

Gestão e Representação do Conhecimento

Hellen Carmo de Oliveira

Cedric Luiz de Carvalho

Technical Report - RT-INF_003-08 - Relatório Técnico
March - 2008 - Março

The contents of this document are the sole responsibility of the authors.
O conteúdo do presente documento é de única responsabilidade dos autores.

Instituto de Informática
Universidade Federal de Goiás
www.inf.ufg.br

Gestão e Representação do Conhecimento

Hellen Carmo de Oliveira *
hellen.sistemas@gmail.com

Cedric Luiz de Carvalho †
cedric@inf.ufg.br

Abstract. *The knowledge is a good that should be properly stored so that its retrieval and distribution can be done efficiently. In order to do this the knowledge must be syntactically and semantically described. In a computational environment, the knowledge should be represented according to logics and rules of a specific language, so that one can use it to solve problems in an efficient and effective way. Thus, knowledge representation can be seen as an intelligent way to organize information to help people solving complex problems. Information Technology provides database to store data, information and knowledge. This text shows how the knowledge can be represented and stored in a computer and the steps to be followed to achieve this, according to the Knowledge Engineering, by using rules, logics, semantic networks, frames, etc.*

Keywords: Knowledge, Information Management, Knowledge Representation

Resumo. *O conhecimento é um bem que deve ser armazenado de forma a permitir sua recuperação e distribuição com organização e objetividade. Para o conhecimento ser armazenado, recuperado e apresentado ele deve estar descrito de forma sintática e semântica. Em um ambiente computacional, o conhecimento deve ser representado de acordo com lógicas e regras de uma linguagem específica, a fim de que se permita sua utilização para solução de problemas de forma eficiente e eficaz. Sendo assim, a representação do conhecimento pode ser vista como uma forma inteligente de se organizar informações para auxílio na solução de problemas complexos. A Tecnologia da Informação contribui com recursos de banco de dados para armazenar dados, informações e conhecimento. Desta forma, este trabalho trata da representação computacional do conhecimento, as etapas para esta representação dentro da Engenharia do Conhecimento utilizando-se de regras, lógicas, redes semânticas, quadros etc.*

Palavras-Chave: Conhecimento, Gestão da Informação, Representação do Conhecimento

1 Introdução

Uma das principais tendências que se pode identificar no momento atual é o deslocamento do paradigma de sociedade industrial para sociedade da informação ou sociedade do conhecimento.

*Mestrando em Ciência da Computação – INF/UFG

†Orientador – INF/UFG

O conhecimento é essencial para tomadores de decisão, no sentido de se planejar e implementar produtos, serviços e sistemas de informação. Nos países mais avançados, o acúmulo de informações e a sua reunião para estruturar esquemas de conhecimentos cada vez maiores e mais sofisticados permitiu mudanças qualitativas no setor econômico. Hoje, o poderio econômico internacional de um país está diretamente relacionado ao fator conhecimento. Por isso, este deve ser exposto para qualquer cidadão para promoção de avanços em todos os setores (econômicos, tecnológicos, científicos, culturais etc).

Na sociedade da informação, onde o conhecimento é o único recurso realmente significativo, é necessário que existam espaços de divulgação de textos, vídeos, imagens e outros que retratem o pensamento dos profissionais, pesquisadores, professores e estudantes. Em todos esses espaços de divulgação é necessária uma ordem e coerência das informações para que o conhecimento seja realmente adquirido.

A tecnologia da informação encontra-se diante do crescente desafio de proporcionar aos tomadores de decisão a apresentação de informações confiáveis, precisas, oportunas e relevantes.

Segundo Lévy[11], atualmente, a informação pode estar registrada em diferentes suportes concomitantemente, seja em documento impresso, documento eletrônico e/ou documento digital. Estes documentos podem estar disponíveis em diversos ambientes da informação: no ambiente atual, no ambiente eletrônico, no ambiente digital e/ou no ambiente virtual.

A *World Wide Web* ou simplesmente, *Web*, é um caminho de acesso à informação na Internet. É um modelo de domínio público de informação que foi construído para a busca de informações, o que pode ocorrer através de diretórios, mecanismos de busca e metabuscadores.

