

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
CURSO DE PÓS-GRADUAÇÃO EM ENGENHARIA E GESTÃO
DO CONHECIMENTO**

Luiz Fernando Lopes

**UM MODELO DE ENGENHARIA DO CONHECIMENTO
BASEADO EM ONTOLOGIA E CÁLCULO PROBABILÍSTICO
PARA O APOIO AO DIAGNÓSTICO**

Tese submetida ao Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento da Universidade Federal de Santa Catarina para a obtenção do Grau de Doutor em Engenharia do Conhecimento

Orientador: Prof. Dr. Alexandre Leopoldo Gonçalves

Coorientador: Prof. Dr. José Leomar Todesco

Florianópolis
2011

Catálogo na fonte elaborada pela biblioteca
da
Universidade Federal de Santa Catarina

L864m Lopes, Luiz Fernando

Um Modelo de engenharia do conhecimento baseado em ontologia e cálculo probabilístico para o apoio ao diagnóstico [tese] Luiz Fernando Lopes; orientador, Alexandre Leopoldo Gonçalves, coorientador, José Leomar Todesco. – Florianópolis, 2011.

233 p.: il., grafs., tabs.; 21cm

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Florianópolis, 2011.

1. Engenharia e gestão do conhecimento. 2. Teoria do conhecimento. 3. Ontologia. 4. Processo decisório. I. Gonçalves, Alexandre Leopoldo. II. Todesco, José Leomar. III. Universidade Federal de Santa Catarina - Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento. IV. Título.

CDU: 659.2

Luiz Fernando Lopes

UM MODELO DE ENGENHARIA DO CONHECIMENTO
BASEADO EM ONTOLOGIA E CÁLCULO PROBABILÍSTICO
PARA O APOIO AO DIAGNÓSTICO

Esta Tese foi julgada adequada para obtenção do Título de
“Doutor em Engenharia do Conhecimento”, e aprovada em sua forma
final pelo Programa de Pós-Graduação em Engenharia e Gestão do
Conhecimento

Florianópolis, 29 de setembro de 2011.

Prof. Paulo Maurício Selig, Dr.
Coordenador do Curso

Banca Examinadora:

Prof. Dr. Alexandre Leopoldo
Gonçalves,
Orientador
Universidade Federal de Santa
Catarina

Prof. Dr. José Leomar Todesco,
Coorientador
Universidade Federal de Santa
Catarina

Prof. Dr. Roberto Carlos dos Santos Pacheco,
Universidade Federal de Santa Catarina

Prof. Dr. Neri dos Santos,
Universidade Federal de Santa Catarina

Prof. Dr. Francisco Antonio Pereira Fialho,
Universidade Federal de Santa Catarina

Prof. Dr. Marcos Augusto Hochuli Shmeil,
Pontifícia Universidade Católica do Paraná

Profa. Dra. Luciana Vieira Castilho-Weinert,
Universidade Federal do Paraná

Prof. Dr. Luiz Fernando Nunes,
Universidade Tecnológica Federal do Paraná

Dedico este trabalho a minha mãe Maria Delurdes, pelas orações e ensinamentos, a minha esposa Fabiola, pelo apoio e cumplicidade, e a meus filhos Gabriel, Juliana e Marcelo, como incentivo ao estudo e à pesquisa.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por sempre iluminar meu caminho, por me dar forças para não desistir, e por permitir alcançar este objetivo;

Ao meu pai Nilo (*in memoriam*), a minha mãe Delurdes, a minha esposa Fabiola, a meus filhos Gabriel, Juliana e Marcelo, e a meus irmãos Ayrton (*in memoriam*), Rosilene, Milton e Rosani pela convivência e apoio ao longo de minha história;

Ao Prof. Dr. Alexandre Leopoldo Gonçalves, como conselheiro e amigo, por sua paciência, apoio, dedicação e, principalmente, pela sua disposição em orientar-me, no início de sua carreira, tornando-o um grande guerreiro e vencedor;

Ao Prof. Dr. José Leomar Todesco, meu coorientador, por suas contribuições em momentos decisivos, que delinearão a estruturação deste trabalho;

Ao amigo e irmão de coração Maurício Capobianco Lopes, a quem devo muito respeito e admiração, por seu companheirismo em todos os momentos que tivemos a oportunidade de compartilhar os desafios impostos ao longo desta caminhada;

Agradeço à Profa Dra Luciana Vieira Castilho Weinert pelas sugestões recebidas, de fundamental importância, mostrando ser uma profissional de alta competência na pesquisa e na arte de ensinar;

Aos colegas e amigos Eloi, Cesar, Nelson, Samuel, Rosane, Severino e tantos outros que me proporcionaram inúmeros momentos de satisfação e alegria ao longo do curso;

Agradeço aos profissionais Fernando, Martins, Inácio, Danielle, Sandra, Matos (*in memoriam*), Marcos, Flora, Andrea e Bárbara que fizeram muito além do que lhes foi solicitado;

Agradeço aos especialistas que participaram desta pesquisa, respondendo aos questionários e apresentando preciosas contribuições e sugestões para o trabalho;

Aos membros da Banca Examinadora desta tese (qualificação e defesa final), pelas correções e indicações de visões futuras.

Agradeço à UFSC e à UTFPR, em especial os professores do Departamento de Matemática, pela oportunidade que me deram de realizar este curso;

Aos Professores do EGC, pelos sábios conhecimentos compartilhados e amizades criadas;

Aos funcionários da secretaria de Pós-Graduação, Michele, Nathana e Airton, pelo pronto atendimento e ajuda nos momentos finais do curso;

Por fim, quero expressar meus profundos agradecimentos a todas as pessoas que, direta ou indiretamente, contribuíram para tornar este sonho uma realidade.

"Conhecer não basta, devemos aplicar.
Desejar não basta, devemos fazer".
(Goethe, 1806)

"A melhor maneira de predizer o futuro é
inventá-lo".
(Alan Curtis Kay, 1971)

RESUMO

Lopes, Luiz Fernando. Um modelo de engenharia do conhecimento baseado em ontologia e cálculo probabilístico para o apoio ao diagnóstico. Florianópolis, 2011. 233p. Tese (Doutoramento) Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento – EGC, UFSC.

O diagnóstico, como tarefa intensiva em conhecimento, é um processo complexo uma vez que existe uma grande variedade de elementos e circunstâncias a serem considerados para uma tomada de decisão. Incertezas geradas pela subjetividade, imprecisão e/ou falta de informações atualizadas existem em quase todos os estágios e interferem quanto à segurança e eficácia no resultado. Os dados e informações úteis, quando coletados e tratados adequadamente (técnica), provenientes de diagnósticos realizados (processo) e que permanecem em estado latente, podem tornar-se uma valiosa fonte de conhecimento se associados à experiência e observação do profissional (humano) que os utiliza. Assim, o objetivo desta pesquisa é propor um modelo de Engenharia do Conhecimento que possibilita a geração de novos conhecimentos para apoiar o processo de diagnóstico. As metodologias, métodos e técnicas da Engenharia do Conhecimento, utilizados neste modelo para apoiar este processo, são: CommonKADS, Ontologias, Cálculo Probabilístico e Sistemas de Descoberta Baseados na Literatura. Através da integração entre esses elementos, o modelo proposto é aplicado em um estudo de caso, o qual possibilita que evidências sejam destacadas e analisadas através de pesquisa literária como possíveis novos conhecimentos. Após a confirmação de um novo conhecimento, envolvendo a comunidade científica, o processo de inferência é atualizado. Para a verificação do aspecto de consistência do modelo, buscou-se o consenso de opiniões em um grupo de especialistas utilizando o método Delphi. Os resultados mostram que a aceitação nos conceitos, métodos e técnicas, que compõem o modelo, fica acima de um mínimo estabelecido para este estudo e os comentários dos especialistas geraram reflexões para compor o resultado final deste trabalho. Conclui-se, portanto, que, através desta pesquisa, o modelo proposto atende os requisitos para a geração de novos conhecimentos e contribui para o aperfeiçoamento da tarefa de diagnóstico.

Palavras-chave: Engenharia do Conhecimento. Sistemas Baseados em Conhecimento. Diagnóstico. Ontologia. Cálculo Probabilístico.

ABSTRACT

Lopes, Luiz Fernando. An engineering model based on ontology and probabilistic calculation to support the diagnosis. Florianópolis, 2011. 233p. Tese (Doutoramento) Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento – EGC, UFSC.

The diagnosis, as knowledge-intensive task, is a complex process since there is a wide variety of elements and circumstances to be considered for a decision-making. Uncertainty generated by the subjectivity, vagueness and/or lack of updated information exist in almost all stages and interfere for the safety and efficacy in the outcome. The data and useful information, when collected and treated appropriately (technical), deriving from diagnosis accomplished and which remain latent (process), can become a valuable source of knowledge if associated with the experience and observation of the individual (human) who uses them. The goal of this research is to propose a model of Knowledge Engineering that allows the creation of new knowledge to support the diagnosis process. The methodologies, methods and techniques of Knowledge Engineering, used on this model to support the process are: CommonKADS, Ontology, Probabilistic Calculation and Discovery Systems Based on Literature. Through the integration of these elements, the proposed model is applied to a case study which allows evidence to be highlighted and analyzed through research literature as potential new knowledge. After the information of a new knowledge, involving the scientific community, the inference process is updated. To verify the consistency aspect of the model, it is sought the consensus of opinions in a group of experts using the Delphi method. The results show that the acceptance of the concepts, methods and techniques that comprise the model are above the minimum established for this study, and comments from the experts generated ideas to compose the final result of this work. It is concluded, therefore, that through this research, the proposed model meets the requirements for the generation of new knowledge, and contributes to the improvement of the diagnostic test.

Keywords: Knowledge Engineering. Knowledge Based Systems. Diagnosis. Ontology. Probabilistic Calculation.

LISTA DE FIGURAS

Figura 1 - Alguns fatores que podem induzir a erros de diagnóstico	38
Figura 2 - Componentes de um SE	40
Figura 3 - Sistema Baseado em Conhecimento “Oncocin”	42
Figura 4 - Projeto EON	43
Figura 5 - Modelo SOMKS	44
Figura 6 - Modelo FENZA	45
Figura 7 - Modelo SWE e KHAM	47
Figura 8 - Modelo ODDIN	48
Figura 9 - Modelo SECI de conversões do conhecimento	52
Figura 10- Metodologia CommonKADS	54
Figura 11- Visão das três categorias de conhecimento	60
Figura 12- Especificação do modelo de conhecimento no CommonKADS	61
Figura 13- Hierarquia dos tipos de tarefas de conhecimento intensivo	61
Figura 14- Inferência Cobrir	65
Figura 15- Inferência Especificar	66
Figura 16- Diagrama de Atividades do Modelo de Tarefa de Diagnóstico	67
Figura 17- Modelo de Tarefa completo	67
Figura 18- Principais métodos de Representação do Conhecimento	70
Figura 19- Principais componentes de uma ontologia	73
Figura 20- Modelo de descoberta ABC de Swanson	76
Figura 21- Processo aberto de descoberta	77
Figura 22- Processo fechado de descoberta	78
Figura 23- Comparação entre um conjunto clássico (a) e um <i>fuzzy</i> (b)	82
Figura 24- Processamento <i>fuzzy</i>	83
Figura 25- Esquema básico de uma rede Bayesiana simples	84
Figura 26- Arquitetura do modelo proposto	97
Figura 27- Diagrama de atividades para apoio ao diagnóstico	101
Figura 28- Diagrama de atividades para a geração de um novo conhecimento	102
Figura 29- Diagrama CESM para o processo de diagnóstico	106
Figura 30- Efeito sistêmico do processo de diagnóstico	107
Figura 31- Classes e subclasses da ontologia	108
Figura 32- Relações entre classes	109
Figura 33- Restrições de classes	110
Figura 34- Exemplo de instância	111
Figura 35- Inferência no <i>Protégé</i>	111
Figura 36- Registro de dados	112
Figura 37- Rede Bayesiana	113
Figura 38- Rede Bayesiana instanciada para o reclamante	114
Figura 39- Rede Bayesiana instanciada para uma solução	115
Figura 40- Consulta ao <i>software</i> BITOLA	116
Figura 41- Relação de artigos com as associações	117
Figura 42- Atualização na base de conhecimento	118

LISTA DE QUADROS

Quadro 1 - Regra típica do sistema MYCIN.....	41
Quadro 2 - Descrição dos modelos do CommonKADS.....	55
Quadro 3 - MO1: Problemas e Oportunidades.....	56
Quadro 4 - MO2: Aspectos Variantes.....	56
Quadro 5 - MO3: Detalhamento dos Processos	57
Quadro 6 - MO4: Ativos de Conhecimento	57
Quadro 7 - MO5: Checklist para a tomada de decisão de viabilidade	59
Quadro 8 - Visão das tarefas intensivas de conhecimento do tipo analítico	62
Quadro 9 - Visão das tarefas intensivas de conhecimento do tipo sintético.....	63
Quadro 10 - Caracterização geral da tarefa de diagnóstico.....	64
Quadro 11 - OTA1: <i>Checklist</i> para Decisão sobre Impactos e Melhorias.....	105
Quadro 12 - Base de Dados	112

LISTA DE TABELAS

Tabela 1 - Base de dados para a descoberta de novas relações	80
Tabela 2 - Resumo da amostra de profissionais	96
Tabela 3 - Porcentagens atribuídas em relação às evidências e aos sintomas	114
Tabela 4 - Relação de Fator Suporte e Confiança	116
Tabela 5 - Contato com algum sistema que auxilia no diagnóstico	119
Tabela 6 - Receios/medos em utilizar um sistema de apoio	120
Tabela 7 - Auxílio de um Sistema Baseado em Conhecimento	120
Tabela 8 - Necessidade de considerar diferentes variáveis para o contexto	121
Tabela 9 - Necessidade de considerar diagnósticos realizados	121
Tabela 10 - Crença na potencialidade dos SBC's	122
Tabela 11 - Possibilidade de representar o conhecimento	122
Tabela 12 - Visão sistêmica	123
Tabela 13 - Conhecimento em estado latente	124
Tabela 14 - Formação de uma base de dados	124
Tabela 15 - Uso de Sistemas de Descoberta Baseado na Literatura	125
Tabela 16 - Conhecimento comum em diferentes especialidades	126

LISTA DE ABREVIATURAS E SIGLAS

SBC - Sistema Baseado em Conhecimento
SE – Sistema Especialista
EC - Engenharia do Conhecimento
KADS - *Knowledge Acquisition and Documentation Structuring*
MO – Modelo de Organização
MC- Modelo de Conhecimento
CESM - *Composition – Environment – Structure – Mechanism*
CP – Cálculo Probabilístico
RB – Rede Bayesiana
DBL - Descoberta Baseada na Literatura
GC – Gestão do Conhecimento
EGC – Engenharia e Gestão do Conhecimento
OPAL - *Design and implementation of an algebraic programming language*
EON - *A Component-Based Approach to Automation of Protocol-Directed Therapy*
SOMKS - Symptoms Ontology for Mapping Diagnostic Knowledge Systems
OWL - *Web Ontology Language*
DL – *Description Logics*
RBC – Raciocínio Baseado em Casos
ODDIN - Ontology-driven differential Diagnosis based on logical Inference
SECI – Socialização – Externalização – Combinação - Internalização
IA – Inteligência Artificial
KDD - *Knowledge Discovery in Databases*
AC – Aquisição de Conhecimento
NB – Naïve Bayes
BD – Base de Dados
BC – Base de Conhecimento
FSup – Fator Suporte
FConf – Fator de Confiança
KDT - *Knowledge Discovery in Text*
AUC – Acne por Umidade e Calor
ACT – Acne por Calor Tóxico
ADQ – Acne por Deficiência do QI
IMC – Índice de Massa Corpórea
Diag_F – Diagnóstico Final
D_sin – Diagnóstico dos Sintomas
D_evi – Diagnóstico das Evidências

SUMÁRIO

1	INTRODUÇÃO.....	27
1.1	CONSIDERAÇÕES INICIAIS.....	27
1.2	JUSTIFICATIVA DO TRABALHO	31
1.3	ADERÊNCIA AO OBJETO DE PESQUISA DO PROGRAMA.....	31
1.4	PRESSUPOSTOS	32
1.5	DELIMITAÇÃO.....	33
1.6	OBJETIVO DO TRABALHO	34
1.6.1	Objetivo Geral.....	34
1.6.2	Objetivos Específicos	34
1.7	ORGANIZAÇÃO DO TRABALHO	35
2	FUNDAMENTAÇÃO TEÓRICA.....	37
2.1	DIAGNÓSTICO	37
2.1.1	Automatização do Processo de Diagnóstico	39
2.1.2	Trabalhos Correlatos	43
2.1.2.1	Modelo SOMKS	43
2.1.2.2	Modelo FENZA	45
2.1.2.3	Modelo SWE e KHAM.....	46
2.1.2.4	Modelo ODDIN	48
2.2	GESTÃO DO CONHECIMENTO	49
2.2.1	Conversão do Conhecimento.....	50
2.3	ENGENHARIA DO CONHECIMENTO	52
2.3.1	Engenharia do Conhecimento e a Metodologia CommonKADS...	53
2.3.1.1	Modelo de Organização no CommonKADS	55
2.3.1.2	Modelo de Conhecimento no CommonKADS	59
2.3.1.2.1	<i>Especificação do Conhecimento para a Tarefa de Diagnóstico</i>	<i>63</i>
2.3.2	Engenharia do Conhecimento e a Modelagem CESM.....	68
2.3.3	Explicitação do Conhecimento.....	68
2.3.4	Representação do Conhecimento (RC).....	69
2.3.4.1	Ontologia.....	71
2.4	DESCOBERTA DO CONHECIMENTO	73
2.4.1	<i>Knowledge Discovery in Databases (KDD)</i>.....	74
2.4.2	Descoberta Baseada na Literatura (DBL).....	76
2.4.2.1	Predicações Semânticas	79
2.5	TRATAMENTO DE INCERTEZAS.....	80
2.5.1	Categorias de Incerteza	81
2.5.2	Lógica Fuzzy	81
2.5.3	Redes Bayesianas.....	83
2.5.3.1	Naïve Bayes	84
2.6	MÉTODO DELPHI	85
2.6.1	Características do Método Delphi.....	86
2.7	CONSIDERAÇÕES SOBRE O CAPÍTULO	89

3	PROCEDIMENTOS METODOLÓGICOS.....	91
3.1	OPERACIONALIZAÇÃO DO MÉTODO DELPHI	92
3.2	ESPECIALISTAS PARTICIPANTES.....	94
4	MODELO PROPOSTO E APLICAÇÃO	97
4.1	ARQUITETURA DO MODELO PROPOSTO.....	97
4.2	DIAGRAMA DE ATIVIDADES.....	100
4.2.1	Fluxo de Apoio ao Diagnóstico	100
4.2.2	Fluxo para a Geração de Novo Conhecimento.....	102
4.3	APLICAÇÃO DO MODELO – ESTUDO DE CASO	103
4.3.1	Análise de Contexto.....	103
4.3.2	Construção da Ontologia	105
4.3.3	Construção da Base de Dados	112
4.3.4	Diagnóstico Probabilístico	113
4.3.5	Novo Conhecimento	115
5	ANÁLISE E DISCUSSÃO DOS RESULTADOS.....	119
5.1	ANÁLISE DOS RESULTADOS.....	119
5.2	DISCUSSÃO DOS RESULTADOS.....	126
5.3	ANÁLISE ENTRE MODELO PROPOSTO E CORRELATOS	128
5.3.1	Visão Sistêmica.....	128
5.3.2	Manipulação de Incertezas.....	128
5.3.3	Resultado Final do Diagnóstico.....	129
5.3.4	Análise de Evidências.....	129
6	CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS.....	131
6.1	CONCLUSÕES.....	131
6.2	SUGESTÕES PARA TRABALHOS FUTUROS.....	134
	REFERÊNCIAS.....	137
	APÊNDICE I – QUESTIONÁRIO.....	147
	APÊNDICE II – COMENTÁRIOS DOS ESPECIALISTAS.....	155
	APÊNDICE III – FERRAMENTAL	165
	APÊNDICE IV - GUIA DE UTILIZAÇÃO DO MODELO.....	169
	APÊNDICE V - PUBLICAÇÃO EM REVISTA NACIONAL	177
	APÊNDICE VI - PUBLICAÇÃO EM CONGRESSO NACIONAL.....	207
	APÊNDICE VII - PUBLICAÇÃO EM CONGRESSO NACIONAL	221

1 INTRODUÇÃO

1.1 CONSIDERAÇÕES INICIAIS

O diagnóstico, como tarefa intensiva em conhecimento, é um caso particular na tomada de decisão e, por ser um processo complexo, depende da execução de várias tarefas cognitivas que envolvem a aplicação da inteligência humana (SHORLIFFE *et al.*, 2001).

Van Bommel e Musen (1997) definem a tarefa de diagnosticar como a descrição de um problema, numa determinada área, em termos de conceitos conhecidos. Já, o processo de diagnóstico é definido como um conjunto de ações necessárias para obter o diagnóstico. Quando estas ações envolvem subjetividade e falta de informações atualizadas, o resultado do diagnóstico fica comprometido devido à inconsistência dos fatores que norteiam a tomada de decisão (KONG, XU & YANG, 2008).

O processo de diagnóstico, por exemplo, na área da educação, da saúde, da qualidade de água, entre outros, envolve grande quantidade de imprecisões e incertezas (UZOKA *et al.*, 2010). Na saúde, por ser uma área onde profissionais e outros envolvidos fazem parte das ações, as incertezas aparecem em quase todos os estágios (KONG, XU & YANG, 2008). Nela, pacientes não sabem descrever exatamente o que estão sentindo; médicos e enfermeiras não relatam precisamente as observações feitas; laboratórios registram resultados analisados por técnicos com algum grau de erro; pesquisadores médicos não caracterizam exatamente como as doenças alteram o funcionamento normal do corpo humano (HAIMOWITZ *et al.*, 1988).

Há que considerar também, segundo Sackett *et al.* (2003), que outros fatores dificultam a tomada de decisão, como: a) a necessidade de informações recentes válidas para o diagnóstico; b) a inadequação das fontes tradicionais para fornecer essas informações, porque estão desatualizadas, são ineficazes e muito volumosas; c) a disparidade existente entre a habilidade diagnóstica, que aumenta com a experiência, e o conhecimento atualizado que diminui com o passar do tempo; e d) a dificuldade do profissional em dedicar algum tempo para encontrar e assimilar evidências, procurando assim atualizar sua base de conhecimento através de leitura e estudo.

Ao longo dos anos, informações e dados coletados através de diagnósticos realizados foram e ainda continuam sendo armazenados em fichas de papel ou, mais recentemente, em fichários eletrônicos.

Estes elementos, que de certa forma mantêm-se encobertos, se colhidos e tratados adequadamente através dos recursos disponíveis, podem tornar-se uma valiosa fonte de conhecimento quando associados à experiência e observação do profissional.

O avanço tecnológico ocorrido nos últimos anos tem impulsionado o desenvolvimento de Sistemas Baseados em Conhecimento (SBC) para auxiliar nas tarefas intensivas em conhecimento (ZHAO, YANXIANG & HUI, 2005; RODRÍGUEZ *et al.*, 2009a). Entretanto, alavancar o potencial destas aplicações, em relação ao processo de diagnóstico, tem sido uma ação crítica e desafiadora no tocante à precisão e eficácia destes sistemas (GARCÍA-CRESPO *et al.*, 2010).

Um dos principais desafios na representação e no raciocínio, a respeito do conhecimento dentro de um domínio específico, é como minimizar estas imprecisões e incertezas de forma que os SBC possam auxiliar os profissionais em suas tomadas de decisão (LIN, HU & LIU SHENG, 2006).

O desenvolvimento de sistemas que utilizam a inteligência computacional para auxiliar o processo de diagnóstico teve seu início por volta dos anos 60, através dos Sistemas Especialistas (SE), e com o passar do tempo, muitos deles encontram-se desatualizados e com tecnologias obsoletas (RODRÍGUEZ *et al.*, 2009b).

Coiera (2003), ao comentar sobre os SBC, enfatiza que as limitações que ocorrem são causadas por serem muito especializados e precisos somente em alguns domínios e irreais em outros ou são extremamente simplificados. Conclui afirmando que esses sistemas não são capazes de detectar quando estão fazendo uma classificação incorreta que pode levar a um diagnóstico incorreto.

Minchin *et al.* (2006) sugerem que o foco de desenvolvimento destes sistemas esteja na construção de modelos em que a estrutura que representa o conhecimento de um domínio específico e o SBC sejam elaborados em conjunto com especialistas no domínio. Além disso, recomendam que as técnicas probabilísticas sejam aplicadas sobre banco de dados, de modo que as regras de decisão possam ser atualizadas constantemente.

Fuentes-Lorenzo *et al.* (2009) ressaltam que as tecnologias semânticas que estão sendo desenvolvidas e melhoradas ao lado do avanço da Web Semântica devem ser exploradas para revelar relações ocultas dentro das informações disponíveis. Salientam ainda que, nas áreas onde a homogeneidade terminológica é particularmente

problemática, grandes benefícios podem ser obtidos com o emprego desta tecnologia.

O desenvolvimento de pesquisas, propondo a inserção de novos conhecimentos à base de conhecimento de tais sistemas, trarão maior precisão e eficiência ao processo de diagnóstico (GARCÍA-CRESPO *et al.*, 2010; RODRÍGUEZ *et al.*, 2009b).

Diante deste contexto, uma nova postura pode ser dada a este processo complexo que envolve um alto grau de raciocínio e uma grande quantidade de incertezas na tomada de decisão. Lançar mão dos métodos e técnicas disponíveis para permitir que possíveis novos conhecimentos sejam gerados e contribuam para um processo dinâmico, que realize atualizações constantemente, é uma das metas a serem alcançadas.

Portanto, o problema atual a ser enfrentado é como utilizar as informações e os dados colhidos, adequadamente, em diagnósticos que vêm sendo realizados, de forma que novos conhecimentos possam ser gerados e contribuam para um aperfeiçoamento desta tarefa.

A Engenharia do Conhecimento (EC) é uma área que procura aplicar métodos científicos para a construção de SBC. Estes sistemas são desenvolvidos para auxiliar os humanos na resolução de problemas nos mais diversos setores e atividades, oferecendo suporte à GC nas atividades intensivas em conhecimento (SCHREIBER *et al.*, 2002).

A metodologia CommonKADS (Knowledge Acquisition and Documentation Structuring) (SCHREIBER *et al.*, 2002) fornece um conjunto de etapas para o desenvolvimento de SBC. No nível de contexto, o Modelo de Organização (MO) desta metodologia suporta a maioria dos recursos da organização para descobrir problemas e oportunidades, estabelecendo sua viabilidade e avaliando os possíveis impactos do desenvolvimento destes sistemas.

No nível de conceito, o Modelo de Conhecimento (MC) desta metodologia procura especificar tarefas intensivas de conhecimento, explicando em detalhes o tipo e a estrutura do conhecimento usado na execução de uma tarefa (SCHREIBER *et al.*, 2002).

Algumas destas etapas da metodologia CommonKADS podem ser aprimoradas utilizando a modelagem proposta por Bunge (2003), através do modelo CESM (*Composition – Environment – Structure – Mechanism*), o qual proporciona uma visão sistêmica e pode contribuir, significativamente, para a geração, elicitação e estruturação do conhecimento em um domínio específico.

No processo de geração do conhecimento, CALHOUN e STARBUCK (2005) comentam que esta pode ocorrer por meio da

organização de conhecimento anterior em novas formas, da combinação de informações relevantes, ou mesmo de *insights* acerca da aplicação de conhecimento existente em novos contextos.

Gómez-Pérez (1999), dentro do processo de elicitación do conhecimento, destaca a técnica da ontologia como forma eficiente de representar o conhecimento através de um conjunto de termos ordenados hierarquicamente para descrever um domínio que pode ser usado como um esqueleto para uma base de conhecimento.

Noy e McGuinness (2001) salientam que, além de ser uma forma eficiente de descrever e estruturar um domínio específico, uma ontologia tem por objetivo definir um vocabulário comum para pesquisadores que necessitam compartilhar informações dentro de uma especialidade.

Como principais características da ontologia tem-se a forma de representar a estrutura organizacional de um grande domínio complexo e a sua capacidade de raciocinar sobre ele. Porém, sua aplicação é limitada devido à inabilidade de lidar com subjetividade e incertezas (KOLLER e PFEFFER, 1998; COSTA *et al.*, 2009).

Esta deficiência encontrada na representação do conhecimento através da ontologia pode ser contornada utilizando o Cálculo Probabilístico (CP) que, através das redes causais (como as redes Bayesianas (RB)), possibilita manipular “variáveis” devido à sua excelente expressividade gráfica e elevado poder computacional (YANG e CALMET, 2005).

Portanto, tem-se na ontologia associada ao cálculo probabilístico uma estrutura adequada complementar que aproveita a potencialidade de cada um deles no apoio ao processo de diagnóstico.

Outro ponto importante a ser observado na rede causal é o refinamento probabilístico feito sobre uma base de dados. Registros de diagnósticos realizados, pelo profissional, podem tornar-se informações relevantes, quando extraídos de forma adequada desta base de dados. Uma evidência destacada, ao ser confirmada por meio dos Sistemas de Descoberta Baseada na Literatura (DBL) (HRISTOVSKI *et al.*, 2006), e comprovada pela comunidade científica, poderá gerar um novo conhecimento a ser inserido à base de conhecimento da ontologia.

Assim entende-se que a Engenharia do Conhecimento, através de conceitos, métodos e tecnologias que estão disponíveis e possibilitam manipular fontes de conhecimento, explorando a potencialidade de cada uma delas, pode contribuir significativamente para o aperfeiçoamento do processo de diagnóstico.

1.2 JUSTIFICATIVA DO TRABALHO

No contexto deste trabalho encontra-se a originalidade na construção de um modelo de Engenharia do Conhecimento, cujo fundamento está na representação formal do conhecimento, no cálculo probabilístico e na descoberta baseada na literatura para o apoio ao diagnóstico no processo da geração de novos conhecimentos. Informações e dados coletados, provenientes de diagnósticos realizados, quando colhidos e tratados adequadamente, podem tornar-se valiosa fonte de conhecimento.

Ao integrar diferentes abordagens técnicas e teóricas em um modelo, inicialmente conceitual, pretende-se cooperar cientificamente de modo efetivo no aprimoramento da tarefa de diagnóstico. O desenvolvimento deste modelo permite também, como contribuição científica, dar aos profissionais, através de uma visão sistêmica ao analisar uma queixa, possibilidade de evidenciar novos conhecimentos para o avanço metodológico deste processo.

Convém lembrar que este trabalho tem sido motivado, também, pela complexidade do processo de diagnóstico e a disponibilidade de novos métodos e técnicas que a Engenharia do Conhecimento tem para apoiá-lo. A construção de um SBC que permita ao profissional trabalhar com as informações e dados obtidos sistemicamente, através de diagnósticos realizados, servirá para explicitação de conhecimento latente que até então permanecia encoberto ou que não vinha sendo utilizado.

1.3 ADERÊNCIA AO OBJETO DE PESQUISA DO PROGRAMA

Inicialmente, a Engenharia do Conhecimento (EC) foi tratada como uma subárea da Inteligência Artificial na construção de sistemas de conhecimento para a solução de problemas específicos; transcendeu esta visão ao considerar todo o contexto sistêmico organizacional das atividades intensivas em conhecimento.

Assim, a EC deve prover um conjunto de métodos, técnicas e ferramentas que forneçam suporte à Gestão do Conhecimento (GC) a partir da formalização e explicitação das atividades intensivas em conhecimento nas organizações (SCHREIBER *et al.*, 2002).

O diagnóstico, como tarefa intensiva em conhecimento, é um processo complexo que visa determinar porque um sistema corretamente projetado não funciona conforme esperado e, ao envolver

subjetividade e falta de informações atualizadas, fica sujeito à insegurança e à ineficácia na produção do resultado final.

A proposta a ser alcançada neste trabalho é utilizar metodologias, métodos e técnicas que a EC disponibiliza para apoiar este processo, através da geração de novos conhecimentos, seja na área da saúde, educação ou qualquer outra área, objetivando a proposta que fundamenta a Engenharia e Gestão do Conhecimento no tocante à interdisciplinaridade entre as áreas do conhecimento.

Este trabalho está inserido na linha de pesquisa de teoria e prática em Engenharia do Conhecimento, que tem como foco estudar as metodologias e técnicas dessa área e suas relações com a gestão do conhecimento.

A aderência desse trabalho ao objeto de pesquisa do programa pode ser reforçada a partir do objetivo de pesquisa e objetivo principal do Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento (EGC):

O objeto de pesquisa do EGC refere-se aos macro-processos de explicitação, gestão e disseminação do conhecimento. Estes incluem os processos de criação (e.g., inovação de ruptura), descoberta (e.g., redes sociais), aquisição (e.g., inovação evolutiva), formalização / codificação (e.g., ontologias), armazenamento (e.g., memória organizacional), uso (e.g., melhores práticas), compartilhamento (e.g., comunidades de prática), transferência (e.g., educação corporativa) e evolução (e.g., observatório do conhecimento) [...] Deste modo, o objetivo do EGC consiste em investigar, conceber, desenvolver e aplicar modelos, métodos e técnicas relacionados tanto a processos/bens/serviços como ao seu conteúdo técnico-científico [...].

1.4 PRESSUPOSTOS

Considerando o problema anteriormente mencionado, os seguintes pressupostos são apresentados:

- Grande parte dos dados e informações, que são gerados no processo de diagnóstico, não está sendo utilizada como fonte de pesquisa. Esta característica sugere que se desenvolva modelo específico para lidar com esses elementos;

- A subjetividade, imprecisões e incertezas, presentes no processo de diagnóstico, quando colhidas e tratadas adequadamente fornecem informações relevantes a serem analisadas;
- O desenvolvimento de métodos para descoberta de conhecimento, a partir da análise das associações e dos relacionamentos semânticos entre elementos textuais de um domínio, está crescendo rapidamente e pode auxiliar na obtenção de novos conhecimentos;
- Um modelo de Engenharia do Conhecimento que associa a representação do conhecimento por meio de ontologia, o cálculo probabilístico através das redes causais e a descoberta baseada na literatura, pode contribuir para geração e validação de novos conhecimentos, com a participação de especialistas, para o apoio ao processo de diagnóstico.

Portanto, o diagnóstico como tarefa intensiva em conhecimento, é um processo complexo que ao considerar, dentro de uma visão sistêmica, as variáveis do mundo real, necessita de um conjunto de conhecimento que possa torná-lo mais seguro e eficaz.

Neste contexto, o trabalho procura explorar o seguinte problema:

- Como manipular dados e informações, obtidos no processo de diagnóstico, de forma que novos conhecimentos possam ser gerados e contribuam para o aperfeiçoamento desta tarefa intensiva em conhecimento?

1.5 DELIMITAÇÃO

O que esta pesquisa propõe é gerar um modelo de Engenharia do Conhecimento para apoio ao profissional no processo de diagnósticos. O modelo é composto por módulos integrados que interagem entre si, proporcionando maior precisão e eficácia na tomada de decisão.

Salienta-se que o trabalho não tem como objetivo o desenvolvimento de uma ferramenta computacional voltada ao usuário final, bem como não tem o propósito de discutir a organização da ontologia, a rede causal e a base de dados a serem utilizadas para gerar novos conhecimentos.

A validação da proposta foi realizada de forma exploratória, pois a argumentação e o convencimento, segundo Wazlawick (2008), são as principais ferramentas do pesquisador em áreas emergentes.

Entrevistas e questionários foram aplicados aos especialistas, através do método Delphi. A análise comparativa entre os modelos correlatos e o modelo proposto, bem como um estudo de caso que propicie uma visão de aplicação prática, também são apresentados.

1.6 OBJETIVO DO TRABALHO

O objetivo do trabalho está na elaboração conceitual de um Sistema Baseado em Conhecimento que, através da integração de métodos, técnicas e ferramentas que a Engenharia do Conhecimento disponibiliza, possibilite o aprimoramento do processo de diagnóstico, tornando-o mais seguro e eficaz.

1.6.1 Objetivo Geral

O objetivo geral deste trabalho consiste na proposição de um modelo de Engenharia do Conhecimento baseado em ontologia e no cálculo probabilístico que possibilita a criação de novos conhecimentos para apoiar o processo de diagnóstico.

1.6.2 Objetivos Específicos

Para se alcançar o objetivo geral, foram estabelecidos os seguintes objetivos específicos:

- Identificar pontos fortes e fracos dos modelos de SBC fundamentado em ontologias que sejam voltados à tarefa de diagnóstico;
- Identificar na literatura o ferramental de Engenharia do Conhecimento que possa contribuir na tarefa de diagnóstico e com a descoberta de novos conhecimentos;
- Propor um modelo voltado à melhoria do processo de diagnóstico através da geração de novos conhecimentos que leve em consideração os objetivos específicos anteriores;
- Demonstrar a viabilidade do modelo proposto por meio do desenvolvimento de um protótipo assim como sua aplicação em um estudo de caso;

- Realizar análise comparativa do modelo proposto com outros modelos correlatos identificados na literatura; e
- Verificar aspectos de consistência do modelo proposto através da avaliação feita por profissionais e especialistas que atuam com a tarefa de diagnóstico.

1.7 ORGANIZAÇÃO DO TRABALHO

Este trabalho é composto de 6 capítulos, além da introdução já apresentada, os demais capítulos são relacionados a seguir.

No capítulo 2 apresenta-se a fundamentação teórica. Inicialmente, são abordados o conceito de diagnóstico, o histórico dos sistemas desenvolvidos em uma determinada área do conhecimento e os trabalhos correlatos atuais a esta pesquisa; a Engenharia do Conhecimento através da metodologia CommonKADS auxiliada pela modelagem CESM estruturam a construção do modelo proposto de Sistema Baseado em Conhecimento; o tratamento dado às incertezas através das redes causais, e os sistemas de descoberta de conhecimento proporcionam a geração de um possível novo conhecimento. Encerrando este capítulo é apresentado o Método Delphi que tem como objetivo verificar a consistência do modelo proposto.

O Capítulo 3 apresenta os procedimentos metodológicos utilizados para a elaboração do Modelo de Conhecimento.

O capítulo 4 exhibe o Modelo de Engenharia do Conhecimento e o aplica num estudo de caso na área da saúde. Utilizaram-se trabalhos desenvolvidos pelo autor, e aprovados pela comunidade acadêmica, para compor este estudo de caso.

O Capítulo 5 apresenta os resultados obtidos através da pesquisa feita entre os especialistas, propondo discussões/reflexões sobre os temas abordados bem como uma análise entre o modelo proposto e os modelos correlatos. No último Capítulo são colocadas as conclusões da tese e as sugestões para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

A Engenharia do Conhecimento tem por objetivo investigar e propor modelos, estabelecendo dentro do contexto sistêmico, metodologias, métodos e ferramentas para atividades intensivas em conhecimento no apoio à gestão do conhecimento organizacional.

Dentre as atividades intensivas em conhecimento nas organizações, destaca-se a tarefa de diagnóstico, como uma tarefa de elevado nível intelectual e cognitivo, responsável pela tomada de decisões dentro de critérios previamente constituídos.

2.1 DIAGNÓSTICO

A palavra “diagnóstico” tem origem grega, significando discernir ou distinguir. Segundo Kleer e Williams (1987), entende-se por diagnóstico o processo que visa determinar porque um sistema corretamente projetado não funciona conforme esperado. Assim, para realizar a tarefa de diagnóstico é necessário observar e levantar hipóteses de modo a chegar a um diagnóstico (resultado) descrevendo o estado do processo observado. De forma simplificada, Schreiber *et al.* (2002) destacam que o objetivo do diagnóstico é encontrar a causa do mau funcionamento de um sistema, por exemplo, na área da saúde, o diagnóstico estabelece a identificação de uma doença pela investigação de suas manifestações.

Taylor, Aitchison e McGirr (1971) apontaram algumas fontes de incertezas, analisando o processo de diagnóstico clínico como um processo de tomada de decisão:

- Os dados relativos à doença do paciente podem estar incompletos, o que pode ocorrer no começo da doença antes que surjam evidências suficientes, ou em qualquer período, se o clínico não obteve informações completas;
- Pode haver uma falta de eficácia na análise sistemática, mesmo que os dados sejam corretos e suficientes para concluir o diagnóstico;
- O profissional pode deixar-se influenciar pela sua hipótese diagnóstica inicial, de modo a deixar de colher dados essenciais ou por interpretar erroneamente as evidências acumuladas.

Além destes fatores, pode haver falta de conhecimento do profissional acerca da doença específica ou da especialidade médica. A figura 1 mostra um esquema simplificado de alguns fatores que podem

induzir a erros no diagnóstico, a partir de informações sobre sintomas, sinais, história clínica e resultados de exames do paciente.

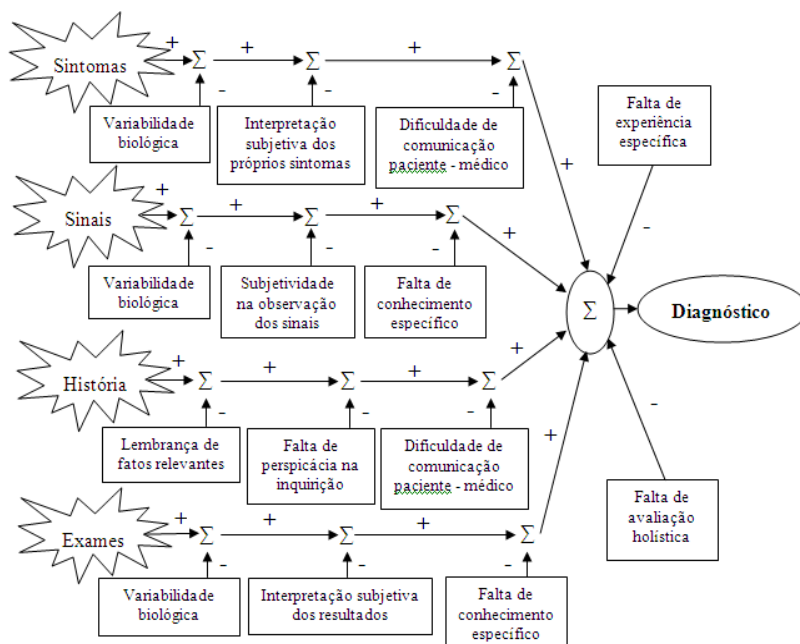


Figura 1 - Alguns fatores que podem induzir a erros de diagnóstico

Fonte: adaptado de LOPES (1996)

O processo de diagnóstico clínico basicamente envolve dois procedimentos: a coleta de fatos e a análise destes fatos. O processo é bastante complexo e os erros no diagnóstico podem ser devido à conduta imperfeita em ambos os procedimentos.

Podgorelec, Grasic e Pavlic (2009) sintetizaram as duas principais etapas envolvidas no processo de diagnóstico médico:

a) Coleta de Fatos: A medição é a base fundamental da ciência. Entretanto, apesar do grande avanço científico, ainda existem dificuldades para o profissional da saúde se expressar em termos quantitativos e objetivos durante o procedimento de coleta de informações do paciente. A coleta de fatos compreende os seguintes tópicos, em ordem cronológica:

- **História:** corresponde à anamnese do paciente, onde se cataloga não só informações periféricas tais como a história familiar, mas também os sintomas de que o paciente se queixa. Entende-

se por sintoma uma sensação subjetiva ou outra observação pessoal que o próprio paciente descreve para o médico;

- Exame físico: sempre que possível, os sinais observados devem ser verificados quantitativa ou qualitativamente. Entende-se por sinal qualquer desvio da normalidade observada pelo médico durante o exame físico;
- Exames auxiliares: a partir de uma análise preliminar se estabelece uma ou mais hipóteses diagnósticas que vão orientar a seleção dos exames;
- Observação da evolução: algumas condições apresentam poucos sintomas e sinais característicos e o seu diagnóstico definitivo só é possível através da observação da evolução da doença no tempo.

b) Análise dos Fatos: Após o profissional terminar de categorizar as evidências coletadas, a etapa seguinte consiste em classificá-las de modo a concluir um diagnóstico.

Em suma, espera-se que ao final do processo de diagnóstico, com base nos achados e nas hipóteses levantadas, o profissional chegue à conclusão de que o paciente possui determinada condição, patologia ou doença.

2.1.1 Automação do Processo de Diagnóstico

Atualmente com o avanço tecnológico, o processo de diagnóstico, seja em que área for, tem sido aprimorado através dos Sistemas Baseados em Conhecimento que buscam a segurança e a eficiência na tomada de decisão (MINCHIN *et al.*, 2006; FENZA *et al.*, 2008; SWE e KHAM, 2009; GARCÍA-CRESPO *et al.*, 2010).

Neles, uma grande quantidade de dados e informações levantados pelo profissional, que até então ficava armazenada em fichários na forma de papel, hoje em dia sendo colocada em fichários eletrônicos aguardando o momento oportuno para ser utilizada como fonte de pesquisa e aprimoramento científico.

Ainda uma meta a ser alcançada é proporcionar meios que possibilitem a utilização destas fontes de conhecimento, dentro do processo de diagnóstico, de forma que o profissional possa estar sempre se atualizando e contribua, através de seu exercício, para o avanço da ciência.

O uso de máquinas para auxiliar no processo de diagnóstico tem sido almejado pelos profissionais mesmo antes do início da era

computacional (MUSEN, SHAHAR e SHORTLIFFE, 2006). Os primeiros artigos sobre esta utilização apareceram por volta de 1950 e, logo em seguida, os primeiros protótipos experimentais.

Tomando-se como ponto de partida para os sistemas que auxiliam na tomada de decisão, na saúde o sistema de Dombal *et al.* (1972), apoiado no teorema de Bayes (1763), foi concebido para realizar o diagnóstico em 304 pacientes com dor abdominal. Expressando relações de causa e efeito, este modelo Bayesiano associa valores de probabilidade entre as variáveis de um determinado domínio do conhecimento em que os dados destes pacientes tornaram-se os atributos para a classificação em uma das sete patologias propostas para tratamento.

Os resultados obtidos pelo sistema mostraram que 91,8% obtiveram acerto nas indicações, em contraste com o diagnóstico efetuado pelos médicos, que acusou um índice de acerto variando entre 65 a 80% (MUSEN, SHAHAR e SHORTLIFFE, 2006).

Na sequência, os Sistemas Especialistas começaram a ser desenvolvidos para resolver problemas em um determinado domínio cujo conhecimento utilizado é obtido a partir de especialistas naquele domínio (LUGER, 2009).

A figura 2 representa o desenvolvimento de um Sistema Especialista onde o Engenheiro do Conhecimento obtém as informações de um Especialista no Domínio, para formar a base de conhecimento (conhecimento estático). Uma máquina de inferência (regras – conhecimento dinâmico do sistema) associada à base de conhecimento elabora a tomada de decisão.

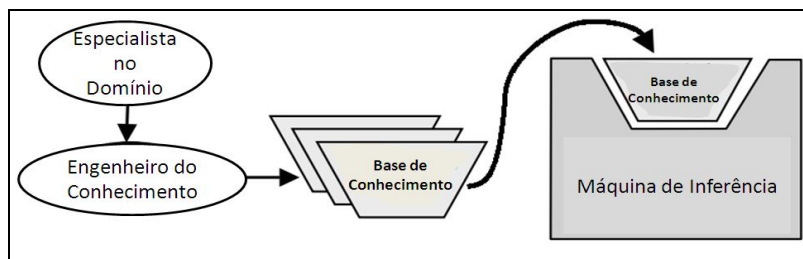


Figura 2 - Componentes de um SE

Fonte: adaptado de Gennari *et al.* (2003)

Salienta-se que esta forma de trabalhar com o conhecimento de um domínio específico não corresponde à conceituação atual da EC. Adiante no item 2.3.1, correspondente à EC, será visto que várias etapas compõem a construção de um Sistema Baseado em Conhecimento.

Dando continuidade, Edward Shortliffe (1976) desenvolveu um Sistema Especialista para a área médica (MYCIN), o qual teve um papel fundamental no desenvolvimento dos futuros Sistemas Especialistas. O sistema MYCIN foi concebido para resolver o problema do diagnóstico e tratamento de doenças infecciosas do sangue através de um conjunto de regras (SHORTLIFFE, 1976).

O conhecimento de doenças infecciosas no MYCIN foi representado através de regras de produção, conforme quadro 1, cada uma contendo um “pacote” de conhecimento derivado a partir de discussões com a colaboração de especialistas. Embora o programa nunca tenha sido usado clinicamente, ele “pavimentou” o modo de pesquisa e desenvolvimento nos anos 80 (MUSEN, SHAHAR e SHORTLIFFE, 2006).

Regra 507 Se:

- | |
|---|
| <ol style="list-style-type: none"> 1. A infecção que requer terapia é meningite, 2. Não foram vistos organismos na marcha da cultura, 3. O tipo de infecção é por bactéria, 4. O paciente não tem um dano detectado na cabeça, e 5. A idade do paciente está entre 15 anos e 55 anos |
|---|

Então: Os organismos que podem estar causando a infecção são <i>diplococcus-pneumoniae</i> e <i>neisseria- meningitidis</i>
--

Quadro 1 - Regra típica do sistema MYCIN

Fonte: adaptado de MUSEN, SHAHAR e SHORTLIFE (2006)

Segundo Buchanan e Duda (1983), diversos problemas apareceram no desenvolvimento dos sistemas especialistas. Iniciando pela dificuldade que o engenheiro do conhecimento tinha para modelar o conhecimento do especialista e a necessidade da sua participação em todas as fases (Identificação, conceitualização, formalização, implementação, teste e revisão) para a construção do sistema.

Ampliando o conceito dos sistemas especialistas, os SBC foram desenvolvidos com o objetivo de separar explicitamente o conhecimento de domínio do resto do sistema. Através da codificação simbólica de conceitos derivados de especialistas em um campo, formava-se uma base de conhecimento utilizada para prover o tipo de análise de problema e conselho que um especialista pode realizar.

A rápida evolução do conhecimento médico tornou a manutenção da base de conhecimento um importante problema. Pesquisadores desenvolveram um programa de computador com Linguagem de

Programação Algébrica (OPAL) para facilitar a interação direta entre a base de conhecimento e o especialista, sem a necessidade de um programador de computador (ERICSSON & MUSEN, 1993; LANZOLA *et al.*, 1995).

O Sistema Oncocin (1981) foi um dos primeiros SBC que tentou modelar decisão e sequenciar ações ao longo do tempo (figura 3). Utilizado em recomendações para terapia de doenças cancerígenas baseadas em protocolo, envolveu o uso de representação do conhecimento, em particular a representação do protocolo de tratamento clínico baseado em regras “se – então” (GENNARI *et al.*, 2003; MINCHIN *et al.*, 2006).

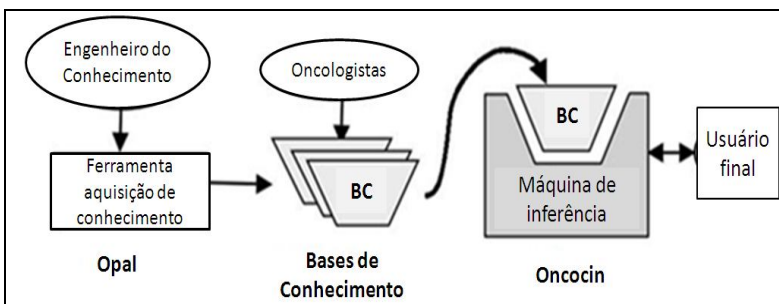


Figura 3 - Sistema Baseado em Conhecimento “Oncocin”

Fonte: adaptado de Gennari *et al.* (2003)

O projeto de abordagem baseada em componentes EON, apresentado na figura 4, veio dar continuidade ao sistema Oncocin, sendo considerado como a segunda geração de Sistema Baseado em Conhecimento. Este projeto é composto por vários programas que planejam a terapia baseada em protocolo e determinam se um paciente é potencialmente qualificado para determinado protocolo de tratamento. O modelo EON, como diretriz de raciocínio, é bastante geral e pode ser usado para qualquer diretriz de tratamento médico (MUSEN *et al.*, 1996).

