR 90] HARMON, Paul; SAYER, Brian. **Expert Systems for Business and Industry**. New York: John Wiley & Sons, 1990.
- [HEI 97] HEIJST, G. van; SCHREIBER, A. T.; WIELINGA, B. J. Using Explicit Ontologies in KBS Development. **The International Journal of Human-Computer Studies**, IJHCS, [S.l.], v.46, n.2/3, p.183-292, Feb. 1997. Disponível em: <<http://www.academicpress.com/ijhcs>>. Acesso em: jan. 2001.
- [KOL 93] KOLODNER, J. **Case-Based Reasoning**. San Mateo: Morgan Kaufmann, 1993.
- [LEA 88] LEÃO, Beatriz de Faria. **Construção da Base de Conhecimento de um Sistema Especialista de Apoio ao Diagnóstico de Cardiopatias Congênitas**. São Paulo: Escola Paulista de Medicina, 1988. Tese de Doutorado.
- [MAR 88] MARCUS, S. (Ed.). **Automating Knowledge Acquisition for Expert Systems**. Boston: Kluwer Academic Publisher, 1988.
- [MAT 91] MATTOS, N. M. **An Approach to Knowledge Base Management**. Berlin: Spring-Verlag, 1991.
- [MIZ 95] MIZOGUCHI, R.; VANWELKENHUYSEN, J.; IKEDA, M. Task Ontology for Reuse of Problem Solving Knowledge. In: MARS, N. J. I. (Ed.). **Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing**. Amsterdam: IOS Press, 1995. p.46-57.
- [MOT 98] MOTTA, Enrico; ZDRAHAL, Zdenek. A Library of Problem-Solving Components Based on the Integration of the Search Paradigm with Task and Method Ontologies. **The International Journal of Human Computer Studies**, IJHCS, [S.l.], v.49, n.4, p.437-470, Oct. 1998. Disponível em: <<http://www.academicpress.com/ijhcs>>. Acesso em: jan. 2001.
- [NEW 82] NEWELL, A. The Knowledge Level. **Artificial Intelligence**, [S.l.], v.18, n.1, p. 87-127, 1982.
- [NON 97] NONAKA, I.; TAKEUCHI, H. **Criação de Conhecimento na Empresa**. São Paulo: Campus, 1997.
- [PAR 89] PARSAYE, K.; CHIGNELL, M.; KHOSHAFIAN, S.; WONG, H. **Intelligent Databases**. New York: John Wiley & Sons, 1989.
- [SCH 94] SCHREIBER, Guus et al. CommonKADS: A Comprehensive Methodology for KBS Development. **IEEE Expert**, [S.l.], v.9, n.6, p.28-37, Dec. 1994.

- [SCH 99] SCHREIBER, Guus et al. **Knowledge Engineering and Management - The CommonKADS Methodology**. Cambridge: The MIT Press, Dec. 1999.
- [SIL 97] SILVA, Luís A. de L. **Banco de Dados Inteligente para Análises Petrográficas**. Porto Alegre: PPGC da UFRGS, 1997. Projeto de Diplomação.
- [SIL 99] SILVA, Luís A. de L. **PetroGrapher: Projeto e Desenvolvimento de um Banco de Dados Inteligente**. Porto Alegre: Instituto de Informática/UFRGS, 1999. (Relatório Final Bolsista DTI-G CNPq. Processo Individual n.360375/97-9 - NV, Processo Institucional n.610019/96-1).
- [SPE 96] SPEEL, Piet-Hein; ABEN, Manfred. Applying a Library of Problem-Solving Methods on a Real-Life Task. In: WORKSHOP ON KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED SYSTEM, KAW, 10., 1996, Banff, Canada. **Proceedings...** [S.l.]: SRDG Publications, 1996. Disponível em: <<http://ksi.cpsc.ucalgary.ca/KAW/KAW96/KAW96Proc.html>>. Acesso em: jan. 2001.
- [STE 95] STEFIK, Mark. **Introduction to Knowledge Systems**. San Francisco: Morgan Kaufmann Publishers, 1995.
- [STU 98] STUDER, Rudi; BENJAMINS, V. Richard; FENSEL, Dieter. Knowledge Engineering: Principles and Methods. **Data & Knowledge Engineering**, [S.l.], v.25, p.161-197, 1998. Disponível em: <<http://www.cs.vu.nl/~dieter/pub.html>>. Acesso em: jan. 2001.
- [STU 99] STUDER, Rudi; FENSEL, Dieter; DECKER, Stefan; BENJAMINS, V. Richard. Knowledge Engineering: Survey and Future Directions. In: GERMAN CONFERENCE ON KNOWLEDGE-BASED SYSTEMS, 5., 1999, Wuerzburg, Germany. **Proceedings...** Berlin: Springer-Verlag, 1999. Disponível em: <<http://www.aifb.uni-karlsruhe.de/WBS/publications/index.html>>. Acesso em: jan. 2001. (Lecture Notes in Artificial Intelligence, LNAI, v.1570).
- [STU 2000] STUDER, Rudi; DECKER, Stefan; FENSEL, Dieter; STAAB, Steffen. Situation and Perspective of Knowledge Engineering. In: CUENA, J. et al. (Ed.). **Knowledge Engineering and Agent Technology**. [S.l.]: IOS Press, 2000. Disponível em: <<http://www.aifb.uni-karlsruhe.de/WBS/publications/index.html>>. Acesso em: jan. 2001.
- [TEI 96] TEIJE, A. ten; HARMELEN, F. van; SCHREIBER, A. Th.; WIELINGA, B. J. Construction of Problem-Solving Methods as Parametric Design. In: WORKSHOP ON KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED SYSTEM, KAW, 10., 1996, Banff, Canada. **Proceedings...** [S.l.]: SRDG Publications, 1996. Disponível em: <<http://ksi.cpsc.ucalgary.ca/KAW/KAW96/KAW96Proc.html>>. Acesso em: jan. 2001.

- [TAN 95] TANIMOTO, Steven L. **The Elements of Artificial Intelligence Using Common Lisp**. 2nd ed. New York: Computer Science Press, 1995.
- [TUR 92] TURBAN, Efraim. **Expert Systems and Applied Artificial Intelligence**. New York: Macmillan Publishing Company, 1992.
- [USC 96] USCHOLD, Mike; GRUNINGER, Michael. Ontologies: Principles, Methods and Applications. **Knowledge Engineering Review**, [S.l.], v.11, n.2, p.93-155, June 1996.
- [VAN 89] VANLEHN, Kurt. Problem-Solving and Cognitive Skill Acquisition. In: POSNER, M. I. (Ed.). **Foundations of Cognitive Science**. Cambridge: The MIT Press, 1989. p.526-579.
- [WIE 98] WIELINGA, Bob J.; AKKERMANS, J. M.; SCHREIBER, A. Th.. A Competence Theory Approach to Problem-Solving Method Construction. **The International Journal of Human Computer Studies**, IJHCS, [S.l.], v.49, n.4, p.315-338, Oct. 1998. Disponível em: <<http://www.academicpress.com/ijhcs>>. Acesso em: jan. 2001.
- [WII 97] WIIG, Karl M. Roles of Knowledge-Based Systems in Support of Knowledge Management. In: LIEBOWITZ, J.; WILCOX, L. C. (Ed.). **Knowledge Management and Its Integrative Elements**. New York: CRC Press, 1997. p.69-87.