Se, por um lado, a tecnologia e o desenvolvimento das redes de informações possibilitam a difusão do conhecimento, por outro lado impulsionam a publicação direta da fonte ao consumidor. Isto gera uma falta de padrões para a disponibilização de documentos/informação na Internet e, conseqüentemente, dificulta a busca e a recuperação da informação nesses ambientes virtuais.

O processamento da informação é uma atividade complexa uma vez que a informação, dependendo do contexto e do domínio do conhecimento, pode ter significados diversos, ou seja, um termo ora representa um conceito ora pode representar outro e constituir relações conceituais diferentes[14].

Neste trabalho, será discutido o potencial dos fundamentos teórico-metodológicos para organização do conhecimento, visando a representação e recuperação da informação na *Web*. Serão apresentadas formas de representar computacionalmente o saber, a crença, ou a suposição. Também se discutirá como o computador pode usar os dados para guiar o seu comportamento, isto é, formas de representar o conhecimento computacionalmente, permitindo que este seja recuperado posteriormente.

Na seção 2, serão apresentados os conceitos ligados ao Gerenciamento do Conhecimento, tais como: o que é o conhecimento, quais são os fundamentos da gestão do conhecimento e como é o processo de transição da gestão da informação à gestão do conhecimento. Na seção 3, serão abordados os processos de representação do conhecimento, de acordo com os métodos da engenharia do conhecimento, no âmbito da Inteligência Artificial. Nesta seção, serão tratados os processos envolvidos na aquisição e na representação do conhecimento. Por fim, na seção 4, examinar-se-á quais os benefícios da gestão do conhecimento e em que a representação do conhecimento poderá ser útil para uma gestão eficiente.

2 Gestão do Conhecimento

A gestão do conhecimento envolve a sua descrição, organização, representação e tratamento. Nas subseções a seguir serão apresentados algumas definições e conceitos sobre estas questões.

2.1 Conhecimento

Segundo Platão, filósofo da Grécia antiga e um dos primeiros pensadores a sistematizar a definição do conhecimento, ele pode ser considerado “a crença verdadeiramente justificada”. Os filósofos mais modernos definem o conhecimento como a representação elaborada pela inteligência, exclusivamente a partir de impressões pessoais, como um processo humano e dinâmico para justificar a crença pessoal com relação à verdade. O conhecimento pode ainda ser considerado como a capacidade de agir. Ou seja, a criação do conhecimento é unicamente feita por indivíduos[23].

Segundo o Dicionário Aurélio [5], as definições de conhecer e conhecimento são:

“**Conhecer**: do Lat. *cognoscere* v. tr., ter conhecimento; ter a idéia ou a noção de; ter relações com alguém; saber, estar certo de; ser muito versado em; distinguir, apreciar; julgar, avaliar; reconhecer, admitir; ter visto ou visitado; v. int., tomar conhecimento.”

“**Conhecimento**: s. m., ato ou efeito de conhecer; idéia; noção; informação; notícia; experiência; discernimento; relações entre pessoas não íntimas; consciência de si próprio; recibo de contribuição paga; (no pl.) instrução; (no pl.) perícia; (no pl.) erudição; (no pl.) cultura;”

Segundo Burnham [3]: “Conhecimento é uma mistura fluída de experiência condensada, valores, informação contextual e introspecção experimentada, a qual proporciona uma estrutura para a avaliação e incorporação de novas experiências e informações.”

O conhecimento pode estar sob dois formatos:

- **Tácito (subjeto)**: conhecimento que o indivíduo adquiriu ao longo da vida, que está na cabeça das pessoas. Geralmente, é difícil de ser formalizado ou explicado a outra pessoa, pois é subjetivo e inerente às habilidades de uma pessoa. É de difícil captura, registro e divulgação;
- **Explícito (objeto)**: se refere ao conhecimento que pode ser transmitido em linguagem formal e sistemática. É o conhecimento formal, claro, regrado, fácil de ser comunicado. Pode ser formalizado em textos, desenhos, diagramas, etc. assim como guardado em bases de dados ou publicações.