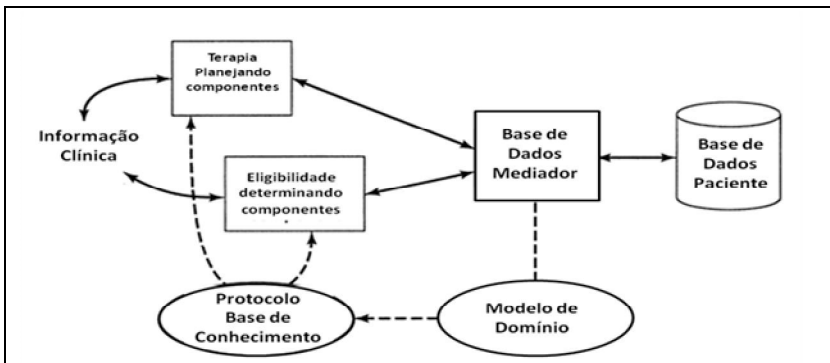


Figura 4 - Projeto EON

Fonte: adaptado de MUSEN *et al.* (1996)

Para a avaliação e o tratamento da hipertensão, foi projetado o sistema Athena (2000) utilizando a arquitetura do modelo EON para controle da pressão sanguínea. ATHENA tem dois componentes principais: uma base de conhecimento que modela o conhecimento da hipertensão independentemente de seu uso e uma diretriz intérprete que cria interface específica para recomendação de tratamento do paciente. O objetivo foi separar o conhecimento declarativo do conhecimento procedural (GOLDSTEIN *et al.*, 2000).

2.1.2 Trabalhos Correlatos

Atualmente, diferentes abordagens têm sido propostas para construir SBC no apoio ao processo de diagnóstico. Entre eles, alguns trabalhos que utilizam tecnologia semântica e propõem formas de lidar com as imprecisões e incertezas, serão analisados na sequência.

2.1.2.1 Modelo SOMKS

O modelo SOMKS (*Symptoms Ontology for Mapping Diagnostic Knowledge Systems*), proposto por Minchin *et al.* (2006), figura 5, é uma abordagem organizada para tratar de domínios abrangentes e complexos. Uma ontologia extensa de sintomas é utilizada para mapear domínios hierárquicos de decisões quando associada com informação estatística de classes de referência e um conjunto de SBC associado a cada classe/domínio que compõem este modelo.

Inicialmente sintomas obtidos são examinados em uma ontologia para definir o menor nível da superclasse que contém todas as classes com os sintomas especificados.

O SBC define os atributos estatísticos que são significantes em uma forma de recomendação para a escolha da classe na camada mais inferior dentro da superclasse. Após a avaliação, o SBC será associado à subclasse recomendada.

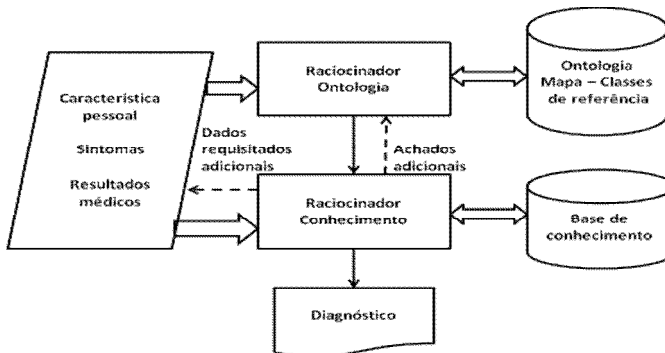


Figura 5 - Modelo SOMKS

Fonte: adaptado de MINCHIN *et al.* (2006)

Numa visão geral deste modelo tem-se a proposta de imitar a situação onde, inicialmente, um clínico geral analisa todos os sintomas e características apresentadas pelo paciente. Em seguida, encaminha-o para o especialista do domínio específico que avaliará com maior rigor as informações recebidas a fim de fornecer o diagnóstico.

Portanto, o modelo SOMKS é um sistema baseado em regras (OWL – *Web Ontology Language* – DL – *Description Logics*) que procura cobrir globalmente as enfermidades (abrangência de escopo: clínico geral) e identifica quais sinais e sintomas são relevantes, levando em consideração o cenário no qual o paciente está inserido, a fim de maximizar a significância das variáveis (focaliza $\Rightarrow$ especialista).

Como sugestão para futuras pesquisas, Minchin *et al.* (2006) propõem que a estrutura da ontologia e os SBC devem ser desenvolvidos em conjunto com especialistas do domínio. As estatísticas das classes de referências necessitam de aprendizagem probabilística ou as regras de decisão devem ser aprendidas a partir de dados registrados pelo especialista.

Uma apreciação deste modelo quanto aos pontos fortes:

- Abrangência de escopo: cobertura global;

- Tratamento dado às incertezas: redes Bayesianas (probabilístico);
- Classes de referência: o paciente é caracterizado dentro de um conjunto de indivíduos com peculiaridades semelhantes (membro típico) a fim de potencializar seus sinais e sintomas.

Uma apreciação deste modelo quanto aos pontos fracos:

- Não possui um mecanismo de aperfeiçoamento contínuo do sistema;
- Não utiliza os registros clínicos coletados para descobrir novos conhecimentos.

2.1.2.2 Modelo FENZA

Proposto por Fenza *et al.* (2008), apresentado na figura 6, o modelo utiliza a modelagem *fuzzy* envolvendo agentes e técnicas orientadas a serviços para aprimorar os dados adquiridos do paciente e treiná-los, considerando o conhecimento colecionado.

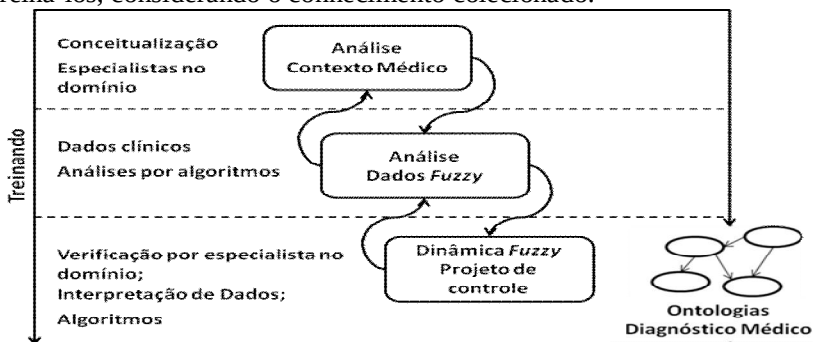


Figura 6 - Modelo FENZA

Fonte: adaptado de Fenza *et al.* (2008)

O uso de ontologias é fundamental para a usabilidade do sistema, pois permite a descrição semântica dos dados e as inter-relações sobre os conceitos conectados, refletindo a granularidade da informação processada. Agentes adaptáveis administram esta ontologia ou os conceitos específicos de informações descobertas para possível extração de relação sobre os dados que são úteis durante a formulação do diagnóstico médico.

Os dados clínicos fornecidos pelo paciente são coletados e dispostos adequadamente para serem processados por meio de técnicas

de agrupamento *fuzzy*. Através de filtragens e normalizações, as informações ganham em precisão e homogeneidade de representação gerando conhecimento em diversos níveis de granularidade de acordo com a profundidade do diagnóstico.

Tem-se na facilidade de anotação dos dados, por meio da ferramenta linguística *fuzzy*, uma forma mais adaptável de interpretação e compreensão humana. As regras *fuzzy* geradas para o processo de raciocínio, que produzem o diagnóstico, estão sujeitas à validação por parte dos especialistas.

Uma apreciação deste modelo quanto aos pontos fortes:

- Tratamento dado às incertezas através da lógica *fuzzy*: forma mais adaptável de interpretação e compreensão humana através de regras;
- Os dados coletados dos pacientes são processados através de filtros e procedimentos de normalização: precisão e homogeneidade de representação da informação;
- novas relações podem ser descobertas através de conceitos vagamente correlacionados: aperfeiçoamento do sistema.

Uma apreciação deste modelo quanto aos pontos fracos:

- É um sistema que potencializa a precisão dos dados coletados e não a descoberta de um novo conhecimento: dados preestabelecidos;
- O mecanismo de aperfeiçoamento contínuo é feito através da análise humana (especialistas) sobre novas relações entre os conceitos já estabelecidos: não ocorre consulta automática aos arquivos literários.

2.1.2.3 Modelo SWE e KHAM

Proposto por Swe e Kham (2009), figura 7, o modelo é baseado em uma estratégia de resolver problemas através do reuso de soluções passadas (casos) focando novos problemas. Uma experiência passada é armazenada na forma de problemas resolvidos em uma “base de casos”. Quando um novo caso surge, o modelo procura caso similar a fim de identificar a solução para o problema atual.

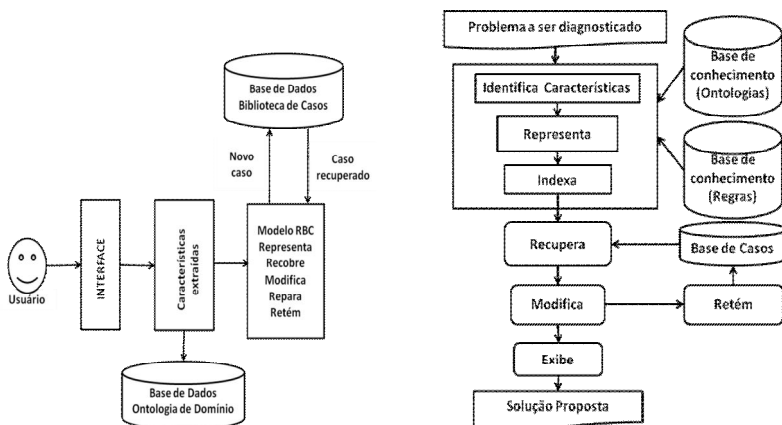


Figura 7 - Modelo SWE e KHAM
Fonte: adaptado de Swe e Kham (2009)

Este modelo consiste de dois componentes principais. O primeiro componente é a ontologia de domínio que é usada para estruturar o caso como características padronizadas. O segundo componente é o módulo RBC usado para armazenar o conhecimento de diagnóstico médico e recuperar o caso mais similar de acordo com o desejo e a necessidade do usuário.

Uma apreciação deste modelo quanto aos pontos fortes:

- Utiliza uma memória de casos passados como fonte de pesquisa para a tomada de decisão em um novo caso;
- Possibilita atualizar o sistema incluindo novos casos à memória de trabalho.

Uma apreciação deste modelo quanto aos pontos fracos:

- É preciso ter uma métrica adequada para medir a similaridade de casos;
- Não tem mecanismo para trabalhar com incertezas;
- O sistema não possibilita realçar uma evidência como um possível novo conhecimento.

2.1.2.4 Modelo ODDIN

O modelo ODDIN (*Ontology-driven differential diagnosis based on logical inference and probabilistic refinements*), proposto por García-Crespo *et al.* (2010), apresentado na figura 8, tem o comportamento de um Sistema Especialista baseado em regras. Utiliza ontologias combinadas com estratégias de computação probabilística de vários fatores e lógica de inferência. O sistema permite análise dos fatores envolvidos, atribuindo pesos aos parâmetros introduzidos. A probabilidade para cada diagnóstico é calculada como uma função dos parâmetros em termos de percentagens que são introduzidas ao sistema.

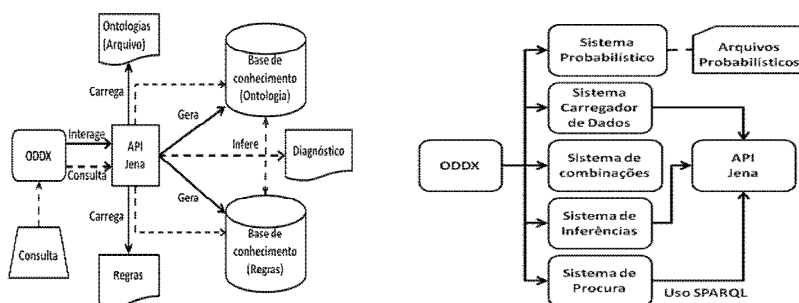


Figura 8 - Modelo ODDIN

Fonte: adaptado de GARCÍA-CRESPO *et al.* (2010)

Para executar a análise probabilística, é estabelecido um peso para cada variável do diagnóstico. O peso indica a importância da variável no diagnóstico final. Por exemplo, na formação de um diagnóstico os sintomas têm um peso de 30%, os testes de laboratório 50%, sexo 3,5%, idade 3,5%, países visitados 10%, transfusões 1,5%, operações 1,5%. Desta forma tem-se um diagnóstico probabilístico elaborado a partir destes parâmetros clínicos.

Como sugestão para futuras pesquisas, García-Crespo *et al.* (2010) propõem a integração de componentes que investiguem o efeito dos medicamentos no diagnóstico através de seus efeitos colaterais, incluindo-os como parâmetros probabilísticos.

Uma apreciação deste modelo quanto aos pontos fortes:

- Permite análise dos fatores envolvidos, atribuindo pesos aos parâmetros introduzidos;

- Um peso para cada variável do diagnóstico é estabelecido.
- Uma apreciação deste modelo quanto aos pontos fracos:
- Não possui um mecanismo de aperfeiçoamento contínuo do sistema;
 - Não utiliza os registros clínicos coletados para descobrir novos conhecimentos.

Portanto, com a análise dos trabalhos anteriormente apresentados, observa-se que atualmente o desenvolvimento dos SBC's, para o apoio ao processo de diagnóstico, utiliza uma composição de módulos e técnicas para a execução da tarefa intensiva em conhecimento.

Dentre estes módulos e técnicas destacam-se a representação do conhecimento através da ontologia, estruturando a base de conhecimento; as estratégias probabilísticas, que possibilitam meios de manipular as incertezas; e a busca por novos conhecimentos, em base de dados, constituindo os elementos fundamentais para o aperfeiçoamento do processo de diagnóstico.

A busca por conhecimentos novos para o processo de diagnóstico deve passar pela Gestão do Conhecimento (GC), pois é através dela que dados, informações e conhecimentos são potencializados para a criação de novos conhecimentos, como será visto na sequência.

2.2 GESTÃO DO CONHECIMENTO

Existem diversas abordagens sobre o tema “Gestão do Conhecimento”; entretanto, Fialho *et al.* (2006) comentam que ainda não existe um conceito padrão a respeito desse termo, que muda conforme o entendimento dos autores.

Segundo Gaertner Group (1998), a GC é uma disciplina que promove, com visão integrada, o gerenciamento e o compartilhamento de todo o ativo de informação possuído pela empresa. Esta informação pode estar em um banco de dados, documentos, procedimentos, bem como em pessoas, através de suas experiências e habilidades.

Para Keeling e Lambert (2000), a meta principal da GC é utilizar a experiência e a compreensão das pessoas de maneira organizada e com o objetivo de enriquecer a propriedade intelectual.

Smith e Farquhar (2000) apresentam a GC como forma de melhorar o desempenho organizacional porque permite a captura, o compartilhamento e a aplicação do conhecimento coletivo para que se

tomem decisões corretas. A manutenção do conhecimento atualizado requer que este conhecimento seja permanentemente revisto.

Portanto, tem-se na GC um modo de aproveitar as informações disponíveis, obtidas nas mais diferentes formas, para a geração, atualização, armazenamento, compartilhamento e aplicação do conhecimento com o objetivo de enriquecer a propriedade intelectual, melhorando o desempenho organizacional e proporcionando uma tomada de decisão de forma mais segura e correta.

2.2.1 Conversão do Conhecimento

Nonaka e Takeuchi (1997) consideram o conhecimento como um processo humano dinâmico de justificar a crença pessoal com relação à verdade. Também pode ser definido como o fato ou condição de conhecer algo familiarmente adquirido através de experiência ou associação (SWE e KHAM, 2009).

Assim, o processo de criação do conhecimento pode ocorrer por meio da organização de conhecimento anterior em novas formas, da combinação de informações relevantes, ou mesmo de *insights* acerca da aplicação de conhecimento existente em novos contextos (CALHOUM e STARBUCK, 2005). Nonaka, Toyama e Konno (2000) complementam afirmando que o conhecimento só existe dentro de um contexto específico que envolve tempo e espaço, portanto é dinâmico. Sem estes requisitos, não passa de uma simples informação. A informação torna-se conhecimento quando é interpretada por indivíduos em um dado contexto e fundamentada em crenças e compromissos.

Segundo esses autores há dois tipos de conhecimento: conhecimento tácito e o conhecimento explícito. O conhecimento tácito é altamente pessoal e difícil de formalizar, o que dificulta sua transmissão e compartilhamento. Palpites subjetivos, *insights*, intuições, valores e emoções fazem parte desta categoria. Por outro lado, o conhecimento explícito pode ser expresso em palavras e números, sendo facilmente comunicado e compartilhado sob a forma de dados brutos, fórmulas científicas, princípios universais, especificações e manuais. O conhecimento explícito pode ser transmitido formalmente e sistematicamente entre indivíduos (NONAKA, TOYAMA e KONNO, 2000).

De acordo Nonaka e Toyama (2003), o conhecimento é criado em uma espiral de “conversão do conhecimento” que surge a partir das interações entre o conhecimento explícito e o tácito. Esta conversão denominada “Processo SECI” é caracterizada por quatro modos de

conversão: Socialização, Externalização, Combinação e a Internalização.

- Socialização é o processo de converter o conhecimento tácito em novo conhecimento tácito através do compartilhamento de experiências individuais e coletivas. Tipicamente a socialização ocorre no aprendizado tradicional (produção artesanal) onde o aprendiz obtém o conhecimento tácito necessário do mestre para realizar o seu trabalho através de experiências manuais ao invés de utilizar livros ou manuscritos (NONAKA, TOYAMA e KONNO, 2000). Como o conhecimento tácito é difícil de ser formalizado em tempo e lugar específico, este conhecimento pode ser adquirido somente através de troca de experiências, por meio da convivência direta (NONAKA e TOYAMA, 2005).

- Externalização é o processo de articulação do conhecimento tácito em explícito. Quando o conhecimento tácito é explicitado, converte-se num conhecimento cristalizado, permitindo assim o seu compartilhamento para outros e torna-se a base de um novo conhecimento. Conceitos criados para o desenvolvimento de novos produtos é um exemplo deste processo de conversão (NONAKA, TOYAMA e KONNO, 2000).

- Combinação é o processo de sistematização do conhecimento explícito em um SBC, onde os indivíduos trocam e combinam conhecimento de forma sistemática (NONAKA e TOYAMA, 2005). O conhecimento explícito é coletado dentro e fora da organização e então combinado, editado ou processado a fim de formar um novo conhecimento (NONAKA, TOYAMA e KONNO, 2000).

- Internalização consiste na incorporação de conhecimento explícito sob a forma de conhecimento tácito. Este processo ocorre através do “aprender fazendo”, pela verbalização, diagramação e modelos mentais (SENGE, 2004). Quando a maioria dos membros da organização compartilha do novo modelo mental, o conhecimento passa a fazer parte da cultura organizacional. E a cada internalização bem sucedida, o ciclo reinicia, levando ao aperfeiçoamento ou à inovação (NONAKA e TOYAMA, 2005).

A figura 9 sintetiza a espiral de conversão do conhecimento destacando a fase da combinação onde os indivíduos trocam e combinam conhecimento de forma sistêmica.

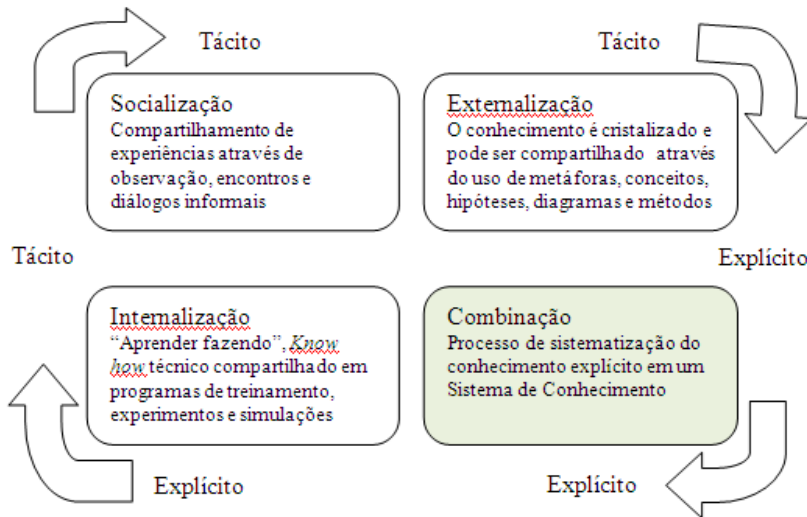


Figura 9 - Modelo SECI de conversões do conhecimento

Fonte: adaptado de NONAKA e TAKEUCHI (1997)

2.3 ENGENHARIA DO CONHECIMENTO

A Engenharia do Conhecimento é uma área que tem por objetivo prover sistemas capazes de efetivar a explicitação e preservação do conhecimento organizacional. Inicialmente tratada como uma subárea da Inteligência Artificial (IA) na construção de SBC para a solução de problemas específicos, a Engenharia do Conhecimento transcendeu esta visão ao considerar todo o contexto sistêmico organizacional das atividades intensivas em conhecimento (SCHREIBER *et al.*, 2002).

Assim, a Engenharia do Conhecimento deve prover um conjunto de ferramentas que forneçam suporte à GC a partir da formalização e explicitação das atividades intensivas em conhecimento nas organizações (SCHREIBER *et al.*, 2002). Para isto, Struder *et al.* (1998) descreve como fundamental a aplicação de metodologias, teorias, métodos, técnicas e ferramentas voltadas à modelagem do conhecimento e sua consequente representação em sistemas de conhecimento.

Neste contexto, a construção de um SBC deve ser realizada através de uma abordagem estruturada, com uma visão organizacional e sistêmica.

2.3.1 Engenharia do Conhecimento e a Metodologia CommonKADS

Uma valiosa fonte de princípios e técnicas de modelagem do conhecimento é a metodologia CommonKADS (SCHREIBER *et al.*, 2002), a qual oferece um conjunto de etapas para o desenvolvimento de Sistemas Baseados em Conhecimento. A metodologia é baseada em quatro princípios que fundamentam a Engenharia do Conhecimento (SCHREIBER *et al.*, 2002):

- A Engenharia do Conhecimento aborda diferentes aspectos dos modelos humanos de conhecimento e não apenas a extração do conhecimento de um especialista;
- A modelagem do conhecimento deve estar focada inicialmente na estrutura conceitual e não nos detalhes de programação dos sistemas de conhecimento;
- O conhecimento tem uma estrutura interna estável, podendo ser categorizado em tipos ou papéis;
- Um projeto de SBC deve ser construído de forma espiral, a partir de uma metodologia que permita este tipo de abordagem.

Além disso, três questões são fundamentais no processo de modelagem de um SBC utilizando-se da metodologia CommonKADS (SCHREIBER *et al.*, 2002): Por quê? O quê? Como? Para responder a estas questões, a metodologia é organizada em três níveis de modelos (SCHREIBER *et al.*, 2002):

- Nível de contexto: tem por objetivo identificar os elementos fundamentais do ambiente e do contexto do Sistema de Conhecimento;
- Nível conceitual: tem por objetivo explicitar e formalizar o conhecimento, bem como destacar de que modo ocorrem as interações entre os agentes envolvidos em sua utilização;
- Nível de projeto: tem por objetivo apresentar todos os detalhamentos necessários para a construção de um Sistema de Conhecimento.

A figura 10 apresenta os modelos de cada um dos níveis do CommonKADS. No Nível de Contexto, os Modelos de Organização, Tarefa e Agente são estruturados. No Nível de Conceito são estruturados os Modelos de Conhecimento e Comunicação. No Nível de Artefato é estruturado o Modelo de Projeto.

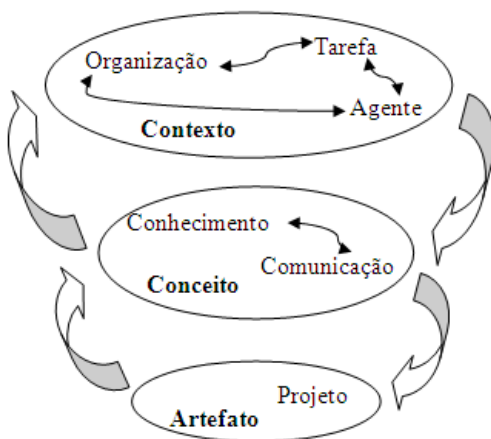


Figura 10- Metodologia CommonKADS

Fonte: adaptado de SCHREIBER *et al.* (2002)

O quadro 2 apresenta uma descrição mais detalhada dos níveis do CommonKADS, para a construção de sistemas baseado em conhecimento, bem como sobre os modelos a serem construídos em cada um dos níveis.

Nível	Modelo	Objetivo
Contexto	Organização	Identificar e analisar o contexto organizacional onde o sistema estará inserido.
	Tarefa	Detalhar todas as tarefas envolvidas no sistema, com foco nas atividades intensivas em conhecimento.
	Agente	Identificar os agentes humanos e não humanos e suas responsabilidades no Sistema de Conhecimento.
Conceito	Conhecimento	Apresentar detalhadamente o modelo conceitual de conhecimento do sistema, destacando as tarefas, as inferências e o domínio para sua posterior implementação.
	Comunicação	Modelar a comunicação entre os agentes envolvidos no funcionamento do sistema.

Artefato	Projeto	Definir as especificações técnicas necessárias para a implementação e funcionamento do sistema, com base nas etapas anteriores.
----------	---------	---

Quadro 2 - Descrição dos modelos do CommonKADS

Como previamente mencionado, CommonKADS é uma metodologia que engloba todos os passos para o projeto e implementação dos sistemas intensivos em conhecimento. No contexto deste trabalho, o modelo de organização (problemas e oportunidades) e o modelo de conhecimento serão abordados como fonte para estruturar o processo de diagnóstico.

O motivo desta escolha é fundamentalmente descobrir os problemas e as oportunidades estabelecendo sua viabilidade e avaliando os possíveis impactos do desenvolvimento do Sistema Baseado em Conhecimento na organização e explicar em detalhes o tipo e estrutura do conhecimento usado na execução da tarefa intensiva em conhecimento relativa ao processo de diagnóstico.

2.3.1.1 Modelo de Organização no CommonKADS

O modelo de organização da metodologia CommonKADS é composto de um conjunto de planilhas com a finalidade de detalhar o contexto do Sistema de Conhecimento na organização. Estas planilhas são apresentadas nos quadros 3 a 7. A primeira planilha é denominada de MO1 – Problemas e Oportunidades (quadro 3). Nesta planilha o problema é tratado com base em uma visão organizacional para além do processo de diagnóstico, uma vez que a tomada de decisão envolve outras etapas.

Modelo de Organização	Planilha OM-1 Problemas e Oportunidades
Problemas e Oportunidades	Lista dos problemas e oportunidades, baseada em entrevistas, <i>brainstorm</i> , encontros visionários, discussão com gerente, entre outros.
Contexto Organizacional	Indicar de maneira concisa as características chave do contexto amplo da organização. Algumas das características consideradas importantes: 1. Missão, visão, metas da organização, 2. Fatores externos importantes de dependência da organização,

	3. Estratégias da organização, e 4. Sua cadeia de valores com os respectivos limites.
Soluções	Possíveis soluções para os problemas e oportunidades, como as sugeridas nas entrevistas e discussões, e características do contexto da organização.

Quadro 3 - MO1: Problemas e Oportunidades

A planilha seguinte é denominada de MO2 – Aspectos Variantes (quadro 4). Nela são descritos elementos organizacionais que devem ser considerados pelo engenheiro do conhecimento. Como o foco do trabalho está no diagnóstico, são considerados apenas os elementos relevantes no contexto desta tarefa.

Modelo de Organização	Planilha OM-2 Diferentes Aspectos
Estrutura	Dar o organograma em termos de departamento, como unidades, seção, grupos.
Processo	Esboço do <i>layout</i> dos processos do negócio.
Pessoas	Indicação dos membros envolvidos como ator ou parte interessada incluindo os gestores, fornecedores, clientes, investidores, usuários. Podem não ser atuais, ou que tiveram papel funcional na organização.
Recursos	Descrição dos recursos utilizados no processo do negócio, tal como: 1. Sistemas de informação/conhecimento, rede, outros <i>software</i> e <i>hardware</i> , 2. Equipamentos e materiais, e 3. Tecnologias, patentes, justiça, privilégio, entre outros.
Conhecimento	O conhecimento representa um recurso chave explorado no processo de negócio no presente contexto, razão do destaque do grupo de recursos, A descrição do componente é feita na planilha OM-4 em ativos do conhecimento.
Cultura e Poder	Tal como listar as regras não escritas de um jogo, estas características não são claras como: o estilo de trabalho e de comunicação, relacionamento social e habilidade de relacionamento interpessoal, relacionamentos formal e informal e de rede.

Quadro 4 - MO2: Aspectos Variantes

A terceira planilha é a MO3 - Detalhamento dos Processos (quadro 5) onde são destacadas as atividades realizadas no contexto da

tarefa analisada. O foco principal é identificar as atividades intensivas em conhecimento e determinar seu significado para o sistema.

Modelo de Organização	Planilha OM-3 Detalhamento do Processo/Tarefas				
Nome da Tarefa	Executada por	Local?	Recurso de Conhecimento	É intensivo de Conhecimento?	Qual é a sua Significância?
Tarefa-1, uma parte de processo descrito em OM-2	Um agente específico: humano, do grupo Pessoas; programa, do grupo Recursos em OM-2.	Algum local específico no grupo Estrutura em OM-2.	Lista dos recursos de conhecimento usados por esta tarefa.	Variável Booleana indicando se a tarefa é considerada intensiva de conhecimento ou não.	Indicar o quanto esta tarefa é significativa por meio de uma escala de cinco posições em termos da frequência, missão, custo, recurso.
Tarefa-2,	...	...	...	...	...

Quadro 5 - MO3: Detalhamento dos Processos

A quarta planilha é a MO4 – Ativos de Conhecimento (quadro 6). Nesta planilha o engenheiro procura identificar quais ativos de conhecimento os agentes têm sobre a tarefa.

Modelo de Organização	Planilha OM-4 Descrição dos Recursos de Conhecimento					
Recursos de Conhecimento	Possuído por	Usado em	Forma é correta?	Local é correto?	Tempo é correto?	Quantidade é correta?
Nome-1 (planilha OM-3)	Agente planilha OM-3	Tarefa planilha OM-3	Sim ou Não, comentário	Sim ou Não, comentário	Sim ou Não, comentário	Sim ou Não, comentário
Nome-2	...	...	...	...	...	...

Quadro 6 - MO4: Ativos de Conhecimento

A quinta e última planilha do modelo de organização é a MO5 – Checklist (quadro 7), em que o engenheiro estabelece os elementos que definem a viabilidade do Sistema de Conhecimento a ser construído.

Viabilidade	Modelo de Organização - Planilha OM-5 Checklist para as possíveis decisões.
Viabilidade do Negócio	Para um dado problema/oportunidade da área e suas respectivas soluções, as seguintes questões devem ser respondidas: 1. Qual é o benefício esperado, dada a solução sugerida? Ambos, a economia tangível e os benefícios intangíveis, são identificados? 2. Qual é a grandeza esperada do acréscimo? 3. Qual é o custo esperado para a solução considerada? 4. Como são comparadas as possíveis soluções? 5. São requeridas mudanças na organização? 6. Qual a extensão da economia e os riscos do negócio e as incertezas envolvidas com respeito à solução considerada?
Viabilidade Técnica	Para um dado problema/oportunidade da área e suas respectivas soluções, as seguintes questões devem ser respondidas: 1. Qual é a complexidade da tarefa, em termos de conhecimento e raciocínio, da solução considerada para o Sistema de Conhecimento? O estado da arte dos métodos e técnicas disponíveis é adequado? 2. Há aspectos críticos envolvidos, relacionados a tempo, qualidade, necessidade de recursos, ou outros fatores? Se sim, como chegar a ele? 3. Ele é claro, qual é a origem da medida e como se testa a validade, qualidade, e o desempenho da suficiência? 4. Qual é a complexidade requerida na interação com o usuário final? O estado da arte dos métodos e técnicas disponíveis é adequado? 5. Qual é a complexidade na interação com outros sistemas de informação e outros possíveis recursos (interoperabilidade, integração de sistemas)? O estado da arte dos métodos e técnicas disponíveis é adequado? 6. Há adicionais tecnologias de riscos e incertezas?
Viabilidade do Projeto	Para um dado problema/oportunidade da área e suas respectivas soluções, as seguintes questões devem ser respondidas:

	1. Existe comprometimento adequado entre os atores e intervenientes (gerentes, especialistas, usuários, investidores, fornecedores, clientes, membros da equipe do projeto) para a próxima etapa do projeto? 2. As necessidades de recurso em termos de tempo, orçamento, equipamento, pessoal podem estar disponíveis? 3. Os conhecimentos requeridos e outras competências são disponíveis? 4. As expectativas ao projeto e os resultados são realistas? 5. O projeto da organização e suas comunicações internas bem como as externas estão adequadas? 6. O avanço do projeto promove riscos e incertezas?
Ações Propostas	Documento de viabilidade das decisões sujeito ao comprometimento gerencial e tomadas de decisão. Ele pondera e integra os resultados da análise prévia dentro da etapa recomendada para a ação: 1. Foco: Qual é o foco recomendado para área identificada do problema/oportunidade? 2. Solução alvo: Qual é a direção da solução recomendada para esta área de foco? 3. Quais são as expectativas do resultado, custo, e benefícios? 4. Que ações do projeto são requeridas para alcançar as metas? 5. Riscos: Se o ambiente interno ou externo da organização mudar, sob que condições é prudente reconsiderar a decisão proposta?

Quadro 7 - MO5: Checklist para a tomada de decisão de viabilidade

2.3.1.2 Modelo de Conhecimento no CommonKADS

O modelo de conhecimento da metodologia CommonKADS é próprio para especificar tarefas intensivas de conhecimento e explica em detalhes o tipo e a estrutura do conhecimento usado na execução de uma tarefa. O modelo é composto por três componentes, em que cada um deles captura uma parte da estrutura do conhecimento:

- Conhecimento de tarefa: define o objetivo da tarefa intensiva em conhecimento (meta), podendo separar em pequenas tarefas (decomposição) e controlá-las (controle). Responde o que será feito e como será feito;

- Conhecimento de inferência: contém as inferências básicas responsáveis pelas tarefas de raciocínio, os papéis de conhecimento que conectam o conhecimento de domínio às inferências e às funções de transferência;
- Conhecimento de domínio: formado pela estrutura de domínio envolvendo os conceitos, as relações e as regras em conjunto com a base de conhecimento.

Estas três categorias são mostradas na figura 11. A coluna da direita mostra um exemplo de elemento de conhecimento no domínio do diagnóstico na área da saúde.

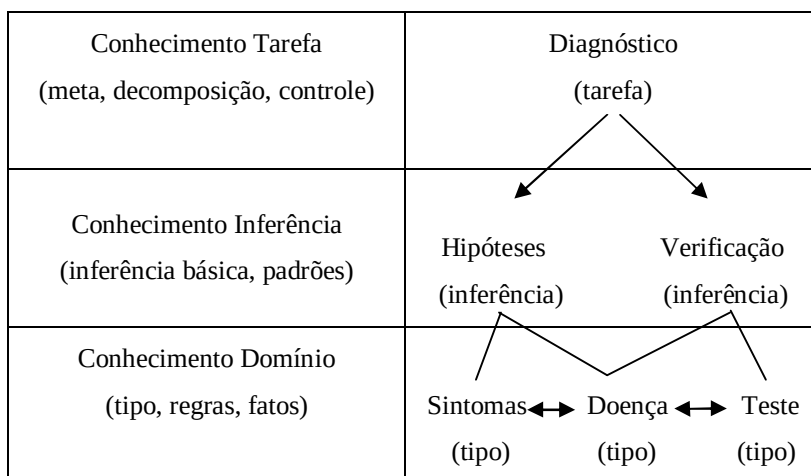


Figura 11- Visão das três categorias de conhecimento

Fonte: adaptado de SCHREIBER *et al.* (2002)

A figura 12 resume de forma esquemática o modelo de conhecimento da metodologia CommonKADS, apresentando os três componentes: as tarefas, a estrutura de inferência e o esquema de domínio com a base de conhecimento.

Em muitas aplicações, a mesma tarefa aparece frequentemente; então, faz sentido que na metodologia CommonKADS tais tarefas sejam reconhecidas e descritas em uma forma de modelo de tarefa, para ser reusada em diferentes aplicações. Um modelo de tarefa é um modelo de conhecimento parcial no qual as inferências e o conhecimento de tarefa são especificados (MUSTAFARAJ, HOOF e FREISLEBEN, 2006).

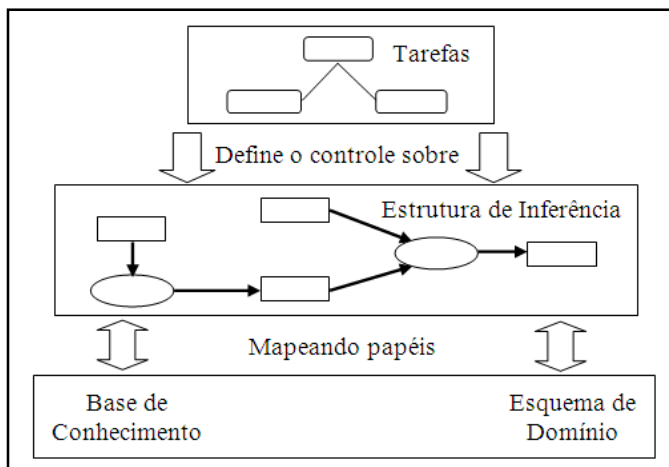


Figura 12- Especificação do modelo de conhecimento no CommonKADS
 Fonte: adaptado de Mustafaraj, Hoof e Freisleben (2006)

Dois grupos de tarefas intensivas em conhecimento, apresentados na figura 13, são distinguidos: tarefas analíticas e tarefas sintéticas. As principais diferenças entre estes dois grupos de tarefas residem na estrutura do sistema que será analisado ou construído.

Para cada uma destas tarefas (*templates*), um projeto é executado de acordo com as necessidades e características do sistema construído (analítica) ou o sistema a ser elaborado (sintética).

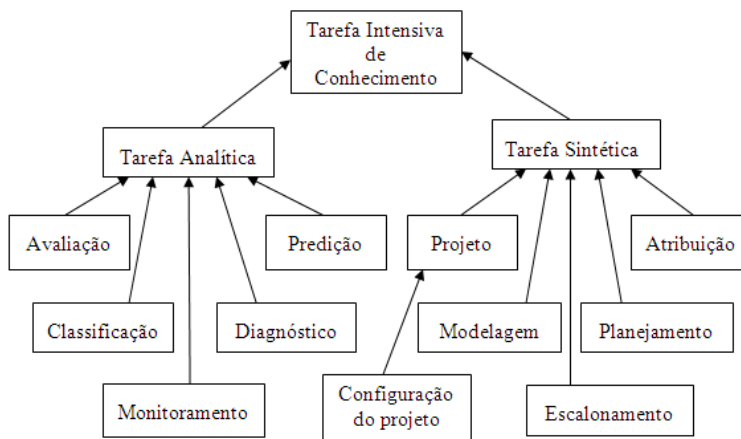


Figura 13- Hierarquia dos tipos de tarefas de conhecimento intensivo
 Fonte: adaptado de SCHREIBER *et al.* (2002)

O quadro 8 mostra as tarefas intensivas em conhecimento do tipo analítico como pode ser visto a seguir.

Tipo de Tarefa	Entrada	Saída	Conhecimento	Características
Analítico	Observações do Sistema	Caracterização do Sistema	Modelo do Sistema	Descrição do Sistema
Classificação	Características do objeto	Classes de objetos	Associação das características com as classes	Conjunto de classes pré-definidas
Diagnóstico	Sintomas e denúncias	Categoria de falhas	Modelo de comportamento do sistema	Variedade das saídas (cadeia causal, estado, componente) e depende do seu modo de uso (diagnose)
Avaliação	Descrição do caso	Classe de decisão	Critérios, normas	A avaliação é feita em um instante do tempo (Monitoramento)
Monitoramento	Dados do sistema	Classe de discrepância	Comportamento normal do sistema	Frequentes mudanças no sistema. Tarefas são executadas repetidamente
Predição	Dados do sistema	Estado do sistema	Modelo de comportamento do sistema	O estado da saída é uma descrição do sistema em um ponto futuro

Quadro 8 - Visão das tarefas intensivas de conhecimento do tipo analítico

O quadro 9 exhibe as tarefas intensivas em conhecimento do tipo sintético.

Tipo de Tarefa	Entrada	Saída	Conhecimento	Características
Sintético	Requisitos do sistema	Estrutura do Sistema	Elementos, restrições, preferências	Descrição do sistema necessita ser gerado
Projeto	Requisitos do sistema	Classe de Objetos	Componentes, restrições, preferências. Associação de recursos e classes	Pode ser incluído projeto criativo de componentes
Configuração do projeto	Requisitos do sistema	Descrição dos artefatos	Componentes, projeto estrutural, restrições, preferências	Subtipo do projeto onde todos os componentes são pré-definidos
Atribuição	Conjunto de dois objetos, requisitos.	Mapeamento-1 → Mapeamento-2 →	Restrições, preferências	O mapeamento não necessita ser um a um

Planejamento	Metas e requisitos	Plano de ação	Ações, restrição, preferências	Ações são (parcialmente) ordenadas no tempo
Escalonamento	Atividades de trabalho, recursos, espaço de tempo, requisitos	Esquema = atividades alocadas para o espaço de tempo do recurso	Restrições, preferências	Caractere orientado no tempo é distinto da atribuição
Modelagem	Requisitos	Modelo	Elementos do modelo, gabarito do modelo, restrições, preferências	Podem ser incluídas sínteses criativas

Quadro 9 - Visão das tarefas intensivas de conhecimento do tipo sintético

Na sequência, a especificação do conhecimento para tarefa de diagnóstico será apresentada e detalhada nos modelos que a compõem.

2.3.1.2.1 Especificação do Conhecimento para a Tarefa de Diagnóstico

O objetivo desta etapa é obter a especificação completa do modelo de conhecimento, envolvendo os modelos de tarefa, inferência e domínio.

- Modelo de Tarefa

A primeira atividade é a seleção do modelo de tarefa intensiva em conhecimento a ser utilizada. A partir dos *templates* apresentados no quadro 8, foi selecionada a tarefa de diagnóstico para análise por ser o foco de pesquisa deste trabalho. Segundo Schreiber *et al.* (2000), trata-se de uma tarefa analítica que difere da tarefa de classificação, pois o que se deseja é encontrar a resposta ao mau funcionamento do sistema.

O quadro 10 exibe um resumo da caracterização geral para o *template* da tarefa de diagnóstico.

Meta	Achar a falha que causa um problema no sistema
Exemplo típico	Diagnóstico de um equipamento técnico, tal como uma copiadora
Terminologia	Reclamação/Sintoma: o dado que inicia o processo de diagnóstico

	Hipótese: uma solução potencial (de acordo com a falha) Diferencial: jogo ativo das hipóteses Achados/Evidências: dados adicionais sobre o sistema sendo diagnosticado Falta: solução encontrada pelo processo de raciocínio do diagnóstico
Entrada	Sintomas e ou reclamações
Saída	Falhas adicionadas às evidências
Características	Em princípio, a tarefa de diagnóstico deve sempre ter o mesmo modelo do comportamento do sistema que está sendo diagnosticado.

Quadro 10 - Caracterização geral da tarefa de diagnóstico

A próxima etapa é a construção completa do modelo de conhecimento, apresentada a seguir. A abordagem utilizada foi a *middle-out*, que parte do modelo de inferência, para, em seguida, propor os modelos de tarefa e domínio. Esta é a abordagem mais indicada quando o modelo de tarefa é bem conhecido (SCHREIBER *et al.*, 2002).

- Modelo de Inferência

O modelo de inferência descreve como a estrutura de informação ou conhecimento pode ser utilizada no processo de raciocínio. O conhecimento de inferência é descrito através das primitivas de inferências, papéis do conhecimento e funções de transferência (SCHREIBER *et al.*, 2002), descritas a seguir:

Primitivas de Inferências - Descrevem o menor nível de decomposição funcional. São unidades básicas de processamento de informações. Executa um passo primitivo de raciocínio, ou seja: cobrir, especificar, comparar.

Papéis do Conhecimento – Inferências são descritas em termos de papéis funcionais: nomes abstratos de objetos que indicam sua função no processo de raciocínio. Tais papéis chamam-se papéis de conhecimento (*knowledge roles*). Distinguem-se dois tipos de papéis: papéis dinâmicos e papéis estáticos, como: reclamação, hipóteses, evidências, entre outros.

Funções de transferência – Comunicam-se com o mundo externo. É a função que transfere uma informação entre o agente de raciocínio

descrito no modelo de conhecimento e o mundo externo, tais como: obter, receber, apresentar, prover.

No modelo de tarefa proposto são utilizadas as seguintes inferências: cobrir, especificar, obter e comparar.

- Cobrir: esta inferência (figura 14) busca, através de um modelo de causas e efeitos, uma possível hipótese para um determinado problema ou reclamação. Assim, a partir de um estado visível – que é observado, será sugerida uma hipótese para um estado não visível – que não pode ser observado, baseada nos modelos causais.

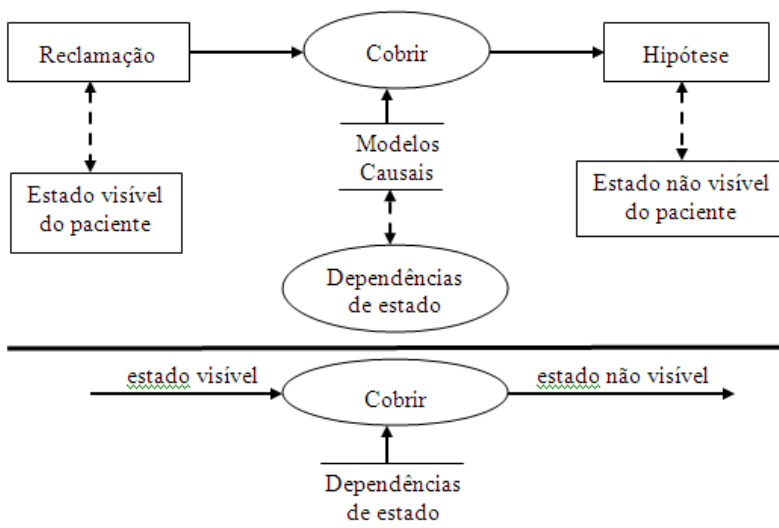


Figura 14- Inferência Cobrir

Fonte: adaptado de SCHREIBER *et al.* (2002)

- Comparar: esta inferência compara o resultado esperado com o desejado. Assim, a partir do resultado da inferência especificar, o profissional deve observar o estado sugerido, através da função de transferência obter, de forma a confirmar ou não a hipótese sendo analisada.
- Obter: trata-se de uma função de transferência, através da qual o profissional observa o estado indicado na inferência especificar;

- Especificar: esta inferência (figura 15) especifica um possível estado observável, a partir de uma hipótese. Assim, a partir de um estado não visível, será sugerido um estado esperado para o estado não visível, considerando os modelos de manifestação para este estado não visível;

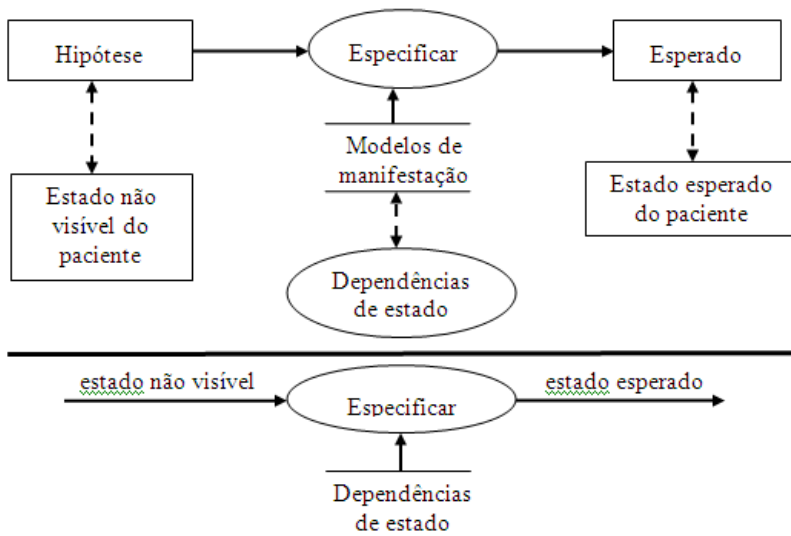


Figura 15- Inferência Especificar

Fonte: adaptado de SCHREIBER *et al.* (2002)

A figura 16 apresenta na forma de diagrama de atividades, a estrutura de controle que representa o modelo de tarefa proposto. Já a figura 17 apresenta o modelo de tarefa completo e anotado para o diagnóstico a partir de geração e teste. Através da figura 17 é possível ter uma compreensão global da execução da tarefa, considerando as inferências propostas.

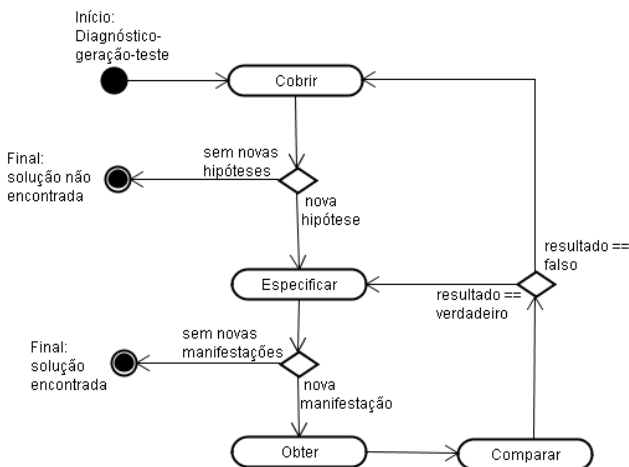


Figura 16- Diagrama de Atividades do Modelo de Tarefa de Diagnóstico
Fonte: adaptado de SCHREIBER *et al.* (2002)

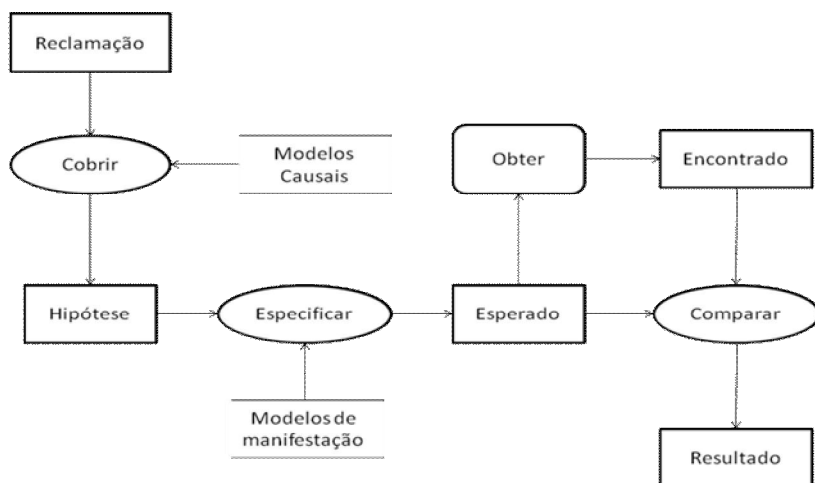


Figura 17- Modelo de Tarefa completo
Fonte: adaptado de Schreiber *et al.* (2002)

- Modelo de Domínio

A próxima atividade é a construção de um modelo inicial de conhecimento de domínio que será apresentada nos tópicos a seguir

abordando a explicitação e a representação do conhecimento de um domínio específico.