Transformar o conhecimento tácito em explícito, tornando-o reutilizável por outras pessoas, não é uma tarefa simples pois, como descrito anteriormente, o conhecimento tácito é pessoal e difícil de ser articulado em uma linguagem formal, já que envolve diversos fatores (emocionais, psicológicos e outros). Para haver a transformação do conhecimento e, conseqüentemente, a expansão deste, é necessária uma interação social entre os dois formatos do mesmo, ou seja, fazer uma “mistura” dos conhecimentos tácitos e explícitos. A partir daí, o conhecimento individual acumulado precisará ser socializado novamente de forma a gerar novos conceitos, quando aplicado à novas necessidades.

A reestruturação das informações existentes através da classificação, do acréscimo, da combinação e da categorização do conhecimento explícito pode levar a novos conhecimentos.

A interação entre o conhecimento explícito e o conhecimento tácito resulta em quatro modos de conversão do conhecimento, discutidos a seguir e apresentados na Figura 1, descrita por Nonaka e Takeuchi[16].

- Socialização - é o processo de compartilhamento de experiências entre os indivíduos de um grupo e que se desenvolve, freqüentemente, por meio da observação, da imitação e da prática. Dessa forma, é possível transferir o conhecimento tácito entre os indivíduos e a associação de um mesmo tipo de conhecimento a diferentes contextos individuais;
- Externalização - é o processo de organização do conhecimento tácito em conhecimento explícito, por meio de metáforas, analogias, conceitos, hipóteses e modelos, permitindo a criação de conceitos novos e explícitos baseando-se no conhecimento tácito;
- Combinação - é o processo de sistematização de conceitos em um sistema de conhecimento. Assim, envolve a combinação de conjuntos de conhecimento explícito (tais como, a classificação, a sumarização, a pesquisa e a categorização das informações) com a utilização da tecnologia de banco de dados, podendo levar à criação de novos conhecimentos;
- Internalização - é o processo de incorporação do conhecimento explícito no conhecimento tácito, ou seja, é o modo pelo qual o conhecimento explícito torna-se ferramenta de aprendizagem, por meio de manuais ou documentos, e volta a assumir um contexto abstrato e subjetivo para cada membro na organização.



Figura 1: O processo cíclico de conversão do conhecimento

2.2 Fundamentos da gestão do conhecimento

O conceito de gestão do conhecimento surgiu no início da década de 1990 e logo se tornou parte da estratégia empresarial[26]. Diversas definições são apresentadas na literatura, mas, em resumo, a gestão do conhecimento é uma coleção de processos responsáveis pela criação, disseminação e utilização do conhecimento, visando atingir plenamente os objetivos da organização[7].

Por ser uma área muito abrangente, a gestão do conhecimento é tratada atualmente sob diferentes enfoques de acordo com a área de atuação e interesse de cada um dos pesquisadores, estudiosos e profissionais de negócio[25]. A seguir, são abordados alguns termos, e respectivas definições, referentes à gestão do conhecimento:

- Memória organizacional - é o conjunto de todo tipo de artefatos e documentos de que se tem registro na organização.

- Aprendizado organizacional - tem seu entendimento baseado num quadro de mudanças organizacionais cuja base de conhecimento está centrada em novos valores e novas formas de entender o mundo, as pessoas e as organizações.
- Ecologia da informação - pode ser considerada como o ramo do conhecimento que estuda o meio no qual a informação se origina, se prolifera, sofre refinamentos e é compartilhada gerando, nesse processo, novas informações.
- Inteligência competitiva - é uma estratégia para a organização descobrir o que se passa no ambiente de negócios do seu setor, e esse conhecimento dá aos gestores condições de tomar atitudes que forneçam à organização vantagem competitiva.

Esses termos estão diretamente ligados à gestão do conhecimento no ambiente organizacional e a ela compete estabelecer a correta conexão entre as estratégias de negócios, o conhecimento e a tecnologia da informação.

Em resumo, a gestão do conhecimento apresenta como objetivos:

- Tornar acessível grande quantidade de informação corporativa;
- Permitir a identificação e mapeamento dos ativos de conhecimento e de informações relacionadas a uma certa instituição ou comunidade;
- Dar apoio à criação de novos conhecimentos, visando benefícios administrativos, em pesquisas e outros;
- Dar utilidade e lógica aos dados, transformando-os em informação compreensível e essencial ao desenvolvimento da instituição.