2.3.2 Engenharia do Conhecimento e a Modelagem CESM

Kendall e Creen (2007) destacam que a Engenharia do Conhecimento deve estar envolvida nas atividades de aquisição, validação, representação, inferência, explicação e justificativa dos elementos de conhecimento presentes em uma determinada tarefa intensiva em conhecimento.

Com base nisto, pressupõe-se que a Engenharia do Conhecimento possa fornecer um conjunto de recursos focados em sistemas de conhecimento de forma a explicitar o conhecimento existente na tarefa de diagnóstico.

2.3.3 Explicitação do Conhecimento

A conceituação de sistemas apresentada por Bunge (2003), por meio de seu modelo CESM, contribui, dentro de uma visão sistêmica, para a explicitação do conhecimento através de uma modelagem mais descritiva e abrangente para os elementos que compõem um domínio específico.

Segundo Bunge (1997), para entender sobre qualquer sistema, seja real ou artificial, é necessário compreender como ele trabalha, isto é, desvendar o seu mecanismo de ação, por exemplo, o movimento molecular explicando a fermentação da uva para obtenção do vinho.

Bunge (2000), em seu sistema filosófico, propõe os seguintes postulados:

- Toda coisa, seja concreta ou abstrata, é um sistema ou um componente ou potencial componente de sistema;
- Os sistemas têm características sistêmicas (emergentes) que seus componentes não têm;
- Todos os problemas deveriam ser abordados de forma sistêmica e não de forma fragmentada;
- Todas as ideias deveriam ser unidas em sistemas (preferencialmente teorias);
- O teste de qualquer coisa, seja ideia, método ou artefato, supõe a validade de outros itens que são tomados como pontos de referência (*benchmark*) provisoriamente.

Para Bunge (2003), o modelo CESM proporciona uma forma de modelar qualquer sistema mediante a quádrupla: *Composition*; *Environment*; *Structure*; *Mechanism*, onde:

- *Composition* (composição): conjunto de todas as partes do sistema;
- *Environment* (ambiente): coleção de itens que não pertencem ao sistema, porém atuam ou sofrem a ação por algum ou por todos os componentes;
- *Structure* (estrutura): coleção de ligações entre componentes e entre esses e itens do ambiente;
- *Mechanism* (mecanismo): coleção de processos que geram mudança qualitativa no sistema.

Kilov e Sack (2008), quanto às ideias de Bunge, destacam que os conceitos de relação e composição, envolvendo os componentes, são a essência no funcionamento de um determinado sistema.

A abordagem sistêmica Bungeana “não é uma teoria para substituir outras teorias” (BUNGE, 2004, p. 91), mas uma estratégia ou visão de mundo para guiar o empreendimento de pesquisa sobre sistemas.

Assim, de acordo com o sistemismo proposto por Bunge (2003), qualquer sistema, para ser compreendido, deve ser visto de uma forma global, isto é, não só o relacionamento dos componentes entre si, mas também, destes com os itens do ambiente onde estão inseridos.

2.3.4 Representação do Conhecimento (RC)

Um dos grandes desafios da Inteligência Artificial (IA) tem sido a representação do conhecimento humano em sistemas computacionais. A representação, através de formalismos simbólicos, pretende facilitar a manipulação do conhecimento envolvido num domínio, de forma que possa ser utilizado para os devidos fins.

Vickery (1986), em uma breve revisão sobre a representação do conhecimento, comenta:

A representação do conhecimento na forma simbólica é uma questão que preocupou o mundo da documentação desde sua origem. O problema é relevante agora em muitas outras situações além dos documentos e índices. A estrutura de registros e arquivos em bases de dados; a estrutura de dados nos programas de computador; a estrutura sintática e semântica da linguagem natural; a

representação do conhecimento em inteligência artificial; os modelos de memória humana: em todos estes campos é necessário decidir como o conhecimento pode ser representado de forma que as representações possam ser manipuladas.

Para Carter (1999), representar conhecimento é proporcionar aos sistemas inteligentes informações a respeito de um domínio específico na forma que possa ser processado eficazmente. Já para Sowa (2000), a Representação do Conhecimento (RC) é a aplicação da lógica na tarefa de construir modelos computacionais sobre algum domínio.

Entre os muitos métodos propostos na literatura para representar o conhecimento, possivelmente os mais frequentemente usados são a lógica, regras, redes semânticas e *frames* (LUGER, 2009). A figura 18 exibe os principais métodos para representação de conhecimento que serão descritos a seguir.

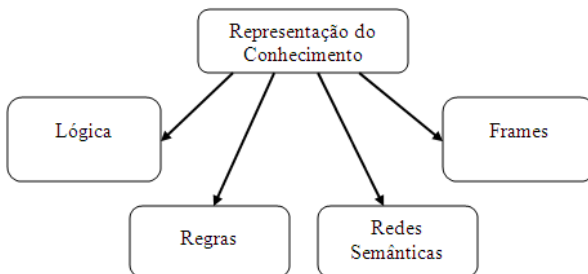


Figura 18- Principais métodos de Representação do Conhecimento

Fonte: adaptado de CASTILHO e LOPES (2010)

- **Lógica:** Tendo origem na Matemática e na Filosofia com o objetivo de caracterizar os princípios de raciocínio correto, ela propõe, através da linguagem de representação formal, deduções consistentes e completas (LUGER, 2009);
- **Regras:** São sequências lógicas compostas por premissa (antecedentes) e conclusões (consequências), representando o domínio do conhecimento através de proposições na forma “SE... ENTÃO...”. De certo modo, as regras iniciaram os modelos de representação do conhecimento, sendo muito usadas na construção de sistemas especialistas ao possibilitar um bom nível de representação; porém, são pouco flexíveis e consomem muito tempo para serem elaboradas;

- Rede Semântica: As redes semânticas são estruturas compostas por um conjunto de sistemas heterogêneos (nós), conectados por arcos indicando alguma relação (é-um e parte-de) entre eles. Apresenta, entre outras características, a facilidade de entendimento na representação visual e a capacidade de manipular o “senso comum”. Como limitação tem-se a dificuldade de representar conhecimento procedimental e a baixa expressividade;
- *Frames*: Segundo Minsky (1977) “Um *frame* é uma estrutura de dados complexa, útil para modelar objetos do mundo real. Assemelham-se a registros de um banco de dados, porém são mais poderosos e expressivos.” Esta estrutura, uma evolução das redes semânticas, é formada por um mecanismo de herança, composta de “gavetas” (*slots* ou escaninhos) no qual é possível adicionar valor (instâncias). É uma maneira bastante organizada e hierarquizada de se representar o conhecimento (NIEVOLA, 1995). Embora tenha evoluído na representação do conhecimento, a falta de uma semântica formal e uma terminologia padrão dificulta o processo de inferência com base no conhecimento armazenado.

Assim, uma base de conhecimento pode ser descrita como um mapeamento entre os objetos e as relações de um domínio de problema e os objetos computacionais e as relações de um programa. Os resultados de inferências na base de conhecimento devem corresponder aos resultados de ações ou observações no mundo. Os objetos, as relações e as inferências computacionais disponíveis para os programadores são mediados pela linguagem de representação do conhecimento (CASTILHO e LOPES, 2010).

2.3.4.1 Ontologia

O termo ontologia, proveniente da filosofia, também está inserido na engenharia e gestão do conhecimento como meio de representação formal do conhecimento, com o propósito de facilitar o compartilhamento deste entre pessoas e sistemas.

Para Neches *et al.* (1991), “Uma ontologia define os termos básicos e relações compreendendo o vocabulário de uma área específica, bem como as regras para combinação entre termos e relações para definir extensões do vocabulário”.

Gruber (1993), afirma que “uma ontologia é uma especificação formal e explícita para um conceito compartilhado”. Da definição de Gruber (1993), Fensel (2002) comenta os termos formal, explícita e compartilhada:

- Formal, pelo fato de que uma ontologia deve ser compreendida e processada por máquinas;
- Explícita significa que os conceitos utilizados, bem como as restrições sobre seu uso, são objetivamente definidos e claros;
- Compartilhado no sentido de que uma ontologia reflete o conhecimento consensual sobre um determinado assunto por uma comunidade de especialistas no domínio.

Outra definição para ontologia, complementar às demais, pois relata informações a respeito de sua estrutura, é a apresentada por Gómez-Pérez (1999): um conjunto de termos ordenados hierarquicamente para descrever um domínio que pode ser usado como um esqueleto para uma base de conhecimento.

A definição de Gómez-Pérez (1999) faz uma distinção importante entre ontologia e base de conhecimento. Uma ontologia forma uma estrutura sobre a qual é possível construir uma base de conhecimento. A ontologia fornece um conjunto de conceitos e termos para descrever um determinado domínio, enquanto a base de conhecimento usa estes termos para descrever uma determinada realidade. Caso esta realidade seja modificada, a base de conhecimento será modificada; porém a ontologia permanecerá inalterada, desde que o domínio seja o mesmo.

O uso de uma ontologia permite, então, a definição de um domínio no qual será possível trabalhar em determinada área específica, possibilitando a melhora no processo de extração de informação e o intercâmbio do conhecimento.

Um dos aspectos fundamentais na construção de uma ontologia é a familiarização com o domínio que tem o objetivo de permitir o completo entendimento sobre o problema, através da consulta às diferentes fontes de conhecimento disponíveis.

O conhecimento é formalizado na ontologia, segundo Gruber (1993), por cinco tipos de componentes, ou seja: conceitos (também chamados de classes), relações, funções (também chamadas de propriedades), axiomas e instâncias.

Gómez-Pérez (1999) comenta cada um destes componentes que definem a estrutura da ontologia:

- Classes são usadas em um sentido amplo. Um conjunto de classes e uma hierarquia entre estas classes formam uma taxonomia. Por exemplo, a classe “mãe” é uma subclasse da classe “mulher”;
- Relações representam um tipo de interação entre as classes de um domínio. Um exemplo de relacionamento entre as classes “pessoa” e “casa” é o relacionamento “eh_proprietario”;
- Função é um caso especial de relacionamento em que um conjunto de elementos tem uma única relação com outro elemento. Um exemplo de função é “ser_pai”, onde a classe “homem” e a classe “mulher” estão relacionados a uma classe “pessoa”;
- Axiomas são regras que são sempre válidas (são regras que possibilitam inferências sobre as classes). Um exemplo de axioma é afirmar que toda pessoa tem uma mãe;
- Instâncias são especializações das classes (representam indivíduos específicos de uma determinada classe).

A figura 19 apresenta os principais componentes de uma ontologia.

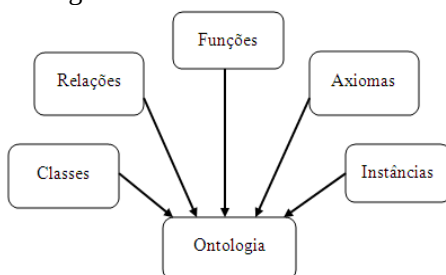


Figura 19- Principais componentes de uma ontologia

Fonte: adaptado de CASTILHO e LOPES (2010)

2.4 DESCOBERTA DO CONHECIMENTO

Fayyad *et al.* (1996) definem o processo de descoberta de conhecimento como um conjunto de várias atividades para dar sentido a um banco de dados.

Em um mundo com aumentos aparentemente ilimitados em conhecimento científico, agrupados em banco de dados, pesquisadores

lutam para manter a perícia e o desenvolvimento do conhecimento em suas áreas (GANIZ *et al.*, 2005).

2.4.1 Knowledge Discovery in Databases (KDD)

Knowledge Discovery in Databases é um processo de Aquisição de Conhecimento (AC) a partir de banco de dados e que consiste num conjunto de atividades que tem por objetivo permitir a extração de conhecimento útil e com significado. Trata-se de um campo que integra diversos conceitos na área de computação, ao utilizar banco de dados, inteligência artificial e algoritmos. Por exemplo, além de sua característica multidisciplinar, via de regra, ela é aplicada para outras áreas de conhecimento, sendo muito útil em processos de gestão do conhecimento. Uma das principais referências sobre o KDD é o trabalho de Fayyad *et al.* (1996).

Segundo Freitas (2000) o conhecimento a ser descoberto em um processo de KDD deve satisfazer principalmente a três propriedades: deve ser correto (tanto quanto possível); deve ser compreensível por usuários humanos; e deve ser interessante, útil ou novo. Além disso, o método de descoberta do conhecimento deve apresentar as seguintes características: deve ser eficiente (acurado), genérico (aplicável a vários tipos de dados) e flexível (facilmente modificável).

O KDD apresenta um ciclo de desenvolvimento baseado em diversas etapas, segundo Silva (2004):

- Pré-processamento: nesta etapa procura-se fazer a preparação dos dados para o processo de extração de conhecimento, uma vez que eles podem apresentar distorções por ausência de informação, falta de padronização ou ruídos, já que os mesmos podem ter origem em diferentes fontes;
- Transformação: também conhecida como etapa exploratória, é onde se faz o processo de integração das diferentes fontes de dados, eliminando redundâncias, dados conflitantes e procurando agregar mais conteúdo semântico aos mesmos. Assim, a transformação deve usar técnicas importantes tais como agregação, generalização, normalização, construção de novos atributos, entre outras;
- Extração de conhecimento: também conhecida como etapa de mineração dos dados, nela deve ser extraído o conhecimento a partir das fontes de dados criadas na etapa anterior. Para a tarefa de mineração deve ser selecionado o algoritmo mais

adequado de acordo com as abordagens/objetivos que se deseja alcançar;

- Pós-processamento: consiste na avaliação do processo de extração de conhecimento, procurando analisar e validar os resultados obtidos. Na avaliação das descobertas são aplicadas métricas para validar os resultados alcançados.

Segundo Li e Ruan (2007), a etapa mais estudada pelos pesquisadores é a de extração de conhecimento, mas eles alertam para o fato de que a chave do sucesso de um processo de KDD pode estar na etapa exploratória, sobretudo na busca de padrões em grandes volumes de dados. De fato, a grande maioria dos trabalhos disponíveis concentra-se nas técnicas de mineração dos dados.

O processo de mineração pode ser aplicado de diversas formas de acordo com o objetivo que se pretende alcançar. No processo de verificação, por exemplo, o usuário sugere uma hipótese e realiza consultas nos dados para validá-la ou negá-la. Já no processo de descoberta cabe ao sistema procurar relações existentes entre as informações disponíveis nas fontes. Este processo ainda pode ser subdividido em predição, onde se procura inferir valores futuros a partir de informações conhecidas, ou descrição, onde se procuram padrões interpretáveis pelos humanos.

As tarefas mais comuns realizadas em processos de KDD são (TURBAN e ARONSON, 2001):

- Classificação: consiste em examinar as propriedades (aspectos, estrutura) de um objeto (dados) e atribuí-lo a uma das classes predefinidas em um determinado contexto. Sua solução pode ser alcançada através da abordagem simbólica (algoritmos baseados em árvores de decisão ou regras de produção), biológica (baseados em redes neurais ou algoritmos genéticos) ou estatística (baseados em classificadores Bayesianos ou de vizinhos mais próximos);
- Agrupamento: consiste em classificação, mas as classes não são conhecidas ou predeterminadas. Neste caso, existem alguns métodos para auxílio a estas tarefas como os de particionamento, hierárquicos, baseados em densidade, baseados em grades e baseados em modelos;
- Associação: consiste em descobrir associações entre informações, de modo a verificar se a presença de um item em uma determinada transação implica a presença de outro item na mesma transação. Nesta tarefa dois parâmetros são

fundamentais: o suporte que caracteriza a frequência e a confiança que caracteriza o percentual de ocorrências que satisfaz a regra.

2.4.2 Descoberta Baseada na Literatura (DBL)

Combinando a informação recuperada com seus próprios experimentos e observações, o cientista cria novos conhecimentos (WEEBER *et al.*, 2001).

A descoberta baseada na literatura é um método para encontrar relações entre conceitos, não abertamente declaradas, embutidas na literatura publicada (AHLERS *et al.*, 2007). Geralmente, tem a forma de relação entre dois conceitos primários; por exemplo, um medicamento que trata uma determinada doença ou um gene como a causa de uma enfermidade (HRISTOVSKI *et al.*, 2006).

Swanson (1986) propôs um paradigma no qual uma relação entre dois conceitos A e C, não diretamente declarados em uma pesquisa literária, pode ser descoberta via um terceiro conceito B.

Por exemplo, depois de notar uma associação entre óleo de peixe (A) e viscosidade sanguínea (B) e outra associação entre viscosidade sanguínea (B) e a doença de Raynaud (C), Swanson propôs que óleo de peixe (A) trata a doença de Raynaud (C).

A figura 20 esquematiza o modelo ABC de Swanson onde as relações AB e BC são conhecidas e registradas na literatura. A relação implícita AC é uma possível nova descoberta.

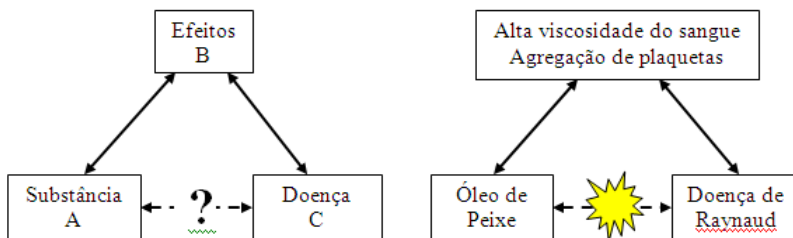


Figura 20- Modelo de descoberta ABC de Swanson

Fonte: adaptado de WEEBER *et al.* (2001)

Segundo Swanson, este modelo pode ser praticado de duas formas:

- Processo aberto de descoberta;
- Processo fechado de descoberta.

O processo aberto de descoberta é caracterizado pela geração de um nova hipótese onde, inicialmente, um conceito de interesse é pesquisado. Seja C uma doença a ser investigada, tentam-se encontrar na literatura potenciais conexões B sobre a doença C, tipicamente processos fisiológicos que possuam um papel relevante nesta enfermidade.

Em seguida, o sistema procura determinar elementos A, normalmente substâncias, que atuam sobre B. No decorrer do processo de descoberta é normal que seja encontrado um grande número de B e A. De fato, o maior desafio para as ferramentas de apoio à descoberta prende-se com a seleção do grande número de possíveis relações.

Como resultado deste processo, o pesquisador pode formar uma hipótese sobre a ação da substância A para o tratamento da doença C, através da relação dada por B. A figura 21 apresenta o processo aberto de descoberta onde as setas em negrito indicam as potenciais vias de interesse no processo.

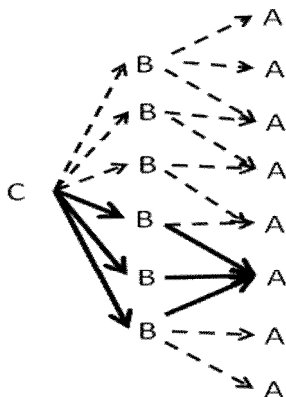


Figura 21- Processo aberto de descoberta

Fonte: adaptado de WEEBER *et al.* (2001)

Gordon & Lindsay (1996) sintetizam, passo-a-passo, o processo aberto de descoberta de Swanson:

1. Escolha um tópico de interesse $\Rightarrow$ Doença de Raynaud;
2. Procure na literatura $C = \{\text{Raynaud}\}$;
3. Prediga qual termo B (por exemplo, o fator sanguíneo pode ser estudado em relação a Raynaud);
4. Procure na literatura $C' = \{\text{Raynaud}\} \cap \{\text{sangue}\}$;

5. Note dois descritores comuns: viscosidade sanguínea e agregação de plaquetas;
6. Procure na literatura $A = \{\text{viscosidade sanguínea}\} \cup \{\text{agregação de plaquetas}\}$;
7. Note o termo “óleo de peixe”;
8. Procure na literatura $A = \{\text{óleo de peixe}\}$;
9. Mostre que $\{\text{óleo de peixe}\} \cap \{\text{Raynaud}\} = \{\}$;
10. Mostre a conexão plausível entre óleo de peixe e Raynaud.

O processo fechado de descoberta é usado para teste e elaboração de hipóteses. Nesta aproximação o pesquisador já formulou uma hipótese sobre um tratamento, como a interação entre uma substância e uma doença e utiliza o sistema com o objetivo de testar ou modificar a sua hipótese.

Desta forma inicia-se, simultaneamente, uma pesquisa com uma doença C e uma substância A. O sistema propõe então várias possibilidades de conexão B para seleção e para uma análise mais detalhada por parte do pesquisador. Quanto maior o número de ligações entre A e C, maior a probabilidade de existir uma relação entre os dois.

A figura 22 apresenta o processo fechado de descoberta no qual as setas contínuas indicam potenciais vias de descoberta, enquanto as tracejadas não sugerem possíveis relações.

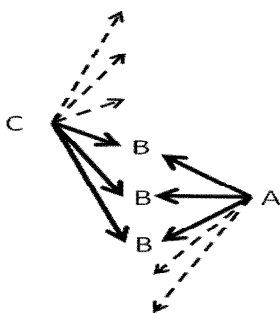


Figura 22- Processo fechado de descoberta
Fonte: WEEBER *et al.* (2001)

Como exemplo deste processo, Swanson estudou a conexão entre enxaqueca (A) e magnésio (C), examinando os títulos de resultado de pesquisas em literaturas disjuntas. Ele encontrou vários termos médicos (B) que incluem epilepsia e obstrução de canal de cálcio que ocorrem frequentemente em ambos os conjuntos disjuntos de títulos literários. A

partir destes termos de conexão, ele pôde concluir que a insuficiência de magnésio está envolvida na enxaqueca, gerando esta hipótese pela combinação de duas literaturas disjuntas via uma conexão particular.

Os sistemas de descoberta baseada na literatura, como o de Swanson (1986), normalmente usam conceitos simples de co-ocorrência de palavras ou frases como seus mecanismos primários. Nenhuma informação semântica a respeito da natureza da relação entre conceitos é provida (AHLERS *et al.*, 2007).

O uso da co-ocorrência tem várias desvantagens, entre elas destacam-se:

- Os pesquisadores devem ler um grande número de citações no Medline quando revisam as relações candidatas;
- Os sistemas tendem a produzir várias relações espúrias (ilegítimas, falsas);
- Não há nenhuma explanação explícita da relação descoberta.

Procurando minimizar estas deficiências, as relações semânticas podem contribuir no sentido de tornar o processo de co-ocorrência mais eficiente.

2.4.2.1 Predicações Semânticas

Hristovski *et al.* (2006), ampliando o paradigma de Swanson (1986), definiram analogamente os conceitos X, Y e Z, e eficientizaram as co-ocorrências através de predicações semânticas que fornecem uma informação específica sobre a natureza da associação.

Com a informação mais específica provida pela predicação semântica, o processo de descoberta torna-se mais compreensível, diminuindo o número de relações a serem avaliadas por humanos e possibilita uma melhor capacidade de esclarecimento (HRISTOVSKI *et al.*, 2006).

Enquanto as descobertas de Swanson estão baseadas na frequência de palavras ou frases em títulos de artigos como forma de medir a força entre conceitos, a predicação semântica utiliza regras de associação como medida de eficácia na relação entre os termos (HRISTOVSKI *et al.*, 2007).

Uma regra de associação constitui-se em um relacionamento $X \Rightarrow Y$ (se X então Y), onde X e Y são conjuntos de itens e a intersecção entre os mesmos estabelece um conjunto vazio. Cada regra de associação é associada a um Fator de Suporte (FSup), utilizado para

indicar a relevância da regra, e a um Fator de Confiança (FConf), utilizado para indicar a validade da regra.

O FSup é dado pela razão do número de tuplas que satisfazem X e Y, sobre o número total de tuplas ($F\text{Sup} = |X \cup Y| / N$); e o FConf é definido pela razão do número de tuplas que satisfazem X e Y sobre o número de tuplas que satisfazem X ($F\text{Conf} = |X \cup Y| / |X|$).

A tabela 1 exemplifica o cálculo destes fatores para a regra de associação “se X então Y” apoiado em uma base de dados contendo 10 identificadores.

Tabela 1 - Base de dados para a descoberta de novas relações

Identificador (Base de dados)	Conceito X	Conceito Y
1	Sim	Não
2	Sim	Não
3	Não	Sim
4	Sim	Sim
5	Sim	Não
6	Sim	Sim
7	Sim	Sim
8	Não	Sim
9	Sim	Não
10	Sim	Sim

$$F\text{Sup}_{XY} = \frac{X \cup Y}{N} = \frac{4}{10} = 0,4 = 40\% \quad F\text{Conf}_{XY} = \frac{X \cup Y}{X} = \frac{4}{8} = 0,5 = 50\%$$

Portanto, a regra de associação “se X então Y” nesta base de dados tem uma relevância de 40% com um grau de confiança de 50%.

2.5 TRATAMENTO DE INCERTEZAS

Segundo Yang (2007), há muitos sistemas que são projetados e desenvolvidos baseados em precisão e certeza. Eles provêm soluções irreais baseadas na suposição de ambientes fechados. Para a maioria das aplicações reais, a incerteza é inevitável e não pode ser ignorada.

2.5.1 Categorias de Incerteza

Smets (1997) destaca que a imperfeição da informação é a maior dificuldade enfrentada em um ambiente aberto. Para ele, geralmente esta imperfeição pode ser agrupada em imprecisão, inconsistência ou incerteza.

- Imprecisão apresenta ambiguidade, indeterminação ou aproximação da informação;
- Inconsistência expressa conclusões contraditórias que podem ser obtidas baseadas nas informações ou declarações;
- Incerteza é causada pela falta de conhecimento sobre o ambiente onde é necessário decidir a verdade de declarações.

Imprecisão e inconsistência são propriedades essencialmente relacionadas ao conteúdo da informação, enquanto a incerteza é uma propriedade da relação entre a informação e o nosso conhecimento sobre o mundo (SMETS, 1997).

Solo e Grupta (2000) propõem duas categorias de incertezas:

- Incertezas devido às informações que surgem do comportamento aleatório dos sistemas físicos;
- Incertezas devido às informações que surgem da percepção humana e processos cognitivos.

Para as incertezas geradas nos sistemas físicos, várias teorias e métodos estatísticos têm sido investigados e usados, porém, para as incertezas da percepção humana, pouca sofisticação nos processos tem sido feito (SOLO e GRUPTA, 2000).

Dentre estes processos, destacam-se, neste trabalho: lógica *fuzzy* e redes Bayesianas.

2.5.2 Lógica Fuzzy

A lógica *fuzzy* foi apresentada em meados dos anos 60 por Zadeh (1965), reconhecendo que a natureza verdadeira e falsa da lógica booleana (0 ou 1) não se aplica em muitos problemas do mundo real. Diferentemente da lógica booleana, a lógica *fuzzy* pretende modelar raciocínios que são aproximados e não exatos.

A teoria da lógica *fuzzy* está sempre relacionada com a modelagem de raciocínios que são aproximados e não exatos, exibindo imprecisões graduais que possam ser modelados matematicamente (ZADEH, 1965).

Para considerar os intervalos infinitos entre o verdadeiro e o falso, Zadeh introduziu a ideia de um conjunto que ele chamou de conjunto *fuzzy*. Um conjunto *fuzzy* consiste de um universo de discurso e funções de pertinência que mapeiam todos os elementos do universo de discurso para um valor de pertinência entre 0 e 1 (ZADEH, 1965).

Para exemplificar a diferença entre um conjunto clássico e um conjunto *fuzzy* é apresentada a figura 23. No primeiro gráfico (24,a), tem-se claramente que as pessoas consideradas altas são as que têm altura igual ou superior a 1,75 metros. No segundo gráfico (24,b), pessoas com altura inferior a 1,75 m também podem pertencer ao conjunto de pessoas altas com algum grau de pertinência.

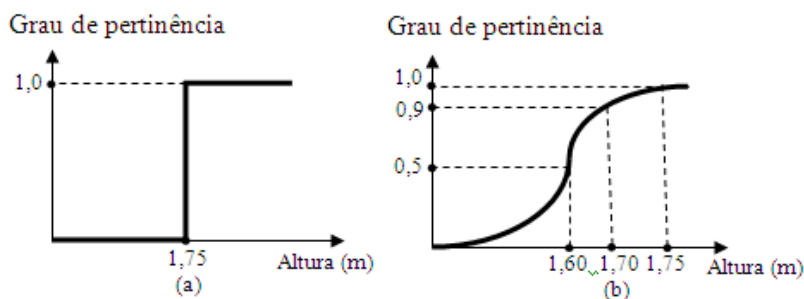


Figura 23- Comparação entre um conjunto clássico (a) e um *fuzzy* (b)

Fonte: adaptado de Collazos, Brasil e Azevedo (2008)

Os sistemas *fuzzy* recebem valores de entrada *crisp* e devem responder com valores de saída *crisp*. O processamento *fuzzy* está dividido em três etapas: fuzzificação, avaliação das regras e desfuzzificação. Como a figura 24 mostra, o sistema controlado envia informações *crisp* que são convertidas em valores *fuzzy* (fuzzificação). Esses valores *fuzzy* são processados pela base de conhecimento (regras e inferência). A saída *fuzzy* é desfuzzificada (convertida para valores *crisp*), que pode mudar as condições de operações do sistema controlado.

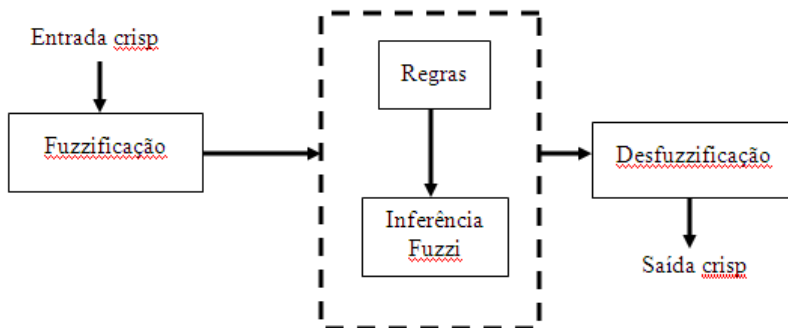


Figura 24- Processamento *fuzzy*

Fonte: Adaptado de Collazos, Brasil e Azevedo (2008)

2.5.3 Redes Bayesianas

Segundo Fred (2004) RB são representações que modelam conhecimento a partir de um modelo probabilístico que expressa relações de causas e efeitos, associando valores de probabilidade entre as variáveis de um determinado domínio do conhecimento. As RB lidam com as probabilidades condicionais do domínio do problema a ser estudado.

As RB destacam-se como uma das melhores técnicas de inteligência computacional, tendo-se tornado um dos principais métodos para se trabalhar com incerteza no campo da inteligência artificial (HUANG *et al.*, 2004).

Russel e Norvig (1995) apresentam as RBs como uma equação simples aplicada em todos os sistemas modernos de inferência probabilística. Elas são baseadas no teorema proposto pelo reverendo Thomas Bayes em 1763 (equação 1).

$$P(A / B) = \frac{P(A) \times P(B / A)}{P(B)}$$

Equação 1 - Teorema de Bayes

A interpretação deste teorema é feita da seguinte forma: a probabilidade condicionada de ocorrer o evento A, dado que ocorreu o evento B – $P(A/B)$, equivale à probabilidade incondicionada do evento A – $P(A)$, multiplicada pela probabilidade condicionada de ocorrer o evento B, dado que ocorreu o evento A – $P(B/A)$, dividindo-se tudo pela probabilidade incondicionada de ocorrer o evento B – $P(B)$.

Assim, as RB podem ser consideradas como diagramas ou grafos acíclicos dirigidos compostos por nós que representam as variáveis (entrada e saída) inter-relacionadas e arcs que representam dependências entre variáveis em um modelo probabilístico (figura 25).

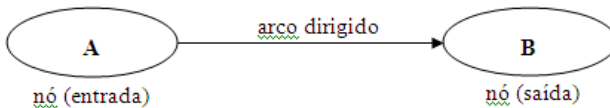


Figura 25- Esquema básico de uma rede Bayesiana simples

Entre os estudos consultados enfatiza-se o trabalho de Mello (2001) que destaca as potencialidades das RB em algoritmos que permitem a aprendizagem sobre um determinado domínio de problema, sendo mais competitivos que os melhores classificadores e apresentando um tempo computacional menor para a aprendizagem.

Sendo um caso particular das Redes Bayesianas, o Naïve Bayes (NB) será apresentado na sequência como facilitador para o cálculo das probabilidades.

2.5.3.1 Naïve Bayes

O Naïve Bayes (classificador ingênuo) é um caso especial de uma rede Bayesiana e é um classificador amplamente utilizado em aprendizagem de máquina.

A principal diferença entre uma RB e uma NB está no fato de que numa rede Bayesiana as variáveis independentes podem ter relação umas com as outras, enquanto no NB estas variáveis assumem total independência entre elas, estando única e exclusivamente relacionadas com a variável dependente.

Para demonstrar, consideram-se as seguintes condições:

- Variáveis Independentes: $X_1, X_2, \dots, X_n$,
- Variável dependente: C .

Então, o que se deseja é determinar a probabilidade de ocorrer o evento C dado as evidências $X_1, X_2, \dots, X_n$, ou seja:

$$P(C|X_1, X_2, \dots, X_n) = ?$$

Aplicando o teorema de Bayes, teremos:

$$P(C|X_1, X_2, \dots, X_n) = \frac{P(C) \cdot P(X_1, X_2, \dots, X_n|C)}{P(X_1, X_2, \dots, X_n)}$$

Como o denominador da equação não depende de C, e no final da análise serão normalizadas todas as probabilidades, então se pode dispensar esta parte da equação, ficando:

$$P(C|X_1, X_2, \dots, X_n) = P(C) \cdot P(X_1, X_2, \dots, X_n|C)$$

Partindo do princípio de que as evidências $X_1, X_2, \dots, X_n$ são independentes uma das outras, tem-se:

$$P(C | X_1, X_2, \dots, X_n) = P(C) \cdot P(X_1|C) \cdot P(X_2|C) \cdot \dots \cdot P(X_n | C)$$

Portanto, a probabilidade de ocorrer o evento “C” sendo dadas as evidências $X_1, X_2, \dots, X_n$, é dado pela probabilidade de ocorrer o evento “C” $\rightarrow P(C)$ multiplicado pelas probabilidades condicionais de cada uma das variáveis em relação à ocorrência do evento “C” $P(X_1/C)$, $P(X_2/C)$, $P(X_3/C)$, $\dots$, $P(X_n/C)$.

Neste trabalho será abordada a rede Bayesiana que, segundo Jensen (2001), devido ao formalismo, tem sido usada para trabalhar em domínios de aplicação na qual a incerteza é predominante.

2.6 MÉTODO DELPHI

O Método Delphi, de modo geral, é definido como uma atividade interativa desenhada para combinar opiniões de um grupo de especialistas objetivando consenso. Baseia-se em um processo de comunicação de grupo a fim de permitir que indivíduos, como um todo, possam lidar e explorar um problema complexo.

No artigo que relata a primeira experiência com o método Delphi, seus criadores, Dalkey e Helmer (1963), afirmam que o objetivo principal é obter o mais confiável consenso entre os especialistas sobre o problema tratado, embora isso nem sempre seja possível.

Este método envolve a aplicação sucessiva de questionários a um grupo de especialistas ao longo de várias rodadas. A pesquisa visa,

basicamente, a prospecção de tendências futuras sobre o objeto em estudo. No intervalo de cada rodada são feitas análises estatísticas das respostas e o resultado é compilado em novos questionários que, por sua vez, são novamente distribuídos ao grupo.

O método surgiu da necessidade de analisar cenários de alta complexidade, a partir da opinião e do consenso entre diferentes especialistas, conferindo confiabilidade ao resultado da aplicação do modelo. Os especialistas constituem a própria fonte dos dados (opiniões), pois são estes os detentores da capacidade de julgamento, para avaliar o desempenho de um determinado objeto em estudo (PORTO, SOUZA e BUARQUE, 1989).

O método Delphi inicia-se com a consolidação do questionário pela equipe de coordenação, para posterior aplicação no grupo de especialistas selecionados. A cada rodada, a equipe de coordenação faz uma análise de respostas fornecidas pelos especialistas, permitindo fornecer um quadro de opiniões sob a forma de dados estatísticos. Os resultados parciais retornam para os especialistas reverem suas opiniões e analisar as divergências do grupo. O objetivo é buscar o consenso de opinião sobre o assunto em questão até que se atinja um "estado estacionário".

2.6.1 Características do Método Delphi

As principais características do Método Delphi são: construção dos questionários, uso de especialistas, amostragem, anonimato, rodadas Delphi, confiabilidade e validade, e consenso.

1) Construção dos Questionários

A base do método envolve um questionário que é elaborado por uma equipe de coordenação e enviado a um grupo de especialistas participantes previamente selecionados. Na elaboração dos itens dos questionários, deve haver a preocupação com a precisão, independência e quantificação dos resultados.

Segundo Wright e Giovinazzo (2000), as principais recomendações de elaboração de questões são:

- Evitar eventos compostos: se o evento contiver uma parte com a qual o especialista concorda e outra com a qual discorda, fica difícil saber o que responder. Neste caso, a solução seria de separar os assuntos para obter respostas corretas;

- Evitar colocações ambíguas: não utilizar jargões técnicos considerados de domínio público ou termos como “comum”, “normal”, “uso geral”, “segmento significante de” e “será uma realidade”, pois as pessoas podem ter diferentes concepções sobre o significado da mesma palavra. Uma solução viável seria a utilização de colocações quantitativas;
- Tornar o questionário simples de ser respondido: o questionário deve ser elaborado para o especialista e não para o organizador da aplicação do método;
- Manter o número de questões: há um limite máximo prático para manter a atenção e a concentração de um especialista sobre um questionário. Este limite depende dos tipos de questão existentes e do perfil dos respondentes, mas não deve ultrapassar o número de 25 questões;
- Esclarecer previsões contraditórias: caso existam eventos excludentes no questionário, o especialista deve ser informado sobre a situação para auxiliar na lógica da resposta e evitar cair numa suposta armadilha por parte do organizador da pesquisa;
- Evitar ordenamento de proposições: evitar o pedido de priorização em uma série de proposições, neste caso pode-se substituir o ordenamento por uma avaliação individual de determinada proposição;
- Permitir complementação dos especialistas: as questões devem permitir contribuições dos especialistas visando enriquecer a pesquisa.

2) Uso de Especialistas

O Método se apoia na experiência, sabedoria e criatividade de um painel de especialistas e no fato de que o julgamento coletivo (quando realizado de maneira estruturada) é supostamente melhor que qualquer julgamento individual (CAGNIN, 2000).

Os especialistas são selecionados em vários setores de atuação (como academia, empresa e governo), para que todos tenham representatividade na pesquisa, permitindo então maior credibilidade ao estudo (SANT ANA, 2005).

Segundo Adams (1980), os especialistas devem sentir-se pessoalmente envolvidos com o problema em debate e motivados para as respostas. Devem possuir informações consistentes, a serem

compartilhadas, e estar convencidos de que os resultados da pesquisa irão fornecer informações valiosas, não acessíveis de outra forma.

3) Amostragem

Não existe uma amostra ideal para obter consenso na conjunção de contribuições de especialistas que trabalham a distância. O Método Delphi está baseado em fundamentos teóricos diferentes das pesquisas sociais convencionais, ou seja: suas amostras estão imunes ao problema de retorno dos questionários que é fundamental quando se trata das amostras representativas das pesquisas tipo *survey* (SANTOS e AMARAL, 2004; QUIRINO e IRIAS, 1998; QUIRINO, LUIZ e DIAS, 1999; FERRAZ, 1993; CRISTO, 2002; WRIGHT e GIOVINAZZO, 2000).

Rowe e Wright (1999) revisaram 27 estudos sobre aplicação do Método Delphi e concluíram que não há relação consistente entre o tamanho da amostra e a eficiência da aplicação do método.

4) Anonimato

O anonimato é um dos pontos mais importantes e menos questionados pela literatura pesquisada (KAYO e SECURATO, 1997). Segundo Santos e Amaral (2004), o anonimato garante aos especialistas a igualdade na expressão de idéias, pois não há possibilidade de pressão psicológica pelos membros mais influentes do grupo.

5) Rodadas Delphi

O número de rodadas nas quais os questionários são enviados para os especialistas depende do tempo necessário para obtenção do consenso entre os participantes. Segundo Cardoso *et al.* (2005), o tempo necessário para a elaboração do questionário e sua aplicação, conforme experiências conhecidas é da ordem de quatro (4) a seis (6) meses.

6) Confiabilidade e Validade

Estudos realizados demonstram que os resultados são confiáveis e que a cada rodada a consistência torna-se maior (KEENEY, HASSON e MCKENNA, 2001).

Por outro lado, como os participantes do Delphi são especialistas de grupos ou de áreas de conhecimento, então a validade do método pode ser assumida.

7) Consenso

A questão do consenso ainda é polêmica. Alguns autores divergem sobre a necessidade de obter consenso. Existe o grupo que considera o consenso como condição essencial do processo, enquanto outros acreditam que o consenso deve ser alcançado, mas pode não ocorrer em todas as questões, porém esta situação não invalida os objetivos da pesquisa (CARDOSO *et al.*, 2005). Com o desenvolvimento do Método Delphi, o consenso deixou de ser exigido dependendo do tipo de informação que se deseja coletar (KAYO e SECURATO, 1997).

2.7 CONSIDERAÇÕES SOBRE O CAPÍTULO

Nota-se que, neste capítulo, foi dada muita ênfase ao processo de diagnóstico na área da saúde por apresentar uma estrutura bem definida que permite aplicar os conceitos, métodos e técnicas abordados neste trabalho; citam-se os seguintes fatos como principais motivos que levaram a esta escolha:

- Importância do processo de diagnóstico na área da saúde (relevância);
- Arquivos literários estruturados que possibilitam pesquisas dirigidas (descoberta baseada na literatura); e
- Especialistas pesquisadores, dentro desta área, que atuam na atividade de diagnóstico (validação de consistência do modelo proposto).

3 PROCEDIMENTOS METODOLÓGICOS

A Engenharia do Conhecimento é uma área que tem por objetivo investigar e propor modelos de representação de conhecimento através de metodologias, métodos e técnicas. Dentre as metodologias destaca-se a CommonKADS (SCHREIBER *et al.*, 2002), a qual oferece um conjunto de etapas para o desenvolvimento de Sistemas Baseados em Conhecimento.

Inicialmente, através desta metodologia, a primeira etapa compreende uma análise de contexto onde será desenvolvido o modelo proposto. Nesta análise abordam-se os diferentes aspectos dos modelos humanos de conhecimento e não apenas a extração do conhecimento de um especialista. Pretende-se com esta análise responder as questões: Por quê? O quê? Como?, necessárias à modelagem de um Sistema de Conhecimento.

Na sequência, a próxima etapa é explicitar e formalizar o conhecimento, bem como destacar como ocorrem as interações entre os agentes envolvidos em sua utilização. A construção da ontologia de domínio, auxiliada pela modelagem CESM, torna o domínio específico mais descritivo e abrangente sob ponto de vista sistêmico.

A integração entre a técnica da ontologia e o cálculo probabilístico, envolvendo incertezas que compõem o domínio específico, forma a estrutura para a construção do modelo.

Com o modelo proposto estruturado, inicia-se a aplicação em um estudo de caso. Salienta-se que, devido à generalidade do modelo elaborado, qualquer área do conhecimento tem aplicação. Neste trabalho, a metodologia será aplicada a uma clínica especializada em Acupuntura e a enfermidade Acne será utilizada como referência para construção do domínio específico.

O estudo de caso, parte desta pesquisa, tem como objetivo apresentar os conceitos, métodos e técnicas empregados na elaboração do modelo. Com o auxílio de especialistas, os assuntos abordados na construção do modelo serão colocados em discussão. Pretende-se através do consenso entre os especialistas verificar a consistência do modelo para o apoio ao diagnóstico através da geração de novos conhecimentos.

Como se trata de um modelo conceitual onde os pacientes e dados são hipotéticos, utilizou-se a prospecção, processo este que se ocupa de procurar, sistematicamente, examinar o futuro de longo prazo da ciência, da tecnologia, da economia e da sociedade, com o objetivo de identificar as áreas de pesquisa estratégica e as tecnologias genéricas

emergentes que têm a propensão de gerar os maiores benefícios econômicos e sociais (MARTIN *et al.*, 1998).

O termo prospecção é usado no sentido de perspectiva e não tem a mesma conotação de previsão, que estaria mais próximo de projeção ou futurologia. A prospecção considera que não existe um único futuro. Dependendo da ação ou da não-ação no presente, são possíveis muitos futuros, mas somente um deles ocorrerá (CUHLS e GRUPP, 2001).

O planejamento prospectivo é, portanto, um processo estruturado e coordenado, que tem como função a formulação de estratégias para se atingir objetivos. Nesse sentido, atingir o futuro previsto passa a ser até secundário, uma vez que o principal objetivo do processo é orientar as decisões e as ações do presente.

O método Delphi tem sido um dos instrumentos mais utilizados na realização de estudos prospectivos e deve ser usado sempre que a informação não pode ser quantificada ou quando os dados históricos não estão disponíveis (MARINHO, 2006). Vale ressaltar que, mesmo com dados históricos, a opinião de especialistas pode ser usada como uma forma de complementar as informações obtidas e de captação de conhecimentos tácitos, sinais fracos e *insights*.

Para verificar a consistência do modelo proposto, o método Delphi foi aplicado, em um grupo de especialistas, para validar os conceitos, métodos e técnicas empregados na elaboração do modelo que apoia o processo de diagnóstico no sentido de possibilitar a geração de novos conhecimentos.

3.1 OPERACIONALIZAÇÃO DO MÉTODO DELPHI

O método Delphi foi escolhido devido à consistência teórica para obtenção de consenso em um grupo de especialistas, onde por meio da aplicação de questionários sucessivos e um processo de *feedback* das respostas, os participantes têm condições de revisar seu ponto de vista embasado nas opiniões do grupo.

Não há uma forma rígida para implementação do método Delphi; os tipos de questionários a serem utilizados, número de rodadas e quantidade de especialistas participantes são exemplos de variáveis com amplo espectro de variação entre as aplicações.

Para a escolha dos especialistas levou-se em consideração os anos de experiência na atividade de diagnóstico e a correlação entre a área de atuação profissional com a área da pesquisa. Foram selecionados professores da pós-graduação de universidades públicas e

particulares, bem como profissionais que atuam em órgãos municipais, estaduais e federais.

A operacionalização do método Delphi será estruturada em três fases:

FASE 1: Esta fase inicia-se com a escolha dos especialistas que participarão da pesquisa definindo-os de forma não – probabilística, mas sim intencionalmente. Foram selecionados quatorze (14) especialistas para compor a amostra empregada neste estudo. Na sequência, elaborou-se o questionário com perguntas referentes aos conceitos, técnicas e ferramentas utilizadas para a montagem do modelo proposto. Inicialmente, o questionário foi disponibilizado para um grupo de três especialistas para verificar as oportunidades de melhoria do instrumento. Após a inserção das sugestões feitas pelos profissionais no pré-teste, o questionário resultante foi levado de forma impressa e enviado por *e-mail* a todos os participantes. Além disso, para facilitar o entendimento das questões, foi feita uma apresentação individual para cada um dos especialistas, envolvendo a aplicação do modelo proposto num estudo de caso. Também foi disponibilizado aos profissionais um guia de utilização do modelo (Apêndice IV), no qual são exemplificadas todas as etapas/fases que envolvem o funcionamento do mesmo. Para verificar o grau de concordância dos especialistas, em relação aos métodos, técnicas e ferramentas, foi utilizada uma escala com três níveis de resposta: concordo; tenho restrições; discordo. A escala foi escolhida por ser uma escala de simples mensuração de atitude, na qual cada especialista atribui seu resultado de forma independente, sendo que os escores obtidos pelas proposições podem ser correlacionados com os totais alcançados. O prazo dado para o retorno dos questionários respondidos delimitou-se em duas semanas.

FASE 2: Após o término da primeira rodada, foi elaborado um relatório contendo as opções desejadas e os comentários, sugestões ou justificativas dos participantes. Também foi elaborado um quadro estatístico para cada questão, contendo as percentagens de escolha de cada item, considerando a totalidade do grupo. A segunda rodada começou com o envio do mesmo questionário; porém leva em seu conteúdo a síntese das considerações de cada especialista e as percentagens de concordância para cada questão. Esta nova etapa visou

subsidiar uma reavaliação sobre os temas abordados, tendo como parâmetro as opiniões geradas pelo do grupo de especialistas dadas na primeira fase. O novo prazo foi de duas semanas para devolução dos questionários.

FASE 3: Quando há um consenso sobre o objetivo da proposta de pesquisa, o processo é finalizado por meio da elaboração de um relatório final. O resultado desta consulta pretendeu gerar uma gama de informações que permitiu validar a sistemática proposta, bem como evidenciar necessidades de um maior detalhamento ou obter novas interpretações por meio dos resultados encontrados. Os comentários e os *insights* gerados podem ser utilizados para fomentar as discussões e auxiliar na tomada de decisão. Além disso, a aplicação dos questionários permite a visualização de alternativas de cruzamento de dados na busca de encontrar diferentes soluções que anteriormente não eram consideradas.

3.2 ESPECIALISTAS PARTICIPANTES

Na condução do método Delphi contou-se com especialistas experientes que, há vários anos de atividades, atuam diretamente no processo de diagnóstico ou estão envolvidos em sistemas automatizados de diagnóstico. Os profissionais foram encontrados em universidades públicas e particulares, hospitais ou clínicas, bem como em órgãos da saúde municipais, estaduais e federais.

Abaixo uma breve descrição da formação e da atuação profissional destes 14 especialistas. Seus nomes serão mantidos em sigilo para garantir-lhes privacidade.

- Médico, Doutor em Gerontologia Biomédica. Atua em Geriatria, Gerontologia, Cuidados Paliativos, e Bioética. Tem 30 anos de experiência na prática de diagnóstico.
- Fisioterapeuta, Mestre em Tecnologia na Saúde. Atua na área de Fisioterapia, com ênfase em Acupuntura e Terapias Manuais. Tem 15 anos de experiência na prática de diagnóstico.
- Médica, Especialista em Oncologia. Atua na área de Oncologia Pediátrica. Tem 36 anos de experiência na prática de diagnóstico.
- Bacharel em Informática, Doutor em Engenharia Eletrotécnica e de Computadores. Atua na área da Ciência da Computação,

com ênfase em Inteligência Artificial e automatização de sistemas.

- Fisioterapeuta, Mestre em Tecnologia na Saúde. Atua na área de Fisioterapia, Acupuntura, RPG e Fisioterapia em Ortopedia. Tem 5 anos de experiência na prática de diagnóstico.
- Médica, Especialista em Cardiologia. Tem 11 anos de experiência na prática de diagnóstico.
- Fisioterapeuta, Doutora em Ciências - Engenharia Biomédica. Atua nas áreas de Reabilitação Neurológica, Neuropediátrica e Informática em Saúde. Tem 9 anos de experiência na prática de diagnóstico.
- Médico, Especialista em Saúde Comunitária. Atua na rede Municipal de Saúde. Tem 25 anos de experiência na prática de diagnóstico.
- Administrador de Empresas, Doutor em Engenharia Eletrotécnica e de Computadores. Atua na área de Ciência da Computação, com ênfase em Sistemas Multiagente.
- Fisioterapeuta, Mestre em Tecnologia na Saúde. Atua nas áreas de Acupuntura, Acupuntura Auricular, Eletroacupuntura e Analgesia. Tem 20 anos de experiência na prática de diagnóstico.
- Psicóloga, Mestre em Informática na Saúde. Atua na área de Psicologia Clínica. Tem 12 anos de experiência na prática de diagnóstico.
- Médico, Especialista em Medicina do Trabalho Ocupacional e Clínica Médica. Atua como Clínico Geral. Tem 28 anos de experiência na prática de diagnóstico.
- Médico, Especialista em Homeopatia e Acupuntura. Atua na rede Municipal de Saúde. Tem 34 anos de experiência na prática de diagnóstico.
- Médico, Doutor em Urologia. Atua na área de Urologia. Tem 25 anos de experiência na prática de diagnóstico.

A tabela 2 resume a amostra de profissionais: quanto à área do conhecimento; número de profissionais que trabalham com diagnóstico e tempo médio de experiência com a tarefa de diagnosticar.

Tabela 2 - Resumo da amostra de profissionais

Área do Conhecimento	Número de Profissionais	Área de Atuação	Tempo médio de experiência
Exatas - Ciência da Computação	2 = 14 %	Sistemas automatizados para diagnóstico	
Saúde	12 = 86 %	Diagnóstico	20 anos

4 MODELO PROPOSTO E APLICAÇÃO

Para início de entendimento do modelo proposto, apresenta-se a sua arquitetura através de seus componentes, as relações entre eles e a análise de cada etapa envolvida. Em seguida, será exibido o diagrama de atividades composto pelo fluxo para apoio ao diagnóstico e o fluxo para a geração de um novo conhecimento, tendo como objetivo mostrar o funcionamento de cada um dos processos.

A aplicação deste modelo é feita em um estudo de caso, na área da acupuntura, englobando as principais etapas para a construção de um Sistema Baseado em Conhecimento. A ontologia utilizada para este estudo é apenas um exemplo de aplicação, pois o modelo é genérico e permite adaptação a qualquer domínio específico que envolva a tarefa de diagnóstico.

4.1 ARQUITETURA DO MODELO PROPOSTO

A arquitetura do modelo é apresentada pela figura 26 e na sequência, as etapas são descritas.

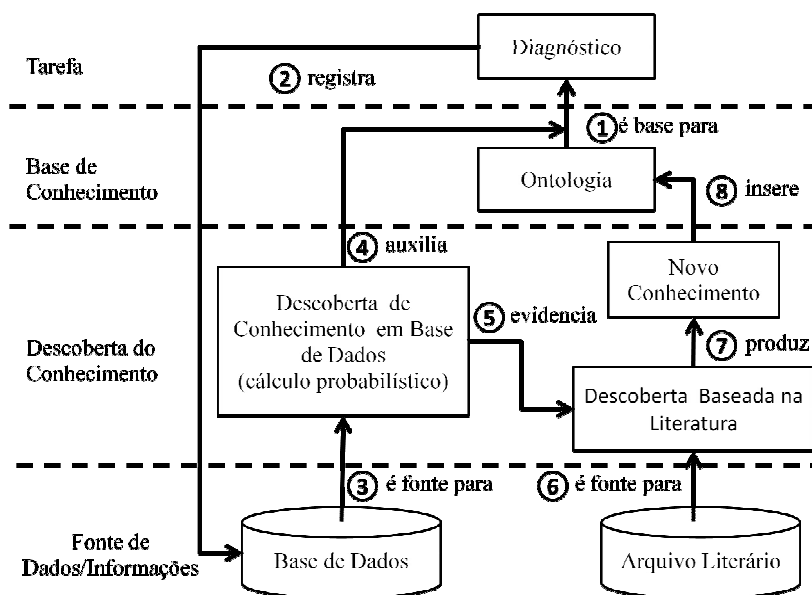


Figura 26- Arquitetura do modelo proposto

- **Tarefa**

A etapa correspondente à tarefa tem a função de realizar o diagnóstico. Segundo Schreiber *et al.* (2002), trata-se de uma tarefa analítica cujo objetivo é encontrar a resposta ao mau funcionamento do sistema. Os passos envolvidos nesta etapa são:

1) Recebe as informações provenientes da base de conhecimento (ontologia) e, através das inferências, procura chegar ao resultado.

Pela figura 26 pode-se observar que em “1” a ontologia tem um relacionamento com o diagnóstico do tipo “é base para”, permitindo, desta forma, concluir que “A ontologia é base para realização do diagnóstico”.

2) Fornece à Base de Dados (BD) os registros dos diagnósticos realizados.

Pela figura 26 pode-se observar que em “2” o diagnóstico tem um relacionamento com a base de dados do tipo “registra”, permitindo, desta forma, concluir que “O diagnóstico registra as informações na base de dados”.

- **Base de Conhecimento**

A etapa correspondente à Base de Conhecimento é formada pela ontologia de um domínio específico. No trabalho de Lopes *et al.* (2010) uma ontologia de domínio é construída e através da modelagem CESM obtém-se uma estrutura sistêmica mais descritiva e abrangente que auxilia a realização do diagnóstico.

Os passos envolvidos nesta etapa são:

1) Provê as informações que servem de base para que a tarefa de diagnóstico seja executada.

2) Permite a inserção de possíveis novos conhecimentos nos modelos de manifestação da tarefa de diagnóstico.

Pela figura 26 pode-se observar que em “8” a ontologia tem um relacionamento com o novo conhecimento do tipo “insere”, permitindo, desta forma, concluir que “A ontologia insere um novo conhecimento em sua base de conhecimento”. Esta inserção será feita pelo especialista após confirmação, através da comunidade científica, da geração de um novo conhecimento.

• Descoberta do Conhecimento

A etapa correspondente à Descoberta do Conhecimento é responsável pelos métodos e técnicas para a extração do conhecimento. As técnicas envolvidas nesta etapa são o cálculo probabilístico (CP) e a Descoberta Baseada na Literatura (DBL). Para o CP estão envolvidas as seguintes atividades:

1) Realiza a busca, na fonte de dados, dos elementos necessários para executar o diagnóstico probabilístico.

2) Auxilia na execução do diagnóstico quando os dados obtidos na base de conhecimento são insuficientes para a realização da tarefa.

Pela figura 26 pode-se observar que em “4” o cálculo probabilístico tem um relacionamento com o diagnóstico do tipo “auxilia”, permitindo, desta forma, concluir que “o cálculo probabilístico auxilia o diagnóstico”.

3) Evidencia dados relevantes para serem submetidos a uma análise literária.

Pela figura 26 pode-se observar que em “5” o cálculo probabilístico tem um relacionamento com a Descoberta Baseada na Literatura do tipo “evidencia”, permitindo, desta forma, concluir que “O cálculo probabilístico evidencia uma informação relevante para a Descoberta Baseada na Literatura”.

Para a DBL estão envolvidos os seguintes passos:

1) Utiliza a evidência gerada pelo cálculo probabilístico como elemento candidato a um novo conhecimento.

2) Realiza a busca no Arquivo Literário elementos necessários para executar a análise da evidência dentro de padrões científicos.

Pela figura 26 pode-se observar que em “6” o Arquivo Literário tem um relacionamento com a Descoberta baseada na Literatura do tipo “é fonte para”, permitindo, desta forma, concluir que “O Arquivo Literário é fonte para a Descoberta Baseada na Literatura”.

3) Comprova cientificamente o sucesso ou fracasso de uma evidência realçada como um possível novo conhecimento. Este novo conhecimento deverá ser confirmado ou rejeitado usando julgamento humano, métodos de laboratório ou investigações clínicas, dependendo da natureza do conceito descoberto.

Pela figura 26 pode-se observar que em “7” a DBL tem um relacionamento com o Novo Conhecimento do tipo “produz”, permitindo, desta forma, concluir que “A DBL produz um Novo Conhecimento”.

- **Fonte de Dados/Informações**

A etapa correspondente à Base de Dados e ao Arquivo Literário compõe a fonte de pesquisa para o sistema probabilístico e o sistema de descoberta baseado na literatura, respectivamente. Os passos envolvidos em relação à Base de Dados são:

1) Recebe e estrutura os registros dos diagnósticos realizados para futuras consultas.

2) Serve de fonte de dados para a realização das inferências do cálculo probabilístico.

Pela figura 26 pode-se observar que em “3” a base de dados tem um relacionamento com o CP do tipo “é fonte para”, permitindo, desta forma, concluir que “A base de dados é fonte para o cálculo probabilístico”.

Em relação ao Arquivo Literário, o passo envolvido é fornecer as informações científicas necessárias para que um novo conhecimento seja identificado.

4.2 DIAGRAMA DE ATIVIDADES

O objetivo do diagrama de atividades é mostrar o fluxo de controle de uma atividade para outra. Inicialmente é apresentado o fluxo para apoio ao diagnóstico que possibilita obter um resultado probabilístico e, em seguida, o fluxo para a geração de novos conhecimentos através de evidências realçadas e confrontadas com arquivo literário.

4.2.1 Fluxo de Apoio ao Diagnóstico

A figura 27 exibe o fluxo do processo de diagnóstico. Neste fluxo, estão associados o diagnóstico proveniente da análise das manifestações, bem como o diagnóstico probabilístico, obtido através da base de dados.

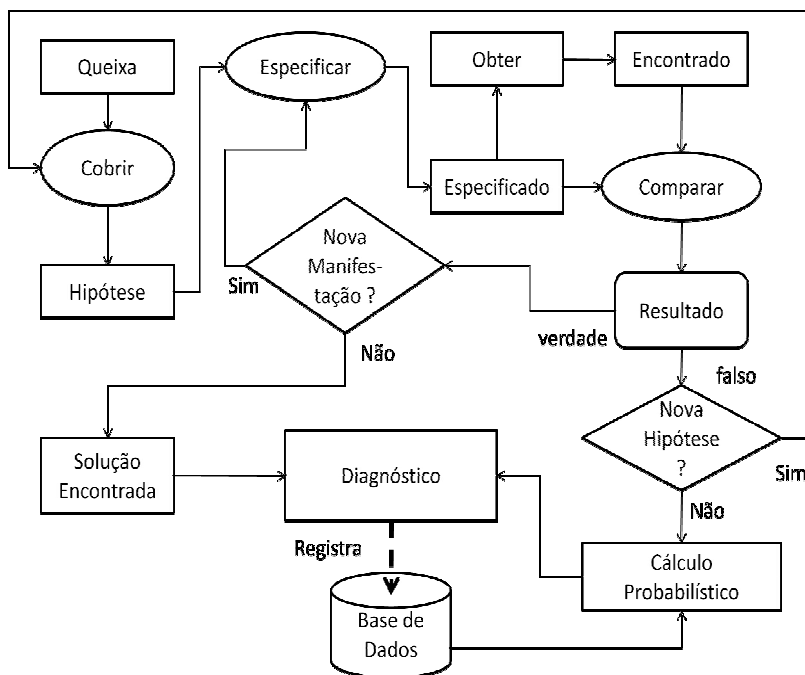


Figura 27- Diagrama de atividades para apoio ao diagnóstico

A “Queixa” dá início ao processo de busca à solução do problema. A inferência “Cobrir” procura encontrar em uma rede causal as possíveis hipóteses para a reclamação. Uma “Hipótese” para solucionar a reclamação é especificada. Na inferência “Especificar” todas as manifestações referentes à hipótese levantada são analisadas.

Sendo especificada uma manifestação, a função de transferência “Obter” obtém o valor atual da observação a fim de testar as manifestações candidatas. A inferência “Comparar” verifica se o valor atual observado é consistente com a manifestação selecionada da hipótese que está sendo testada.

Sendo “verdade” o resultado, uma nova manifestação da hipótese será testada. Se todas as manifestações forem consistentes, a solução para a “Queixa” foi encontrada.

Portanto, com a obtenção do diagnóstico, inicia-se o tratamento à “Queixa”. O sucesso do tratamento fornece as informações que são inseridas em uma base de dados para utilização posterior.

Se alguma das manifestações que compõem a hipótese não for consistente, o resultado da comparação será “falso” e, uma nova

hipótese será testada, reiniciando o ciclo de comparação entre os valores especificados e os valores encontrados.

No caso em que nenhuma das hipóteses consiga cobrir integralmente a queixa, o cálculo probabilístico fornecerá um diagnóstico aproximado, tendo como fonte de informação a base de dados dos registros coletados.

4.2.2 Fluxo para a Geração de Novo Conhecimento

Dando sequência ao diagrama de atividades, a figura 28 mostra o fluxo complementar para a geração de um novo conhecimento.

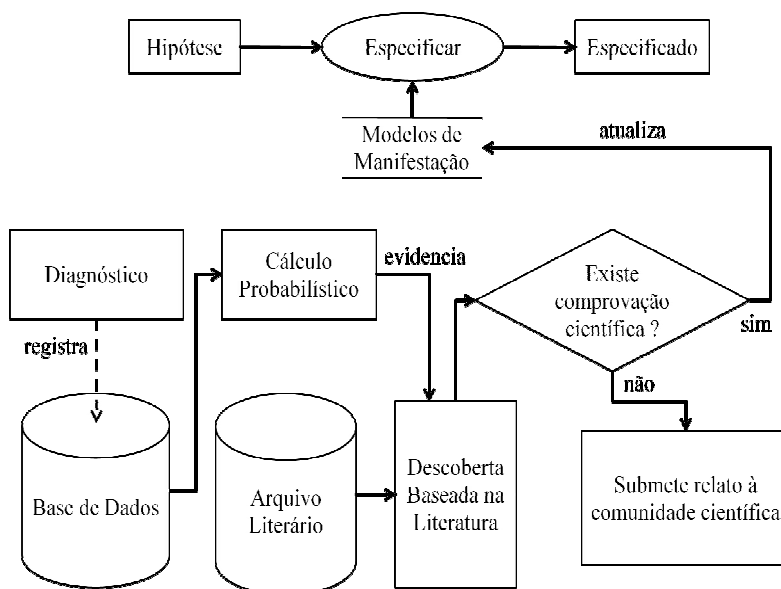


Figura 28- Diagrama de atividades para a geração de um novo conhecimento

Utilizando as informações obtidas pelos diagnósticos realizados e armazenadas na base de dados, através do processo de inferência, o cálculo probabilístico sugere evidências candidatas à geração de um possível novo conhecimento.

Assim, uma evidência realçada será confrontada com o arquivo literário envolvendo o sistema de descoberta baseado na literatura. Caso seja confirmado cientificamente este novo conhecimento, o mesmo será inserido nos modelos de manifestação que caracterizam uma determinada patologia.

O Apêndice IV, deste trabalho, apresenta um “Guia de utilização do modelo proposto” cuja finalidade é orientar o usuário na construção das etapas que compõem a arquitetura do modelo. Salienta-se que no Apêndice III são apresentadas algumas “ferramentas” úteis que permitem a escolha da construção de cada etapa do modelo.

4.3 APLICAÇÃO DO MODELO – ESTUDO DE CASO

Nesta seção apresenta-se um estudo de caso, na área da acupuntura, que aplica o modelo proposto para a construção de um Sistema Baseado em Conhecimento utilizando metodologia, métodos e técnicas disponíveis na Engenharia do Conhecimento para a realização do processo de diagnóstico.

4.3.1 Análise de Contexto

A metodologia CommonKADS, dentro da Engenharia do Conhecimento, foi utilizada na estruturação de um Sistema de Conhecimento para diagnóstico em Acupuntura que é apresentado no trabalho de Lopes *et al.* (2011) (Apêndice V).

Neste trabalho, esta metodologia fornece uma visão de contexto sistêmico sobre a organização e o problema a ser tratado. Isto a torna menos focada na construção de um artefato, mas sim, na elaboração de um modelo de conhecimento que possa gerar resultados de forma a minimizar as incertezas e riscos quanto à implantação de um Sistema de Conhecimento para a organização.

Dentre as planilhas que fazem parte deste trabalho e que compõem o nível de contexto, por exemplo, é apresentada no quadro 11 a planilha denominada por OTA1 - *Checklist* para Decisão sobre Impactos e Melhorias; ela resume um conjunto de informações que auxiliam o processo decisório de implantação de um Sistema de Conhecimento tornando-o mais seguro, eficaz e abrangente às necessidades de quem irá utilizá-lo.

Modelo de Organização, Tarefa e Agente	OTA1 – <i>Checklist</i> para Decisão sobre Impactos e Melhorias
Impactos e mudanças na organização	Estrutura: não há impactos e mudanças necessárias na estrutura da organização. Processos: diagnóstico será realizado apoiado por um Sistema de Conhecimento, proporcionando maior eficácia e qualidade no processo. Recursos: será necessário a utilização de um Sistema de Conhecimento disponível em um computador. Pessoas: o acupunturista terá a seu dispor um Sistema de Conhecimento, não existindo outros impactos ou mudanças para as pessoas na organização. Conhecimento: o conhecimento que anteriormente estava apenas no domínio do especialista, passa a estar formalizado e explicitado através de um Sistema de Conhecimento e pode ser compartilhado entre especialistas, profissionais e estudantes da área. Cultura e poder: o Sistema de Conhecimento pode impactar positivamente para a organização na medida em que ele possibilitará gerar novos conhecimentos sobre os pacientes e sobre os processos de diagnóstico e tratamento. Permitirá maior compartilhamento de informações, possibilitando que um paciente possa ser acompanhado por mais de um profissional.
Impactos e mudanças na tarefa e nos agentes	O acupunturista, de posse de um Sistema de Conhecimento que auxilie no diagnóstico, poderá tomar decisões mais rapidamente e com maior precisão, reduzindo o grau de incerteza e o tempo de execução da tarefa. Com isso, com o mesmo tempo que ele tem disponível para atendimento na situação atual, ele poderá atender um maior número de pacientes, a partir da implantação da solução proposta. O sistema também pode ser utilizado como fonte de conhecimento e formação para novos profissionais que desejam atuar na área. Para os pacientes, espera-se que o tratamento torne-se mais eficaz, melhorando mais rapidamente sua qualidade de vida.
Atitudes e compromissos	O acupunturista deve se comprometer a utilizar o Sistema de Conhecimento, como fonte primordial para o seu processo decisório, fazendo sugestões e críticas para revisão e aperfeiçoamento do mesmo.

Modelo de Organização, Tarefa e Agente	OTA1 – Checklist para Decisão sobre Impactos e Melhorias
Ações propostas	Melhoria: implantar um Sistema de Conhecimento que auxilie o diagnóstico para o tratamento de um paciente de acupuntura. Acompanhamento: para a implantação adequada do sistema é fundamental a validação do sistema com especialistas e treinamento para utilização. Resultados esperados: espera-se agilizar o processo de tratamento, gerando com isto ganhos para a organização já destacados anteriormente. Os custos de implantação são relativamente baixos, em função da disponibilidade técnica e tecnológica que viabilizam a construção do Sistema de Conhecimento proposto.

Quadro 11 - OTA1: Checklist para Decisão sobre Impactos e Melhorias

4.3.2 Construção da Ontologia

A partir da construção do nível de contexto apresentado em Lopes *et al.* (2011) (Apêndice V), foi elaborada em Lopes *et al.* (2009) (Apêndice VI) uma ontologia de apoio ao diagnóstico aplicado à Acupuntura. Para a estruturação e documentação da ontologia foi utilizada a ferramenta OntoKEM devido ao detalhamento que proporciona e à possibilidade de reuso (RAUTENBERG *et al.* 2008).

Na sequência, em Lopes *et al.* (2010) (Apêndice VII), a ontologia foi expandida através do modelo CESM, o qual proporcionou, dentro de uma visão sistêmica, a explicitação do conhecimento através de uma modelagem mais descritiva e abrangente para os elementos que compõem um domínio específico.

A figura 29 ilustra o diagrama CESM para o sistema de diagnóstico, destacando:

- 1) os componentes – coleção de todas as partes do sistema – são compostos pelo paciente e seus sintomas e sinais, os resultados dos exames realizados em laboratório e o profissional que executa o diagnóstico;
- 2) o ambiente – coleção dos itens que não pertencem ao sistema, mas atuam ou sofrem a ação por algum ou todos os componentes do sistema - está representado pelas atividades profissionais, as atividades físicas, os tipos de alimentação e os sentimentos ou emoções, sendo estes alguns dos elementos escolhidos que podem interferir diretamente no resultado do diagnóstico;

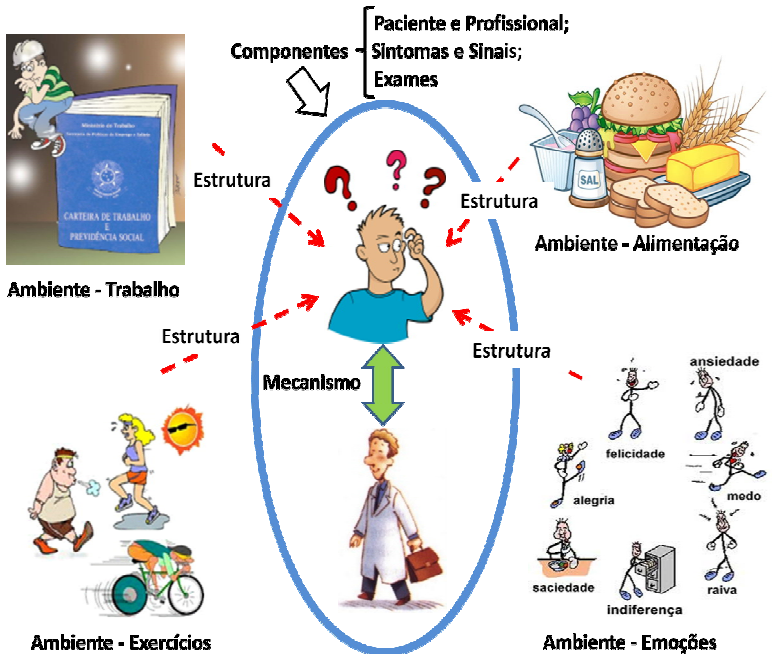


Figura 29- Diagrama CSM para o processo de diagnóstico

3) a estrutura, no modelo CSM, é definida como a coleção de ligações/relações entre componentes e entre esses e itens do ambiente. Como exemplo, pode-se citar as ligações da forma “O paciente tem sintoma” ou “O paciente faz exercícios” ou ainda “A ansiedade é uma emoção do paciente”; e

4) o mecanismo - coleção de processos que geram mudanças qualitativas no sistema – pode ser interpretado como as restrições que definem as limitações nas relações entre os componentes e o ambiente e são usados para realizar o processo de consulta ou raciocínio sobre o sistema.

Baseado na quádrupla CSM, a figura 30 apresenta o diagrama que envolve o processo de diagnóstico dos profissionais nas áreas da saúde. Neste sistema, o paciente acometido de uma enfermidade busca tratamento para sua patologia. O profissional analisa, de forma sistêmica, os sintomas, sinais, exames e as evidências levando em conta os fatores internos (componentes) e os fatores externos (ambiente) do paciente, que, ao serem relacionados (estrutura), são limitados pelas

restrições, caracterizando determinada patologia para tratamento (mecanismo).

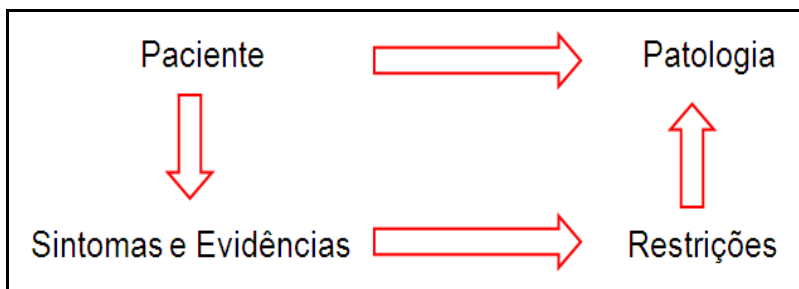


Figura 30- Efeito sistêmico do processo de diagnóstico

- Classes e Subclasses

Os procedimentos de Aquisição do Conhecimento, adquiridos através de consulta a especialista, manuais e livros e a Representação do Conhecimento permitem a formalização de uma base de conhecimento formada pelas classes e subclasses da ontologia que compõem o modelo hierárquico da mesma.

A figura 31 apresenta a hierarquia da ontologia proposta para a enfermidade Acne, gerada a partir da ferramenta *Protégé* em que são consideradas as classes: Paciente, Órgão, Atividade Física, Emoção, Alimentação e Ocupação.

Foram considerados como órgãos: Pulso, Língua, Pele e Face, que apresentam sintomas e sinais para evidenciar o tipo de patologia em um paciente acometido de acne. Os órgãos foram hierarquizados em subclasses:

- Subclasses da classe Pele: Pele seca e Pele oleosa;
- Subclasses da classe Língua: Língua pálida, Língua saburra amarela e Língua saburra pegajosa;
- Subclasses da classe Face: Face vermelha e Face inchada;
- Subclasses da classe Pulso: Pulso encharcado, Pulso deslizante e Pulso rápido.

Dentre as possíveis patologias para o tratamento de Acne, foram consideradas as classes para o diagnóstico: Acne por Umidade e Calor (AUC), Acne por Calor Tóxico (ACT) e Acne por Deficiência do QI (ADQ) (energia vital).



Figura 31- Classes e subclasses da ontologia

Foram selecionados alguns elementos do ambiente que influenciam o processo de diagnóstico, para compor as classes externas. Os elementos escolhidos foram: os tipos de alimentação, os tipos de atividades físicas, os tipos de ocupação do paciente no trabalho e os tipos de emoção.

As classes externas foram hierarquizadas em subclasses:

- a) Subclasses da classe Alimentação: Vegetariana, Carnívora e Balanceada;
 - b) Subclasses da classe Atividade Física: Ativo, Semiativo e Sedentário;
 - c) Subclasses da classe Ocupação: Insalubre, Estressante e Normal;
 - d) Subclasses da classe Emoção: Ansiedade, Tristeza, Medo e Raiva.
- Relações entre Classes

As relações existentes na ontologia são necessárias para estabelecer os diferentes relacionamentos entre as classes e subclasses. A figura 32 apresenta as relações existentes entre as classes Paciente, Órgão, Alimentação, Atividade Física, Ocupação e Emoção. A classe Paciente tem um relacionamento com a classe Alimentação do tipo “temAlimentação”; assim, pode-se concluir que “Paciente tem Alimentação”. Nota-se também que a relação “temAlimentação” possui a relação inversa “nutre”, da forma “Alimentação nutre o Paciente”.

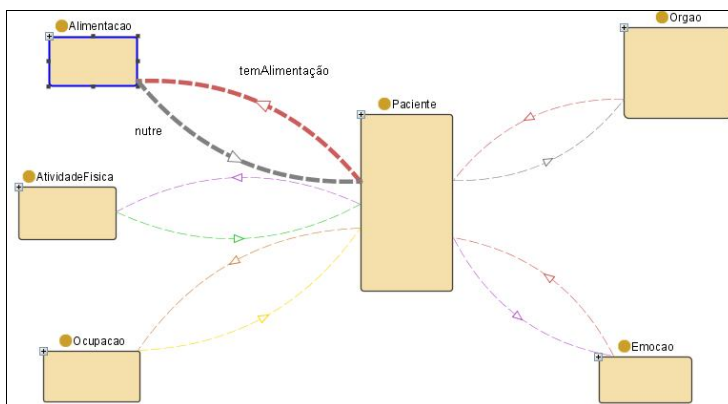


Figura 32- Relações entre classes

- Regras ou axiomas

Após construir a hierarquia de classes com suas respectivas relações, foram definidas as restrições, que são as regras que restringem os relacionamentos e permitem estabelecer se um indivíduo pertence ou não a uma determinada classe.

Pela figura 33 pode-se observar que para um paciente ter ACT é necessário e suficiente que apresente os seguintes sintomas: Face vermelha; Pulso deslizante ou Pulso rápido; Língua saburra amarela ou Língua saburra pegajosa.

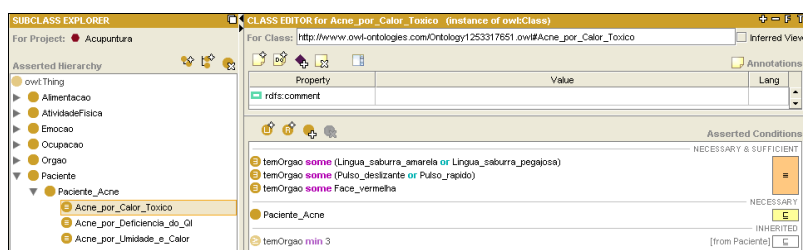


Figura 33- Restrições de classes

Para a patologia ADQ foram definidas as seguintes restrições: Pele oleosa, Pulso encharcado, Língua pálida ou Língua saburra pegajosa. Para a patologia AUC foram definidas as seguintes restrições: Face inchada, Pele oleosa, Língua saburra amarela ou Língua saburra pegajosa. Já para a relação “temOrgão” foi definido como necessário e suficiente o paciente ter no mínimo três órgãos sintomáticos.

- Instâncias

A construção das classes permite criar as instâncias para a base da ontologia. As instâncias representam indivíduos específicos de uma determinada classe.

Os pacientes, devidamente registrados clinicamente, formam as instâncias que são submetidas ao processo de inferência.

A figura 34 apresenta a relação dos pacientes hipotéticos e destaca a paciente Ana que tem os sintomas Face inchada, Língua saburra pegajosa, Pele oleosa e Pulso deslizante. Apresenta também as seguintes características: Etnia - negra, Idade - adulta, Índice de Massa Corpórea (IMC) – obesa, Sexo – feminino, Ocupação – insalubre,

Alimentação – vegetariana, Atividade Física – sedentária e a Emoção – medo.

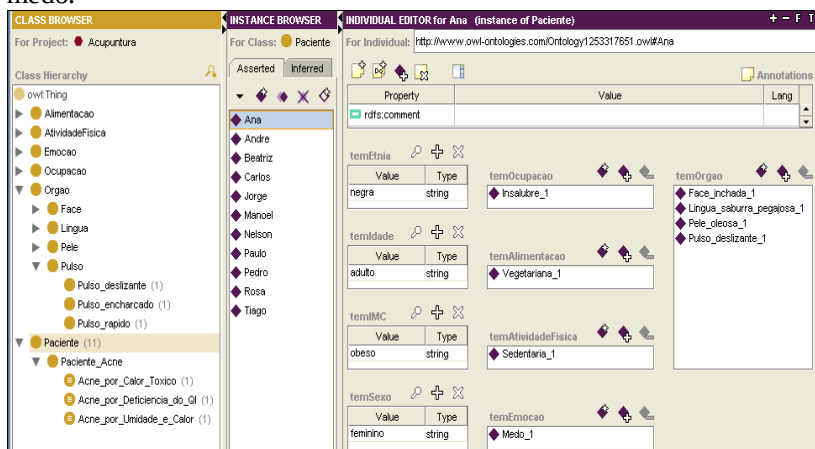


Figura 34- Exemplo de instância

• Inferências

A inferência ou o processo de raciocínio procura classificar o paciente dentro de uma determinada patologia.

O resultado do processo de inferência é apresentado na figura 35. Os pacientes André, Jorge, Manoel, Nelson e Tiago foram diagnosticados com a patologia ACT. A patologia ADQ está presente nos pacientes Beatriz, Carlos e Rosa. Foram diagnosticados com a patologia AUC os pacientes Ana e Pedro.



Figura 35- Inferência no *Protégé*

Dos onze pacientes submetidos ao diagnóstico, somente Paulo não foi diagnosticado. Pelos registros clínicos, apresentados na figura 36, Face vermelha, Língua saburra pegajosa, Pele seca e Pulso encharcado, não foi possível caracterizar este paciente em uma das patologias.

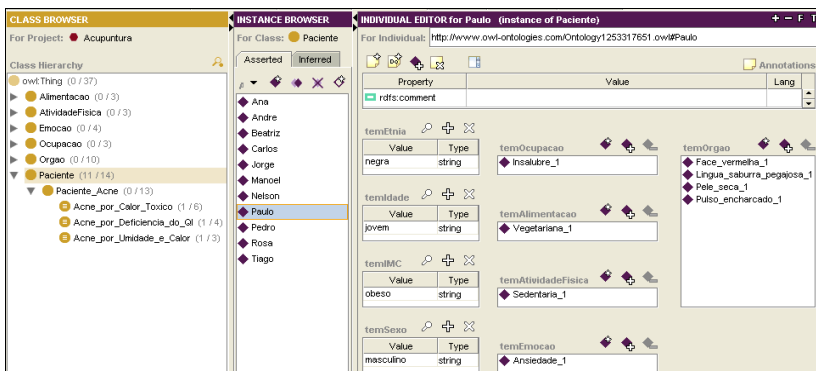


Figura 36- Registro de dados

Conclui-se, portanto, que os dados informados ao sistema, referentes ao paciente Paulo, são insuficientes ou incertos. Diante desta ocorrência, é elaborado para o paciente um diagnóstico probabilístico que tem como suporte a base de dados gerada pelo sistema.

4.3.3 Construção da Base de Dados

Através dos registros clínicos dos pacientes que foram diagnosticados e obtiveram sucesso em seus tratamentos, organizou-se a base de dados apresentada no quadro 12, que serve como fonte para o cálculo probabilístico.

Diag_F	D_sin	D_evi	etnia	idade	sex	imc	alim	ocup	emoc	ativ	face	lingua	pele	pulso
ACT	act	act	bran	adul	mas	magr	bal	norm	raiva	seden	verm	amar	oleo	desl
ACT	act	act	amar	jov	mas	obes	veg	insal	raiva	semia	verm	amar	oleo	desl
ACT	act	act	negr	jov	mas	sobr	carn	insal	ansie	seden	verm	amar	seca	rap
ACT	act	act	negr	idos	mas	obes	carn	estres	raiva	seden	verm	pegaj	seca	rap
ACT	act	act	bran	adul	mas	magr	carn	norm	raiva	ativo	verm	amar	seca	desl
ADQ	adq	adq	negr	adul	fem	magr	carn	insal	ansie	semia	inch	palid	oleo	ench
ADQ	adq	adq	amar	jov	mas	obes	veg	norm	ansie	ativo	verm	pegaj	oleo	ench
ADQ	adq	adq	negr	jov	fem	obes	bal	estres	ansie	seden	inch	pegaj	oleo	ench
AUC	auc	auc	negr	adul	fem	obes	veg	insal	medo	seden	inch	pegaj	oleo	desl
AUC	auc	auc	bran	adul	mas	magr	bal	estres	medo	seden	inch	amar	oleo	rap

Quadro 12 - Base de Dados

Fonte: Dados fictícios

Neste quadro estão registrados os sinais e sintomas (face, língua, pele e pulso), as evidências (sexo, idade, emoção, ocupação, trabalho, IMC, alimentação e atividade física) e o resultado do Diagnóstico Final (Diag_F), Diagnóstico dos Sintomas (D_sin) e Diagnóstico das Evidências (D_evi) de cada paciente tratado.

4.3.4 Diagnóstico Probabilístico

Para realizar a análise probabilística, foi construído um protótipo simplificado no *software* Netica (NORSYS NETICA, 2008), utilizando redes Bayesianas, que são representações que mapeiam conhecimento a partir de um modelo probabilístico, expressando relações de causas e efeitos, através da associação de probabilidade em um determinado domínio do conhecimento (RUSSEL e NORVIG, 1995).

A figura 37 apresenta a rede Bayesiana construída a partir da base de dados. Os sinais e sintomas (Pele, Face, Língua e Pulso) possibilitam a formação do resultado referente ao diagnóstico D_Sin. As evidências (sexo, idade, emoção, ocupação, trabalho, IMC, alimentação e atividade física) permitem a formação do resultado referente ao diagnóstico D_Evi.

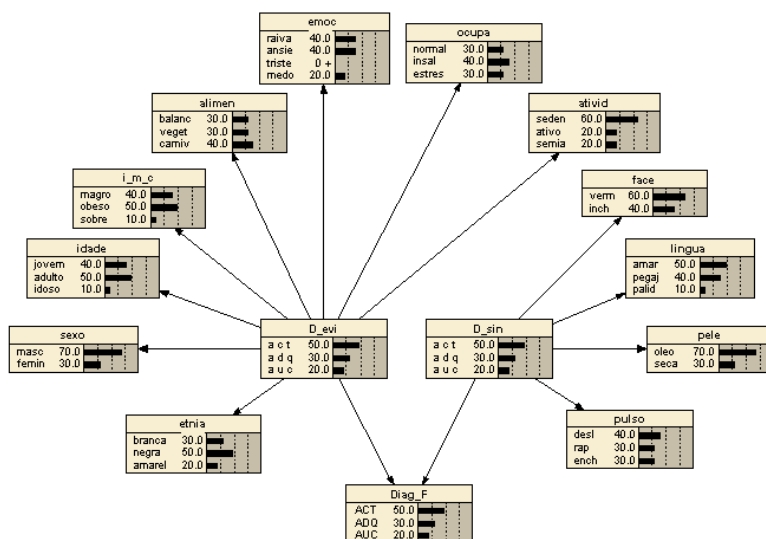


Figura 37- Rede Bayesiana

O diagnóstico final é uma composição entre o D_Sin e D_Evi, que leva em consideração o grau de significância de cada um para a formação do resultado final.

A tabela 3 apresenta as porcentagens atribuídas a cada uma das partes acima citadas. Ao diagnóstico proveniente dos sintomas e ao teste de laboratório foi, aleatoriamente, atribuído um peso de 95%, enquanto para o diagnóstico proveniente das evidências, também aleatoriamente, foi aplicado um peso de 5%.

Tabela 3 - Porcentagens atribuídas em relação às evidências e aos sintomas

D_evi	D_sin	ACT	ADQ	AUC
a c t	a c t	100.000	0.000	0.000
a c t	a d q	5.000	95.000	0.000
a c t	a u c	5.000	0.000	95.000
a d q	a c t	95.000	5.000	0.000
a d q	a d q	0.000	100.000	0.000
a d q	a u c	0.000	5.000	95.000
a u c	a c t	95.000	0.000	5.000
a u c	a d q	0.000	95.000	5.000
a u c	a u c	0.000	0.000	100.000

Depois de concluída a estruturação da rede Bayesiana, foi possível inserir os dados referentes ao paciente Paulo. A figura 38 exibe a rede instanciada.

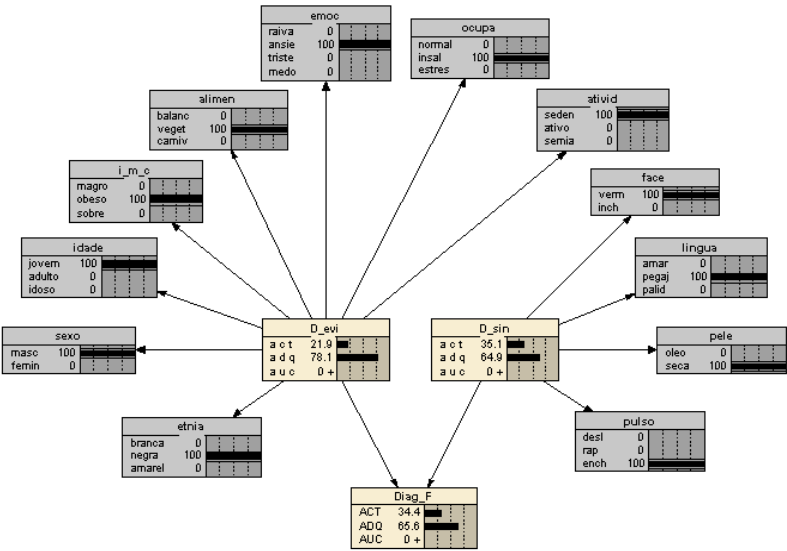


Figura 38- Rede Bayesiana instanciada para o reclamante

De acordo com os resultados obtidos, tem-se o diagnóstico probabilístico para o paciente Paulo que atribui um valor de 65,6% para a patologia ADQ e 34,4% para a patologia ACT.

Portanto, com 65,6% de probabilidade de acerto, a patologia ADQ é a melhor indicação de tratamento para o paciente.

4.3.5 Novo Conhecimento

A base de dados armazena o registro clínico dos pacientes tratados. Muitas informações relevantes podem surgir com o auxílio do cálculo probabilístico na manipulação destes dados. Uma evidência realçada pode tornar-se um novo conhecimento quando confrontada com o arquivo literário através do sistema de descoberta baseada na literatura.

Por exemplo, na figura 39 instancia-se a patologia “ADQ”. Em destaque, observa-se também que a evidência “ansiedade” (ansie) aparece em 100% dos pacientes com esta patologia. É uma informação que, ao ser confrontada com a literatura científica, pode originar um novo conhecimento para esta enfermidade.

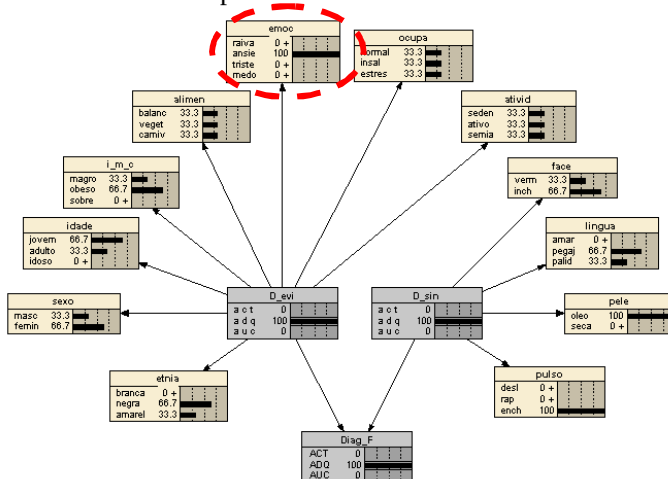


Figura 39- Rede Bayesiana instanciada para uma solução

Diversas ferramentas de pesquisa apoiam este processo de descoberta, entre elas o *software* Bitola (HRISTOVSKI *et al.*, 2001) que é um sistema interativo e funciona como suporte de descoberta baseado na literatura. A proposta do sistema é ajudar os pesquisadores a

fazer novas descobertas através das predicções semânticas e regras de associação.

A figura 40 exemplifica uma consulta feita no *software* Bitola, procurando relacionar a patologia “Acne” (conceito X) com tipo semântico “doença”, e a evidência “Ansiedade” (conceito Z) com tipo semântico “processo mental”. O elemento intermediário (conceito Y), que faz a ligação entre os conceitos X e Z, tem como tipo semântico “função fisiológica”.

Estas informações, inseridas no sistema, possibilitam que se pesquise, em arquivo literário, todos os artigos que envolvem estes conceitos.

BITOLA - Biomedical Discovery Support System: Closed Discovery

(Program authors: Dimatar Hristovski, Borut Peterlin)

Find Starting Concept X

Concept: Acne Vulgaris

Semantic Types: Disease or Syndrome/

Find End Concept Z

Concept: Anxiety

Semantic Types: Mental Process/

Chr.Loc.:

CUI: C0001144

Chr.Loc.:

CUI: C0003467

The starting concept X and the end concept Z co-occur in 15 Medline documents (FreqXZ=15).

Limit Ys

Contains:

Find Intermediate Ys

Semantic Group: Physiologic

Semantic Type: Physiologic Function

ConfXY*ConfYZ

Order by (Ys)

Descending

FreqXY >=0

FreqYZ >=0

ConfXY >=0

ConfYZ >=0

Page size: 200

Figura 40- Consulta ao *software* BITOLA

O primeiro resultado fornecido pelo sistema é exibido na tabela 4 que apresenta o Fator Suporte (frequência de co-ocorrência) da associação, dado pelas colunas “FreqXY” e “FreqZY”. Também é apresentado o Fator de Confiança (percentagem de registros) da associação, dado pelas colunas “ConfXY” e “ConfYZ”.

A última coluna contendo o produto entre “ConfXY” e “ConfYZ” estabelece a força da associação entre os conceitos X, Y e Z pesquisados.

Tabela 4 - Relação de Fator Suporte e Confiança

Intermediate Concepts Y: (first 10 of 10)							
Concept Name	Semantic Type	FreqXY	ConfXY(%)	FreqYZ	ConfYZ(%)	FreqXY*FreqYZ	ConfXY*ConfYZ
Menopause	Physiologic Function	23	0.336	110	0.621	2530	0.208
Physiologic pulse	Physiologic Function	1	0.015	166	1.119	166	0.016
Premenopause	Physiologic Function	1	0.015	16	0.484	16	0.007
Immunity	Physiologic Function	8	0.117	2	0.054	72	0.006
Thirst	Physiologic Function	1	0.015	2	0.358	9	0.005
Pigmentation physiologic function	Physiologic Function	17	0.248	1	0.018	17	0.004
Electrophysiology (science)	Physiologic Function	1	0.015	71	0.108	71	0.002
Energy Metabolism	Physiologic Function	1	0.015	26	0.063	26	0.001
Vascular resistance	Physiologic Function	1	0.015	17	0.054	17	0.001
Biological Transport, Active	Physiologic Function	1	0.015	1	0.004	1	0.000

Na sequência, conforme interesse numa determinada função fisiológica, obtêm-se os artigos relacionados aos conceitos previamente estabelecidos para análise e geração de possíveis novos conhecimentos, (figura 41).

☐	Clinical implications of lipid peroxidation in acne vulgaris: old wine in new bottles.
1.	Bowe WP, Logan AC. Lipids Health Dis. 2010 Dec 9;141. PMID: 21143923 [PubMed - indexed for MEDLINE] Free PMC Article Free full text Related citations
☐	Psychological impact of isotretinoin treatment in patients with moderate and severe acne.
2.	Simić D, Situm M, Letica E, Penavić JZ, Zivković MV, Tomić T. Coli Antropol. 2009 Dec;33 Suppl 2:15-9. PMID: 20120397 [PubMed - indexed for MEDLINE] Related citations
☐	Comparison of depression, anxiety and life quality in acne vulgaris patients who were treated with either isotretinoin or topical agents.
3.	Kaymak Y, Taner E, Taner Y. Int J Dermatol. 2009 Jan;48(1):41-6. PMID: 19126049 [PubMed - indexed for MEDLINE] Related citations

Figura 41- Relação de artigos com as associações

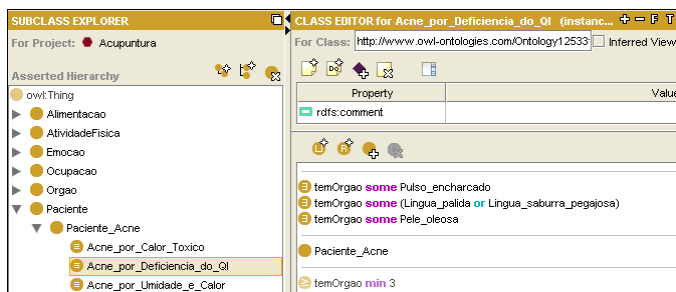
Após análise manual dos artigos envolvidos e diante da confirmação de um novo conhecimento, a atualização na base de conhecimento do domínio específico, apresentada na figura 42, pode ser executada.

Esta atualização envolve a inserção da nova descoberta entre as restrições exigidas para compor as manifestações de determinada patologia. Por exemplo, a patologia “ADQ” tinha como restrições as seguintes manifestações:

- Pulso encharcado;
- Língua pálida ou Língua saburra pegajosa; e
- Pele oleosa.

Após a confirmação do novo conhecimento, as restrições passam a contar com novo integrante, ou seja, a manifestação “ansiedade”:

- Pulso encharcado;
- Língua pálida ou Língua saburra pegajosa;
- Pele oleosa; e
- Ansiedade.



antes

depois

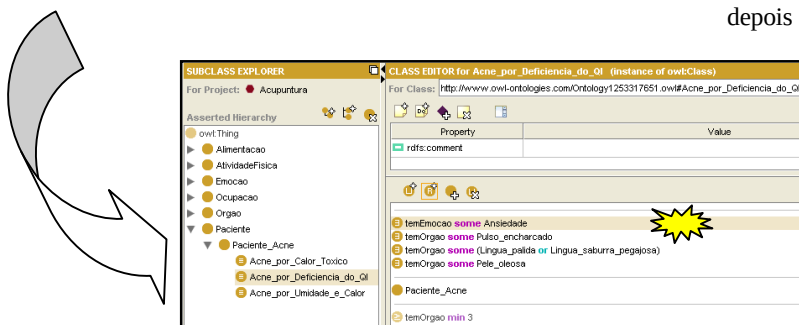


Figura 42- Atualização na base de conhecimento

Atualmente, já existem estudos sobre o uso de técnicas envolvendo raciocínio incremental (PARSIA, HALASCHEK-WIENER e SIRIN, 2006; GRAU, B. C.; HORROCKS, I.; KAZAKOV, Y. e SATTLE, U., 2007) que permitem realizar atualizações automaticamente na ontologia.

O próximo capítulo diz respeito aos resultados obtidos através da metodologia da pesquisa utilizada para verificação da consistência do modelo proposto e, em seguida, uma discussão é levantada, em que se questionam e se discutem as principais contribuições oferecidas pelo modelo. Terminando este capítulo, uma análise entre o modelo proposto e os modelos correlatos é feita. Procura-se, nesta análise, levantar as principais características do modelo proposto em relação aos demais.

5 ANÁLISE E DISCUSSÃO DOS RESULTADOS

5.1 ANÁLISE DOS RESULTADOS

Nesta seção, apresentam-se os resultados da verificação da consistência do modelo, através dos conceitos, métodos e técnicas utilizados para promover o aprimoramento no processo de diagnóstico. Para tal propósito, conforme citado nos procedimentos metodológicos deste trabalho, foram aplicados questionários (Apêndice I), baseados no método Delphi, sendo estes respondidos por quatorze especialistas.

Inicialmente, foram feitas cinco perguntas objetivas para verificar o encadeamento e a coerência dos profissionais ao responder mais de uma pergunta sobre o mesmo assunto. Já na segunda parte do questionário, as perguntas formuladas permitiam que os especialistas manifestassem a sua concordância ou discordância bem como as sugestões, justificativas e comentários sobre os assuntos abordados, os quais se encontram no Apêndice II.

ETAPA 1 – PERGUNTAS PRELIMINARES

Na primeira pergunta, procurou-se identificar se o especialista já teve contato ou experiência com algum tipo de sistema que o auxiliasse na realização do diagnóstico.

Pelos resultados apresentados na tabela 5, constata-se que 78,57% dos especialistas já tiveram contato com algum tipo de sistema que os auxiliou no processo de diagnóstico. Este resultado demonstra que a amostra de participantes tem noção básica sobre o assunto que será tratado.

Tabela 5 - Contato com algum sistema que auxilia no diagnóstico

Opção	Sim	Não	Não optou
Percentual de respostas	78,57%	21,43%	0%

Na segunda pergunta, procurou-se identificar quais receios/medos o especialista tem ao utilizar um sistema de apoio.

Pelos resultados apresentados na tabela 6, observa-se que 71,43% dos especialistas apontam a confiabilidade suplantando receios/medos na utilização de um sistema de apoio. Este resultado demonstra a necessidade de promover desenvolvimento dos sistemas para

proporcionar uma maior confiabilidade por parte dos especialistas na sua utilização bem como investir em adequado *design* de interface, melhorando a usabilidade, navegação, entre outros.

Tabela 6 - Receios/medos em utilizar um sistema de apoio

Opção	Confiabilidade do sistema	Pouco conhecimento em informática	Interface pouco amigável	Não optou
Percentual de respostas	71,43%	21,43%	7,14%	0%

Na terceira pergunta, procurou-se identificar se o especialista utilizaria um Sistema Baseado em Conhecimento como auxílio à sua atividade.

Pelos resultados apresentados na tabela 7, constata-se que 100% dos especialistas utilizariam um Sistema Baseado em Conhecimento, para assistência em suas atividades. Com esta implicação é importante notar que, mesmo com a desconfiança nos sistemas atuais, vislumbra-se a forte tendência (intenção) de utilizá-los como auxílio no processo de diagnóstico.