2.3 Da gestão da informação à gestão do conhecimento

A tecnologia da informação (TI) possibilita coletar, processar e armazenar dados pela utilização de sistemas de informação, objetivando proporcionar aos tomadores de decisão a apresentação de informações confiáveis, precisas, oportunas e relevantes[24].

A TI representa um papel importante no processo de armazenamento e disseminação eficaz das informações. É nesse contexto que a gestão do conhecimento atua com o objetivo de agregar valores às informações.

Drucker[6] refere-se à gestão do conhecimento como o passo seguinte à gestão da informação. Com base no pressuposto de que a gestão da informação é condição básica e fundamental para a gestão do conhecimento, é necessário diferenciá-las (Figura 2).



Figura 2: Relação entre gestão da informação e gestão do conhecimento

Dados e informação não são conhecimentos, embora muitas vezes sejam considerados como tal. Dados representam um conjunto de fatos discretos e objetivos sobre eventos, podendo ser entendidos numa organização como a matéria-prima básica da informação. A informação, por sua vez, constitui-se de um fluxo de dados interpretados, dotados de relevância e propósito[6].

Conforme Setzer[27], tem-se as seguintes definições:

- **Dado:** é a informação bruta, envolvendo a descrição de um objeto ou de um evento. Em si não é dotado de relevância, mas é a matéria-prima para a criação da informação.
- **Informação:** são os dados interpretados, dotados de relevância e objetivo, capazes de gerar conhecimento.
- **Conhecimento:** é derivado da informação, sendo uma mistura de elementos, podendo ser comparado a um sistema vivo, pois cresce e sofre modificações conforme interage com o meio.

A descoberta do conhecimento é caracterizada como um processo composto por várias etapas operacionais. Na Figura 2 são apresentadas as três etapas essenciais. Em resumo, a etapa de Processamento de Dados compreende as funções relacionadas à captação, à organização e tratamento dos dados. Esta etapa tem como objetivo a preparação dos dados para os algoritmos da etapa seguinte, a gestão da informação. Durante a etapa de gestão da informação é realizada a busca por conhecimentos úteis no contexto da aplicação de descoberta do conhecimento em base de dados. A etapa de gestão do conhecimento, tem como objetivo viabilizar a avaliação da utilidade do conhecimento descoberto [9].

Na próxima seção, será apresentado como viabilizar ou representar todo o conhecimento avaliado pela etapa de gestão do conhecimento.

3 Representação do conhecimento

Nas seções anteriores foram definidos os conceitos de dados, informações e conhecimento. Foram apresentados os motivos da busca pela aquisição do conhecimento na atualidade. Também foi mostrado que para obter, armazenar e organizar tanta informação e abstrair das mesmas o conhecimento útil é necessário executar algumas etapas, sendo a principal delas a gestão do conhecimento como foi apresentado na Seção 2.1.

Os dados são geralmente armazenados em bancos de dados. O conhecimento pode ser armazenado em estrutura semelhantes. Entretanto, essas estruturas, denominadas de “bases

de conhecimento”, não são como os bancos de dados tradicionais, são bases que constituem partes de programas/sistemas computacionais inteligentes ou especialistas. Elas não são uma simples coleção de informações. A base tradicional de dados com dados, arquivos, registros e seus relacionamentos estáticos é aqui substituída por uma base de regras e fatos e também heurísticas que correspondem ao conhecimento do especialista (ou dos especialistas) no domínio sobre o qual foram construídos os sistemas[13]. Esses sistemas utilizam-se de várias técnicas para a representação do conhecimento na memória do computador (regras de produção, redes semânticas, frames etc).

Programas de computador são baseados em modelos abstratos do mundo. Esses modelos são onipresentes não só em programas de computador, mas também no pensamento e na linguagem humana. Quando se programa um computador, faz-se uma descrição explícita de um modelo, codificando na linguagem adequada, as regras e fatos, que se considera relevantes ao comportamento do programa. Este trabalha seguindo estas regras e manipulando estas representações. De fato, os computadores diferem de outras máquinas porque funcionam manipulando as representações que são formuladas sempre nos termos dos modelos. Resume-se assim que “não há computação sem representação”[10].

A representação que o computador manipula a fim de alcançar as soluções representa um modelo mental de um certo domínio ou modelo do mundo. Essa representação preocupa-se com três tipos de entidades[10], como mostrado na Figura 3.