Tabela 7 - Auxílio de um Sistema Baseado em Conhecimento

Opção	Sim	Não	Não optou
Percentual de respostas	100%	0%	0%

Na quarta pergunta, procurou-se identificar se os especialistas julgam necessário considerar diferentes variáveis que compõem o contexto, não somente o agente envolvido, para explorar relações entre as informações coletadas e assim concluir o diagnóstico.

Pelos resultados apresentados na tabela 8, verifica-se que 100% dos especialistas julgam necessário considerar diferentes variáveis na composição do contexto, no qual o agente está incluído, para concluir o diagnóstico. Conclui-se, portanto, que a visão sistêmica é de fundamental importância na coleta de variáveis que compõem o cenário dentro do processo de diagnóstico.

Tabela 8 - Necessidade de considerar diferentes variáveis para o contexto

Opção	Sim	Não	Não optou
Percentual de respostas	100%	0%	0%

Na quinta pergunta, procurou-se identificar se os especialistas julgam importante registrar os dados de diagnósticos realizados para direcionar novos diagnósticos embasando-se em casos anteriores.

Pelos resultados apresentados na tabela 9, constata-se que 92,86% dos especialistas acreditam na importância de registrar dados e informações de diagnósticos realizados para servirem como fonte para novos diagnósticos. Este resultado demonstra que a formação de uma base de dados, proveniente de diagnósticos realizados, será de grande valia para novos diagnósticos visando gerar informações relevantes para o aprimoramento do processo.

Tabela 9 - Necessidade de considerar diagnósticos realizados

Opção	Sim	Não	Não optou
Percentual de respostas	92,86%	0%	7,14%

Dando sequência à análise dos resultados obtidos, através dos questionários, apresentam-se as estatísticas do percentual de respostas e opiniões dos especialistas quanto aos conceitos, métodos e técnicas que compõem a elaboração e integração do modelo proposto.

ETAPA 2 – SISTEMA BASEADO EM CONHECIMENTO

Nesta etapa, procurou-se identificar se o especialista acredita no potencial destes sistemas como apoio e orientação ao processo de diagnóstico.

Com os resultados apresentados na tabela 10, verificou-se que 85,71% dos especialistas acreditam na potencialidade dos SBC. Nas respostas dos especialistas que também concordam, mas que manifestam alguma restrição, há o percentual de 14,29%.

Tabela 10 - Crença na potencialidade dos SBC's

Opção	Concordo	Tenho Restrições	Discordo	Não optou
Percentual de respostas	85,71%	14,29%	0%	0%

Aqui tem-se o consenso no potencial dos sistemas como suporte e apoio à decisão; nota-se que alguns profissionais têm receio, por temer que os usuários fiquem presos somente às alternativas geradas pelo conhecimento representado no sistema. Portanto, há necessidade de orientar as pessoas quanto ao papel do sistema de apenas auxiliar e nunca substituir o especialista a quem cabe a responsabilidade final do diagnóstico, o que reitera a nomenclatura dos modelos como sistemas de apoio à decisão.

ETAPA 3 – REPRESENTAÇÃO DO CONHECIMENTO

Na segunda questão, procurou-se identificar se o especialista acredita na possibilidade de representar o conhecimento específico do domínio de forma que seja analisado por um sistema.

Pelos resultados apresentados na tabela 11, constata-se que 64,29% dos especialistas acreditam na possibilidade de representar o conhecimento de forma que possa ser interpretado por um sistema. Nas respostas dos especialistas que também concordam, mas que manifestam alguma restrição, há o percentual de 35,71%.

Tabela 11 - Possibilidade de representar o conhecimento

Opção	Concordo	Tenho Restrições	Discordo	Não optou
Percentual de respostas	64,29%	35,71%	0%	0%

Pelas opiniões apresentadas tem-se o consenso de que grande parte do conhecimento adquirido por um especialista pode ser representado de maneira sistemática e formal. Na medida em que o processo de aquisição do conhecimento tenha completude e consistência suficientes para cobrir determinado domínio, o conhecimento adquirido e formalmente representado será capaz de modelar o domínio integralmente, podendo ser analisado por um sistema. Salienta-se que, dentro das opiniões dadas, a intuição e a vivência profissional são

fatores primordiais no processo de diagnóstico e que dificilmente serão representadas.

ETAPA 4 – VISÃO SISTÊMICA

Na questão três, procurou-se identificar se o especialista acredita na possibilidade da inquirição ir além das manifestações apresentadas e relatadas pelo reclamante. Fatores externos que compõem o ambiente, no qual o reclamante está inserido, contribuem para tornar o processo de diagnóstico mais real e compreensivo.

Os resultados da tabela 12, mostram que 92,86% dos especialistas acreditam que a investigação sobre uma queixa deve ir além das manifestações apresentadas pelo reclamante. Nas respostas dos especialistas que também concordam, mas que manifestam alguma restrição, há o percentual de 7,14%.

Tabela 12 - Visão sistêmica

Opção	Concordo	Tenho Restrições	Discordo	Não optou
Percentual de respostas	92,86%	7,14%	0%	0%

Observando a tabela, vê-se que o processo de diagnóstico é muito complexo, envolvendo muitas variáveis a serem consideradas, e não apenas o que o reclamante relata ou apresenta em um único momento, avaliação ou consulta. Portanto, é imprescindível uma visão sistêmica envolvendo fatores externos, isto é, fatores que de alguma maneira possam estar relacionados com a queixa, para compor o processo de diagnóstico. Estes elementos, que a princípio só fazem parte de um contexto, levam o reclamante a ser inserido dentro de uma realidade e possibilitam, quando analisados em conjunto, extrair informações que podem auxiliar na tomada de decisão.

ETAPA 5 – CONHECIMENTO EM ESTADO LATENTE

Na quarta questão, procurou-se identificar se o especialista reforça o fato de que muito conhecimento, proveniente de dados e informações coletados no interrogatório, continua em estado latente.

Pelo apresentado na tabela 13, constata-se que 85,71% dos especialistas acreditam que muitos dados e informações colhidos na inquirição permanecem em estado latente. Nas respostas dos

especialistas que também concordam, mas que manifestam alguma restrição, há o percentual de 14,29%.

Tabela 13 - Conhecimento em estado latente

Opção	Concordo	Tenho Restrições	Discordo	Não optou
Percentual de respostas	85,71%	14,29%	0%	0%

Aqui, o consenso é de que, realmente, o armazenamento de informações pode proporcionar o levantamento de relacionamentos que pareçam ocultos aos olhos do especialista em um primeiro momento. Assim, será papel atribuído ao sistema, correlacionar informações e dados embutidos em uma ampla variedade de casos ou diagnósticos que muitas vezes se perdem no próprio histórico da evolução do tratamento.

ETAPA 6 – BASE DE DADOS

Neste item, procurou-se identificar se o especialista acredita que, com a formação de uma base de dados, a partir de diagnósticos realizados, será possível extrair informações relevantes que gerem novos conhecimentos.

Com os resultados da tabela 14, constata-se que 92,86% dos especialistas acreditam. Nas respostas dos especialistas que também concordam, mas que manifestam alguma restrição, há o percentual de 7,14%.

Tabela 14 - Formação de uma base de dados

Opção	Concordo	Tenho Restrições	Discordo	Não optou
Percentual de respostas	92,86%	7,14%	0%	0%

Com estes dados, verifica-se que a prática baseada em evidências científicas só será possível se houver um registro sistematizado e rigoroso das ações, para possibilitar compilações posteriores. Hoje, com as metodologias de KD (descoberta de conhecimento), torna-se viável o armazenamento e a utilização de dados e informações para descobrir novas relações sistêmicas e associações gerando evidências que, após analisadas, poderão tornar-se novos conhecimentos.

ETAPA 7 – DESCOBERTA BASEADA NA LITERATURA

Nesta última questão, procurou-se identificar se o especialista acredita que, fazendo uso de sistemas de descoberta baseada na literatura, é possível obter indicações que revelem a existência de correlações entre elementos até então encobertos e que sejam úteis no aprimoramento de um domínio específico.

Pelos resultados apresentados na tabela 15, constata-se que 92,86% dos especialistas acreditam. Nas respostas dos especialistas que também concordam, mas que manifestam alguma restrição, há o percentual de 7,14%.

Tabela 15 - Uso de Sistemas de Descoberta Baseado na Literatura

Opção	Concordo	Tenho Restrições	Discordo	Não optou
Percentual de respostas	92,86%	7,14%	0%	0%

Os percentuais, aqui, mostram que a descoberta em literatura pode ser parte importante de um sistema de apoio à decisão por chamar a atenção do especialista para informações que ele pode não estar considerando no momento de realizar o diagnóstico de um determinado reclamante. A atualização de conhecimentos na literatura é tão rápida que muitas vezes impossibilita o profissional de acompanhar a evolução de novas descobertas dentro de sua especialidade.

Ainda, dando sequência à sexta etapa, que trata sobre os sistemas de descoberta baseado na literatura, formulou-se a sétima questão, em que se procurou identificar se o especialista acredita num sistema que possibilite rastrear relações (associações) entre elementos, dentro de trabalhos em diferentes especialidades e que lhe dará maior facilidade em descobrir novos conhecimentos.

Pelos resultados obtidos na tabela 16, constata-se que 78,57% dos especialistas acreditam. Nas respostas dos especialistas que também concordam, mas que manifestam alguma restrição, há o percentual de 21,43%.

Tabela 16 - Conhecimento comum em diferentes especialidades

Opção	Concordo	Tenho Restrições	Discordo	Não optou
Percentual de respostas	78,57%	21,43%	0%	0%

Pelas opiniões apresentadas tem-se o consenso de que o diagnóstico de um reclamante por diferentes especialidades pode sim ter suas características em comum descobertas por um sistema de modo que este auxilie na determinação de fatores globais e relevantes para cada caso. Cabe, aqui, salientar que, ao permitir uma visão sistêmica e holística sobre determinado tema, pode-se chegar à compreensão, ao entendimento do mesmo e às possíveis ações sobre ele, de maneira eficaz. Hoje, para efeito de solução, não há como desconsiderar os vários modos de interpretação sobre o fenômeno que está sendo analisado.

5.2 DISCUSSÃO DOS RESULTADOS

Este trabalho teve como objetivo principal o desenvolvimento de um modelo de Engenharia do Conhecimento capaz de contribuir para a sistematização e aprimoramento do processo de diagnóstico, que envolve incertezas e subjetividades. Assim, questionam-se e discutem-se quais contribuições de fato foram estas.

Os modelos baseados em ontologias são muito importantes para representação sistemática e formal do conhecimento, com completude e consistência suficientes para cobrir determinado domínio (CASTILHO-WEINERT, 2010).

Este conhecimento representado, quando aliado a outras técnicas, como o cálculo probabilístico e a descoberta baseada em literatura, pode ajudar o profissional no processo de diagnóstico, apresentando melhores opções de conclusão diagnóstica e orientando nas perguntas que podem ser feitas durante o processo.

Além disto, cabe salientar a importância do modelo para a descoberta de novos conhecimentos, pois os profissionais, muitas vezes, ficam restritos a determinadas estratégias de raciocínio e não conseguem visualizar novas alternativas para a solução de um caso.

Desta forma, a existência de um modelo sistêmico que vislumbre novas estratégias de ação para se chegar a um diagnóstico e que gere

novos conhecimentos e o presente ao profissional, é de grande importância, inclusive para capacitação e atualização do mesmo. Ainda, ao descobrir novos conhecimentos, o modelo proposto pode melhorar e favorecer a habilidade de diagnóstico de um profissional que não seja tão especialista ou experiente no domínio em questão.

O uso de associação e o estabelecimento de correlações, entre um novo diagnóstico e aqueles já conhecidos, é semelhante à forma de raciocínio de muitos especialistas, comparando algo novo com algo aprendido ou com casos antigos. Assim, a descoberta baseada na literatura fornece a capacidade de se comparar um caso novo até mesmo com casos ou informações que não se teve a oportunidade de vivenciar em sua rotina, aumentando o conhecimento e melhorando a capacidade de decisão do profissional.

Outra importante contribuição do modelo está na possibilidade de sistematização e registro das atividades de diagnóstico de um profissional. A rotina faz com que se mantenham poucos registros da atividade cotidiana; isto corrobora com a necessidade de se extrair e arquivar dados sobre diagnóstico de forma sistemática e acessível, para contribuir com novos diagnósticos e gerar novos conhecimentos. Desta forma, possibilitará não somente o “reuso” das informações, mas também o seu compartilhamento.

Quanto à validação do modelo, foi aplicado o método Delphi que possibilita, através de debate entre especialistas em um determinado domínio do conhecimento, chegar a um consenso de opiniões envolvendo a metodologia, métodos e técnicas empregados na elaboração do modelo. Salienta-se que a emergência de consenso entre os profissionais foi possibilitada devido ao forte embasamento na literatura das estratégias utilizadas para a elaboração do modelo.

O estudo de caso apresentado, dentro da área da acupuntura, trouxe a oportunidade de trabalhar com um assunto que desperta cada vez mais interesse pelas chamadas práticas complementares ou alternativas nos serviços de saúde no Brasil. No caso específico deste trabalho, a enfermidade Acne foi utilizada para demonstrar, de forma didática, a construção de um domínio do conhecimento e, através do modelo proposto, a possibilidade de gerar novos conhecimentos.

5.3 ANÁLISE ENTRE MODELO PROPOSTO E CORRELATOS

A análise entre o modelo proposto e os trabalhos correlatos, já apresentados na seção 2.1.2 (SOMKS, FENZA, SWE e KHAM e ODDIN), neste item envolve os seguintes aspectos:

- Visão sistêmica;
- Manipulação das incertezas;
- Resultado final do diagnóstico; e
- Análise de evidências.

5.3.1 Visão Sistêmica

Entende-se, neste trabalho, por visão sistêmica a capacidade do especialista em observar detalhes e situações, num determinado domínio específico, que podem gerar novas informações úteis com a função de contribuir para o aperfeiçoamento deste domínio.

Esta visão sistêmica, para o processo de diagnóstico, além do modelo proposto, não é característica dos outros modelos. A principal deste é proporcionar, na elaboração da ontologia, condições para gerar novos conhecimentos provenientes das observações e experiências dos profissionais que atuam nesta atividade.

5.3.2 Manipulação de Incertezas

Entre os trabalhos apresentados, as formas de manipular as incertezas diferem em relação aos métodos e técnicas adotados. O modelo proposto, bem como os modelos SOMKS e ODDIN, segue o processo probabilístico, ou seja, cada variável está sujeita a um determinado “peso” para compor o resultado final. Este peso pode ser dado pelo profissional como no modelo proposto e no modelo ODDIN, ou deixar que o próprio sistema se encarregue de atribuir a importância à variável como no modelo SOMKS.

O modelo FENZA utiliza para esta manipulação a técnica da lógica *fuzzy*, o que permite dar às variáveis maior precisão para a elaboração do diagnóstico. Já o modelo SWE E KHAM procura adaptar as variáveis dentro de casos ocorridos procurando a similaridade entre elas.

5.3.3 Resultado Final do Diagnóstico

De acordo com as técnicas adotadas para a manipulação das incertezas, tem-se a forma de apresentar o resultado final para o diagnóstico. Os modelos que utilizam técnicas probabilísticas oferecem resultados com o grau de acerto para cada solução a ser utilizada, enquanto os modelos FENZA e SWE E KHAM só permitem um único resultado.

Salienta-se a importância dos modelos probabilísticos para quem trabalha com a tarefa de diagnosticar, pois permite atribuir ao problema mais de uma solução com probabilidades de acerto.

5.3.4 Análise de Evidências

Um dado ou informação relevante pode tornar-se um novo conhecimento após a análise literária, envolvendo a participação de especialistas no domínio que buscam comprovação científica para a nova descoberta.

Dos trabalhos citados, ‘apenas’ o modelo proposto apresenta técnicas para realçar e confrontar com arquivos literários uma informação relevante de forma eficaz, pois o cálculo probabilístico e o sistema de descoberta baseado na literatura proporcionam meios de simplificar este trabalho de análise.

De modo geral, pode-se verificar que os cinco modelos apresentados têm na ontologia a forma de representação do conhecimento de um domínio específico e procuram através de métodos e técnicas equacionar as incertezas que fazem parte do processo de diagnóstico.

Cada um desses modelos apresenta uma característica peculiar, que o diferencia dos demais na tentativa de explorar as deficiências até então fazendo parte desta atividade. No modelo proposto, esta característica é apresentada, inicialmente, na construção da ontologia ao sugerir, de forma sistêmica, a inclusão de dados e informações úteis que possam levar a novas descobertas.

O modelo também apresenta, como característica própria, o processo de evidenciar um novo conhecimento ao propor um sistema de descoberta baseado em literatura que explora os conceitos de predicação semântica e regras de associação para facilitar a pesquisa literária.

Portanto, tem-se um modelo que se diferencia dos demais na proposição e análise para a geração de um novo conhecimento com a finalidade de aprimorar o processo de diagnóstico.

6 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS

6.1 CONCLUSÕES

O trabalho apresenta um modelo de Engenharia do Conhecimento baseado em ontologia e no cálculo probabilístico que possibilita a criação de novos conhecimentos para apoio à tarefa de diagnóstico, ajudando a decisão humana.

Com a análise da problemática, envolvendo o processo de diagnóstico, observam-se oportunidades de aprimorar os Sistemas Baseados em Conhecimento, recentemente desenvolvidos, que procuram auxiliar esta tarefa intensiva em conhecimento.

Os trabalhos correlatos analisados, fundamentados em ontologias, mostraram eficiência na manipulação das incertezas envolvidas no processo de diagnóstico aplicando técnicas apropriadas. Entretanto, verificou-se a deficiência na geração de novos conhecimentos, a fim de atualizar a base de conhecimento utilizada para a realização desta atividade.

Assim, o problema enfrentado foi como trabalhar, de forma adequada, com as informações e dados obtidos de diagnósticos realizados, de forma que novos conhecimentos possam ser gerados e contribuam para o aprimoramento deste processo.

Para atingir o objetivo deste trabalho, metas foram traçadas para que o resultado final fosse alcançado e serão descritas a seguir.

Inicialmente, os modelos correlatos foram analisados no sentido de verificar as características (pontos fortes e pontos fracos) encontradas para a realização de diagnóstico. Observa-se, entre os modelos, a inexistência de um exemplar que apresente um método sistêmico semi-automático propondo pesquisa para a geração de novos conhecimentos.

Na busca por metodologia, métodos e técnicas, o segundo objetivo específico do trabalho foi identificar na literatura subsídios, dentro da Engenharia e Gestão do Conhecimento, necessários para elaborar um modelo que atenda esta necessidade.

A metodologia, métodos e técnicas encontrados e utilizados foram: CommonKADS, Ontologias, Cálculo Probabilístico, modelagem CISM e Sistemas de Descoberta baseados na Literatura.

A metodologia CommonKADS forneceu todas as etapas para a construção de um modelo de conhecimento. Salienta-se que esta metodologia demonstrou ser adequada na medida em que apresenta uma

visão de contexto sistêmico sobre a organização e o problema a ser tratado. Isto a torna menos focada na construção de um artefato, mas sim, na construção de um modelo de conhecimento onde o ser humano é o ator principal.

Tem-se no uso de Ontologias, como forma de representar o conhecimento de um domínio específico, sob o olhar da modelagem CESM, e o Cálculo Probabilístico, no trato às incertezas deste domínio, uma estrutura adequada complementar que fornece diagnósticos probabilísticos, bem como evidencia informações relevantes para a geração de possíveis novos conhecimentos.

Os Sistemas de Descoberta Baseados na Literatura trazem como inovação as predicações semânticas e as regras de associação. Por intermédio destas técnicas, artigos científicos, envolvendo os elementos de pesquisa, são apresentados para análise. A utilização de tais sistemas tem a importância de levar ao profissional um material útil, proveniente de diversas especialidades, e focado na pesquisa em questão, evitando leituras desnecessárias.

Atendendo ao terceiro e quarto objetivos específicos, através da integração entre esses métodos e técnicas, o modelo de Engenharia do Conhecimento foi idealizado, seguindo a metodologia CommonKADS, e aplicado em um estudo de caso. Optou-se, informalmente, pela área da saúde, em uma clínica que exercita a técnica da Acupuntura. O modelo de organização, tarefa e agentes envolvidos foram analisados, possibilitando estabelecer os elementos que definem a viabilidade do Modelo de Conhecimento a ser construído de forma a minimizar o impacto e o fracasso de implantação.

Um domínio específico dentro da Acupuntura, mais especificamente a enfermidade Acne foi representada sistemática e formalmente, com o auxílio de especialista, livros e manuais, dentro de uma visão sistêmica, o qual possibilitou formar uma Base de Conhecimento. Utilizando esta Base de Conhecimento, pacientes hipotéticos, apresentando sinais, sintomas e evidências foram diagnosticados.

Os resultados dos diagnósticos realizados, bem como os atributos / características dos pacientes envolvidos permitiu a construção de uma Base de Dados, sendo esta responsável por “alimentar” a Rede Causal que tem como suporte conceitual o Cálculo Probabilístico. Esta Rede, quando instanciada, fornece diagnósticos probabilísticos e permite realçar informações relevantes candidatas a gerar novos conhecimentos.

Pelo Sistema de Descoberta Baseado na Literatura, o profissional é levado a analisar, de forma simples e objetiva, entre artigos científicos

de diferentes especialidades, se a informação relevante merece atenção da comunidade científica quanto à indicação de um novo conhecimento a ser inserido na Base de Conhecimento do domínio em questão.

Salienta-se ainda que este estudo embasa a suposição que o modelo de conhecimento desenvolvido nesta tese pode ser generalizado e aplicado em quaisquer outras áreas do conhecimento.

Como forma de verificar a contribuição desta tese no sentido de inovação, foi realizado o quinto objetivo específico que trata da análise entre o modelo proposto e os trabalhos correlatos. Esta comparação permitiu destacar as características principais de inovar em relação a obter e comprovar as evidências para gerar possíveis novos conhecimentos. Além da possibilidade de fornecer um diagnóstico probabilístico, o que é muito útil em diversas áreas do conhecimento que envolvem subjetividade, incerteza e falta de conhecimento, o modelo permite utilizar registros de diagnósticos realizados para fazer novas descobertas.

O último objetivo específico procurou validar o Modelo Proposto. Para esta validação foi utilizado o método Delphi que tem a característica de promover um debate, entre especialistas de um determinado domínio do conhecimento, sobre a metodologia, métodos e técnicas empregados na construção do modelo.

Questionários foram elaborados, com auxílio de profissionais, envolvendo os temas abordados e aplicados em um grupo de quatorze especialistas de forma individual, mantendo o anonimato entre eles. Para cada uma das etapas que compõem o modelo, os especialistas manifestaram sua concordância, restrição ou discordância, bem como, comentários, sugestões ou justificativas.

O consenso entre os especialistas foi atingido já na primeira rodada, o que demonstra o forte embasamento conceitual em todas as etapas que compõem o modelo. As opiniões apresentadas pelos profissionais permitiu formular reflexões sobre os assuntos abordados e possibilitou traçar temas para trabalhos futuros.

Assim, tem-se um modelo de Engenharia do Conhecimento que emprega metodologia, métodos e técnicas disponíveis e possibilita, através da integração entre eles, construir um Sistema de Conhecimento capaz de gerar e confirmar possíveis novos conhecimentos para apoio ao diagnóstico.

A contribuição deste trabalho para com o EGC está na visão de um mundo preocupado com o ser humano envolvendo a interdisciplinaridade das áreas do conhecimento, na Gestão e na

Engenharia do Conhecimento para construção de Sistemas de Conhecimento.

Concluindo, a sistemática adotada para realizar o trabalho atendeu às metas traçadas, tanto do ponto de vista do alcance de seus objetivos, quanto no aspecto da contribuição científica junto ao Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento.

6.2 SUGESTÕES PARA TRABALHOS FUTUROS

Como sugestões para trabalhos futuros de pesquisa, que possam ser embasados no modelo de Engenharia do Conhecimento proposto, indicam-se as seguintes oportunidades de aprimoramento:

- Transformação da base de dados categóricos para uma base de dados contínua, de forma que a precisão das informações também seja considerada (sistema híbrido difuso – probabilístico) ou ainda o uso da lógica paraconsistente no tratamento das incertezas envolvendo contradições;
- A base de dados também pode ser substituída por uma base de casos possibilitando através de correlações de semelhança a construção de um modelo mais robusto;
- Utilização dos dados e informações de diagnósticos realizados que não foram bem sucedidos; estes elementos podem e devem ser aproveitados para a geração de novos conhecimentos;
- Construção de uma rede social para compartilhar informações úteis, dentro de um domínio específico, envolvendo profissionais de diferentes especialidades e a comunidade em geral;
- Implementação deste modelo, com interface amigável, utilizando técnicas de raciocínio incremental (PARSIA, HALASCHEK-WIENER e SIRIN, 2006; GRAU; HORROCKS; KAZAKOV e SATTler, 2007) e extensões à linguagem OWL, tais como a OntoBayes (YANG e CALMET, 2005), PR-OWL (COSTA, LASKEY e LASKEY, 2005) que facilitem a integração automática das ferramentas necessárias para a realização do diagnóstico, bem como a pesquisa em novos conhecimentos;

- Fazer estudo de caso para outras áreas do conhecimento; por exemplo, área jurídica tendo como arquivo literário as jurisprudências existentes, entre outras.

REFERÊNCIAS

- ADAMS, L. A. Delphi forecasting: future issues in grievance arbitration. **Technological Forecasting and Social Change**. v. 18, p. 151-160. 1980.
- AHLERS, C. B. *et al.* Using the literature-based discovery paradigm to investigate drug mechanisms. In: Proceedings of AMIA Annual Symposium, p. 6-10, 2007.
- BUNGE, M. Mechanism and explanation. **Philosophy of the Social Sciences**, v. 27, n. 4, p. 410-465, 1997.
- BUNGE, M. Systemism: the alternative to individualism and holism. **Journal of Socio-Economics**, v. 29, p. 147-157, 2000.
- BUNGE, M. **Emergence and Convergence: Qualitative Novelty and the Unity of Knowledge**. University of Toronto Press, 2003. 330 p.
- BUNGE, M. How does it work? The search for explanatory mechanisms. **Philosophy of the Social Sciences**, v. 34, n. 2, p. 182-210, 2004.
- BUCHANAN, B. G.; DUDA, R. O. Principles of rule-based expert systems, **Advances in Computers, California**, v.22, p. 163-216, 1983.
- CAGNIN, C. H. **Fatores relevantes na implementação de um sistema de gestão ambiental com base na norma ISO 14001**. 2000. 200 f.. Dissertação (Mestrado em Engenharia de Produção) – Departamento de Engenharia de Produção e Sistemas, Universidade Federal de Santa Catarina (UFSC), Florianópolis, 2000.
- CALHOUN, M. A.; STARBUCK, W. H. Barriers do creating knowledge. In. EASTERBY-SMITH, M.; LYLES, M.. **Handbook of organizational learning and knowledge management**, p. 473-492, 2005.
- CARDOSO, L. R. A. *et al.* Prospecção de futuro e método delphi: uma aplicação para a cadeia produtiva da construção habitacional. Associação Nacional de Tecnologia do Ambiente Construído. **Ambiente Construído**. v. 5, n. 3, p. 63-78. Porto Alegre. Jul./Set. 2005.
- CARTER, J. H. **Clinical decision support systems**. E. S. Berner. New York: Springer-Verlag, p. 169-198, 1999.
- CASTILHO-WEINERT, L. V. **Ontologias e Técnicas de Inteligência Artificial Aplicadas ao Diagnóstico em Fisioterapia Neuropediátrica**, 2010. 251f.. Tese de Doutorado. Programa de Pós-

graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba, 2010.

CASTILHO-WEINERT, L. V.; LOPES, H. S. **An Ontology-Based System for Knowledge Management and Learning in Neuropediatric Physiotherapy**. In. E. Szczerbicki & N.T. Nguyen (Eds.): *Smart Information & Knowledge Management*, p. 283-307, 2010.

COIERA, E. **The Guide to Health Informatics**. 2. ed. London: Arnold, 2003.

COLLAZOS, K. L.; BRASIL, L. M.; AZEVEDO, F. M. Sistemas Especialistas na Saúde. In. **BRASIL, L. M. Informática em Saúde**. Brasília: Editora Universa - UCB, p. 153-205, 2008.

COSTA, P.C.G.; LASKEY, K.B.; CHANG, K. **PROGNOS: Applying Probabilistic Ontologies to Distributed Predictive Situation Assessment in Naval Operations**. Proceedings of the Fourteenth International Command and Control Research and Technology Conference (ICCRTS 2009). Washington, D.C., USA, p. 15-17, 2009.

COSTA, P.C.G.; LASKEY, K.B.; LASKEY, K.J. **PR-OWL: a Bayesian ontology language for the Semantic Web**, Proceedings URSW-2005, Galway, Ireland, p. 23-33, 2005.

CRISTO, C. M. P. N. Prospectiva estratégica: instrumento para a construção do futuro e para a elaboração de políticas públicas. **Anais do VII Congresso Internacional del CLAD sobre la Reforma del Estado y de la Administración Pública**. Portugal: Lisboa, 2002.

CUHLS, K.; GRUPP, H. Alemanha: abordagens prospectivas nacionais. **Parcerias Estratégicas**. v.10, p. 76-104, 2001.

DALKEY, N.; HELMER, O. **An experimental application of the Delphi method to the use of experts**. *Management Science*, Maryland, v.9, n.9, p.458-467, 1963.

DE DOMBAL, F.T. *et al.* **Computer-aided Diagnosis of Acute Abdominal Pain**. *British Medical Journal*, 2, p. 9-13, 1972.

ERICSSON, H.; MUSEN, M. A. Metatools for knowledge acquisition. **IEEE Software**, v. 10, n. 3, p. 23-29, 1993.

FAYYAD, U. M. *et al.* **Advances in knowledge discovery and data mining**. AAAI Press, Califórnia: Menlo Park, 1996.

FENSEL, D. Ontology - Based Knowledge Management, **IEEE Computer**, v. 35, n. 11, p. 56-59, 2002.

FENZA, G.; LOIA, V.; SENATORE, S. Fuzzy modeling of medical diagnosis through a service-oriented architecture. **IEEE, Annual Conference of the North American Fuzzy Information Processing Society – NAFIPS**, p. 1-6, 2008.

FERRAZ, R. M. O. L. **Ligação aérea Rio - São Paulo: análise da demanda pelo método delphi**. Dissertação de Mestrado. Departamento de Engenharia de Transportes, Universidade de São Paulo (USP), São Paulo, 1993.

FIALHO, F. A. P. *et al.* **Gestão do conhecimento e aprendizagem: as estratégias competitivas da sociedade pós-industrial**. Florianópolis: Visual Books, 2006.

FREITAS, A. A. **Uma Introdução a Data Mining**. Informática Brasileira em Análise, Centro de Estudos e Sistemas Avançados do Recife, Recife, ano II, n. 32, 2000.

FUENTES-LORENZO, D.; MORATO, J.; GÓMEZ, J. M. Knowledge management in biomedical libraries: A Semantic Web approach. **Information Systems Frontiers**, v. 11, n. 4, p.471–480, 2009.

FRED, A. L. N. **Redes bayesianas**. [S.l.], 2004. Disponível em: <<http://www.lx.it.pt/~afred/docencia/Percepcao/acetatos/rb.pdf>>. Acesso em: 08 mai. 2009.

GAERTNER GROUP. **Tecnologia da Informação, Administração do Conhecimento e Tecnologia: chave do sucesso**. Encarte especial da Revista Exame, nº 669, ago. 1998.

GANIZ, M.C.; POTTENGER, W.M.; JANNECK, C.D. Recent advances in literature-based discovery, **Technical Report, Lehigh University**, LU-CSE-05-027, 2005.

GARCÍA-CRESPO, A. *et al.* ODDIN: Ontology-driven differential diagnosis based on logical inference and probabilistic refinements. **Expert Systems with Applications**, v. 37, n. 3, p. 2621-2628, 2010.

GENNARI, J. H. *et al.* The evolution of *Protégé*: an environment for knowledge-based systems development. **Int. J. Human-Computer Studies**, v.58, p. 89-123, 2003.

GOLDSTEIN, M. K. *et al.* Implementing clinical practice guidelines while taking account of evidence: ATHENA, an easily modifiable decision-support system for management of hypertension in primary care. **Proceedings of the Annual AMIA Fall Symposium**, p. 300-304, 2000.

GÓMEZ-PÉREZ, A. Ontological engineering: a state of the art. **British Computer Society**, v. 2, p. 33-43, 1999.

GORDON, M.D.; LINDSAY, R.K. Toward discovery support systems: A replication, re-examination, and extension of Swanson's work on literature-based discovery of a connection between Raynaud's and fish oil. **Journal of the American Society for Information Sciences**, v.47, p. 116-128, 1996.

GRAU, B. C.; HALASCHEK-WIENER, C.; KAZAKOV, Y. History matters: Incremental ontology reasoning using modules. In. **Proceedings of the 6th International Semantic Web Conference (ISWC 2007)**, Busan, South Korea, Springer, v. 4825, p. 183-196, 2007.

GRUBER, T. R. A translation approach to portable ontology specifications. **Knowledge Acquisition**, v. 5, p. 199-220, 1993.

HAIMOWITZ, I., PATIL, R.; SZOLOVITS, P. Representing medical knowledge in a terminological language is difficult, In. **Proceedings of the Annual Symposium on Computer Application in Medical Care**, p. 101-105, IEEE Computer Society Press, 1988.

HRISTOVSKI, D. *et al.* Supporting discovery in medicine by association rule mining in Medline and UMLS. In **Proceedings of MedInfo Conference**, London, England, v.10, n. 2, p.1344-1348, 2001.

HRISTOVSKI, D. *et al.* Exploiting Semantic Relations for Literature-Based Discovery. **AMIA 2006 Symposium Proceedings**, p. 349 - 353, 2006.

HRISTOVSKI, D. *et al.* Literature based discovery support system and its application to disease gene identification, In. **Computational Discovery**, Springer-Verlag Berlin Heidelberg, LNAI 4660, p. 307-326, 2007.

HUANG, H. *et al.* A comparatively research in incremental learning of Bayesian networks, In. **Proceedings of the fifth world congress on intelligent control and automation Intelligent Control and Automation**, v. 1, Hangzhou, China; 2004, p. 4260-4265.

JENSEN, F. V. **Bayesian Network and Decision Graphs**. New York: Springer-Verlag, 2001.

KAYO, E. K.; SECURATO, J. R. Método delphi: fundamentos, críticas e vieses. **Cadernos de Pesquisas em Administração**. v.1, n.4, p. 51-61. São Paulo. 1997.

- KEELING, C.; LAMBERT, S. Knowledge management in the NHS: positioning the healthcare librarian at the knowledge intersection. **Health Libraries Review**, v.17, p. 136-143, 2000.
- KEENEY, Sinead. HASSON, Felicity. MCKENNA, Hugh P. A critical review of the delphi technique as a research methodology for nursing. **International Journal of Nursing Studies**. v.38, p. 195-200. 2001.
- KENDAL, S. L.; CREEN, M. **An introduction to knowledge engineering**. London: Springer-Verlag, 2007.
- KILOV, H.; SACKS, I. Mechanisms for communication between business and IT experts. **Computer Standards & Interfaces**, v. 31, p. 98-109, 2008.
- KLEER, J.; WILLIAMS, B. Diagnosing multiple faults. **Artificial Intelligence** , v. 32, p. 97-130, 1987.
- KOLLER, D.; PFEFFER, A. Probabilistic frame-based systems. In Proceedings of the 15th. **National Conference on Artificial Intelligence (AAAI'98)**. Madison: AAAI Press, p. 580–587, 1998.
- KONG, G.; XU, D. L.; YANG, J. B. Clinical Decision Support Systems: A Review on Knowledge Representation and Inference Under Uncertainties. **International Journal of Computational Intelligence Systems**, Manchester, v. 1, n. 2, p. 159-167, 2008.
- LANZOLA, G.; QUAGLINI, S.; STEFANELLI, M. Knowledge-acquisition tools for medical knowledge-based systems. **Methods of Information in Medicine**, v.34, p. 25-29, 1995.
- LI, T.; RUAN, D. An extended process model of knowledge discovery in database. **Journal of Enterprise Information Management**, v. 20, n. 2, p. 169-177, 2007.
- LIN, L.; HU, P. J.-H.; LIU SHENG, O. R. A decision support system for lower back pain diagnosis: Uncertainty management and clinical evaluations. **Decision Support Systems**, v. 42, p. 1152–1169, 2006.
- LOPES, H. S. **Analogia e aprendizado evolucionário: Aplicação em diagnóstico clínico**. 1996. 159 f.. Tese (Doutorado em Engenharia Elétrica – Sistemas de Informação), Universidade Federal de Santa Catarina, Florianópolis, 1996.
- LOPES, L. F. *et al.* **Sistema de Conhecimento para diagnóstico em acupuntura: uma modelagem usando o CommonKADS**. Gestão e Produção (UFSCAR. Impresso). v. 18, n. 2, p. 351-366, 2011.

LOPES, L. F. *et al.* **Uma ontologia de apoio ao diagnóstico aplicado à acupuntura.** In: 8º. Congresso da Sociedade Brasileira de Gestão do Conhecimento, 2009, Salvador (BA). Anais do KM Brasil 2009. Salvador (BA), 2009.

LOPES, L. F., GONCALVES, A. L., LOPES, M. C. **Uma ontologia para diagnóstico em acupuntura baseada no modelo CESM.** Anais do KM Brasil 2010. Porto Alegre (RS), 2010.

LUGER, G.F. **Artificial Intelligence: Structures and Strategies for Complex Problem Solving**, 6th ed. Boston: Addison-Wesley Pearson Education, 2009.

MARINHO, S. V. **Uma Proposta de Sistemática para Operacionalização da Estratégia Utilizando o *Balanced Scorecard*.** Tese de Doutorado. Programa de Pós-graduação em Engenharia de produção. Universidade Federal de Santa Catarina (UFSC), 2006.

MARTIN, B.R.; ANDERSON, J.; MACLEAN, M. Identifying Research Priorities in Public-Sector Funding Agencies: Mapping Science Outputs onto User Needs. **Technology Analysis and Strategic Management**, v. 10, 1998.

MELLO, A. G. **Sistema neuronal de reconhecimento de placas em veículos automotores particulares.** Dissertação de Mestrado, Universidade Federal do Rio de Janeiro, COPPE/UFRJ & CBPF, agosto, 2001.

MINCHIN, R. *et al.* Symptoms Ontology for Mapping Diagnostic Knowledge Systems. In. **Proceedings of the 19th IEEE Symposium on Computer-Based medical Systems**, p. 593-598, Washington, DC, USA, 2006.

MINSKY, M. A Framework for Representing Knowledge. In P. Winston (Ed), **The Psychology of Computer Vision**. New York: McGraw-Hill, p. 170–195, 1977.

MUSEN, M. A. *et al.* EON: A component-based approach to automation of protocol-directed therapy. **Journal of the American Medical Informatics Association**, v. 3, n.6, p. 367-388, 1996.

MUSEN, M.; SHAHAR, Y.; SHORTLIFFE, E. Clinical decision-support systems, In. E.H. Shortliffe, J.J. Cimino (Eds.), **Biomedical Informatics: Computer Applications in Health Care and Biomedicine**, 3rd ed., Springer, New York, 2006.

MUSTAFARAJ, E. H.; HOOFF, M.; FREISLEBEN, B. Learning Knowledge Roles for Populating Lightweight Application Ontologies. **The Electronic Journal of Knowledge Management**, v. 4, p. 75-82, 2006.

NECHES, R. *et al.* Enabling technology for knowledge sharing. **AI Magazine**, v. 12, n. 3, p. 16-36, 1991.

NIEVOLA, J.C. **Sistema Inteligente para Auxílio ao Ensino em Traumatologia Crânio -encefálica**, 1995. 168 f.. Tese (Doutorado em Engenharia Elétrica – Sistemas de Informação) – Grupo de Pesquisas em Engenharia Biomédica, Universidade Federal de Santa Catarina, Florianópolis, 1995.

NOY, N.; MCGUINNESS, D. L. **Ontology Development 101: a guide to creating your first ontology**. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, 2001.

NONAKA, I.; TAKEUCHI, H. **Criação de conhecimento na empresa: como as empresas japonesas geram a dinâmica da inovação**. Rio de Janeiro : Campus, 1997.

NONAKA, I.; TOYAMA, R. The knowledge-creating theory revisited: knowledge creation as a synthesizing process. **Knowledge Management Research & Practice**, v. 1, n. 9, p. 2-10, 2003.

NONAKA, I.; TOYAMA, R. The theory of the knowledge-creating firm: subjectivity, objectivity and synthesis. **Industrial and Corporate Change**, v. 14, p. 419-436, 2005.

NONAKA, I.; TOYAMA, R.; KONNO, N. SECI, ba and leadership: a unified model of dynamic knowledge creation. **Long Range Planning**, v. 33, p. 5-34, 2000.

NORSYS NETICA: **bayesian networks graphical application**, 2001. Disponível em: <<http://www.norsys.com/netica.html>>. Acesso em: 02 mai. 2008.

PARSIA, B., HALASCHEK-WIENER, C.; SIRIN, E. Towards incremental reasoning through updates in OWL-DL. In: Proceedings of the workshop at 15th international, World Wide Web conference on reasoning on the Web, 2006.

PODGORELEC, V.; GRASIC, B.; PAVLIC, L. Medical diagnostic process optimization through the semantic integration of data resources.

Computer Methods and Programs in Biomedicine, v. 95, p. 55-67, 2009.

PORTO, C. A.; SOUZA, N. M.; BUARQUE, S. C. **Construção de cenários e prospecção de futuros: treinamento conceitual.**

Cláudio Porto & Consultores Associados Ltda. 1989.

PROTÉGÉ 3.3.1. Stanford: California, 2007. Disponível em: <<http://protégé.stanford.edu>>. Acesso em: 02 fev. 2009.

QUIRINO, T. R. IRIAS, L. J. M. Globalização, agricultura e degradação ambiental no Brasil: perspectivas para a pesquisa e para políticas agropecuárias. **Cadernos de Ciência e Tecnologia**. v.15, n.1, p. 93-124. Brasília, 1998.

QUIRINO, T. R. L.; Alfredo J. B. DIAS, E. C. Tecnologias agropecuárias e impacto ambiental: uma agenda para a pesquisa. **Cadernos de Ciência e Tecnologia**. v.16, n.2, pp. 69-96. Brasília, 1999.

RAUTENBERG, S. *et al.* ontoKEM: uma ferramenta para construção e documentação de ontologias. In. Seminário de Pesquisa em Ontologias no Brasil. **Anais...** Niterói, 2008.

RODRÍGUEZ, A. *et al.* ADONIS: Automated Diagnosis System based on Sound and Precise Logical Descriptions. Computer-Based Medical Systems. **22nd IEEE International Symposium**, Madrid, p. 1-8, 2009a.

RODRÍGUEZ, A. *et al.* MEDBOLI: Medical Diagnosis based on Ontologies and Logical Inference. Proceeding - **International Conference on eHealth, Telemedicine, and Social Medicine, TELEMED 2009**, Cancun, México, p. 233-238, 2009b.

ROWE, G.; WRIGHT, G. The Delphi technique as a forecasting tool: issues and analysis. **International Journal of Forecasting**. v. 15, p. 353-375, 1999.

RUSSEL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. New Jersey: Prentice-Hall, 1995.

SANT'ANA, P. H. M. **Análise prospectiva de tecnologias de energia: validação e análises de uma consulta delphi com especialistas do Brasil.** Dissertação de Mestrado. Pós-Graduação em Engenharia Mecânica, Universidade Estadual de Campinas (UNICAMP). 2005.

SANTOS, L. D.; AMARAL, L. Estudos Delphi com Q-Sort sobre a web a sua utilização em Sistemas de Informação. **Proceedings of the**

Conferência da Associação Portuguesa de Sistemas de Informação. Lisboa, 2004.

SACKETT, D. L. *et al.* **Medicina baseada em evidências: prática e ensino.** Trad. van Carlquist. 2. ed. Porto Alegre: Artmed, 2003. 270p.

SCHREIBER, G. *et al.* **Knowledge engineering and management: the commonKADS methodology**, Massachusetts: MIT Press, 2002.

SENGE, P. M. **A Quinta Disciplina: Arte e Prática da Organização que Aprende** 16. ed. (Tradução: OP Traduções). São Paulo: Editora Nova Cultural, 2004.

SHORTLIFFE, E. H. **Computer-based Medical Consultations: MYCIN.** New York: Elsevier, 1976. 264p.

SHORTLIFFE, E. H. *et al.* **Medical Informatics: Computer Applications in Health Care and Biomedicine.** 2. ed. New York: Springer-Verlag, 2001. 770p.

SILVA, M.P.S. Mineração de Dados: Conceitos, Aplicações e Experimentos com Weka. **Relato Técnico. Livro da Escola Regional de Informática Rio de Janeiro – Espírito Santo.** Porto Alegre: Sociedade Brasileira de Computação, v.1, 2004.

SMETS, P. Imperfect information: imprecision - uncertainty. In Morro A. and Smets Ph. eds. **Uncetminty Management in Information Systems.** Kluwer, Boston, p. 225-254. 1997.

SMITH, R. G.; FARQUHAR, A. The Road Ahead for Knowledge Management: An AI Perspective, **AI Magazine**, Winter, v. 21, n. 4, p.17-40, 2000.

SOLO, A. M.; GRUPTA, M. M. Perspectives on computational perception and cognition under uncertainty. In Proceedings of IEEE International Conference on Industrial Technology 2000. **IEEE Computer Society**, Goa, v. 1, p. 221-224, 2000.

SOWA, J. F. **Knowledge Representation: Logical, Philosophical and Computational Foundations**, Pacific Grove: Brooks/Cole, 2000. 594p.

STRUDER, V. R.; BENJAMINS, R.; FENSEL, D. Knowledge enginnering: principles and methods, **Data & Knowledge Engineering**, v. 25, p. 161-197, 1998.

SWANSON, D. R. Fish oil, Raynaud's syndrome, and undiscovered public knowledge. **Perspectives in Biology and Medicine**, v.30, n. 1, p. 7-18, 1986.

SWE, M.; KHAM, M. Ontology-Based Medical Diagnostic Knowledge Structuring Using Case-Based Reasoning Methodology. **ACM International Conference Proceeding**, p.745-749, 2009.

TAYLOR, T. R.; AITCHISON, J.; MCGIRR, E. M. Doctors as decision-makers: a computer-assisted study of diagnosis as a cognitive skill. **British Medical Journal**, v. 3, p. 35-40, 1971.

TURBAN, E.; ARONSON, J. E. **Decision support systems and intelligent systems**. 6^a ed. New Jersey: Prentice-Hall, 2001, 867p.

UZOKA, F-M. E. *et al.* An experimental comparison of fuzzy logic and analytic hierarchy process for medical decision support systems. **Computer Methods Programs Biomedicine**, v. 103, p. 10-27, 2010.

VAN BEMMEL, J. H.; MUSEN, M. A. **Handbook of Medical Informatics**, Springer-Verlag, 1997, 638p.

VICKERY, B. C. Knowledge Representation: a brief review, **Journal of Documentation**, v. 42, p. 145-159, 1986.

WAZLAWICK, R. S. **Metodologia de pesquisa para ciência da computação**. Rio de Janeiro: Elsevier, 2008, 159p.

WEEBER, M. *et al.* Using concepts in literature-based discovery: Simulating Swanson's Raynaud-Fish oil and migraine-magnesium discoveries. **Journal of the American Society for Information Science and Technology**, v. 52, p. 548-557, 2001.

WRIGHT, J. T. C.; GIOVINAZZO, R. A. Delphi: uma ferramenta de apoio ao planejamento prospectivo. **Caderno de Pesquisas em Administração**. v.1, n.12, p. 54-65, São Paulo. 2000.

YANG, Y.; CALMET, J. OntoBayes: an ontology-driven uncertainty model, In. Conf. on Intelligent Agents, Web Technology and Internet Commerce (**IAWTIC-2005**), IEEE Press, v. 1, p. 457-463, 2005.

YANG, Y. **A Framework for Decision Support Systems Adapted to Uncertain Knowledge**. Tese (Doutorado em Ciência da Computação). Universität Karlsruhe (TH), 2007.

ZADEH, L. A. Fuzzy sets. **Information and Control**, v. 8, p. 338-353, 1965.

ZHAO, W.; YANXIANG, H.; HUI, J. A model of intelligent distributed medical diagnosis and therapy system based on mobile agent and ontology. In. **Proceedings of the 8th international conference on high-performance computing in the Asia-Pacific Region**, p. 582-587, 2005.

APÊNDICE I – QUESTIONÁRIO

Informações iniciais

Nome completo:.....

e-mail:.....

Instituição:..... Função:.....

Instituição:..... Função:.....

Instituição:..... Função:.....

Graduação:.....

Instituição:.....

Graduação:.....

Instituição:.....

Pós-Graduação:.....

Instituição:.....

Pós-Graduação:.....

Instituição:.....

Pós-Graduação:.....

Instituição:.....

Outros:.....

Instituição:.....

Outros:.....

Instituição:.....

Outros:.....

Instituição:.....

Anos de experiência na atividade de diagnóstico:.....

Área de atuação:.....

Perguntas preliminares

1) Você já teve contato ou experiência com algum sistema que auxiliasse no diagnóstico?

☐ sim ☐ não

2) Que receios/medos você teria de utilizar um sistema de apoio?

- ☐ Interface pouco amigável
- ☐ Pouco conhecimento em informática
- ☐ Confiabilidade do sistema

3) Você julga que um sistema baseado em conhecimento* poderia auxiliá-lo em sua atividade?

☐ sim ☐ não

4) Você julga necessário considerar diferentes variáveis que compõem o contexto, não somente o paciente, para explorar relações entre as informações coletadas e assim concluir o diagnóstico?

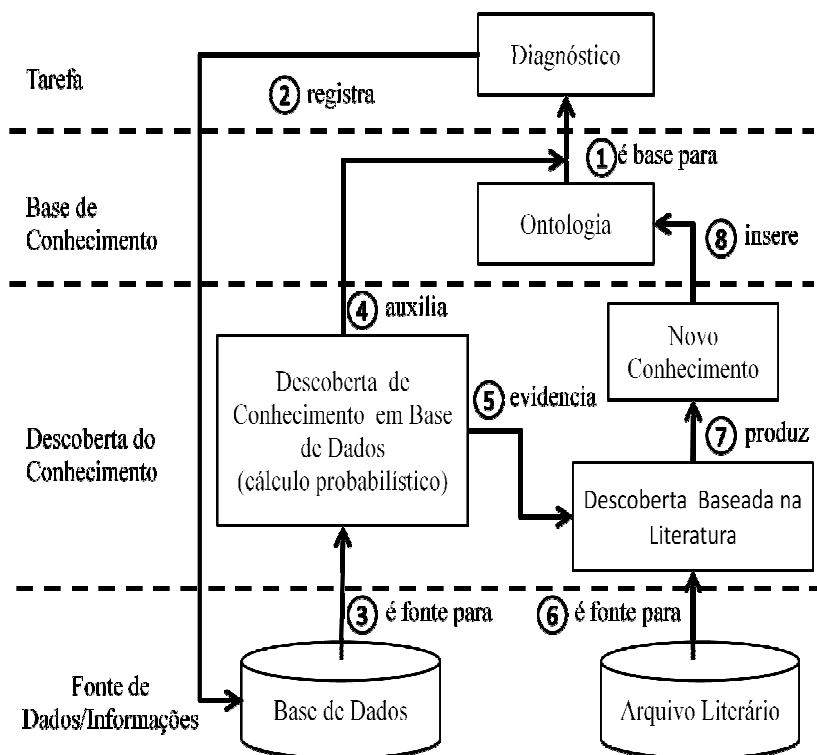
☐ sim ☐ não

5) Você julga importante registrar os dados de diagnósticos realizados para direcionar novos diagnósticos embasando-se em casos anteriores?

☐ sim ☐ não

*Os Sistemas Baseados em Conhecimento (SBC) representam uma área do conhecimento que está englobada na grande área da Inteligência Artificial (IA). Sendo sistemas computacionais, procuram auxiliar os humanos na resolução de problemas nos mais diversos setores e atividades.

Verificação da consistência do modelo para apoio ao processo de diagnóstico



Obs: As questões a seguir, deverão ser respondidas de acordo com seu nível de entendimento, procurando contribuir para o aperfeiçoamento deste processo.

- Devido ao avanço tecnológico ocorrido nos últimos anos, vários sistemas baseados em conhecimento estão sendo desenvolvidos e aprimorados para apoiar o processo de diagnóstico (García-Crespo *et al.*, 2010). Cada vez mais se incrementa estes sistemas com novos métodos e técnicas promovendo maior segurança e eficácia na sua utilização. Portanto:

1) Acredita-se no potencial destes sistemas como forma de auxiliar os profissionais no apoio ao processo de diagnóstico. Salienta-se que estes sistemas não têm por objetivo suprimir o papel do especialista, mas sim orientá-lo em sua atividade.

() Discordo

() Tenho Restrições

() Concordo

Comentários - Sugestões - Justificativas:

.....

- Uma ontologia é definida como sendo um conjunto de termos ordenados hierarquicamente para representar um domínio específico. Ela pode ser usada como um esqueleto para uma base de conhecimento onde são executados processos de inferência (raciocínio). O uso de uma ontologia permite então, a definição de um domínio no qual será possível trabalhar em determinada área específica, possibilitando a melhora no processo de extração de informação e o intercâmbio do conhecimento (Gómez-Pérez, 1999). Assim sendo:

2) Acredita-se que o conhecimento implícito do profissional (aquele que está na cabeça / raciocínio das pessoas), que envolve uma especialidade, pode ser representado (explicitado) através de uma ontologia, a qual permite reuso, interoperabilidade, compartilhamento e integração do mesmo. Além disto, possibilita que novos conhecimentos sejam inseridos para aprimorar a representação e o raciocínio dentro de um domínio específico. O conhecimento específico do profissional pode ser representado de forma que seja analisado por um sistema.

() Discordo

() Tenho Restrições

() Concordo

Comentários - Sugestões - Justificativas:

.....

- A conceituação de sistemas apresentada por meio do modelo CESM (Composition – Environment – Structure – Mechanism) contribui, dentro de uma visão sistêmica, para a explicitação do conhecimento através de uma modelagem mais descritiva e abrangente para os elementos que compõem um domínio específico (Bunge, 2003). Portanto:

3) Acredita-se que quando da interrogação sobre uma determinada queixa, outros fatores podem compor a investigação. A inquirição deve ir além das manifestações apresentadas e relatadas pelo reclamante. Fatores externos que compõem o ambiente, no qual o reclamante está inserido, contribuem para tornar o processo de diagnóstico mais real e compreensivo.

☐ () Discordo

☐ () Tenho Restrições

☐ () Concordo

Comentários - Sugestões - Justificativas:

.....

- Atualmente, com o intuito de obter os sinais, sintomas e exames (testes) que podem levar à solução de uma queixa, muitos dados e informações coletados do reclamante são armazenados em fichários (papel ou eletrônico). Permanecem neste estado, até uma nova consulta, para compor o cenário histórico. Portanto:

4) Acredita-se que muito conhecimento continua em estado latente. Estas informações, quando tratadas adequadamente, podem revelar relações ocultas entre causa e efeito, podendo sugerir possíveis novos conhecimentos para um domínio específico. As informações armazenadas de maneira adequada e sistemática podem contribuir para analisar relações entre os dados obtidos do reclamante, bem como visualizar novas formas de investigá-lo.

☐ () Discordo

☐ () Tenho Restrições

☐ () Concordo

Comentários - Sugestões - Justificativas:

.....

- Descoberta de Conhecimento numa Base de Dados é o processo de aquisição de conhecimento a partir de banco de dados. Consiste num conjunto de técnicas que tem por objetivo permitir a extração de conhecimento útil e com significado que se encontram implícito em base de dados (Fayyad et al., 1996). Assim sendo:

5) Acredita-se que com a formação de uma base de dados, obtida a partir de diagnósticos realizados e bem sucedidos, informações relevantes sejam extraídas e possibilitem que o profissional realize pesquisas dirigidas em arquivos literários com a intenção de descobrir possíveis novos conhecimentos.

() Discordo

() Tenho Restrições

() Concordo

Comentários - Sugestões - Justificativas:

.....

- A Descoberta Baseada na Literatura é um método para encontrar relações entre conceitos, não abertamente declaradas, embutidas na literatura publicada (AHLERS et al., 2007). Geralmente, tem a forma de relação entre dois conceitos primários, por exemplo, um medicamento que trata uma determinada doença ou um gene como a causa de uma enfermidade. As regras de associação e a tecnologia semântica fazem parte das técnicas que auxiliam na extração destas relações (HRISTOVSKI et al., 2006). Portanto:

6) Acredita-se que fazendo uso de sistemas de descoberta baseada na literatura é possível obter indicações que revelem a existência de correlações entre elementos até então encobertos e que sejam úteis no aprimoramento de um domínio específico. Assim, para um dado diagnóstico pode-se descobrir vários elementos/características relacionados a ele.

() Discordo

() Tenho Restrições

() Concordo

Comentários - Sugestões - Justificativas:

.....

- A quantidade de informações disponíveis nos meios de comunicação cresce de forma exponencial. Entretanto, devido à especialização em cada setor, a troca de informações está cada vez mais difícil, isto é, pesquisadores de diferentes áreas dificilmente trocam conhecimento adquirido em suas especialidades (GANIZ et al., 2005). Assim sendo:

7) Acredita-se que um sistema que possibilite rastrear relações (associações) entre elementos, dentro de trabalhos em diferentes especialidades, terá maior facilidade em descobrir novos conhecimentos.

() Discordo

() Tenho Restrições

() Concordo

Comentários - Sugestões - Justificativas:

.....

Referências

- AHLERS, C. B. *et al.* Using the literature-based discovery paradigm to investigate drug mechanisms. AMIA Symposium, p. 6 -10, 2007.
- BUNGE, M. Emergence and Convergence: Qualitative Novelty and the Unity of Knowledge. [S.l.]: University of Toronto Press, 2003. 330 p.
- FAYYAD, U. M. *et al.* Advances in knowledge discovery and data mining. MenloPark: Califórnia: [s.n.], 1996.
- GARCÍA-CRESPO, A. *et al.* ODDIN: Ontology-driven differential diagnosis based on logical inference and probabilistic refinements. Expert Systems with Applications. Madrid, p. 2621 - 2628, 2010.
- GÓMEZ-PÉREZ, A. Ontological engineering: a state of the art. British Computer Society, v. 2, p. 33 - 43, 1999.
- HRISTOVSKI, D. *et al.* Exploiting Semantic Relations for Literature-Based Discovery. AMIA 2006 Symposium Proceedings, p. 349 - 353, 2006.
- GANIZ C. M., POTTENGER M. W. and JAMECK D. C. Recent advances in Literature Based Discovery. Technical report, Lehigh University, 2005.

APÊNDICE II – COMENTÁRIOS DOS ESPECIALISTAS

Na primeira questão, procurou-se identificar se o especialista acredita no potencial destes sistemas como apoio e orientação ao processo de diagnóstico.

Entre as sugestões, justificativas e comentários apresentados, destacam-se:

- “Os sistemas baseados em conhecimento podem sim ajudar o profissional no processo de diagnóstico, acredito que ele pode recordar o profissional de todas as opções de conclusão diagnóstica, bem como orientar nas perguntas que podem ser feitas durante o processo de diagnóstico, pois pode direcionar características (biopsicossociais), sinais e sintomas a serem investigados”.
- “Acredito que quem já trabalha na clínica, atendendo pacientes, já tem um roteiro memorizado para chegar ao diagnóstico. O receio é com os jovens estudantes de medicina que estão começando e que podem confiar “demais” nestes sistemas e esquecer o lado da dedicação aos estudos e ao paciente”.
- “Acredito que os sistemas baseados em conhecimento são uma ferramenta para o apoio do profissional e que pode ser utilizado como suporte ao diagnóstico, porém não necessariamente como uma ferramenta de orientação. Sendo um sistema, e portanto limitado e baseado em um conhecimento que foi representado por um humano, pode ter falhas, bem como não necessariamente é capaz de absorver a visão sistêmica e subjetiva, podendo gerar uma orientação restrita e consequentemente, “contaminada”.
- “A concordância será efetivada se forem atendidos os seguintes pontos:”
 - “a) que o sistema contenha todo o conhecimento com o qual o especialista tomaria a decisão. Creio que em alguns casos até possa ter mais, porém em última análise quem “assina” é o ser humano;”
 - “b) Na decisão humana, há tomada de decisão baseada em conhecimento “tácito” e conhecimento “explícito”. Nem sempre o “tácito” é levado para meios eletrônicos e em muitos casos nem o especialista consegue explicitá-lo, porém o utiliza

na tomada de decisão;”

”c) A tomada de decisão, no domínio da saúde, é fortemente probabilística e incompleta”.

- “Diagnóstico e tomada de conduta são processos muito complexos em várias situações clínicas como a do paciente em UTI ou em uso de múltiplas medicações, ou estado crítico, ou de quadros clínicos sem diagnóstico preciso, necessitam de comandos de alerta, lembrança, apoios para serem conduzidos de forma mais segura e minimizando erros ao longo dos procedimentos”.

Na segunda questão, procurou-se identificar se o especialista acredita na possibilidade de representar o conhecimento específico do domínio de forma que seja analisado por um sistema.

Entre as sugestões, justificativas e comentários apresentados, destacam-se:

- “Acredito sim na possibilidade de representação sistemática e formal do conhecimento, com completude e consistência suficientes para cobrir determinado domínio. Creio que este conhecimento pode se tornar semelhante ao do especialista humano, porém interpretável por um computador. Isto deve ser feito por meio de métodos fortemente embasados na literatura, e incluir o trabalho conjunto do engenheiro do conhecimento e um conjunto de especialistas na área”.
- “Parcialmente sim, o sistema contribui para o raciocínio, auxilia em pesquisa em banco de dados, mas todo o conhecimento, a vivência profissional não”.
- “Pode ser representado parcialmente. Nem todo conhecimento pode ser representado”.
- “Acredito que sim, o difícil é sistematizar estes conhecimentos para alimentar o sistema. Algumas vezes o diagnóstico vem por observar o paciente, ou por intuição”.
- “Tenho restrições, pois o conteúdo na área da saúde é muito amplo, sendo difícil ser totalmente suprido por um sistema. Porém, um sistema de apoio à decisão é importante para auxiliar o profissional a organizar todo conhecimento e, também para agregar novos conceitos e conteúdos”.
- “Acredito que a ontologia seja capaz de representar o conhecimento implícito de um profissional, entretanto a eficácia

da ontologia depende em muito da forma que o conhecimento é representado e por quem. Não necessariamente a ontologia é universal, mesmo dentro de um determinado domínio, já que pode ser dependente da capacidade do especialista representar objetivamente seu conhecimento, que muitas vezes é subjetivo. Entretanto, se adequadamente representado, acredito que o conhecimento pode ser analisado por um sistema”.

- “Creio que não apenas o explícito assim o possa. Cabe também ao implícito tal estruturação. Quanto à possibilidade de ser computado, o conhecimento estruturado e baseado em uma ontologia, diminui as questões semânticas associadas aos termos e seus significados, elegendo uma forma definida, conceituada e computável, ex. lógica de primeira ordem, árvores ontológicas, etc,...”.
- “Porém, é necessário que haja por parte do usuário do sistema a compreensão cognitiva e técnica para absorver e “processar” as informações”.
- “Contando que o profissional tenha tido sua formação com o rigor metodológico que a ontologia esteja propondo”.

Na terceira questão, procurou-se identificar se o especialista acredita na possibilidade da inquirição ir além das manifestações apresentadas e relatadas pelo reclamante. Fatores externos que compõem o ambiente, no qual o reclamante está inserido, contribuem para tornar o processo de diagnóstico mais real e compreensivo.

Entre as sugestões, justificativas e comentários apresentados, destacam-se:

- “Sim, o diagnóstico é um processo global, que se torna deficiente se não considerarmos o indivíduo como um todo e inserido em uma determinada realidade. O fato de considerarmos o ambiente e a influência familiar pode trazer pistas bastante conclusivas para fechar um diagnóstico. Assim, considerar fatores externos além dos sinais e sintomas apresentados pelo paciente nos auxiliam durante o processo de diagnóstico e, também, após o diagnóstico nos auxilia a confirmá-lo e a fazer diagnóstico diferencial, excluindo outras possibilidades”.

- “Com certeza o modo de vida, o ambiente, a alimentação e até o quadro emocional são importantes para analisar o indivíduo como um todo”.
- “Em um processo de diagnóstico a clínica deve ser sempre soberana e por clínica entende-se as relações sistêmicas em que o indivíduo se encontra nos mais amplos aspectos: biológico, social, psicológico, ou seja, o indivíduo dentro da sua visão ecológica”.
- “Em uma visão sistêmica sobre o universo, seja de maneira pragmática ou apenas teórica, aceitamos que há uma associação entre os objetos que o compõem (sejam físicos ou conceituais). Estas associações em alguns casos são diretas e em outros casos indiretas (há outros objetos que se interrelacionam). Frente a esta afirmação, fatores (variáveis) que possam pertencer, aparentemente, a outros domínios, que não o que está sendo analisado, podem, diretamente ou indiretamente estarem influenciando o fenômeno. Para tal há que testá-los. Fica em aberto a questão como escolhê-los. Temos que aprender, desconfia-se, escolhe-se, cria-se a série histórica, observa-se e conclui-se”.
- “Um detalhamento das queixas aliado ao conhecimento de fatores externos atuantes sobre um indivíduo, bem como suas pré-disposições geneticamente determinadas conduzem a maior precisão diagnóstica”.
- “Em especial para o diagnóstico em Medicina Tradicional Chinesa (MTC) onde a acupuntura está inserida. O ambiente e os fatores externos, não só contribuem para o processo de diagnóstico como são fatores imprescindíveis para sua realização adequada, uma vez que eles são possíveis causas etiológicas das doenças (queixas)”.
- “Realmente é necessário avaliar as variantes mais importantes em relação a qualquer queixa para ordenar racionalmente as seleções de hipóteses diagnósticas ou tomadas de decisão”.

Na quarta questão, procurou-se identificar se o especialista reforça o fato de que muito conhecimento, proveniente de dados e informações coletados no interrogatório, continua em estado latente.

Entre as sugestões, justificativas e comentários apresentados, destacam-se:

- “Creio que enquanto profissionais acabamos ficando bitolados em determinadas estratégias de ação e pensamento e deixamos de visualizar outras formas de se analisar o caso de um paciente, e, portanto, deixamos de gerar novos conhecimentos. Desta forma, a existência de um sistema que vislumbre novas estratégias de ação para se chegar a um diagnóstico, e, que gere novos conhecimentos e o apresente ao profissional, será de grande importância inclusive para capacitação e atualização do mesmo”.
- “Acho difícil obter “conhecimentos novos” somente a partir de informações armazenadas sistematicamente”.
- “Concordo, apesar de ter restrições na expressão “muitos dados e informações coletados do reclamante são armazenados em fichários (papel ou eletrônico)”. “Permanecem neste estado, até uma nova consulta, para compor o cenário histórico.” O raciocínio clínico continua além das anotações no fichário e muitas vezes são temporariamente suspensos da mente do profissional para posteriormente serem retornados, muito antes da próxima consulta. Isso seria a “pausa criativa”, onde a mente tenta desfocar os dados, trabalhando “em um nível de consciência diferenciado” a fim de conseguir organizar as informações e chegar a uma solução diagnóstica”.
- “Sem dúvida há muito conhecimento em estado latente. Hoje com as metodologias de KD (Descoberta de Conhecimento), torna-se viável o armazenamento e a utilização do mesmo para aprender-se “associações”, “classes”, “determinantes”, etc....”.
- “Quanto a dizer que vai “visualizar novas formas de investigá-lo”. Creio que mais do que sugerir novos conhecimentos, as informações tratadas adequadamente, resgatam informações que muitas vezes se perdem no próprio histórico da evolução do tratamento. Detalhes que serão, sem dúvida, úteis”.
- “Realmente em estudos de *Follow-up* principalmente as informações armazenadas podem demonstrar evidências que não se revelariam de outra maneira”.

Na quinta questão, procurou-se identificar se o especialista acredita que, com a formação de uma base de dados, a partir de

diagnósticos realizados, será possível extrair informações relevantes que gerem novos conhecimentos.

Entre as sugestões, justificativas e comentários apresentados, destacam-se:

- “A rotina apressada faz com que os profissionais mantenham pouco registro de sua atividade, isto corrobora com a necessidade de extrair e arquivar dados sobre diagnóstico, de forma sistemática e facilmente acessível, para contribuir com novos diagnósticos e gerar novos conhecimentos. Isto permitirá não somente o reuso das informações, mas também o seu compartilhamento”.
- “Banco de Casos permitem a extração de informações relevantes, porém isto não é suficiente para que profissionais realizem pesquisas dirigidas”.
- “Entretanto para isso há necessidade de existir uma base de dados modelada adequadamente, que realmente seja capaz de representar uma realidade, principalmente considerando-se as relações sistêmicas envolvidas nos dados de um diagnóstico”.
- “Creio que não apenas em casos bem sucedidos. O contra exemplo (casos de insucessos) também representa conhecimento, o qual em sua natureza também permite descobrir (aprender) novos. Banco de dados, constituídos a partir de operações em um domínio, constituem uma base de dados (positiva ou negativa) a qual é pragmática tanto para a descoberta de conhecimento como para o raciocínio baseado em casos (CBR)”.
- “Concordo em caso de patologias raras e/ou “novas”, quanto ao fato de gerar conhecimento novo. Em casos clínicos clássicos o banco de dados pode tanto dar segurança ao profissional quanto ao diagnóstico, quando gerar uma acomodação ao estudante da especialidade que no processo de formação do conhecimento encontra pronto as respostas, não precisando pensar as informações. Penso que “novos conhecimentos” podem acontecer quando de patologias raras ou “novas”, onde os dados se repetem, mas pela escassez de incidência o profissional especialista “esquece” de que o fato, o dado já ocorreu anteriormente”.
- “Com certeza. Por exemplo: estão surgindo evidências de benefício no uso de aspirina para redução de mortalidade por

câncer. Evidências estas surgidas de análise de banco de dados”.

Na sexta questão, procurou-se identificar se o especialista acredita que, fazendo uso de sistemas de descoberta baseada na literatura, é possível obter indicações que revelem a existência de correlações entre elementos até então encobertos e que sejam úteis no aprimoramento de um domínio específico.

Entre as sugestões, justificativas e comentários apresentados, destacam-se:

- “O uso de associação e o estabelecimento de correlações me parece muito útil, pois se assemelha a forma de raciocínio do especialista, que sempre compara algo novo com algo aprendido ou com casos antigos. Assim, a descoberta baseada na literatura fornece a capacidade de compararmos um caso novo até mesmo com casos ou informações que ainda não tivemos a oportunidade de vivenciar na rotina clínica, enriquecendo nosso conhecimento e melhorando nossa capacidade de decisão”.
- “Somente baseado em literatura penso que seria pouco, para aprimorar um domínio específico, a vivência profissional é muito importante”.
- “A atualização dos conhecimentos é tão rápida que temos dificuldades em acompanhar toda a literatura médica. Se estes dados podem estar atualizados num sistema, facilitaria muito a nossa vida”.
- “Acredito que a capacidade de análise e cruzamento de dados, incluindo fontes diferentes, pode ampliar o universo que o especialista se encontra. Essa ampliação se dá de forma mais rápida e com fidedignidade, o que permite a descoberta de novos conhecimentos, de possíveis interações, que até então não foram observadas, pensadas ou mesmo denotadas pelo especialista, já que a capacidade de inter-relação da máquina é muito superior do que a da mente humana. Entretanto, a máquina não possui capacidade de julgamento sobre a qualidade da referência utilizada, e mesmo que só se use literatura de alto conceito, os métodos e metodologias utilizadas em um estudo, mesmo que corretos, podem não ser adequados para uma conclusão aplicável a realidade do caso avaliado pelo

especialista. Da mesma maneira, questões culturais também não têm como ser avaliadas dessa forma. Por isso, as descobertas podem ser indícios, mas não necessariamente uma verdade”.

- “A literatura nos permite, como registro de conhecimento obtido e da formalidade do mesmo, disponibilizar, para uso da sociedade do produto da dinâmica do pensar e do fazer em áreas do conhecimento. Dado o autor, a fonte, etc..., ela é aceita como referência de assuntos em dado domínio. Utilizar a mesma em processos de KD ou mesmo em ES (*expert system*) para sustentar uma colocação, é útil e viável, pois além de reforçar (enriquecer) a análise/síntese, permite validar o conhecimento registrado”.
- “No diagnóstico da Acupuntura isto é relevante, pois o número de dados e variáveis que se inter-relacionam são inúmeros e muitas vezes se perdem (são “esquecidos” pelo profissional)”.

Na sétima questão, procurou-se identificar se o especialista acredita num sistema que possibilite rastrear relações (associações) entre elementos, dentro de trabalhos em diferentes especialidades e que lhe dará maior facilidade em descobrir novos conhecimentos.

Entre as sugestões, justificativas e comentários apresentados, destacam-se:

- “Com certeza, este sistema terá capacidade de burlar os obstáculos de um cotidiano corrido, ou a distância que algumas vezes dificulta o contato de dois especialistas. E ao descobrir novos conhecimentos, ainda atento para o fato de que o sistema pode melhorar e favorecer a habilidade de diagnóstico de um profissional que não seja tão especialista ou experiente no domínio em questão”.
- “Acredito que ainda exista troca de conhecimento entre pesquisadores, mas claro, um sistema que contribua esta troca é válida”.
- “A capacidade de análise de um sistema para descoberta de conhecimento é extremamente superior a de que um homem, neste caso o especialista, é capaz. Claro que o sistema em si não é capaz de analisar a qualidade, aplicabilidade e utilidade dos conhecimentos descobertos, situação esta em que o especialista terá que aplicar seu conhecimento implícito para selecionar os novos conhecimentos descobertos. Dessa forma, acredito na

capacidade de novos conhecimentos, porém tenho restrições quando a afirmação fala sobre a descoberta na interdisciplinaridade. Claro que o sistema será capaz de descobrir, porém não sei o quanto o conhecimento descoberto será capaz de ser aplicável se o envolvimento do especialista que fizer a utilização da descoberta também não for interdisciplinar”.

- “Permitir uma visão sistêmica e holística sobre determinado tema, leva a compreensão, ao entendimento do mesmo e as possíveis ações sobre ele, de uma maneira eficaz. Hoje, para efeito de solução, não há como desconsiderar os vários modos de interpretações sobre o fenômeno que está sendo analisado”.
- “O conhecimento realmente precisa ser associativo caso contrário ele torna-se pouco eficiente. Por exemplo: ortopedista tratando uma infecção óssea e sem domínio adequado de microbiologia ou do uso de antibióticos terá menos eficiência na sua intervenção. Os exemplos na prática médica são infindáveis”.

APÊNDICE III – FERRAMENTAL

➤ FERRAMENTAS PARA DESENVOLVIMENTO DE ONTOLOGIAS

O número de ambientes e ferramentas para a construção de ontologias tem crescido muito. Estas ferramentas têm o objetivo de auxiliar no processo de desenvolvimento da ontologia e também no seu reuso.

Dentre as pesquisadas, as mais relevantes são:

- Ontolingua Server, Ontosaurus e WebOnto que foram os primeiros editores de ontologias;
- OntoKEM é uma ferramenta para Engenharia do Conhecimento baseada no processo de desenvolvimento da metodologia 101 e nos artefatos documentais das metodologias para ontologias On-to-Knowledge e Methontology devido ao detalhamento que proporciona na estruturação das ontologias e possibilidade de reuso (RAUTENBERG *et al.*, 2008).
- Protégé, WebODE e OntoEdit são uma nova geração de ambientes para o desenvolvimento de ontologias;
- OILED e DUET são ferramentas que emergiram para o desenvolvimento de ontologias na Web Semântica.

O Protégé foi desenvolvido pelo Grupo de Informática Médica da Universidade de Stanford e possui versões atualizadas constantemente. Seu núcleo é o editor de ontologias e possui uma vasta biblioteca de *plugins* que adicionam maior funcionalidade ao ambiente.

Atualmente os *plugins* disponibilizam a importação e a exportação de linguagens de ontologias (FLogic, Jess, OIL, XML e Prolog), acesso a bases de dados, criação e execução de restrições e fusão de ontologias.

A ferramenta Protégé tem seu código aberto, com uma interface gráfica para edição de ontologias baseadas em conhecimento. Isto contribui para o entendimento do sistema por parte do seu desenvolvedor e possibilita o compartilhamento da estrutura de informação entre especialistas, desenvolvedores e máquinas, além de facilitar a reutilização do conhecimento de domínio, explicitar conceitos consensuais, separar o domínio do conhecimento do conhecimento operacional e analisar o domínio (NOY e McGUINNESS, 2001).

➤ FERRAMENTAS PARA DESENVOLVIMENTO DE REDES BAYESIANAS

Citam-se, como exemplos de ferramentas para construção de redes Bayesianas, os *softwares*:

- BAYESWARE: www.bayesware.com

BAYESWARE foi o primeiro programa de Descoberta de Conhecimento comercial baseado em redes Bayesianas. O *software* Bayesware é uma ferramenta automática de modelar que transforma um banco de dados em uma rede de dependências, procurando o modelo mais provável para os dados observados.

- JAVABAYES: www.cs.cmu.edu/~javabayes/Home

O *software* JavaBayes é formado por um conjunto de ferramentas para a criação e manipulação de redes Bayesianas. O sistema compõe-se de um editor gráfico, uma máquina de inferência central e um conjunto de “*parsers*”.

- NETICA: www.norsys.com

O *software* Netica, desenvolvido pela Norsys, contém a estrutura para a construção de rede Bayesiana mais usada no mundo. Esta rede foi projetada para ser executada de forma simples, segura e com alto desempenho.

➤ FERRAMENTAS PARA DESCOBERTA BASEADA NA LITERATURA

Citam-se, como exemplos de ferramentas para a Descoberta Baseada na Literatura, os seguintes sistemas:

- ARROWSMITH: <http://arrowsmith.psych.uic.edu>

Este sistema procura artigos ou conceitos que podem estar em comum entre dois jogos distintos de artigos. Em outro contexto, este modo de procura pode ser utilizado quando se quer achar uma informação que está presente em um domínio do conhecimento e que pode ser pertinente a outro campo de investigação.

A procura será feita por artigos biomédicos no banco de dados do MEDLINE: a primeira procura define “literatura A” e a segunda define a “literatura C”. O programa gera uma “lista B” de palavras e frases encontradas nos títulos de ambas as literaturas.

A “lista B” pode ser exibida de acordo com a relevância e restrita às categorias semânticas, por exemplo, medicamentos, doenças, funções anatômicas. Para cada termo de interesse é possível ver os artigos que contém “A e B” ou da mesma forma, os artigos que contém “B e C”. Desta forma pode-se avaliar a presença de algo em comum, biologicamente significante na relação entre dois artigos.

- BITOLA: <http://ibmi.mf.uni-lj.si/bitola/>

BITOLA é um sistema interativo de apoio à descoberta biomédica baseada na literatura desenvolvido por Hristovski *et al.* (2005). A finalidade do sistema é ajudar os pesquisadores a fazer potencialmente novas descobertas através de relações entre um conceito inicial de interesse e outro conceito. A procura por esta nova relação ocorre através da mineração na base de dados bibliográfica MEDLINE.

A característica deste sistema é a utilização de regras de associação entre pares de conceitos como método para descobrir novas relações de acordo com a literatura médica.

- BIOSEARCH: <http://luar.dcc.ufmg.br/BioSearch/>

BIOSEARCH é um sistema de descoberta baseado na literatura desenvolvido pelo W. Maciel e colegas de trabalho no Programa de Pós-Graduação em Bioinformática da Universidade Federal de Minas Gerais (UFMG). Este sistema foi projetado para ajudar os pesquisadores acharem interações entre entidades biológicas pela procura em uma base de dados. O banco de dados atual da BIOSEARCH é formado pela reunião entre a patente dos Estados Unidos e o escritório de comércio e marca (USPTO).

APÊNDICE IV - GUIA DE UTILIZAÇÃO DO MODELO

1 – Construa a ontologia de domínio

Segundo Gómez-Pérez (1999), a ontologia (conhecimento explícito formalizado) é composta por conceitos (classes e subclasses), relações, propriedades, regras e instâncias que formam uma estrutura para a base de conhecimento. Construa então a ontologia com auxílio de especialistas, livros, manuais e apontamentos.

Utilize a conceituação proposta por Bunge (2003), através da modelagem CSM (*Composition – Environment – Structure – Mechanism*), para obter uma visão sistêmica do domínio. A explicitação do conhecimento ocorrerá de forma mais descritiva e abrangente para compor o domínio específico.

A figura1, por exemplo, apresenta o diagrama CSM com alguns elementos para compor a ontologia, dentro de um domínio na área da saúde.

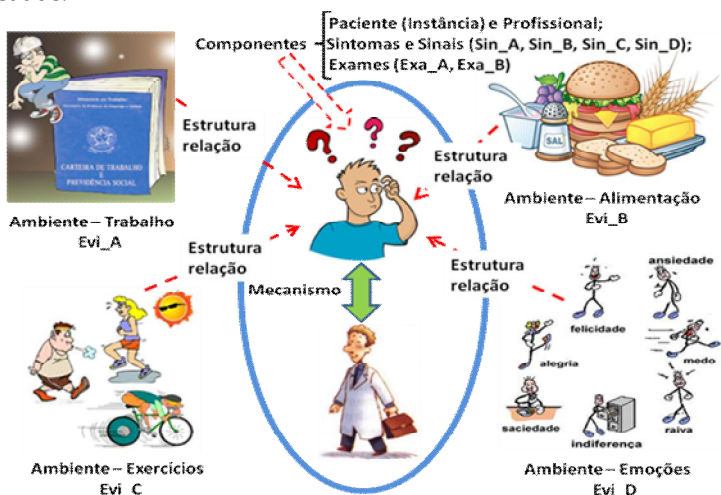


Figura 1: Diagrama CSM

A hierarquia da ontologia gerada a partir da ferramenta *Protégé* é apresentada na figura 2. Nesta ontologia, as instâncias representam os indivíduos que são diagnosticados em Solução X – Sol_X, Solução Y – Sol_Y e Solução Z – Sol_Z, de acordo com os Sintomas (Sin_A, Sin_B, Sin_C e Sin_D) e Exames (Exa_A e Exa_B) apresentados.

As Evidências (Evi_A, Evi_B, Evi_C e Evi_D) que caracterizam cada instância (paciente), permitem que novas descobertas sejam feitas para gerar novos conhecimentos.

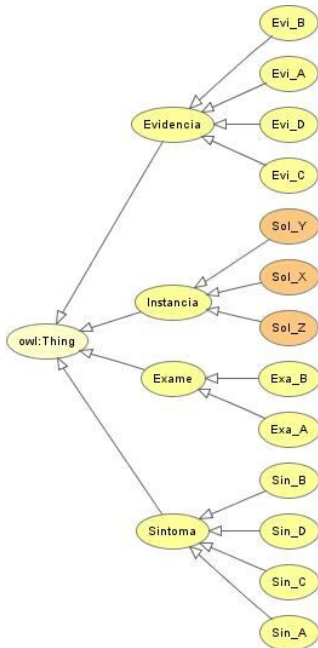


Figura 2: Classes e subclasses da ontologia

Ferramenta que pode ser utilizada para construção da ontologia:

- *Protégé* - <http://protégé.stanford.edu/download/download.html>

2 – Realize o diagnóstico utilizando a base de conhecimento

A ontologia estruturada permite que as instâncias sofram inferências, isto é, que sejam diagnosticadas ou classificadas em uma determinada classe. Portanto, ao inserir uma instância na base de conhecimento, tem-se, a partir do processo de inferência, o diagnóstico. Se não for possível obtê-lo, a partir da base de conhecimento, o diagnóstico será alcançado através da base de dados que está sendo formada com auxílio dos diagnósticos realizados.

3 – Construa a base de dados

É necessário, aqui, confirmar se o diagnóstico foi bem sucedido para uma instância. Caso confirmado, os exames, sintomas e evidências da instância devem ser inseridos numa base de dados (quadro 1) para proporcionar consultas futuras. Caso não seja confirmado, o item 5 deste manual oferece uma solução complementar.

Diag_F	D_Sin	D_Evi	Sin_A	Sin_B	Sin_C	Sin_D	Exa_A	Exa_B	Evi_A	Evi_B	Evi_C	Evi_D
Pat_X	Pat_X	Pat_X	sim	sim	sim	nao	sim	nao	sim	nao	sim	sim
Pat_X	Pat_X	Pat_X	sim	sim	sim	nao	sim	nao	sim	nao	sim	sim
Pat_Y	Pat_Y	Pat_Y	sim	nao	sim	nao	nao	sim	nao	sim	sim	nao
Pat_Y	Pat_Y	Pat_Y	sim	nao	sim	nao	sim	sim	nao	sim	sim	nao
Pat_Y	Pat_Y	Pat_Y	sim	nao	sim	nao	sim	sim	sim	sim	nao	sim
Pat_Y	Pat_Y	Pat_Y	sim	nao	sim	nao	nao	sim	nao	sim	nao	sim
Pat_Y	Pat_Y	Pat_Y	sim	nao	sim	sim	sim	sim	nao	sim	nao	sim
Pat_Z	Pat_Z	Pat_Z	nao	sim	nao	sim	nao	sim	nao	sim	sim	nao
Pat_Z	Pat_Z	Pat_Z	nao	sim	nao	sim	nao	sim	nao	sim	nao	sim
Pat_Z	Pat_Z	Pat_Z	nao	sim	nao	sim	nao	sim	sim	nao	sim	sim

Quadro 1 – Base de Dados

4 – Construa a rede causal

A rede causal para as classes e subclasses da ontologia é apresentada na figura 3. Esta rede, alimentada pela base de dados, quando instanciada, permite obter diagnósticos probabilísticos bem como destacar informações relevantes para pesquisa. O diagnóstico final (Diag_F) desta rede, é dado pela composição ponderada entre o diagnóstico dado pelos sintomas e exames (D_Sin) e pelo diagnóstico dado pelas evidências (D_Evi).

Ferramenta que pode ser utilizada para construção da rede causal:

- *Netica* – disponível em: <http://www.norsys.com/download.html>

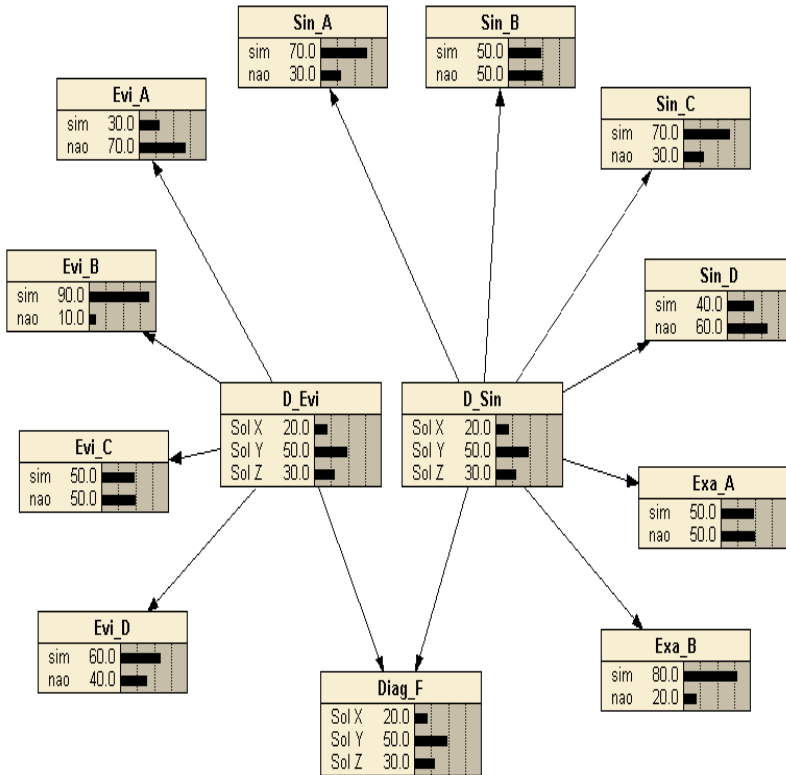


Figura 3: Rede causal (Bayesiana)

5 – Realize o diagnóstico utilizando a base de dados

A rede causal será utilizada se o processo de raciocínio da ontologia não for capaz de fornecer o diagnóstico para uma instância. Com os dados e informações obtidos das instâncias, a rede causal auxilia o processo de diagnóstico fornecendo um resultado probabilístico (causa → efeito).

A figura 4 exibe a rede instanciada para um determinado indivíduo que apresenta: Sin_A, Exa_A, Evi_A, Evi_B e Evi_D.

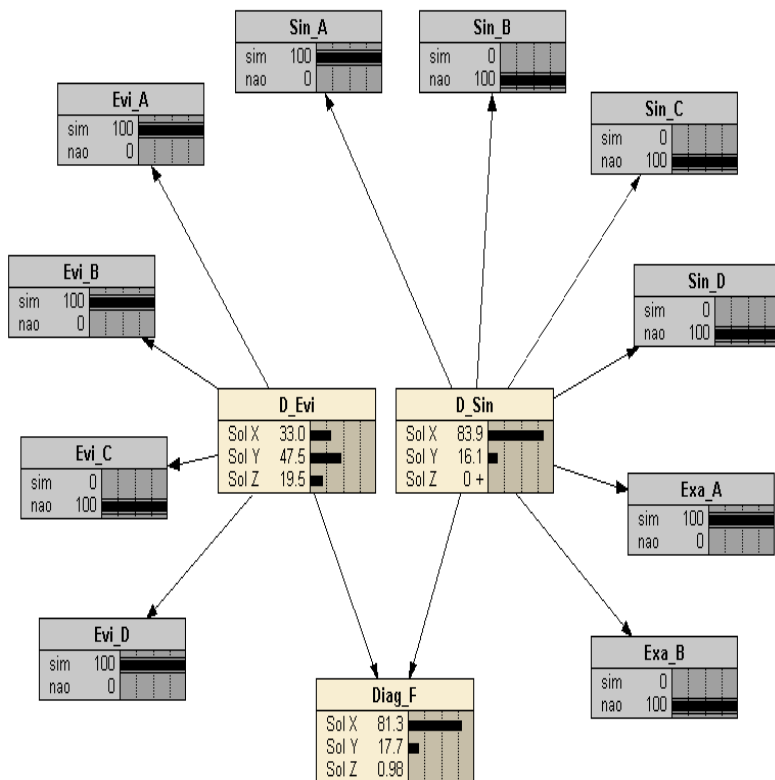


Figura 4: Diagrama probabilístico

6 – Realce evidências

Faça o caminho inverso do diagnóstico, isto é, do efeito para a causa, utilizando a base de dados. Desta forma, a rede causal fornece informações probabilísticas relevantes sobre os dados armazenados.

Na figura 5 está sendo apresentada a Evi_B, em destaque, quando a rede é instanciada para a Sol_X. Esta informação revela que em todos os casos de diagnóstico Sol_X foi encontrada a característica Evi_B.

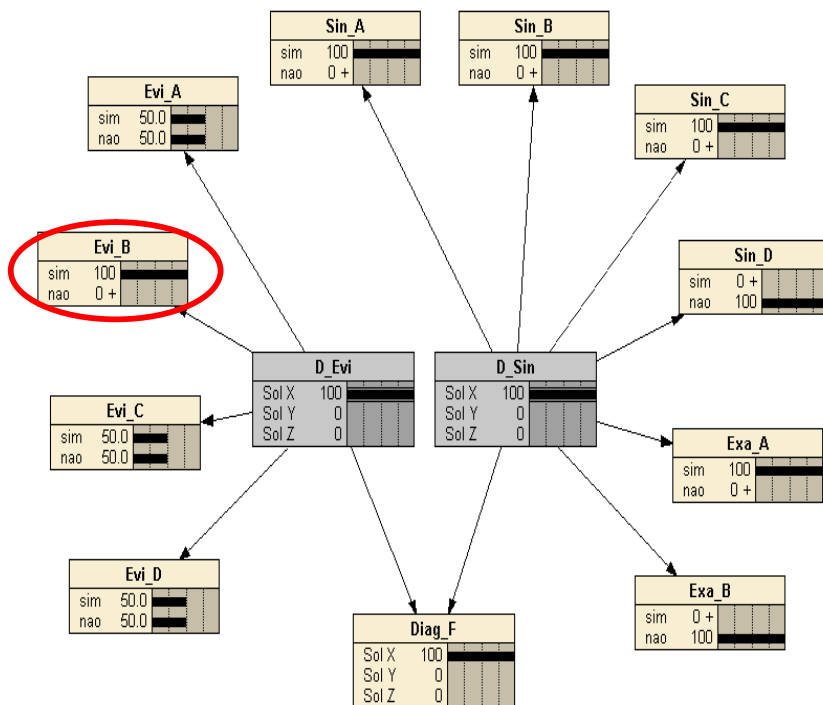


Figura 5: Rede instanciada

7 – Pesquisa em bases textuais

De posse de uma evidência para uma determinada solução, inicie a pesquisa para comprovar a descoberta de um possível novo conhecimento dentro do domínio. Esta evidência (Evi_B) será associada semanticamente com a solução (Sol_X) em arquivos literários para verificar a força de relação existente entre elas (AHLERS *et al.*, 2007; HRISTOVSKI *et al.*, 2006). Como resultados desta busca, serão apresentados trabalhos que deverão ser analisados para confirmar ou não a descoberta de um novo conhecimento.

Ferramenta que pode ser utilizada, por exemplo, na área da saúde, para descobertas baseada na literatura (figura 6):

- Bitola – disponível em: <http://ibmi.mf.uni-lj.si/bitola/>

BITOLA - Biomedical Discovery Support System: Closed Discovery

(Program authors: Dimitar Hristovski, Borut Peterlin)

Find Starting Concept X

Concept:

Semantic Types:

Find End Concept Z

Concept:

Semantic Types:

Figura 6: Tela inicial da ferramenta BITOLA

8 – Atualize a base de conhecimento

Depois de confirmada a descoberta de um novo conhecimento, este deverá ser inserido nas regras que definem as restrições para que um indivíduo pertença a uma determinada classe. A figura 7 mostra a atualização feita na base de conhecimento.

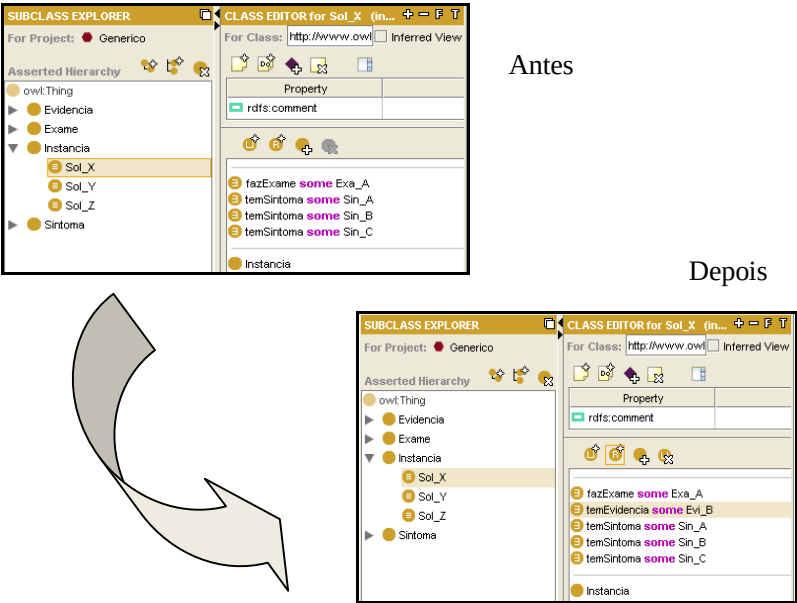
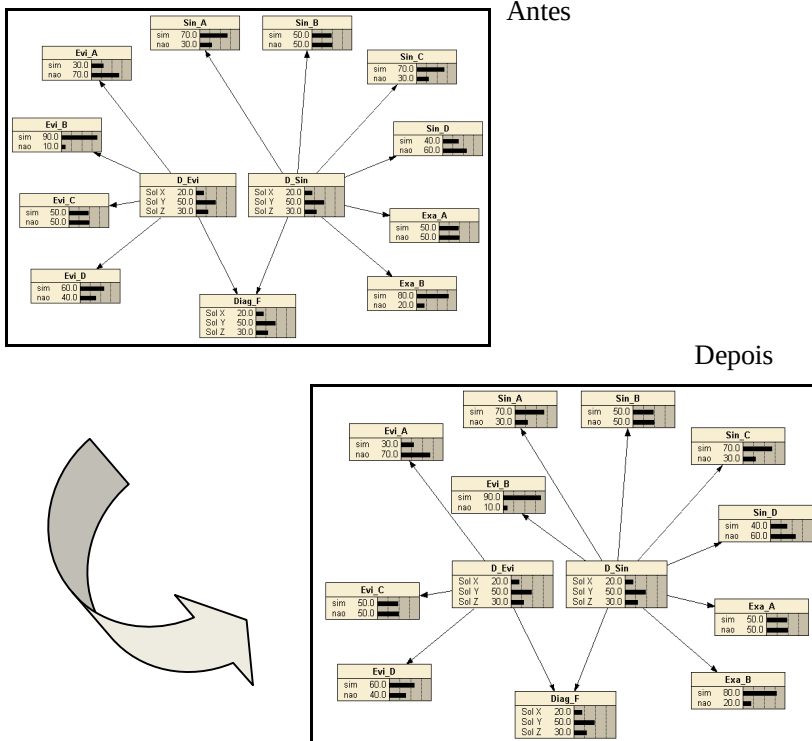


Figura 7: Atualização da base de conhecimento

9 – Atualize a Rede Causal

A rede causal também deverá ser atualizada com a descoberta de um novo conhecimento, para tanto, basta incluí-lo no diagnóstico parcial que contém maior grau de influência no resultado final. A figura 8 exibe esta atualização na rede.



APÊNDICE V - PUBLICAÇÃO EM REVISTA NACIONAL

**SISTEMA DE CONHECIMENTO PARA DIAGNÓSTICO EM
ACUPUNTURA: UMA MODELAGEM USANDO O
COMMONKADS**

*Knowledge System for acupuncture diagnosis:
a modeling using Commonkads*

Luiz Fernando Lopes¹
Maurício Capobianco Lopes¹
Francisco Antonio Pereira Fialho, Dr.¹
Alexandre Leopoldo Gonçalves, Dr.¹

1 Programa de Pós-Graduação em Engenharia e Gestão do
Conhecimento
Universidade Federal de Santa Catarina
Campus Universitário - Caixa Postal 476 - CEP 88040-900 -
Florianópolis, SC – Brasil

lflopes@utfpr.edu.br, mau.capo@egc.ufsc.br, fapfialho@gmail.com,
a.l.goncalves@stela.org.br

Resumo: A tomada de decisão em acupuntura é um processo complexo, uma vez que existe uma grande variedade de elementos e circunstâncias a serem considerados em um tratamento. Uma tarefa intensiva em conhecimento fundamental para o sucesso do tratamento é o diagnóstico. Assim, este artigo tem por objetivo apresentar uma proposta de Sistema de Conhecimento, utilizando a metodologia CommonKADS, com o objetivo de auxiliar na tarefa de diagnóstico em acupuntura, minimizando o grau de incerteza do tratamento e melhorando as condições de saúde e bem-estar do paciente. O modelo proposto é validado pela construção de um protótipo para diagnóstico de acne utilizando redes Bayesianas.

Palavras-chave: Engenharia do Conhecimento. Metodologia CommonKADS. Acupuntura. Ontologia. Diagnóstico. Sistemas de Conhecimento.

Abstract: *Decision making in acupuncture is a complex process, once there is a great variety of elements and circumstances to be considered during the treatment. The diagnosis, a knowledge-intensive task, is fundamental for the success of the treatment. Therefore, this study proposes a knowledge system based on the CommonKADS methodology aiming to support diagnosis tasks and to minimize the level of uncertainty in order to improve patient health conditions and welfare. The proposed model is validated through acne diagnosis prototype using Bayesian networks.*

Keywords: *Knowledge engineering. CommonKADS methodology. Acupuncture. Ontology. Diagnosis. Knowledge systems*

1. INTRODUÇÃO

A Medicina Tradicional Chinesa oferece diversas práticas de saúde voltadas ao bem estar do organismo, tanto no que diz respeito ao físico quanto ao mental. Uma das práticas mais conhecidas e utilizadas é a acupuntura que tem por finalidade a restauração do funcionamento equilibrado do organismo e a prevenção de patologias e afecções. Trata-se de “um método que estimula determinados pontos do corpo com agulhas, com o intuito de restaurar ou manter a saúde” (YAMAMURA, 2001).

A acupuntura proporciona uma visão sistêmica complexa sobre o organismo humano, composta de uma grande quantidade de agentes que se interrelacionam e que sofrem influência de elementos internos e externos. Este contexto sistêmico exige um conjunto de conhecimentos que muitas vezes dificulta a realização de um diagnóstico acurado de forma a oferecer um tratamento mais preciso e rápido aos pacientes.

Neste contexto, a Engenharia e a Gestão do Conhecimento dispõem de métodos e técnicas utilizadas na explicitação, uso e disseminação do conhecimento, que podem prover ao acupunturista um conjunto de recursos focados no apoio ao processo de tratamento de seus pacientes. Lukman *et al.* (2007), apresentam um estudo detalhado sobre métodos computacionais aplicados à Medicina Tradicional Chinesa, destacando bases de dados, ferramentas de mineração e sistemas de diagnóstico. Estes recursos podem apoiar tanto o diagnóstico quanto o acompanhamento do tratamento, visando restabelecer ou manter o equilíbrio sistêmico do organismo.

A necessidade de sistemas que apoiem a atuação do acupunturista é corroborada por Camilloti *et al.* (2006) que chamam a atenção para a necessidade do profissional em dispor de ferramentas baseadas em recursos da tecnologia da informação e comunicação. Estas ferramentas têm o propósito de facilitar o processo decisório no contexto do tratamento de um paciente, uma vez que a acupuntura está fortemente baseada na subjetividade, inexistindo indicadores precisos que permitam avaliar sua eficácia (CAMILLOTTI *et al.*, 2006).

Assim, a construção de um Sistema de Conhecimento neste contexto, deve ser realizada através de uma abordagem estruturada, com uma visão organizacional e sistêmica. Um dos recursos disponibilizados pela Engenharia do Conhecimento para esta finalidade é a metodologia CommonKADS (SCHREIBER *et al.*, 2002).

Sendo assim, o objetivo deste artigo é apresentar a modelagem de um Sistema de Conhecimento para apoio ao diagnóstico de pacientes de acupuntura, utilizando a metodologia CommonKADS. Este sistema visa oferecer ao acupunturista ferramentas adequadas para realizar o diagnóstico com maior eficácia e precisão, acelerando este processo e, consequentemente proporcionando um tratamento mais adequado e oferecendo uma melhor qualidade de vida e saúde ao paciente. Para a validação do modelo proposto é apresentado um protótipo para diagnóstico de acne utilizando redes bayesianas.

2. ACUPUNTURA

A acupuntura desperta cada vez mais interesse pelas chamadas práticas complementares ou alternativas nos serviços de saúde no Brasil e encontra-se regulamentada, em termos da sua utilização, pela portaria nº 971/2006 do Ministério da Saúde.

A palavra acupuntura é originária do latim, no qual *acus* significa agulha e *punctura* significa, literalmente, picar. Consiste, portanto, em inserir agulhas em pontos estratégicos na pele para obter uma resposta terapêutica eficiente.

Utilizando uma linguagem que associa as leis da natureza e a perfeita harmonia entre as partes do organismo, a acupuntura busca a integridade entre o homem e o meio em que vive. O fundamento está nas teorias do *Yin-Yang* (divisão do mundo em duas forças interpretando todos os fenômenos complementares), dos cinco movimentos (água, madeira, fogo, terra e metal) e dos sistemas internos. As forças *Yin-Yang* são reguladas por um princípio chamado *Tao* (taoísmo - caminho) onde,

através de uma interação dinâmica entre elas, procura-se um equilíbrio, ou seja, a saúde (YAMAMURA, 2001).

De acordo com o taoísmo, quando uma das forças se destaca sobre a outra, ocorre um desequilíbrio e surgem condições para o adoecimento (ERNST e WHITE, 2001). Diversos fatores podem ocasionar o desequilíbrio, como por exemplo: condições climáticas extremas, emoções excessivas (alegria, tristeza ou raiva), alimentação não balanceada, excesso de trabalho, sedentarismo, atividade física extenuante, excesso ou falta de sexo, etc. (CHONGHUO, 1993).

Para que este desequilíbrio não ocorra, a acupuntura utiliza um conjunto de procedimentos terapêuticos que visam introduzir estímulos, geralmente pela inserção de agulhas filiformes metálicas, em certos lugares anatomicamente definidos, a fim de obter do organismo a recuperação da saúde ou a prevenção de doença (NASIR, 2002).

A acupuntura é um tratamento de baixo custo e de grandes resultados benéficos, promovendo homeostasia, causando uma melhora na circulação do sangue no cérebro e controlando a dor no sistema nervoso central (SIEDENTOPF, 2005). Como desvantagens ou problemas do tratamento com acupuntura, destacam-se o medo da perfuração por agulhas e da contaminação e o preconceito ainda existente da acupuntura como ciência.

Entretanto, com o crescente aumento do uso deste tratamento pela população, cresce também o interesse da comunidade científica no sentido de comprovar suas propriedades terapêuticas, bem como a validação dos processos metodológicos de diagnóstico e tratamento.

O diagnóstico é uma das tarefas mais complexas na Medicina Tradicional Chinesa (MTC). Segundo Maciocia (2005), o processo de diagnóstico “baseia-se no princípio fundamental de que sintomas e sinais refletem as condições dos órgãos internos e dos canais”, podendo não estar necessariamente relacionados com o processo real de uma determinada doença. Portanto, a investigação para o processo de diagnóstico deve estender-se para além de sintomas e sinais relacionados com a queixa (MACIOCIA, 2005).

Para Lemos (2006), no processo de diagnóstico, o interrogatório do paciente, ou anamnese é uma etapa fundamental, pois busca sintomas subjetivos explanados pelo doente. Por ser uma tarefa de uso intensivo do conhecimento, a Engenharia e Gestão do Conhecimento tem muito a contribuir para aprimorar ainda mais este processo em acupuntura.

3. ENGENHARIA DO CONHECIMENTO E A METODOLOGIA COMMONKADS

A Engenharia do Conhecimento é uma área que tem por objetivo investigar e propor modelos de representação de conhecimento. Inicialmente tratada como uma subárea da Inteligência Artificial na construção de sistemas de conhecimento para a solução de problemas específicos, a Engenharia do Conhecimento transcendeu esta visão ao considerar todo o contexto sistêmico organizacional das atividades intensivas em conhecimento.

Assim, a Engenharia do Conhecimento deve prover um conjunto de ferramentas que forneçam suporte à gestão do conhecimento a partir da formalização e explicitação das atividades intensivas em conhecimento nas organizações (SCHREIBER *et al.*, 2002). Para isto, Studer *et al.* (1998) descreve como fundamental a aplicação de metodologias, teorias, métodos, técnicas e ferramentas voltadas à modelagem do conhecimento e sua conseqüente representação em sistemas de conhecimento.

Neste contexto, destaca-se a metodologia CommonKADS (SCHREIBER *et al.*, 2002), a qual oferece um conjunto de etapas para o desenvolvimento de sistemas baseados em conhecimento. A metodologia é baseada em quatro princípios que fundamentam a nova Engenharia do Conhecimento (SCHREIBER *et al.*, 2002):

- a) a Engenharia do Conhecimento aborda diferentes aspectos dos modelos humanos de conhecimento e não apenas a extração do conhecimento de um especialista;
- b) a modelagem do conhecimento deve estar focada inicialmente na estrutura conceitual e não nos detalhes de programação dos sistemas de conhecimento;
- c) o conhecimento tem uma estrutura interna estável, podendo ser categorizado em tipos ou papéis;
- d) um projeto de Sistema de Conhecimento deve ser construído de forma espiral, a partir de uma metodologia que permita este tipo de abordagem.

Além disso, três questões são fundamentais no processo de modelagem de um Sistema de Conhecimento utilizando-se da metodologia CommonKADS (SCHREIBER *et al.*, 2002): Por quê? O que? Como?. Para responder a estas questões, a metodologia é organizada em três níveis de modelos (SCHREIBER *et al.*, 2002):

- a) **nível de contexto:** tem por objetivo identificar os elementos fundamentais do ambiente e do contexto do Sistema de Conhecimento;
- b) **nível de conceito:** tem por objetivo explicitar e formalizar o conhecimento, bem como destacar como ocorrem as interações entre os agentes envolvidos em sua utilização;
- c) **nível de projeto:** tem por objetivo apresentar todos os detalhes necessários para a construção de um Sistema de Conhecimento.

A figura1 apresenta os modelos de cada um dos níveis do CommonKADS.

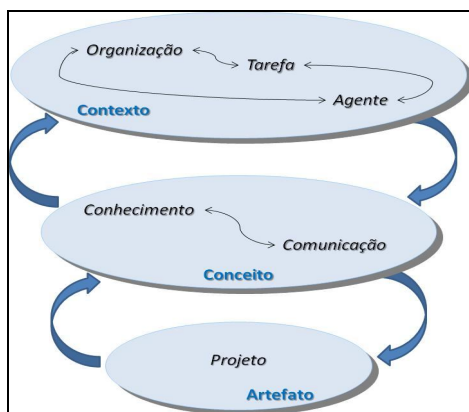


Figura1: Metodologia CommonKADS

Fonte: adaptação (SCHREIBER *et al.*, 2002)

O quadro 1 apresenta uma descrição mais detalhada dos níveis do CommonKADS, para a construção de sistemas de conhecimento, bem como sobre os modelos a serem construídos em cada um dos níveis.

Nível	Modelo	Objetivo
Contexto	Organização	Identificar e analisar o contexto organizacional onde o sistema estará inserido.
	Tarefa	Detalhar todas as tarefas envolvidas no sistema, com foco nas atividades intensivas em conhecimento.
	Agente	Identificar os agentes humanos e não humanos e suas responsabilidades no Sistema de Conhecimento.

Conceito	Conhecimento	Apresentar detalhadamente o modelo conceitual de conhecimento do sistema, destacando as tarefas, as inferências e o domínio para sua posterior implementação.
	Comunicação	Modelar a comunicação entre os agentes envolvidos no funcionamento do sistema.
Artefato	Projeto	Definir as especificações técnicas necessárias para a implementação e funcionamento do sistema, com base nas etapas anteriores,

Quadro 1 - Descrição dos modelos do CommonKADS (SCHREIBER *et al.*, 2002)

Dos três níveis apresentados, destaca-se o nível de contexto. Este é a principal contribuição do CommonKADS para a Engenharia do Conhecimento, uma vez que não está focado apenas na construção de um sistema, mas, sobretudo no entendimento do contexto e das relações sistêmicas e organizacionais onde o mesmo estará inserido. Uma das tarefas fundamentais e que de certa forma conclui este nível é o estudo de viabilidade e impacto tanto no que diz respeito ao negócio quanto às questões técnicas envolvidas.

A seguir será apresentado o desenvolvimento do Sistema de Conhecimento para acupuntura baseado na metodologia CommonKADS.

4. SISTEMA DE CONHECIMENTO PARA DIAGNÓSTICO EM ACUPUNTURA

Muitos já são os trabalhos desenvolvidos de sistemas de conhecimento para diagnóstico (BAUMEISTER, 2004; MILNE, 1987; LUO *et al.*, 2006). Segundo Kleer e Williams (1987), entende-se por diagnóstico o processo que visa determinar porque um sistema corretamente projetado não funciona conforme esperado. Assim, para realizar a tarefa de diagnóstico é necessário observar e levantar hipóteses de modo a chegar a um diagnóstico (resultado) descrevendo o estado do processo observado. De forma simplificada, Schreiber *et al* (2002) destaca que o objetivo do diagnóstico é encontrar a causa do mau funcionamento de um sistema.

Neste sentido, este trabalho apresenta um Sistema de Conhecimento para o diagnóstico em acupuntura, seguindo as etapas da metodologia CommonKADS. São apresentados, entretanto, apenas o

modelo de contexto (organização, tarefas e agentes) e o modelo de conceito (conhecimento e comunicação). O modelo de projeto não está contemplado neste trabalho.

4.1 MODELO DE ORGANIZAÇÃO

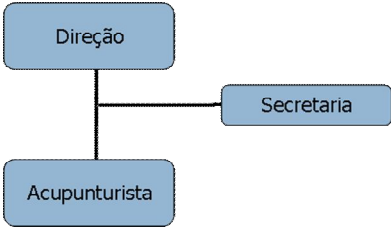
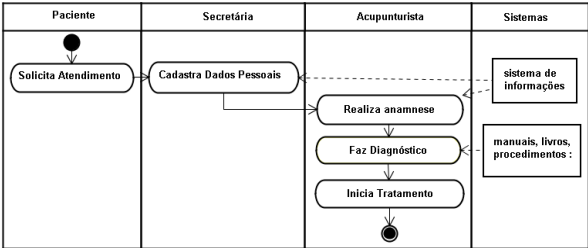
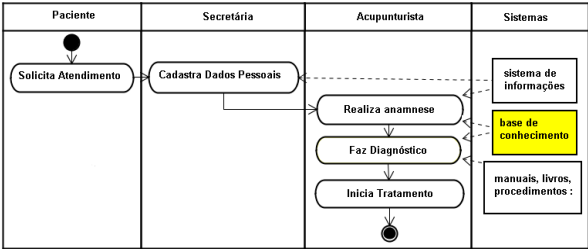
O modelo de organização é composto de um conjunto de planilhas com a finalidade de detalhar o contexto do Sistema de Conhecimento na organização. Estas planilhas serão apresentadas nos Quadros 2 a 6. A primeira planilha é denominada de MO1 – Problemas e Oportunidades (quadro 2). Nesta planilha o problema é tratado com base em uma visão organizacional para além do processo de diagnóstico, uma vez que a acupuntura envolve outras etapas.

Modelo Organizacional	MO1 - Problemas e Oportunidades
Problemas e oportunidades	Problemas: Dificuldade no diagnóstico e acompanhamento para o tratamento baseado na acupuntura. Grande quantidade de pontos existentes para aplicação das agulhas relacionados a diferentes patologias. Oportunidades: Maior precisão no diagnóstico diminuindo o grau de incerteza no tratamento. Definição de indicadores de acompanhamento do tratamento, visando um monitoramento mais preciso das condições globais de saúde do paciente Definição de indicadores globais para a organização onde o acupunturista atua de modo a se fazer a gestão do conhecimento sobre todo o processo organizacional envolvido como o tratamento dos pacientes. Disponibilização de uma base de conhecimento sobre tratamentos através da acupuntura. Melhoria da qualidade de vida do paciente. Auxílio na capacitação de novos profissionais nesta área.

Modelo Organizacional	MO1 - Problemas e Oportunidades
Contexto organizacional	A Acupuntura é normalmente realizada em clínicas autorizadas, por profissionais habilitados com registro no seu órgão de classe dependendo de sua área de formação. No Brasil, a atividade de acupuntura vem sendo reivindicada pelos diversos órgãos de classe que tratam da saúde, no sentido de ser uma especialidade específica de sua área. Assim, cada conselho adota seus critérios e normas específicos para regulamentar a atuação do profissional. Muitas clínicas de saúde utilizam a acupuntura como um tratamento alternativo ou complementar aos tratamentos convencionais, buscando conquistar novos clientes, interessados nesta técnica. Um acupunturista pode atender vários pacientes em uma mesma sessão, o que por um lado amplia suas possibilidades de ganho financeiro, mas por outro lado, dificulta o acompanhamento individual dos pacientes.
Soluções	Estruturar um Sistema de Conhecimento para o profissional de Acupuntura.

Quadro 2 - MO1: Problemas e oportunidades

A planilha seguinte é denominada de MO2 – Aspectos Variantes (quadro 3). Nela são descritos elementos organizacionais que devem ser considerados pelo engenheiro do conhecimento. Como o foco do trabalho é no diagnóstico, são considerados apenas os elementos relevantes no contexto desta tarefa.

Modelo Organizacional	MO2 – Aspectos Variantes
Estrutura	A estrutura de uma organização onde atua um profissional acupunturista pode ser representada conforme a figura 2. A Direção só é relevante no contexto de uma clínica.  Figura 2 - Estrutura básica de uma organização onde atua um profissional acupunturista
Processos	São apresentados aqui dois processos do diagnóstico. Inicialmente é apresentado o processo desconsiderando a disponibilidade de um Sistema de Conhecimento (figura 3) e, em seguida, é apresentada a proposta de processo considerando a existência do mesmo (figura 4).  Figura 3 - Processo de Diagnóstico (atual)  Figura 4 - Processo de Diagnóstico (proposto)

Modelo Organizacional	MO2 – Aspectos Variantes
Pessoas	Os principais agentes envolvidos nos processos relatados com suas respectivas atribuições estão descritos a seguir: • Paciente: gera a demanda (problema/necessidade) para o acupunturista). • Secretária: obtém as informações preliminares do paciente. • Acupunturista: realiza o diagnóstico.
Recursos	Os recursos necessários na atividade do acupunturista são manuais e livros. Na proposta apresentada pretende-se incluir computador, sistema de informações e Sistema de Conhecimento.
Conhecimento	Os conhecimentos fundamentais para as atividades de acupuntura estão relacionados com as patologias e seus respectivos tratamentos. São conhecimentos complexos que envolvem um contexto sistêmico relacionado ao organismo humano. Normalmente é adquirido através de estudos profundos e da prática de sua utilização.
Cultura e Poder	O acupunturista deve buscar a confiança e empatia com o paciente na busca de uma relação harmônica, pois muitos pacientes têm receio do tratamento, uma vez que ele age diretamente sobre o corpo, com um instrumento associado a dor.

Quadro 3 - MO2: Aspectos Variantes

A terceira planilha é a MO3-Detalhamento dos Processos (quadro 4) onde são destacadas as atividades realizadas no contexto da tarefa analisada. O foco principal é identificar as atividades intensivas em conhecimento e determinar seu significado para o sistema.

Modelo Organizacional		MO3 - Detalhamento dos Processos				
Número (id)	Tarefa (Nome da tarefa)	Realizada por (agente)	Local?	Ativo de conhecimento	É intensiva em conhecimento?	Significado
1	Atendimento	Secretária	Sala de Espera		Não	

Modelo Organizacional		MO3 - Detalhamento dos Processos				
2	Anamnese	Acupunturista	Sala de Atendimento	Conhecimentos sobre os fatos relacionados com a doença e com o paciente.	Sim	
3	Diagnóstico	Acupunturista	Sala de Atendimento	Conhecimentos sobre as relações entre os sintomas do paciente e as possibilidades de tratamento.	Sim	

Quadro 4 - MO3: Detalhamento dos Processos

A quarta planilha é a MO4 – Ativos de Conhecimento (quadro 5). Nesta planilha o engenheiro procura identificar quais ativos de conhecimento os agentes têm sobre a tarefa.

Modelo organizacional		MO4 – Ativos de Conhecimento				
Ativo	Possuído por	Usado em	Forma correta?	Local correto?	Tempo correto?	Qualidade adequada?
Conhecimentos sobre os fatos relacionados com a doença e com o paciente.	Paciente	Anamnese Diagnóstico	Sim	Sim	Sim	Sim
Conhecimentos sobre as relações entre os sintomas do paciente e as possibilidades de tratamento.	Acupunturista	Diagnóstico	Sim	Sim	Sim	Depende da experiência do especialista

Quadro 5 - MO4: Ativos de Conhecimento

A quinta e última do modelo de organização é a MO5 – Checklist (quadro 6) onde o engenheiro estabelece os elementos que definem a viabilidade do Sistema de Conhecimento a ser construído.

Modelo Organizacional	MO5 - Checklist para a tomada de decisão de viabilidade
Viabilidade de Negócio	A acupuntura apresenta um mercado em expansão, com o crescente interesse de profissionais da área de saúde, pesquisadores e comunidade científica, os quais têm realizado diversos trabalhos científicos sobre a técnica, comprovando sua eficácia. Além disso, a técnica desperta interesse cada vez maior da população como uma forma de tratamento natural alternativo aos métodos tradicionais baseados na alopatia. Assim, a análise do contexto organizacional permite verificar a possibilidade de ganhos através da realização de um diagnóstico mais preciso e ágil, o que permite acelerar o tratamento e acompanhar mais pacientes de uma forma mais eficiente. Além disso, com a agilidade no tratamento, o paciente obterá melhor qualidade de vida com condições de recomendar o tratamento para outras pessoas.
Viabilidade Técnica	Tecnicamente, a construção de um Sistema de Conhecimento de apoio ao profissional acupunturista é viável, a partir da evolução da Engenharia do Conhecimento e das tecnologias de informação e comunicação. Como exemplos podem ser citados os modelos de representação do conhecimento e as redes bayesianas que possibilitam a modelagem e construção de um Sistema de Conhecimento com um nível de complexidade exigido pela área de acupuntura.
Viabilidade do Projeto	A execução do projeto implica no contato com especialistas para a explicitação e formalização do conhecimento. Como analisado anteriormente na viabilidade técnica, os recursos tecnológicos estão disponíveis requerendo recursos orçamentários de pequena monta para sua execução.
Ações Propostas	A partir do conhecimento de especialistas em acupuntura, pretende-se desenvolver um Sistema de Conhecimento para auxiliar no diagnóstico dos pacientes. O diagnóstico pode ser auxiliado através da criação e manutenção de bases de conhecimento.

Quadro 6 - MO5: Checklist para a tomada de decisão de viabilidade

4.2 MODELO DE TAREFA

O modelo de tarefa é composto por um conjunto de planilhas com a finalidade de detalhar as tarefas intensivas em conhecimento identificadas no modelo de organização. Neste trabalho o foco é apenas na tarefa de diagnóstico. A primeira planilha deste modelo é a TM1 – Análise de Tarefa (quadro 7) que destaca as principais características da tarefa analisada.

Modelo de Tarefa	TM1 - Análise de Tarefa
Tarefa	Diagnóstico
Organização	Corresponde a terceira etapa do processo de atendimento do paciente e é executada pelo acupunturista.
Objetivo ou valor agregado	Identificar as condições de saúde e os desequilíbrios do organismo do paciente.
Dependências e Fluxo	Tarefa de Entrada: anamnese Tarefa de Saída: tratamento
Objetos manuseados	Objetos de Entrada: dados do paciente, resultados da anamnese, livros e estudos sobre acupuntura. Objeto de Saída: diagnóstico da(s) patologia(s) e indicações de tratamento. Objetos internos: experiência do acupunturista.
Tempo e Controles	O diagnóstico é executado previamente ao tratamento, mas pode ser retomado se o tratamento não produzir os resultados esperados. Não há um tempo pré-estabelecido tanto para a duração quanto para a frequência em que ocorre esta tarefa. Pré-condições: a anamnese conseguir identificar adequadamente as condições do paciente Pós-condições: condições de iniciar o tratamento
Agentes	O Acupunturista faz o diagnóstico
Conhecimento e competência	O principal ativo de conhecimento para a realização desta tarefa é o conhecimento e a experiência do acupunturista. Outro fator relevante é a habilidade humana do acupunturista que deve procurar estabelecer empatia e conquistar a confiança do paciente.
Recursos	Os principais recursos quantificáveis utilizados para a realização da tarefa são os manuais e livros sobre acupuntura.

Modelo de Tarefa	TM1 - Análise de Tarefa
Qualidade e Desempenho	A medida de qualidade e performance desta tarefa depende da tarefa seguinte que é o tratamento, ou seja, se o tratamento alcançar os resultados esperados dentro de sua prescrição, a tarefa de diagnóstico foi executada adequadamente.

Quadro 7 - TM1: Análise de Tarefa

Na planilha TM2 – Itens de Conhecimento (quadro 8) o engenheiro faz um checklist sobre o conhecimento da tarefa e identifica pontos possíveis de gargalos e melhorias.

Modelo de Tarefa	TM2 - Itens de Conhecimento	
Nome: Realizar Diagnóstico.	Usado em: Tarefa 3 – Diagnóstico	
Pertence ao: Acupunturista	Domínio: Acupuntura	
Natureza do Conhecimento		Gargalos/Melhorias
Formal, Rigoroso		
Empírico, quantitativo	X	X
Heurístico, Regras	X	
Altamente especializada	X	
Baseado em experiência	X	X
Baseado em atividades		
Incompleto		
Incerto, pode estar incorreto	X	
Mudando rapidamente		
Difícil de verificar		
Tácito, difícil de transmitir		
Formato do conhecimento		Gargalos/Melhorias
Mente	X	X
Papel	X	
Formato eletrônico	X	
Habilidade na ação	X	
Disponibilidade do conhecimento		Gargalos/Melhorias
Limitações de tempo		
Limitações de espaço		
Limitações de acesso		
Limitações de qualidade		
Limitações de forma		

Quadro 8 - TM2: Itens de Conhecimento

4.3 MODELO DE AGENTE

O modelo de agente é composto por uma planilha (quadro 9) que tem a finalidade de apresentar os agentes envolvidos na tarefa. No problema em questão o único agente tratado é o acupunturista.

Modelo de Agentes	AM1 – Agente
Nome	Acupunturista
Organização	O Acupunturista normalmente está subordinado a uma direção em uma organização de saúde, mas tem autonomia para realizar a tarefa de diagnóstico. Trata-se de um agente humano.
Está envolvido em	2 – Anamnese 3 – Diagnóstico
Comunica-se com	Secretária Pacientes
Conhecimento	Acupuntura (diagnóstico)
Outras competências	Habilidade humana procurando estabelecer empatia e conquistar a confiança do paciente.
Responsabilidades e Regras	Responsabilidades: fazer o diagnóstico. Restrições: submete-se às normas e procedimentos definidos tanto pela organização onde está inserido quanto pelos conselhos de classe competentes.

Quadro 9 - AM1: Agente

Para consolidar o modelo de contexto, a planilha OTA1 (quadro 10) apresenta um conjunto de informações para auxiliar o processo decisório sobre a implantação do Sistema de Conhecimento proposto.

Modelo de Organização, Tarefa e Agente	OTA1 – Checklist para Decisão sobre Impactos e Melhorias
Impactos e mudanças na organização	Estrutura: não há impactos e mudanças necessárias na estrutura da organização. Processos: diagnóstico será realizado apoiado por um Sistema de Conhecimento, proporcionando maior eficácia e qualidade no processo. Recursos: será necessário a utilização de um Sistema de Conhecimento disponível em um computador. Pessoas: no acupunturista terá a seu dispor um Sistema de Conhecimento, não existindo outros impactos ou mudanças para as pessoas na organização. Conhecimento: o conhecimento que anteriormente estava apenas no domínio do especialista, passa a estar formalizado e explicitado através de um Sistema de Conhecimento e pode ser compartilhado entre especialistas, profissionais e estudantes da área. Cultura e poder: o Sistema de Conhecimento pode impactar positivamente para a organização na medida em que ele possibilitará gerar novos conhecimentos sobre os pacientes e sobre os processos de diagnóstico e tratamento. Permitirá maior compartilhamento de informações, possibilitando que um paciente possa ser acompanhado por mais de um profissional.
Impactos e mudanças na tarefa e nos agentes	O acupunturista, de posse de um Sistema de Conhecimento que auxilie no diagnóstico, poderá tomar decisões mais rapidamente e com maior precisão, reduzindo o grau de incerteza e o tempo de execução da tarefa. Com isso, com o mesmo tempo que ele tem disponível para atendimento na situação atual, ele poderá atender um maior número de pacientes, a partir da implantação da solução proposta. O sistema também pode ser utilizado como fonte de conhecimento e formação para novos profissionais que desejam atuar na área. Para os pacientes, espera-se que o tratamento torne-se mais eficaz, melhorando mais rapidamente sua qualidade de vida.
Atitudes e compromissos	O acupunturista deve se comprometer a utilizar o Sistema de Conhecimento, como fonte primordial para o seu processo decisório, fazendo sugestões e críticas para revisão e aperfeiçoamento do mesmo.

Modelo de Organização, Tarefa e Agente	OTA1 – Checklist para Decisão sobre Impactos e Melhorias
Ações propostas	Melhoria: implantar um Sistema de Conhecimento que auxilie o diagnóstico para o tratamento de um paciente de acupuntura. Acompanhamento: para a implantação adequada do sistema é fundamental a validação do sistema com especialistas e treinamento para utilização. Resultados esperados: espera-se agilizar o processo de tratamento, gerando com isto ganhos para a organização já destacados anteriormente. Os custos de implantação são relativamente baixos, em função da disponibilidade técnica e tecnológica que viabilizam a construção do Sistema de Conhecimento proposto.

Quadro 10 - OTA1: Checklist para Decisão sobre Impactos e Melhorias

A partir do modelo de contexto apresentado até aqui, serão propostos os modelos de conhecimento e comunicação de modo a viabilizar a implantação do sistema.

4.4 MODELO DE CONHECIMENTO

A construção do modelo de conhecimento envolve três etapas (SCHREIBER *et al.*, 2002) que serão detalhadas a seguir.

4.4.1 Identificação do Conhecimento

Na identificação do conhecimento duas atividades são fundamentais: a familiarização com o domínio e a lista de potenciais componentes de modelo para reuso. Para isto é fundamental que a tarefa intensiva em conhecimento e os principais itens de conhecimento estejam devidamente identificados, os quais foram apresentados em detalhe nas seções anteriores.

A familiarização com o domínio tem o objetivo de permitir o completo entendimento sobre o problema, através da consulta às diferentes fontes de conhecimento disponíveis.

A tarefa intensiva em conhecimento aqui tratada é o diagnóstico de patologias em um paciente de acupuntura, visando determinar a melhor forma de tratamento. Assim, para a realização deste trabalho, as fontes de conhecimento selecionadas foram um especialista em acupuntura e manuais e livros sobre o tema. Com o especialista foram

feitas entrevistas não estruturadas, de forma a se identificar os principais conhecimentos tácitos aplicados na tarefa de diagnóstico. Os manuais e livros foram utilizados para se conhecer os principais elementos e técnicas envolvidas na atividade de acupuntura, com ênfase nas informações mais importantes do processo de diagnóstico.

Na atividade seguinte realizou-se a identificação dos potenciais modelos de reuso para o diagnóstico nas dimensões de domínio e tarefa. Para o modelo de domínio optou-se por sua representação baseado em uma simplificação dos diagramas de classe da UML (*Unified Modeling Language*), pela facilidade de compreensão que os mesmos proporcionam. Para o modelo de tarefa foram estudados os *templates* propostos nos capítulos 5 e 6 de Schreiber *et al.* (2002). A partir destas definições, iniciou-se a etapa de especificação do conhecimento.

4.4.2 Especificação do Conhecimento

O objetivo desta etapa é obter a especificação completa do modelo de conhecimento, envolvendo tarefa, inferência e domínio. A primeira atividade é a seleção do modelo de tarefa a ser utilizado. A partir do estudo dos *templates* apresentados, foi selecionado o *template* do capítulo 5 proposto por Schreiber *et al.* (2002), sendo, necessário, entretanto, alguns ajustes detalhados a seguir.

A atividade seguinte é a de construção de um modelo inicial de conhecimento de domínio, que não será descrita aqui, pois em seguida, apresentaremos o modelo de domínio mais detalhado.

A próxima etapa é a construção completa do modelo de conhecimento, a qual será apresentada a seguir. A abordagem utilizada foi a *middle-out*, que, parte do modelo de inferência, para, em seguida, propor os modelos de tarefa e domínio. Esta é a abordagem mais indicada quando o modelo de tarefa é bem conhecido (SCHREIBER *et al.*, 2002).

Modelo de Inferência

No modelo de tarefa proposto são utilizadas as seguintes inferências:

- **cobrir:** esta inferência busca, através de um modelo de causas e efeitos, uma possível hipótese para um determinado problema ou reclamação. Assim, a partir de um estado

visível, será sugerida uma hipótese para um estado não visível, baseada nos modelos causais. A figura 5 apresenta a representação desta inferência de forma anotada para o exemplo de acupuntura. No exemplo, a partir da reclamação do paciente que tem um problema de acne (estado visível), esta inferência pode trazer uma hipótese que é a existência de acne em função do calor no estômago e pulmões (estado não visível);

- **especificar:** esta inferência especifica um possível estado observável, a partir de uma hipótese. Assim, a partir de um estado não visível, será sugerido um estado esperado para o estado não visível, considerando os modelos de manifestação para este estado não visível. A figura 6 apresenta a representação desta inferência de forma anotada para o exemplo de acupuntura. No exemplo, a partir da hipótese da inferência cobrir (acne em função do calor no estômago e pulmões), é gerado um resultado esperado (pele seca = verdadeiro) que esteja relacionado com a hipótese;
- **obter:** trata-se de uma função de transferência, através da qual o acupunturista observa o estado indicado na inferência especificar. A partir do exemplo da figura 6, ele deverá verificar se o paciente tem pele seca;
- **comparar:** esta inferência compara o resultado esperado com o desejado. Assim, a partir do resultado da inferência especificar, o acupunturista deve observar o estado sugerido, através da função de transferência obter, de forma a confirmar ou não a hipótese sendo analisada.

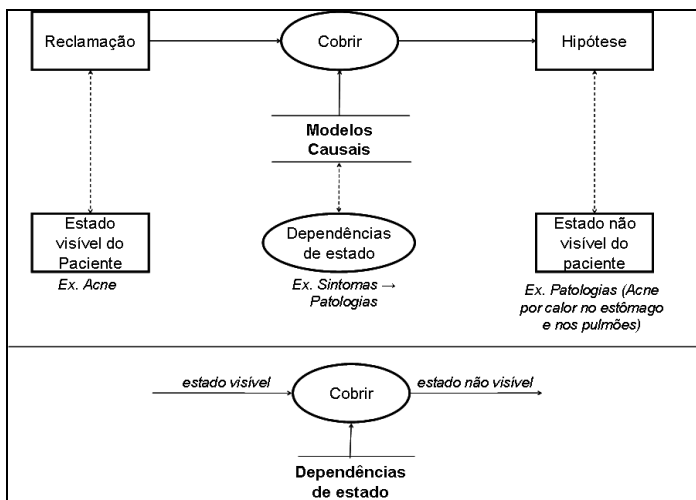


Figura 5 - Inferência Cobrir

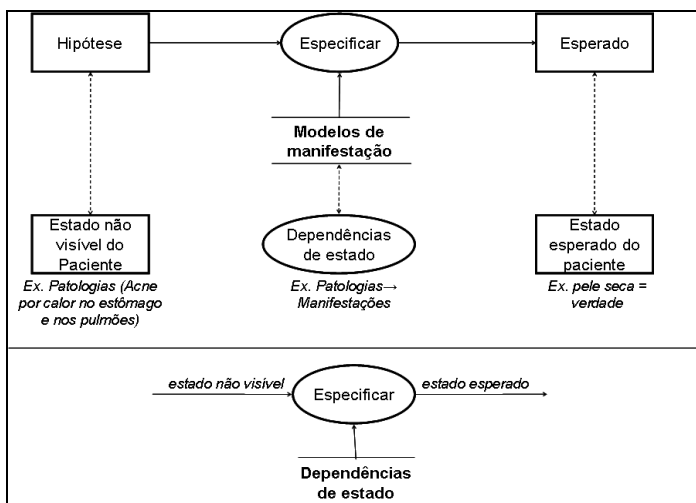


Figura 6 - Inferência Especificar

Modelo de Tarefa

O modelo de tarefa aqui utilizado apresenta modificações em relação ao proposto por Schreiber *et al.* (2002). A primeira alteração está na troca da inferência prever pela inferência especificar, tratada

anteriormente, uma vez que aquela é pouco adequada ao diagnóstico que se pretende realizar. Outra alteração foi proposta na estrutura de controle com a inserção de um laço de repetição para esta nova inferência. O método de tarefa de diagnóstico a partir de geração e teste está apresentado no quadro 11 com estas alterações destacadas em negrito.

```

MÉTODO-TAREFA diagnóstico-geração-e-teste;
  REALIZA: diagnóstico-paciente;
  DECOMPOSIÇÃO:
    INFERÊNCIAS: cobrir, especificar, comparar;
    FUNÇÃO DE TRANSFERÊNCIA: obter;
  PAPÉIS:
    INTERMEDIÁRIO:
      hipótese: "uma patologia provável";
      esperado: "A manifestação da patologia ser verdadeira";
      encontrado: "manifestação efetivamente observada";
      resultado: "O resultado da comparação da manifestação";
    ESTRUTURA DE CONTROLE
      ENQUANTO NOVA-SOLUÇÃO cobrir
(reclamação→hipótese);
      FAÇA
        resultado == verdadeiro;
        ENQUANTO NOVA-SOLUÇÃO especificar
(hipótese→esperado);
        FAÇA
          obter (esperado→encontrado);
          evidência := evidência + encontrado;
          comparar (esperado+encontrado→resultado);
          SE resultado == falso;
            ENTÃO "parar a repetição";
          FIM SE
        FIM ENQUANTO
        SE resultado == verdadeiro;
          ENTÃO "parar a repetição";
        FIM SE
      FIM ENQUANTO
      SE resultado == verdadeiro
        ENTÃO problema-identificado := hipótese;
        SENÃO "nenhuma solução encontrada";
      FIM SE
    FIM MÉTODO-TAREFA diagnóstico-geração-e-teste;

```

Quadro 11 - Modelo de Tarefa de Diagnóstico a partir de geração e teste

A figura 7 apresenta na forma de diagrama de atividades, a estrutura de controle que representa o modelo de tarefa proposto.

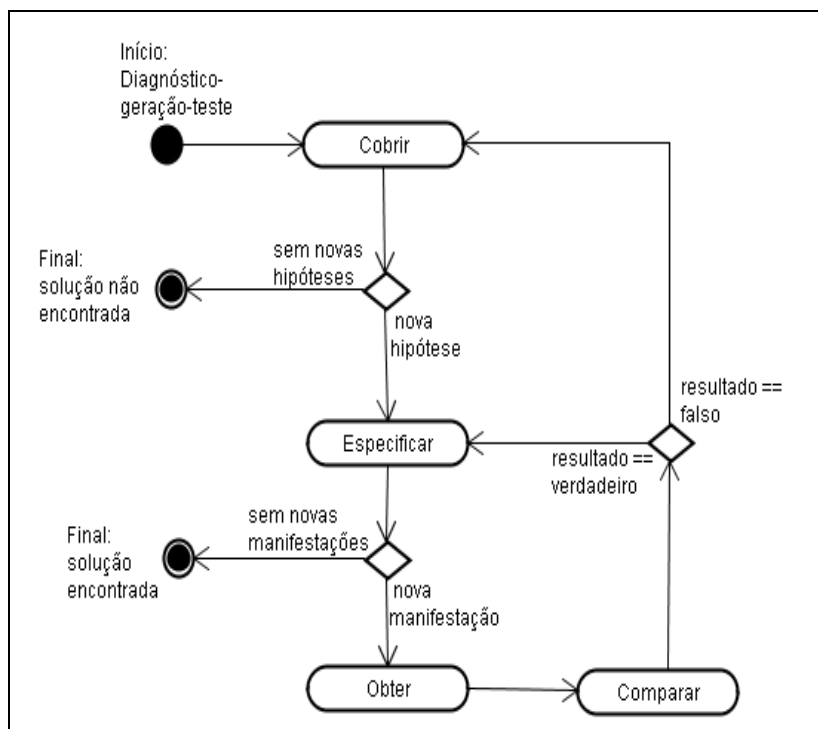


Figura 7 - Diagrama de Atividades do Modelo de Tarefa de Diagnóstico a partir de geração e teste

A figura 8 apresenta o modelo de tarefa completo e anotado para o diagnóstico a partir de geração e teste. Através da figura é possível ter uma compreensão global da execução da tarefa, considerando as inferências propostas.

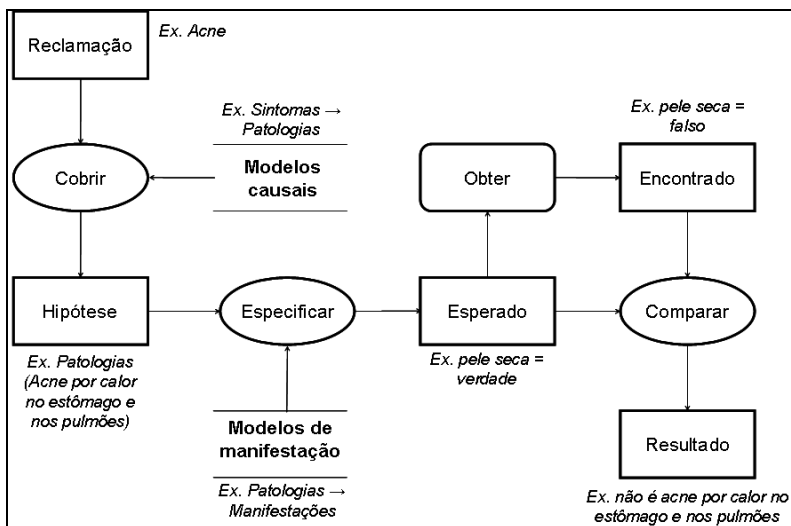


Figura 8 - Modelo de Tarefa Anotado

Modelo de Domínio

A figura 9 apresenta o modelo de domínio simplificado, sem considerar os atributos de cada classe existente no sistema.

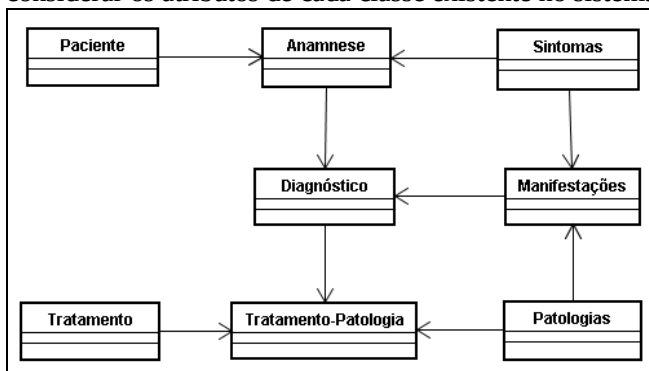


Figura 9 - Modelo de Domínio simplificado

As classes fundamentais para a tarefa proposta estão descritas a seguir:

- Paciente: dados e informações pessoais e características do paciente;

- b) Sintomas: estados observáveis ou sintomas que podem ser tratados na acupuntura;
- c) Anamnese: relaciona paciente e sintomas. São os estados observáveis apresentados por um determinado paciente;
- d) Patologias: estados do paciente que podem ser tratados com a acupuntura;
- e) Manifestações: relaciona sintomas com patologias. São as manifestações de uma patologia a partir de determinados sintomas;
- f) Diagnóstico: relaciona anamnese com manifestações. A partir da anamnese realizada com o paciente e das manifestações existentes para as patologias, constrói-se o diagnóstico;
- g) Tratamento: tratamentos realizados com a acupuntura; e
- h) Tratamento-Patologia: relaciona tratamentos com patologias e com o diagnóstico, indicando para um determinado diagnóstico, qual é o tratamento mais indicado.

As classes de Tratamento e Tratamento-Patologia não têm aplicação no diagnóstico, mas estão representadas apenas para contextualizar o diagnóstico em relação à etapa posterior de acompanhamento do tratamento.

4.4.3 Refinamento do Conhecimento

A etapa de refinamento do conhecimento tem por objetivo validar os modelos de conhecimento propostos, além de preencher as bases de conhecimento. A validação do modelo de conhecimento consiste em identificar se o mesmo está consistente, sem ambigüidades e se atende às necessidades para o qual foi concebido.

Para verificar a validade do modelo, foi construído um protótipo simplificado no software Nética (NORSYS NETICA, 2008), utilizando redes bayesianas, que são representações que mapeiam conhecimento a partir de um modelo probabilístico, expressando relações de causas e efeitos, através da associação de probabilidade em um determinado domínio do conhecimento (RUSSEL E NORVIG, 1995).

O exemplo construído tratou do diagnóstico de acne, apresentado nos diagramas da seção anterior. O protótipo construído contempla apenas uma pequena amostra de todo o conhecimento envolvido no processo de diagnóstico de acne para tratamento através da acupuntura.

A figura 10 apresenta o protótipo construído para validação do Sistema de Conhecimento.

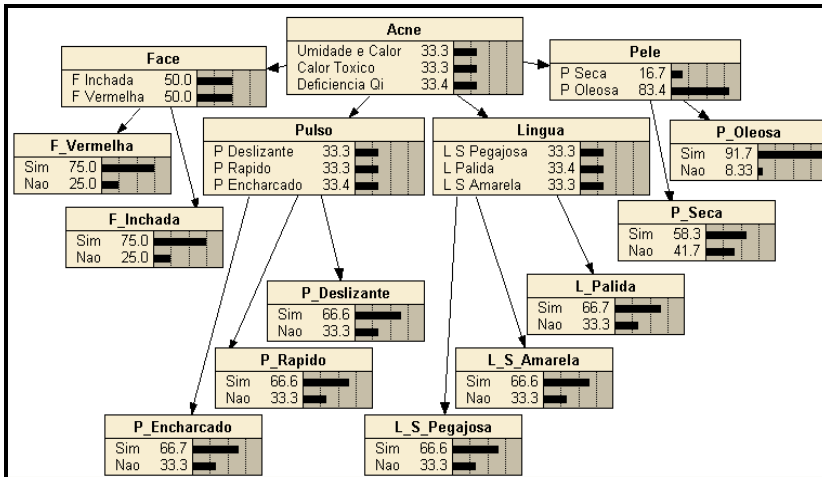


Figura 10 - Exemplo de Diagnóstico de Acne para Tratamento através de Acupuntura usando Redes Bayesianas

Os nós folhas representam as informações obtidas através da anamnese, ou seja, são as manifestações do organismo que podem causar o desequilíbrio que levam a uma determinada patologia, usados na inferência especificar. No caso da rede bayesiana proposta, as patologias, usadas na inferência cobrir, foram definidas como sendo os atributos do problema reclamado pelo paciente (no caso, Acne). Esta abordagem permite, portanto, a partir da anamnese, inferir o tipo de patologia mais provável do paciente, ou seja, com menor grau de incerteza, de forma a indicar o melhor tratamento.

4.5 MODELO DE COMUNICAÇÃO

O modelo de comunicação define como ocorre a troca de informações entre os agentes humanos e não humanos no Sistema de Conhecimento. Nesta tarefa ele é realizado de forma simples, uma vez que há apenas a interação do acupunturista com o Sistema de Conhecimento. Desta forma, optou-se por sua representação através de um diagrama de casos de uso da UML, conforme figura 11 e o quadro 12.



Figura 11 - Diagrama de Caso de Uso para a tarefa de Diagnóstico

Obter o diagnóstico a partir do estado observável apresentado pelo paciente, podendo indicar o tratamento mais adequado.	
Pré-condição:	Ter disponível os dados pessoais do Paciente Ter realizado a anamnese Ter a base de conhecimentos que possibilite diagnosticar a possível patologia do paciente.
Ator:	Acupunturista
Fluxo Principal:	1. O Sistema solicita informações sobre o Paciente 2. O Acupunturista insere as informações solicitadas no Sistema 3. O Sistema executa as inferências a partir das informações repassadas 4. O Sistema informa o possível Diagnóstico
Fluxo Alternativo:	1. O Sistema não chega a um resultado conclusivo

Quadro 12 - Comunicação da tarefa de Diagnóstico

A planilha CM1 – Descrição de transação (quadro 13) é utilizada no CommonKADS para apresentar a síntese do modelo de comunicação.

Modelo de comunicação	CM1 - Descrição de transação
Transação	Obter Diagnóstico
Objeto de informação	Diagnóstico entre as tarefas de anamnese e tratamento.
Agentes envolvidos	Acupunturista e Sistema de Conhecimento.
Plano de Comunicação	Ver Quadro 12
Restrições	Ver Quadro 12
Especificação de troca de informação	É uma transação do tipo pergunta e resposta.

Quadro 13 - CM1: Descrição da Transação

5. CONSIDERAÇÕES FINAIS

O trabalho proposto teve por objetivo apresentar o desenvolvimento de um Sistema de Conhecimento utilizando a metodologia CommonKADS. A metodologia demonstrou ser adequada para os objetivos propostos na medida em que apresenta uma visão de contexto sistêmico sobre a organização e o problema a ser tratado. Isto a torna menos focada na construção de um artefato, mas sim, na construção de um modelo de conhecimento que possa gerar resultados de forma a minimizar as incertezas e riscos e maximizar os ganhos para a organização.

No caso específico deste trabalho, a metodologia foi aplicada a uma clínica especializada em acupuntura, que trata de tarefas intensivas em conhecimento uma vez que o objetivo desta técnica é buscar o equilíbrio do organismo humano visando dar melhores condições de saúde e bem estar aos pacientes. O protótipo apresentado utilizando redes bayesianas fornece indicativos sobre a viabilidade de construção de um Sistema de Conhecimento completo para esta finalidade. Esta etapa deverá ser completada futuramente com a execução do modelo de artefato do CommonKADS.

Quando totalmente implementado, pretende-se que o sistema proposto auxilie o acupunturista na tarefa de diagnóstico, visando resultados mais precisos. O sistema também poderá ser estendido para contemplar a tarefa de monitoramento, através da criação de indicadores que permitam acompanhar o desenvolvimento e a adequação do tratamento, bem como a precisão do diagnóstico.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- BAUMEISTER, Joachim. *Agile development of diagnostic knowledge systems*. IOS Press, 2004.
- CAMILLOTI, B. M., ISRAEL, V. L., e CHI, A., *Necessidade de um sistema de apoio à decisão em acupuntura*, 2006 Disponível em: < www.sbis.org.br/cbis/arquivos/914.pdf >. Acesso em: 08 set. 2008.
- CHONGHUO, T. *Tratado de medicina chinesa*. São Paulo: Roca, 1993.
- ERNST, E., WHITE, A. *Acupuntura: uma avaliação científica*., São Paulo: Manole, 2001.
- KLEER, J., WILLIAMS, B. *Diagnosing multiple faults*. Artificial Intelligence 32 (1): 97-130, 1987.

- LEMOS, Simone F. *Significados de acupuntura por usuários de um serviço de atendimento em saúde*. Dissertação de Mestrado. Pós Graduação em Ciências da Saúde. UFG, UnB, UFMS. Goiânia, 2006.
- LUKMAN, Suryani, HE, Yulan, HUI, Siu-Cheung. Computational methods for traditional Chinese medicine: a survey. *Computer Methods and Programs in Biomedicine*. Volume 88, Issue 3, December 2007, pages 283-294.
- LUO, Jianhui, TU, Haiying, PATTIPATI, K., QIAO, Liu, CHIGUSA, S. Diagnosis knowledge representation and inference. *Instrumentation & Measurement Magazine, IEEE* Volume: 9, Issue: 4. Aug. 2006. Pages 45-52
- MACIOCIA, G. *Diagnóstico na medicina chinesa: um guia geral*. São Paulo: Roca, 2005.
- MILNE, Robert. Strategies for diagnosis. *Systems, Man and Cybernetics, IEEE Transactions*. Volume: 17, Issue: 3, May 1987
- NASIR, L. S. Acupuncture. *Primary Care: Clinics in Office Practice*, Volume 29, Issue 2, P. 393 - 405, 2002.
- NORSYS NETICA: bayesian networks graphical application. [S.l.] 2008. Disponível em: <<http://www.norsys.com/netica.html>>. Acesso em: 22 set. 2008.
- RUSSEL, S., NORVIG, P. *Artificial intelligence: a modern approach*. New Jersey: Prentice-Hall, 1995.
- SCHREIBER, G., AKKERMANS, H., ANJEWIERDEN, A., HOOG, R., SHADBOLT, N., VELDE, W. V., *et al.* *Knowledge engineering and management: the CommonKADS methodology*. Cambridge, Massachusetts: MIT Press, 2002.
- SIEDENTOPF, C. M., *et al.* Laser acupuncture induced specific cerebral cortical and subcortical activations in humans. *Laser in Medical Science*, v. 20, p. 68-73, 2005.
- STUDER, V. R., BENJAMINS, R., & FENSEL, D. Knowledge engineering: principles and methods. *Data & Knowledge Engineering*, 25, Issues 1-2, 161-197, 1998.
- YAMAMURA, Y. *Acupuntura tradicional: a arte de inserir*. 2ª ed., São Paulo: Roca, 2001

APÊNDICE VI - PUBLICAÇÃO EM CONGRESSO NACIONAL

UMA ONTOLOGIA DE APOIO AO DIAGNÓSTICO APLICADO À ACUPUNTURA

Luiz Fernando Lopes¹
Maurício Capobianco Lopes¹
Alexandre Leopoldo Gonçalves, Dr.¹
Francisco Antonio Pereira Fialho, Dr.¹

1 Programa de Pós-Graduação em Engenharia e Gestão do
Conhecimento
Universidade Federal de Santa Catarina
Campus Universitário - Caixa Postal 476 - CEP 88040-900 -
Florianópolis, SC – Brasil

lflopes@utfpr.edu.br, mau.capo@egc.ufsc.br, a.l.goncalves@stela.org.br,
fapfialho@gmail.com

Resumo:

O diagnóstico em acupuntura é um processo complexo que envolve diversas etapas e muitas informações a serem consideradas. Desta forma é fundamental a existência de ferramentas que auxiliem o especialista nesta tarefa. Assim, este artigo tem por objetivo apresentar um protótipo de um modelo de conhecimento baseado na construção de uma ontologia para apoiar o diagnóstico de tratamento de Acne. A ontologia foi construída a partir da metodologia 101, usando as ferramentas Ontokem e *Protégé*. O modelo construído demonstrou a viabilidade do uso de ontologias neste contexto, pois os resultados apresentados atenderam as expectativas definidas para o trabalho, encorajando os autores a aprofundar as pesquisas, de forma a usar as ontologias na busca de novos conhecimentos em acupuntura.

Palavras-chave: Engenharia do Conhecimento; Ontologia; Diagnóstico; Gestão do Conhecimento; Acupuntura

1. INTRODUÇÃO

A Medicina Tradicional Chinesa oferece diversas práticas de saúde voltadas ao bem estar do organismo, tanto no que diz respeito ao físico quanto ao mental. Uma das práticas mais conhecidas e utilizadas é a acupuntura que tem por finalidade a restauração do funcionamento equilibrado do organismo e a prevenção de patologias e afecções. Segundo Yamamura (2001), trata-se de “um método que estimula determinados pontos do corpo com agulhas, com o intuito de restaurar ou manter a saúde”.

A acupuntura proporciona uma visão sistêmica complexa sobre o organismo humano, composta de uma grande quantidade de agentes que se interrelacionam e que sofrem influência de elementos internos e externos. Este contexto sistêmico exige um conjunto de conhecimentos que muitas vezes dificulta a realização de um diagnóstico acurado de forma a oferecer um tratamento mais preciso e rápido aos pacientes.

Neste contexto, a Engenharia e a Gestão do Conhecimento dispõem de métodos e técnicas utilizadas na explicitação, uso e disseminação do conhecimento, que podem prover ao acupunturista um conjunto de recursos focados no apoio ao processo de tratamento de seus pacientes, seja no diagnóstico ou no acompanhamento do tratamento, visando restabelecer ou manter o equilíbrio sistêmico do organismo.

Uma das técnicas providas pela Engenharia do Conhecimento são as ontologias que possibilitam a estruturação, uso e disseminação do conhecimento de um domínio de forma explícita. Segundo Gruber (1993), “uma ontologia é uma hierarquia estruturada de um conjunto de termos para descrever um domínio que possa ser usado como estrutura de uma base de conhecimento”.

Assim sendo, o objetivo deste artigo é apresentar um modelo de conhecimento voltado ao diagnóstico de pacientes de acupuntura, mais especificamente da análise de dados envolvendo as patologias sobre acne. Este modelo visa oferecer ao acupunturista ferramentas adequadas e precisas para realizar o diagnóstico, proporcionando um tratamento mais adequado de modo a oferecer uma melhor qualidade de vida e saúde ao paciente.

Neste contexto, entende-se por diagnóstico o processo que visa determinar porque um sistema corretamente projetado (neste caso, o organismo humano) não funciona conforme esperado (apresenta patologias). De Kleer e Williams (1987) destacam que uma tarefa de diagnóstico é realizada por uma entidade capaz de observar, levantar

hipóteses e chegar a um diagnóstico (resultado), descrevendo o estado do processo observado.

2. ONTOLOGIA

O termo ontologia, proveniente da filosofia, também está inserido na engenharia e gestão do conhecimento como meio de representação formal do conhecimento, com o propósito de facilitar o compartilhamento deste entre pessoas e sistemas.

Para Neches *et al.* (1991), “Uma ontologia define os termos básicos e relações compreendendo o vocabulário de uma área específica, bem como as regras para combinação entre termos e relações para definir extensões do vocabulário”.

Gruber (1993) afirma que “uma ontologia é uma especificação formal e explícita para um conceito compartilhado”. Da definição de Gruber (1993), Fensel (2002) ressalta os termos formal, explícita e compartilhada. Formal, pelo fato de que uma ontologia deve ser compreendida e processada por sistemas. Explícita significa que os conceitos utilizados, bem como as restrições sobre seu uso, são explicitamente definidos. Compartilhado no sentido de que uma ontologia reflete o conhecimento consensual sobre um determinado assunto por uma comunidade.

O uso de uma ontologia permite então, a definição de um domínio no qual será possível trabalhar em determinada área específica, possibilitando a melhora no processo de extração de informação e o intercâmbio do conhecimento.

Portanto, um dos aspectos fundamentais na construção de uma ontologia é a familiarização com o domínio que tem o objetivo de permitir o completo entendimento sobre o problema, através da consulta às diferentes fontes de conhecimento disponíveis.

A construção de ontologias envolve metodologias para sua edificação. No desenvolvimento deste trabalho optou-se pela utilização do Ontokem (2008) que é uma ferramenta para Engenharia do Conhecimento baseado no processo de desenvolvimento da metodologia 101 e nos artefatos documentais das metodologias para ontologias Onto-Knowledge e Methontology devido ao detalhamento que proporciona na estruturação das ontologias e possibilidade de reuso (RAUTENBERG *et al.*, 2008). Também foi utilizada a ferramenta *Protégé* (PROTÉGÉ, 2007), pois apresenta arquitetura expansível, fácil

usabilidade, certo nível de detalhamento nas ontologias e vem sendo utilizada na área da saúde.

Deste modo, a partir do conjunto de metodologias e ferramentas disponíveis, desenvolveu-se esta ontologia através das seguintes etapas:

- a) definição das perguntas de competência: nesta etapa foram definidas as perguntas para as quais a ontologia deveria apresentar suas respostas, definido também os principais termos e relações relacionados a cada pergunta;
- b) definição das classes e subclasses: nesta etapa foi definido o modelo hierárquico da ontologia, definindo-se as classes a partir dos termos associados às perguntas de pesquisa e criando-se as subclasses necessárias para a representação completa do domínio;
- c) definição das relações: nesta etapa foram estabelecidas as ligações entre as classes, de modo a representar como elas se relacionam no contexto do domínio;
- d) definição das restrições: nesta etapa foram definidas as restrições associadas às relações entre as classes, estabelecendo os axiomas que as compõem;
- e) criação das instâncias: nesta etapa foram criadas as instâncias das classes, ou seja, os indivíduos com determinadas características requeridas pelas classes/subclasses. As instâncias são elementos fundamentais para testar a ontologia;
- f) consultas e inferências: nesta etapa foram realizadas consultas (*queries*) e inferências (*reasoning*) para verificar se a ontologia estava apta a responder às perguntas de competência.

3. CONSTRUÇÃO DA ONTOLOGIA

Para a realização deste trabalho, as fontes de conhecimento selecionadas para a construção da ontologia foram um especialista em acupuntura, manuais e livros sobre o tema. Com o especialista foram feitas entrevistas não estruturadas, de forma a se identificar os principais conhecimentos tácitos aplicados na tarefa de diagnóstico. Os manuais e livros foram utilizados para se conhecer os principais elementos e técnicas envolvidas na atividade de acupuntura, com ênfase nas informações mais importantes do processo de diagnóstico.

3.1 PERGUNTAS DE COMPETÊNCIA

Para delimitar o domínio e o escopo da ontologia foram formuladas perguntas de interesse para a abordagem do problema proposto, tendo, para isto, o auxílio do especialista e as outras fontes de conhecimento citados anteriormente.

Estas questões forneceram de uma maneira preliminar, os termos e as relações utilizados na construção da ontologia. Como exemplo de pergunta de competência, citamos: **“Quais órgãos compõem o paciente? X?”**, onde destacamos os termos: **órgão** e **paciente** e a relação: **compõe**.

A figura 1 apresenta, em uma tela do Ontokem, as perguntas de competência que auxiliaram a construção do domínio e o escopo da ontologia acne. Estas perguntas forneceram os principais termos e as relações documentadas formalmente na ontologia. Através da figura é possível observar também na aba “Termos”, os termos relacionados à pergunta. Já na aba “Relações” é possível cadastrar as relações que dizem respeito à pergunta.

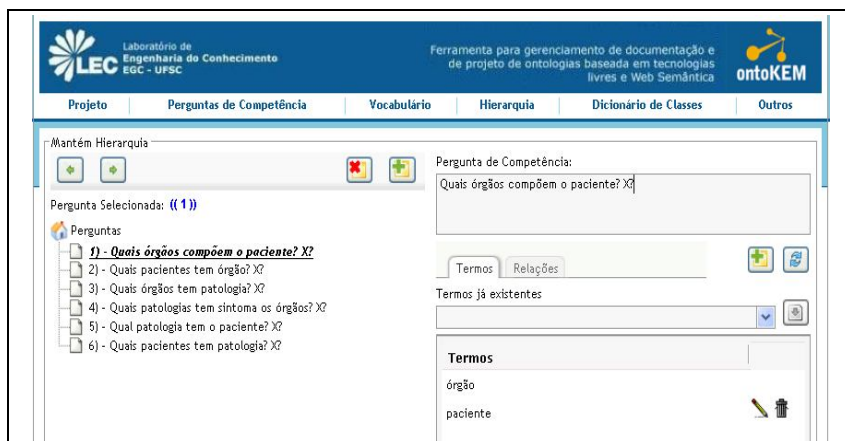


Figura 1 – Perguntas de Competência

É importante destacar que as perguntas de competência formuladas são genéricas, não estando diretamente associadas às patologias de acne. Desta forma, estas perguntas podem ser utilizadas para o diagnóstico de outras patologias.

3.2 CLASSES E SUBCLASSES

Como visto anteriormente, o Ontokem permite explicitar os termos, juntamente com as perguntas de competência. Com isto é possível definir de forma mais rápida as classes e as subclasses existentes na ontologia.

Na ferramenta Ontokem, ao se cadastrar uma classe deve ser feita uma descrição completa sobre suas propriedades. Também devem ser descritas as regras que definem as condições para um indivíduo pertencer à classe. Esta descrição é usada posteriormente na definição das restrições da classe.

As classes e subclasses da ontologia compõem o modelo hierárquico da mesma. A figura 2 apresenta a hierarquia da ontologia proposta, gerada a partir do Protégé.

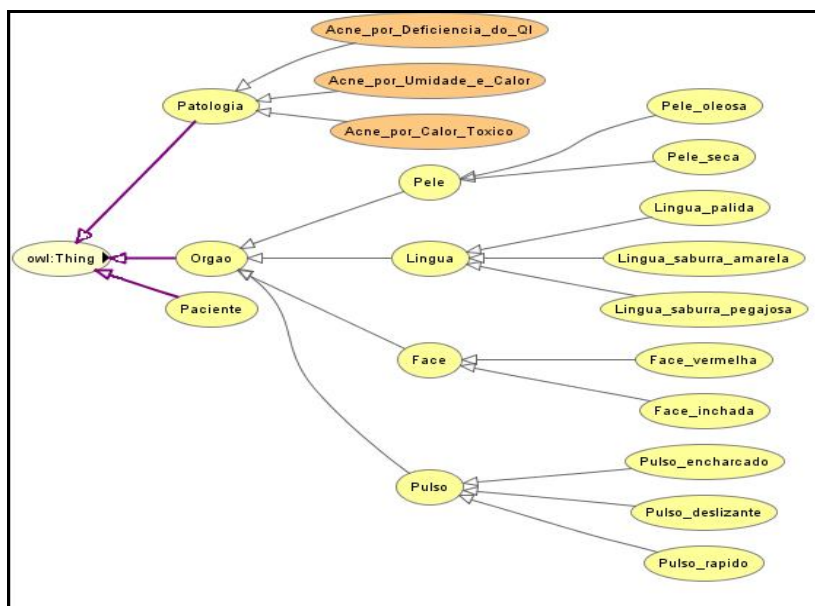


Figura 2 – Classes e subclasses da ontologia

A tabela 1 apresenta a descrição das principais classes do sistema.

Tabela 1 - Descrição das principais classes e subclasses da ontologia

Classe/Subclasse	Descrição
Paciente	Classe dos indivíduos acometidos pela patologia
Órgão	Classe dos órgãos que compõem um indivíduo
Patologia	Classe das patologias a serem diagnosticadas
Face	Classe dos tipos de face de um órgão
Pele	Classe dos tipos de pele de um órgão
Língua	Classe dos tipos de língua de um órgão
Pulso	Classe dos tipos de pulso de um órgão

Nesta ontologia foram considerados os órgãos: Pulso, Língua, Pele e Face, que apresentam sintomas para evidenciar o tipo de patologia em um paciente acometido de acne. Os órgãos foram hierarquizados em subclasses:

- a) para a classe **Pele** foram definidas as subclasses: Pele seca e Pele oleosa;
- b) para a classe **Língua** foram definidas as subclasses: Língua pálida, Língua saburra amarela e Língua saburra pegajosa;
- c) para a classe **Face** foram definidas as subclasses: Face vermelha e Face inchada;
- d) para a classe **Pulso** foram definidas as subclasses: Pulso encharcado, Pulso deslizante e Pulso rápido.

Para representar as patologias envolvendo acne, foram criadas apenas três subclasses, com fins didáticos, delimitando a abrangência do trabalho. As subclasses definidas foram: Acne por Deficiência do QI, Acne por Umidade e Calor e Acne por Calor Tóxico.

Ressalta-se que também neste caso, o modelo hierárquico proposto pode ser usado como base para a modelagem de diagnóstico envolvendo outros tipos de patologias.

3.3 RELAÇÕES ENTRE CLASSES

Além de especificar os termos junto com as perguntas de competência, o Ontokem permite explicitar também as relações existentes entre os mesmos.

Assim, na figura 3, são apresentadas as relações entre as classes da ontologia. Apenas com a definição destas relações é possível fazer com que a ontologia responda às perguntas formuladas anteriormente.

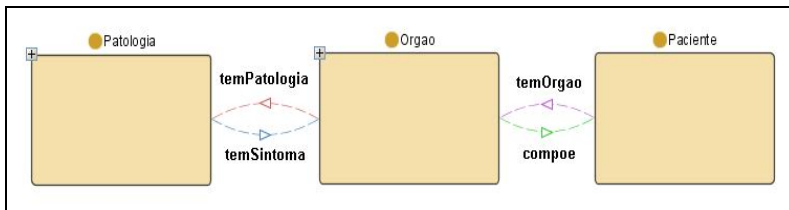


Figura 3 – Relação entre Classes

Pela figura pode-se observar que a classe **Paciente** tem um relacionamento com a classe **Órgão** do tipo “**temOrgão**”, permitindo, desta forma, concluir que “**Paciente tem Órgão**”. Esta relação tem uma inversa denominada “**compõe**”. Assim pode-se concluir que “**Órgão compõe o Paciente**”. Esta mesma analogia pode ser feita entre as classes Órgão e Patologia.

Estas relações também são conhecidas como propriedades de objetos. Outro tipo de propriedade existente nas classes são as propriedades de dados, ou seja, são atributos que caracterizam a classe. Na classe Paciente, por exemplo, poderiam ser criados atributos tais como, nome, idade, sexo, entre outros, mas estes atributos não estão sendo considerados neste momento.

As relações também podem ser inferidas através de restrições ou axiomas, que são regras que possibilitam inferências sobre as classes. Por meio dos axiomas é possível elaborar sentenças verdadeiras entre classes que não possuam uma relação direta. Através das regras poderemos verificar, por exemplo, que Paciente tem determinada Patologia, sem que exista uma relação direta entre as classes Paciente e Patologia.

3.4 RESTRIÇÕES DE CLASSES

Após construir a hierarquia de classes com suas respectivas relações, devem ser definidas as restrições, que são as regras que restringem os relacionamentos e permitem estabelecer se um indivíduo pertence ou não a uma determinada classe.

A figura 4 apresenta as restrições das classes, utilizando as relações definidas anteriormente.

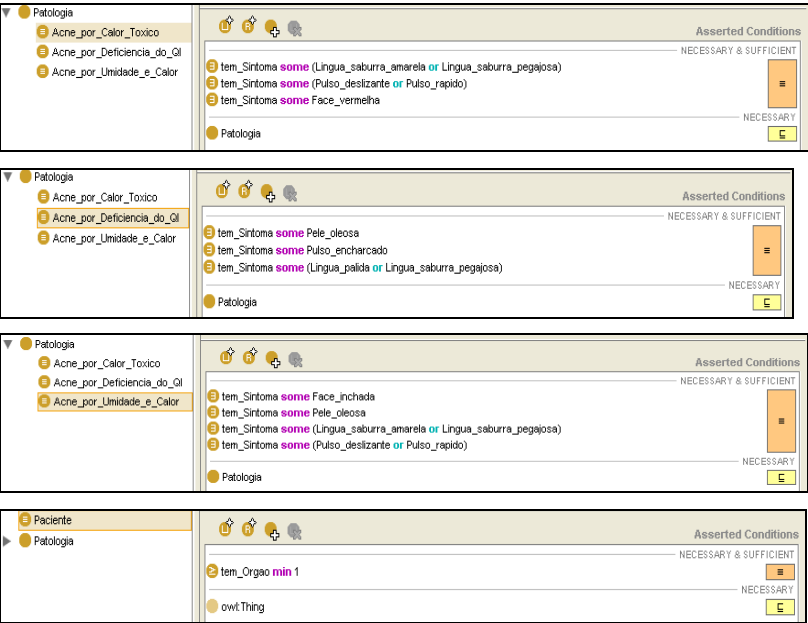


Figura 4 – Restrições de Classe

Pela figura pode-se observar que para um paciente ter Acne por Calor Tóxico é necessário e suficiente que apresente os seguintes sintomas: Face vermelha; Pulso deslizante ou Pulso rápido; Língua saburra amarela ou Língua saburra pegajosa. Esta mesma observação pode ser feita para as demais patologias. Já para a relação **tem_Orgão** foi definido como necessário e suficiente o paciente ter no mínimo um órgão.

3.5 INSTÂNCIAS

Após a construção das classes, subclasses, propriedades e restrições, criaram-se as instâncias para a base da ontologia. As instâncias representam indivíduos específicos de uma determinada classe.

Na formalização das instâncias, citamos como exemplo o Paciente **Gabriel**, que apresenta os seguintes sintomas: Pulso rápido, Língua saburra amarela e Face vermelha. O cadastro desta instância é mostrado na figura 5.

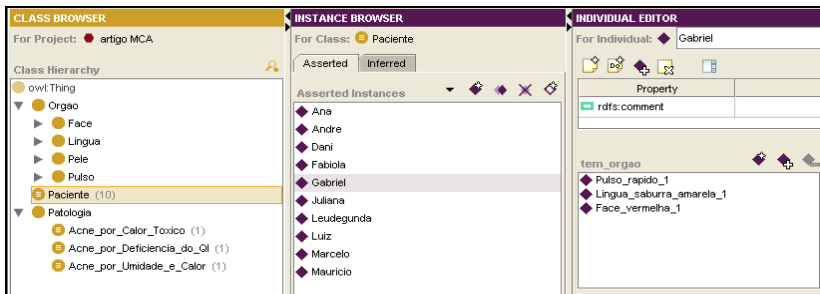


Figura 5 – Exemplo de instância

3.6 CONSULTAS E INFERÊNCIAS

A validação do modelo de conhecimento consiste em identificar se o modelo está consistente, sem ambiguidades e se atende às necessidades para o qual foi concebido.

Para esta validação são apresentados dois casos: uma consulta (*query*) que permite obter uma resposta direta a uma pergunta; e uma inferência, que permite o raciocínio sobre as restrições definidas para as relações entre as classes. No *Protégé* as consultas podem ser realizadas através de uma aba denominada *Queries*. Já as inferências são obtidas através de um raciocinador. Neste caso, foi utilizado o pellet 1.4.

A figura 6 apresenta o exemplo de uma consulta na qual se busca o Paciente que possui órgãos com os seguintes sintomas: Pulso Rápido, Língua Saburra Amarela e Face Vermelha. Como resposta à esta solicitação, temos o paciente **Gabriel**.

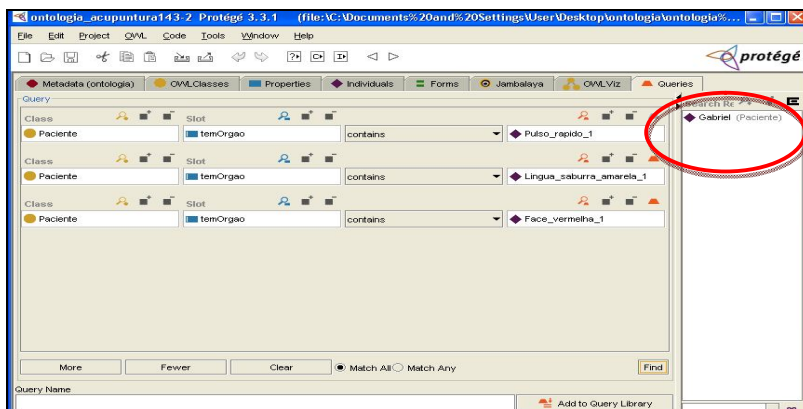


Figura 6 – Consulta no *Protégé*

A figura 7 apresenta o exemplo de inferência realizada através do *Protégé*. Pela figura é possível observar que os pacientes Maria e Pedro foram inferidos como tendo Acne por Umidade e Calor.

De fato, Maria possui os órgãos: Língua saburra amarela, Face inchada, Pele oleosa e Pulso rápido; já Pedro possui os órgãos: Língua saburra amarela, Face inchada, Pele oleosa e Pulso deslizante. Ambos atendem às restrições da patologia, conforme pode ser observado na figura 4.

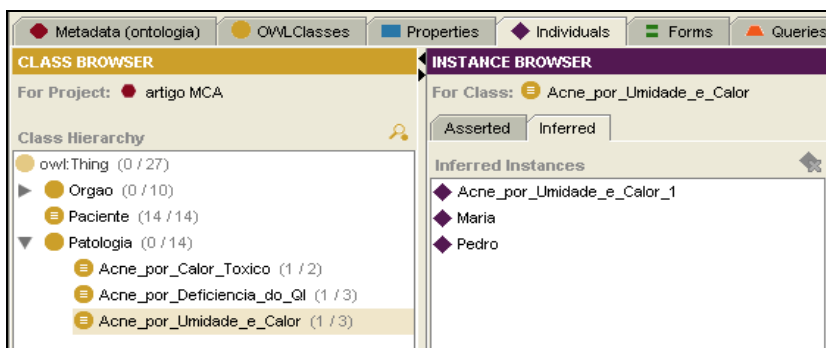


Figura 7 – Inferência no *Protégé*

4. CONSIDERAÇÕES FINAIS

Este artigo apresentou a construção de uma ontologia para apoiar o diagnóstico de acne, sob o ponto de vista da acupuntura. A partir da explicitação envolvendo manuais, livros e especialista foi possível construir e formalizar o domínio de forma experimental, para o compartilhamento de conhecimento. Por se tratar de um protótipo, os dados referentes aos pacientes foram levantados de forma didática, não estando a ontologia pronta para ser utilizada em nível profissional.

Esta abordagem permitiu, entretanto, a partir do diagnóstico de alguns sintomas, inferir o tipo de patologia mais provável do paciente, de forma a indicar o melhor tratamento. Além disso, através dos estudos realizados neste artigo, nota-se que o desenvolvimento de ontologias é de grande importância para a estruturação e padronização de conhecimento e pode auxiliar de forma efetiva em tarefas de diagnóstico.

Destacamos, finalmente, que na Medicina Tradicional Chinesa, além dos sinais e sintomas, outros fatores devem ser considerados para um diagnóstico mais preciso, tais como, o contexto pessoal e social do paciente, levantados no processo de anamnese e que não foram tratados neste artigo.

Com o acréscimo destas informações, acreditamos que o uso de ontologias pode levar à construção de novos conhecimentos ainda não explicitados em livros, manuais ou por especialistas em acupuntura, vinculando elementos externos ao contexto da patologia que podem estar relacionados às suas causas.

Assim, pretende-se aprofundar os estudos aqui apresentados, buscando novos conhecimentos a partir da ontologia, de forma a permitir um diagnóstico mais preciso, diminuindo o grau de incerteza do tratamento e proporcionando um perfeito equilíbrio entre o organismo do paciente e o meio em que vive, de modo a lhe proporcionar uma vida mais saudável e duradoura.

5. REFERÊNCIAS

- DE KLEER, J.; WILLIAMS, B. Diagnosing Multiple Faults. *Artificial Intelligence* 32 (1): 97-130, 1987.
- FENSEL, Dieter. Ontology-Based Knowledge Management. *IEEE Computer*. November, 2002.
- GRUBER, T.R., A translation approach to portable ontology specifications. *Knowledge Acquisition*, 1993, p. 199-220.

- NECHES R., FIKES R.E., FININ T., GRUBER T.R., SENATOR T., SWATOUT W.R., Enabling technology for knowledge sharing. 1991. *AI Magazine* 12(3):36–56.
- ONTOKEM. [S.l.], (2008). Disponível em: <<http://ontokem.egc.ufsc.br>>. Acesso em: 02 fev. 2009.
- PROTÉGÉ* 3.3.1. Stanford: California, 2007. Disponível em: <<http://protégé.stanford.edu>>. Acesso em: 02 fev. 2009.
- RAUTENBERG, Sandro; GAUTHIER, Fernando A. O.; LOTTIN, Poline; DUARTE, Cleiton E. J.; TODESCO, José L. *ontoKEM: uma ferramenta para construção e documentação de ontologias*. In. Seminário de Pesquisa em Ontologias no Brasil. Niterói, 2008. Disponível em <<http://www.uff.br/ontologia/artigos/27.pdf>>. Acesso em 12 jan. 2009.
- YAMAMURA, Y. *Acupuntura tradicional: a arte de inserir*. 2ª ed., São Paulo: Roca, 20

APÊNDICE VII - PUBLICAÇÃO EM CONGRESSO NACIONAL

UMA ONTOLOGIA PARA DIAGNÓSTICO EM ACUPUNTURA BASEADA NO MODELO CESM

Luiz Fernando Lopes¹
Alexandre Leopoldo Gonçalves, Dr.¹
Maurício Capobianco Lopes¹

1 Programa de Pós-Graduação em Engenharia e Gestão do
Conhecimento

Universidade Federal de Santa Catarina

Campus Universitário - Caixa Postal 476 - CEP 88040-900 -
Florianópolis, SC - Brasil

lflopes@utfpr.edu.br, a.l.goncalves@stela.org.br, mau.capo@egc.ufsc.br

Resumo: A tomada de decisão em Acupuntura é um processo complexo, uma vez que existe uma grande variedade de síndromes e padrões a serem consideradas no diagnóstico. Um diagnóstico preciso é fundamental para o sucesso do tratamento. O processo de diagnóstico deve ser feito de forma sistêmica, de modo que, além dos sintomas apresentados pelo paciente, outros fatores possam ser evidenciados. Este artigo tem por objetivo apresentar um modelo de conhecimento voltado ao diagnóstico de pacientes de acupuntura, mais especificamente a ontologia da patologia acne. A metodologia utilizada baseia-se no conceito de sistemismo proposto por Bunge (2000), sob o olhar da modelagem CESM. Esta ontologia possibilitou a representação do conhecimento de forma simples e concisa, propiciando um melhor entendimento sobre o conhecimento modelado e proporcionando um processo de diagnóstico mais eficiente.

Palavras-chave: Engenharia do Conhecimento, Ontologia, CESM

1. INTRODUÇÃO

A Medicina Tradicional Chinesa (MTC) oferece diversas práticas de saúde voltadas ao bem estar do organismo, tanto no que diz respeito ao físico quanto ao mental. Uma das práticas mais conhecidas e utilizadas é a acupuntura que tem por finalidade a restauração do funcionamento equilibrado do organismo e a prevenção de patologias e afecções. Segundo Yamamura (2004), trata-se de um método que estimula determinados pontos do corpo com agulhas, com o intuito de restaurar ou manter a saúde.

A acupuntura proporciona uma visão sistêmica complexa sobre o organismo humano, composta de uma grande quantidade de agentes que se inter-relacionam e que sofrem influência de elementos internos e externos (LUKMAN, YULAN & SIU-CHENG, 2007; MACIOCIA, 2005; YAMAMURA, 2004; ROSS, 1994). Este contexto sistêmico exige um conjunto de conhecimentos que muitas vezes dificulta a realização de um diagnóstico acurado de forma a oferecer um tratamento mais preciso e rápido aos pacientes (CAMILLOTI, ISRAEL & CHI, 2006).

A melhoria do processo de diagnóstico em acupuntura baseada em sistemas de conhecimento tem sido objeto de estudo de poucos autores (RADU, 2001; CAMILLOTI, ISRAEL & CHI, 2006). Uma das principais dificuldades na construção destes tipos de sistema é que normalmente o conhecimento sobre o processo de diagnóstico em acupuntura está internalizado de forma tácita nos profissionais acupunturistas. Neste sentido, este trabalho apresenta uma proposta que busca responder à seguinte questão: como a Engenharia do Conhecimento pode explicitar o conhecimento existente no processo de diagnóstico em acupuntura, de modo a torná-lo mais ágil e preciso?

O diagnóstico é uma tarefa intensiva em conhecimento estudado pela Engenharia do Conhecimento (SCHREIBER, 2002). Segundo Schreiber *et al.* (2002) este campo de estudo oferece muitos conceitos e métodos úteis à Gestão do Conhecimento, visto que toda a análise, modelagem e estudos necessários para gerar sistemas de conhecimento estão fundamentados na mesma.

Nesta mesma linha Kendall e Green (2007) destacam que a Engenharia do Conhecimento deve estar envolvida nas atividades de aquisição, validação, representação, inferência, explicação e justificativa dos elementos de conhecimento presentes em uma determinada tarefa intensiva em conhecimento.

Com base nisto, pressupõe-se que a Engenharia e a Gestão do Conhecimento possam fornecer ao acupunturista um conjunto de recursos focados no apoio ao processo de diagnóstico, tratamento e acompanhamento de seus pacientes. Ressalta-se que este trabalho está focado apenas no processo de diagnóstico.

2. OBJETIVO

Lopes *et al.* (2009) já apresenta uma discussão inicial propondo a construção de uma ontologia para o diagnóstico em acupuntura. Este artigo amplia a proposta anterior, tendo como objetivo apresentar um modelo de conhecimento voltado ao diagnóstico de pacientes de acupuntura, mas que esteja adaptado à conceituação proposta por Bunge. Mais especificamente será apresentada a ontologia da patologia acne utilizando a modelagem CSM.

A conceituação de sistemas apresentada por Bunge (2003), através de seu modelo CSM (*Composition – Environment – Structure – Mechanism*) contribui para uma modelagem mais descritiva e abrangente dos processos que envolvem o diagnóstico na acupuntura, sendo descrita a seguir.

3. METODOLOGIA - MODELAGEM CSM

CSM é um sistema filosófico proposto por Mario Bunge, argentino radicado no Canadá, que dedicou sua vida ao desenvolvimento desta teoria científica. Para ele, todas as coisas são sistemas ou componentes de um sistema, exceto o Universo que é um supersistema de todos os demais.

Segundo Bunge (1997), para entender sobre qualquer sistema, seja real ou artificial, é necessário compreender como ele trabalha, isto é, desvendar o seu mecanismo de ação. Por exemplo, o movimento molecular pode explicar a fermentação da uva para a obtenção do vinho, ou uma alimentação inadequada pode provocar mau funcionamento do organismo e levar a doenças.

A abordagem sistêmica bungeana “não é uma teoria para substituir outras teorias” (BUNGE, 2004, p. 91), mas uma estratégia ou visão de mundo para guiar o empreendimento de pesquisa sobre sistemas.

Bunge (2000), em seu sistema filosófico, propõe os seguintes postulados:

- Toda coisa, seja concreta ou abstrata, é um sistema ou um componente ou potencial componente de sistema;
- Os sistemas têm características sistêmicas (emergentes) que seus componentes não tem;
- Todos os problemas deveriam ser abordados de forma sistêmica em vez de em forma fragmentada;
- Todas as idéias deveriam ser unidas em sistemas (preferencialmente teorias);
- O teste de qualquer coisa, seja idéia, método ou artefato, supõe a validação de outros itens que são tomados como pontos de referência.

Para Bunge (2003), o modelo CESM proporciona uma forma de modelar qualquer sistema mediante a quádrupla: *Composition; Environment; Structure; Mechanism*, onde:

- a) *Composition* (composição): conjunto de todas as partes do sistema;
- b) *Environment* (ambiente): coleção de itens que não pertencem ao sistema, porém atuam ou sofrem a ação por algum ou por todos os componentes;
- c) *Structure* (estrutura): coleção de ligações entre componentes, e entre esses e itens do ambiente;
- d) *Mechanism* (mecanismo): coleção de processos que geram mudanças qualitativas no sistema.

Kilov e Sack (2008), corroborando as idéias de Bunge, destacam que os conceitos de relação e composição, envolvendo os componentes, são a essência do funcionamento de um sistema.

De acordo com o sistemismo proposto por Bunge (2003), qualquer sistema, para ser compreendido, deve ser visto de uma forma global, isto é, não só o relacionamento dos componentes entre si, mas também, destes com os itens do ambiente onde estão inseridos.

A figura 1 apresenta um exemplo de sistema que envolve o processo de diagnóstico médico, baseado na quádrupla CESM, proposta por Bunge. Neste sistema, o Paciente acometido de uma enfermidade busca tratamento para sua Patologia. O profissional analisa, de forma sistêmica, os Sintomas e as Evidências levando em conta os fatores internos (componentes) e os fatores externos (ambiente) do paciente, que, ao serem relacionados (estrutura), são limitados pelas Restrições, caracterizando determinada Patologia para tratamento (mecanismo).

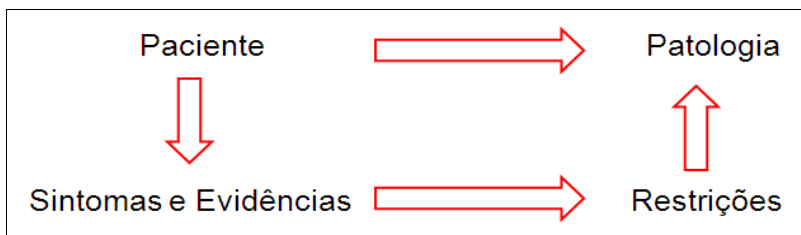


Figura 1. Efeito sistêmico do processo de diagnóstico médico

Com base nestes conceitos, será apresentada, na próxima seção, a construção de uma ontologia baseada no modelo CESM para o diagnóstico em acupuntura.

4. CONSTRUÇÃO DA ONTOLOGIA BASEADA NO MODELO CESM

Gruber (1993) destaca que uma ontologia é uma especificação formal e explícita para um conceito compartilhado. As ontologias são muito apropriadas nos processos requeridos na Engenharia do Conhecimento, pois permitem representar e validar o conhecimento, além de possibilitar inferências e justificativas sobre novos conhecimentos gerados. O uso de uma ontologia permite a definição de um domínio no qual será possível trabalhar em determinada área específica, possibilitando a melhora no processo de extração de informação e o intercâmbio do conhecimento.

Para Neches *et al.* (1991) uma ontologia define os termos básicos e relações compreendendo o vocabulário de uma área específica, bem como as regras para combinação entre termos e relações para definir extensões do vocabulário. Noy e McGuinness (2001) destacam que a construção de uma ontologia pode ter diferentes finalidades entre as quais se destacam: permitir o reuso do conhecimento de um determinado domínio, separar o conhecimento do domínio do conhecimento operacional, e construir um conjunto compartilhado de conhecimentos sobre um domínio.

Uma ontologia é criada baseada em classes estruturadas em forma hierárquica, que representam o domínio de um determinado problema. Estas classes podem conter relações entre si, definidas por restrições. Com base neste modelo é possível fazer inferências para a busca ou geração de conhecimento.

O conhecimento a ser representado em uma ontologia, pode ser tratado de forma sistêmica, levando em consideração a necessidade de analisar todos os elementos que interferem no processo em estudo.

As definições apresentadas permitem verificar muitas semelhanças entre o potencial oferecido pelas ontologias no processo de representação, uso e disseminação do conhecimento, com o modelo CESM de Bunge (2003).

O quadro 1 apresenta de forma simplificada a relação possível entre os elementos do modelo CESM e os conceitos existentes nas ontologias. Cada um dos itens será detalhado nas próximas seções.

Composição	Classes internas: Paciente; Patologia
Ambiente	Classes externas: Alimentação; Atividade física; Ocupação
Estrutura	Hierarquia de classes e relações entre as classes
Mecanismo	Regras e restrições

Quadro 1. Elementos do diagnóstico em acupuntura baseado no modelo CESM e sua relação com as ontologias

4.1 COMPOSIÇÃO

Bunge (2003) define a composição como a coleção de todas as partes do sistema. Na composição, foram considerados como “classes internas” os órgãos do paciente e as possíveis patologias, devido ao fato de que os pacientes, com seus respectivos órgãos, estão diretamente relacionados às enfermidades.

Foram considerados como órgãos: Pulso, Língua, Pele e Face, que apresentam sintomas para evidenciar o tipo de patologia em um paciente acometido de acne. Os órgãos foram hierarquizados em subclasses:

- a) subclasses da classe Pele: Pele seca e Pele oleosa;
- b) subclasses da classe Língua: Língua pálida, Língua saburra amarela e Língua saburra pegajosa;
- c) subclasses da classe Face: Face vermelha e Face inchada;
- d) subclasses da classe Pulso: Pulso encharcado, Pulso deslizante e Pulso rápido.

Dentre as possíveis patologias para o tratamento de Acne, foram consideradas as classes para o diagnóstico: Acne por Umidade e Calor, Acne por Calor Tóxico e Acne por Deficiência do QI (energia vital). No trabalho de Lopes *et al.* (2009) somente estas classes, anteriormente citadas, foram contempladas, servindo de base para a construção da ontologia.

Na proposta aqui realizada, foram incorporados os seguintes atributos para o paciente: sexo, idade, IMC (índice de massa corpórea) e etnia, como fatores que podem contribuir para análise neste processo.

Para o atributo sexo, estão sendo considerados os gêneros: masculino e feminino; para o atributo idade, estão sendo considerados os grupos etários: jovem, adulto e idoso; para o atributo IMC, estão sendo consideradas as constituições físicas: magro, sobrepeso e obeso; para o atributo etnia, estão sendo consideradas as raças branca, negra e amarela.

Destaca-se que este modelo é apenas ilustrativo, com finalidade didática, não estando contemplados, até o momento, todos os elementos necessários para ser utilizado em nível profissional.

4.2 AMBIENTE

Conforme visto anteriormente, Bunge (2003) propõe o ambiente como sendo a coleção de itens que não pertencem ao sistema, mas que, atuam ou sofrem a ação por algum ou todos os componentes do sistema.

Neste contexto, Maciocia (2005) entende que o processo de diagnóstico na medicina chinesa, baseia-se no princípio fundamental de que sintomas e sinais refletem as condições dos órgãos internos e dos canais, podendo não estar necessariamente relacionados com o processo real de uma determinada doença. A investigação para o processo de diagnóstico deve estender-se para além de sintomas e sinais relacionados com a queixa, compondo o ambiente sistêmico do diagnóstico.

Neste trabalho, foram selecionados alguns elementos do ambiente que influenciam o processo de diagnóstico, para compor as classes externas. Os elementos escolhidos foram: os tipos de alimentação, os tipos de atividades físicas e os tipos de ocupação do paciente.

As classes externas foram hierarquizadas em subclasses:

- e) subclasses da classe Alimentação: Vegetariana, Carnívora e Balanceada;
- f) subclasses da classe Atividade Física: Ativo, Semiativo e Sedentário;

- g) subclasses da classe Ocupação: Insalubre, Estressante e Normal.

4.3 ESTRUTURA

No modelo CESM, a estrutura é definida como a coleção de ligações entre componentes, e entre esses e itens do ambiente.

As classes e subclasses estão definidas segundo uma hierarquia conforme apresentado na figura 2. A ontologia foi construída na ferramenta *Protégé* (2007). Todas as figuras foram extraídas do modelo desenvolvido na ferramenta.



Figura 2. Classes e a hierarquia de classe

Além disso, é necessário estabelecer os diferentes relacionamentos entre as classes, representados na figura 3. A figura 3 apresenta, em destaque, a representação entre Paciente e Órgão, para ser

utilizado como exemplo. A classe Paciente tem um relacionamento com a classe Órgão do tipo **temÓrgão**; assim, pode-se concluir que **Paciente tem Órgão**. Nota-se, também, que a relação **temÓrgão** possui a relação inversa **compõe**, da forma **Órgão compõe o Paciente**.

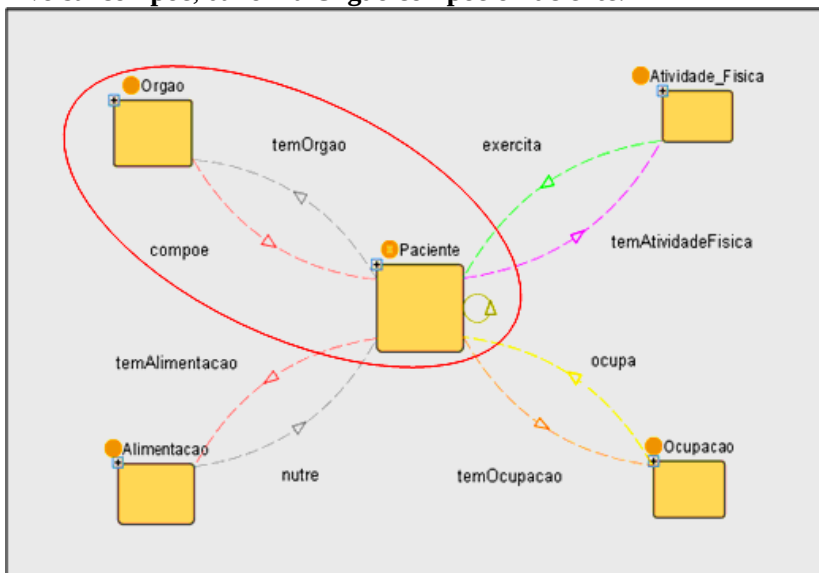


Figura 3. Relações entre Classes

4.4 MECANISMO

Para Bunge (1997), o mecanismo é uma coleção de processos que geram mudanças qualitativas no sistema fazendo com que o sistema evolua ou retroceda.

Para que o sistema possa gerar esta mudança a respeito do processo de diagnóstico, as restrições para uma classe são formalizadas. As restrições definem limitações nas relações entre as classes e são usadas nos processos de consulta e raciocínio sobre a ontologia.

Um exemplo de restrição aplicada a uma classe é apresentado na figura 4. Neste exemplo, foram definidas condições necessárias e suficientes para a classe Acne por Calor Tóxico. Nesta classe os órgãos devem apresentar os seguintes sintomas: Língua saburra amarela ou Língua saburra pegajosa; Pulso deslizante ou Pulso rápido; Face vermelha. Outra restrição desta classe é herdada da classe paciente que estabelece o mínimo de 3 (três) órgãos para caracterizar a patologia.

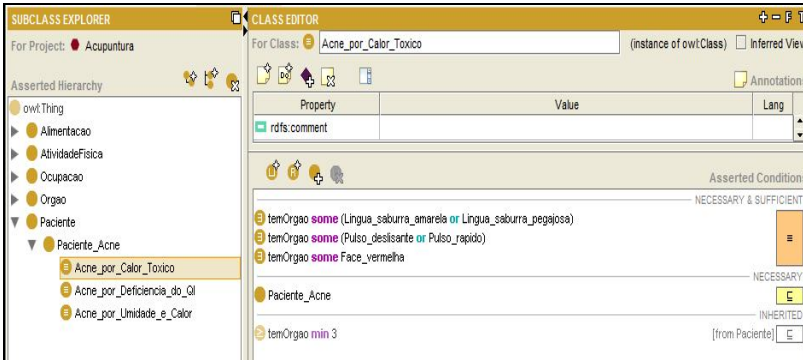


Figura 4. Restrições de Classe

5. RESULTADOS

Depois de realizada a modelagem da ontologia, baseada no modelo CESM, o propósito desta seção é demonstrar os possíveis estados do sistema, dentro da ferramenta *Protégé*. Para isto, é fundamental a criação de indivíduos (instâncias), formando uma base de dados, com os quais serão realizadas as consultas e raciocínios.

A figura 5 apresenta 11 instâncias da classe Paciente.

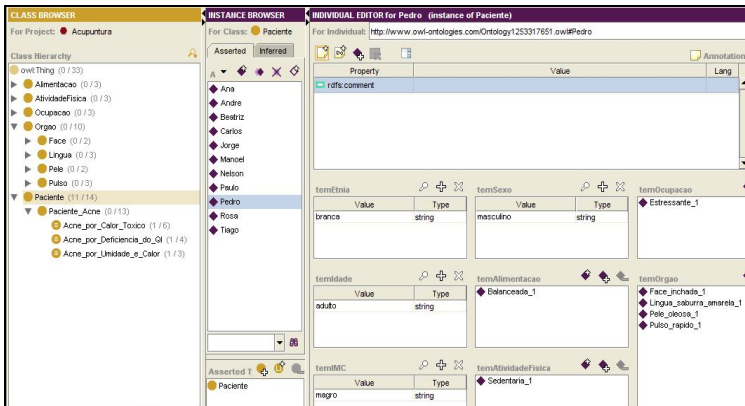


Figura 5. Exemplos de instâncias da classe Paciente

Pela Figura 5, observa-se que o paciente **Pedro**, em destaque, apresenta os seguintes dados: **temEtnia: branca**, **temIdade: adulta**, **temSexo: masculino**, **temIMC: magro**, **temAlimentação: balanceada**, **temOcupação: estressante** e **temAtividadeFisica: Sedentaria_1**.

sedentária. Pedro apresenta ainda os sintomas: **face_inchada**, **língua_saburra_amarela**, **pele_oleosa** e **pulso_rápido**.

Na ontologia é possível fazer consultas ou raciocínios sobre a mesma. As consultas são questões formuladas à ontologia para as quais ela apresenta suas respostas com base nas relações entre os indivíduos. O raciocínio permite inferir novas associações.

A figura 6 apresenta um exemplo de consulta onde estão sendo exibidos todos os pacientes com pele oleosa.

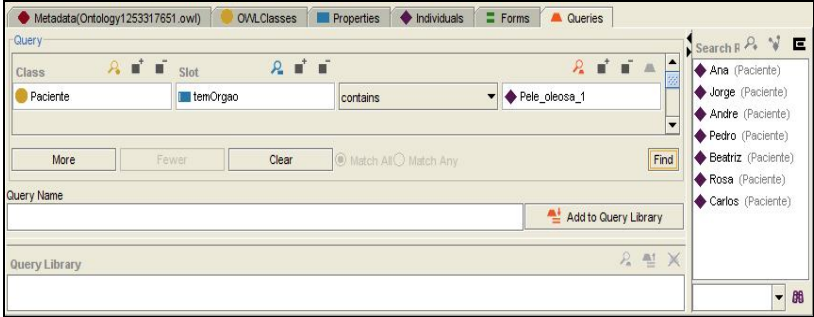


Figura 6. Exemplo de consultas sobre a ontologia Acne

A figura 7 formaliza um exemplo de inferência realizado pela ontologia. O processo de raciocínio inferiu que **Ana** e **Pedro** satisfazem as restrições da patologia **Acne por Umidade e Calor**, sendo, portanto, diagnosticados contendo esta patologia.

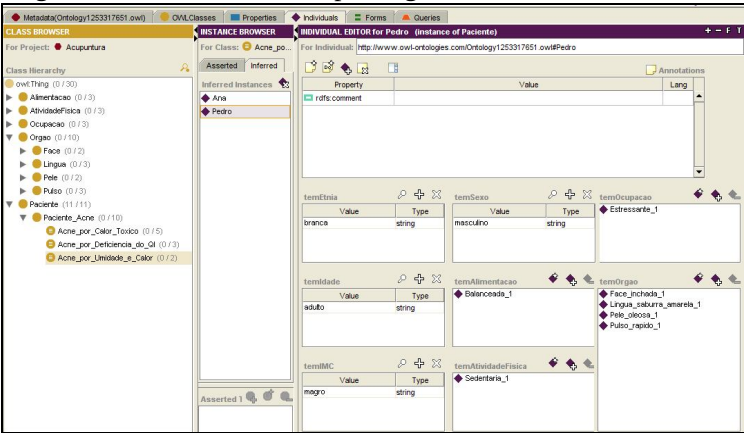


Figura 7. Exemplo de raciocínio sobre a ontologia Acne

6. CONCLUSÃO

A Medicina Tradicional Chinesa tem uma visão ampla e irrestrita ao analisar um organismo. A acupuntura é um tratamento que busca um perfeito equilíbrio do mesmo, devendo, no processo de diagnóstico, investigar-se as mais variadas informações que possam contribuir para um exame mais eficiente.

A principal contribuição deste trabalho é aplicar modelos e ferramentas da Engenharia do Conhecimento que permitam uma melhor representação dos elementos envolvidos no processo de diagnóstico, facilitando sua compreensão e possível expansão.

A união entre ontologias e a conceituação proposta por Bunge (2003), através do modelo CESM, possibilitou que o processo de diagnóstico fosse representado e visualizado de forma sistêmica. Através da quádrupla: componentes, ambiente, estrutura e mecanismo, e dos elementos que compõem uma ontologia, foi possível fazer uma representação objetiva do sistema de diagnóstico da enfermidade acne na acupuntura, estabelecendo-se um processo claro e conciso. Ao aliar estes elementos, a Engenharia do Conhecimento pode contribuir de forma simples para auxiliar na explicitação de conhecimento tácito, sem se utilizar de modelos baseados em caixa-preta, como ocorre geralmente nos sistemas de informação tradicional.

Como trabalhos futuros, pretende-se aprofundar os estudos aqui apresentados, efetuando-se aplicações em ambientes reais. Também se pretende buscar novos conhecimentos a partir do modelo sistêmico, de forma a permitir um diagnóstico mais preciso, diminuindo o grau de incerteza do tratamento. Um perfeito equilíbrio, entre o organismo do paciente e o meio em que vive, pode proporcionar uma vida mais saudável e duradoura.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- BUNGE, M. (2003). *Emergence and convergence: Qualitative novelty and the unity of knowledge*. University of Toronto Press, Toronto – Canadá.
- BUNGE, M. (2004). How does it work? The search for explanatory mechanisms. *Philosophy of the Social Sciences* , 34 (2), 182 - 210.
- BUNGE, M. (1997). Mechanism and explanation. *Philosophy of the Social Science* , 27 (4), 410-465.
- BUNGE, M. (2000). Systemism: the alternative to individualism and holism. *Journal of Socio-Economics* , 29, 147 - 157.

- CAMILLOTI, B. M., ISRAEL, V. L. & CHI, A. (2006). Necessidade de um Sistema de Apoio À Decisão em Acupuntura, Pontifícia Universidade Católica do Paraná (PUCPR), Universidade Federal do Paraná – Litoral (UFPR), Brasil.
- GRUBER, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 6(2): 199–221.
- KENDALL, S. L. & CREEN, M. (2007). *An introduction to knowledge engineering*. London: Springer_Verlag.
- KILOV, H. & SACKS, I. (2009). Mechanisms for communication between business and IT experts. *Computer Standards & Interfaces*, 31 (1): 98–109.
- LOPES, L. F., LOPES, M. C., GONÇALVES, A. L. & FIALHO, F. A. (2009). Uma Ontologia de Apoio ao Diagnóstico Aplicado à Acupuntura. In: *O Brasil no Contexto da Gestão do Conhecimento para a Inovação*, Salvador – Brasil.
- LUKMAN, S., YULAN, H. & SIU-CHENG, H. (2007). Computational methods for traditional Chinese medicine: a survey. *Computer Methods and Programs in Biomedicine*, 88 (3), 283 - 294.
- MACIOCIA, G. (2005). *Diagnóstico na medicina chinesa: um guia geral*. São Paulo: Roca.
- NECHES, R., FIKES, R. E., FININ, T., GRUBER, T. R., SENATOR, T. & SWATOUT, W. R. (1991). Enabling technology for knowledge sharing. *AI Magazine*, 12 (3), 36-56.
- NOY, N. & MCGUINNESS, D. L. (2001). *Ontology Development 101: a guide to creating your first ontology*. Acesso em 15 de nov de 2009, disponível em <http://protégé.stanford.edu/publications/ontology>
- PROTÉGÉ. (2007). <http://protégé.stanford.edu>. Acesso em 2 de setembro de 2009
- RADU, A. A.T. (2001). CompAc information system for traditional Chinese medicine. *Studies in Health Technology and Informatics*, 81, 401-403.
- ROSS, J. (1994). *Zang-Fu: Sistema de Órgãos e Visceras da Medicina Tradicional* (2ª ed.). São Paulo: Roca.
- SCHREIBER, G., AKKERMANS, H., ANJEWIERDEN, A., HOOG, R., HADBOLT, N., VELDE, W.V. & WIELINGA, B. Knowledge engineering and management: the CommonKADS methodology. [S.l.]: Massachusetts: MIT Press, 2002.
- YAMAMURA, Y. (2004). *Acupuntura tradicional: a arte de inserir* (2 ed.). São Paulo: Roca.