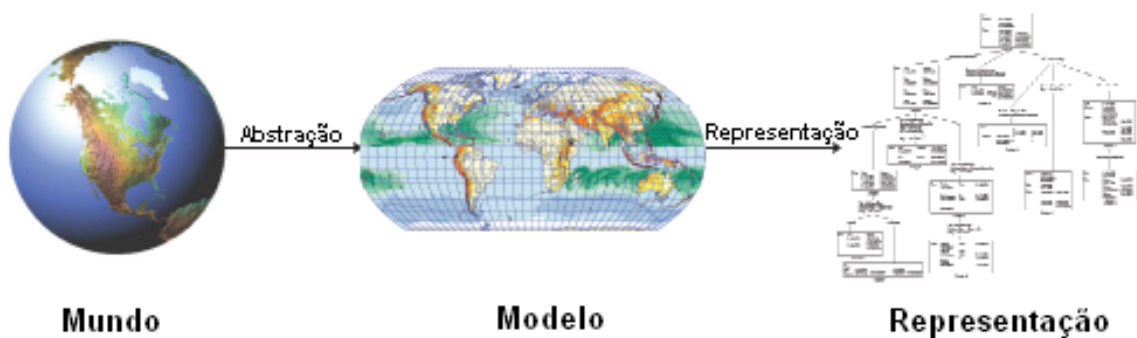


Figura 3: Relação entre mundo, modelo e representação

O mundo possui uma infinita riqueza e complexidade, das quais só se percebe parte. Um modelo do mundo é uma abstração que captura as suas aparências relevantes acerca de um determinado problema ou tarefa. A representação do modelo ou domínio do mundo é um conjunto de regras e afirmações que expressam o modelo usando uma linguagem.

A seguir, são apresentadas algumas definições, de acordo com a Engenharia do Conhecimento, no âmbito da Inteligência Artificial.

A **Inteligência artificial (IA)** é a área da computação que estuda o desenvolvimento de programas que se comportam de maneira inteligente. É orientada ao entendimento, construção e validação de sistemas inteligentes, isto é, que exibem, de alguma forma, características associadas ao que se chama inteligência [10][18]; ou “Inteligência Artificial é a parte da ciência da computação que compreende o projeto de sistemas computacionais que exibam características associadas, quando presentes no comportamento humano, à inteligência”[1].

O desenvolvimento real da IA é, na maior parte, devido ao conceito de um sistema computadorizado capaz de usar uma parcela do conhecimento, explicitamente representado, raciocinando sobre seus objetivos, sobre seu ambiente, sobre outros agentes, e mesmo sobre a si próprio[10]. O campo da IA vai ainda mais além: ele tenta não apenas compreender, mas também construir entidades inteligentes. Russell e Norvig[20] definem a IA como o estudo de

agentes que recebem percepções do ambiente e executam ações. Para esses autores, ela sistematiza e automatiza tarefas intelectuais e, portanto, é potencialmente relevante para qualquer esfera da atividade intelectual humana, ou seja, ela é verdadeiramente um campo universal.

O processo de Engenharia do Conhecimento tem como objetivo capturar e incorporar o conhecimento fundamental de um especialista do domínio ou modelo, bem como seus prognósticos e sistemas de controle. Este processo envolve reunir informação, familiarização do domínio, análise e esforço no projeto. Além disso, o conhecimento acumulado deve ser codificado, testado e refinado[17].

3.1 Engenharia do Conhecimento

A Engenharia do Conhecimento pode ser dividida em duas fases: a “Aquisição do Conhecimento” e a Representação do Conhecimento [8].

3.1.1 Aquisição do Conhecimento

A aquisição do Conhecimento - AC - é uma das atividades da Engenharia do Conhecimento e sua meta é obter conhecimento detalhado utilizado pelo especialista¹ para solucionar problemas e então transformar e transferir esta informação para um programa de computador.

“Aquisição do conhecimento é a transferência e transformação da habilidade ou perícia para resolver problemas contida em alguma fonte de conhecimento para um programa.”[21]

Vários autores citam diferentes fases do processo de aquisição do conhecimento que podem ser resumidas de acordo com a Figura 4.

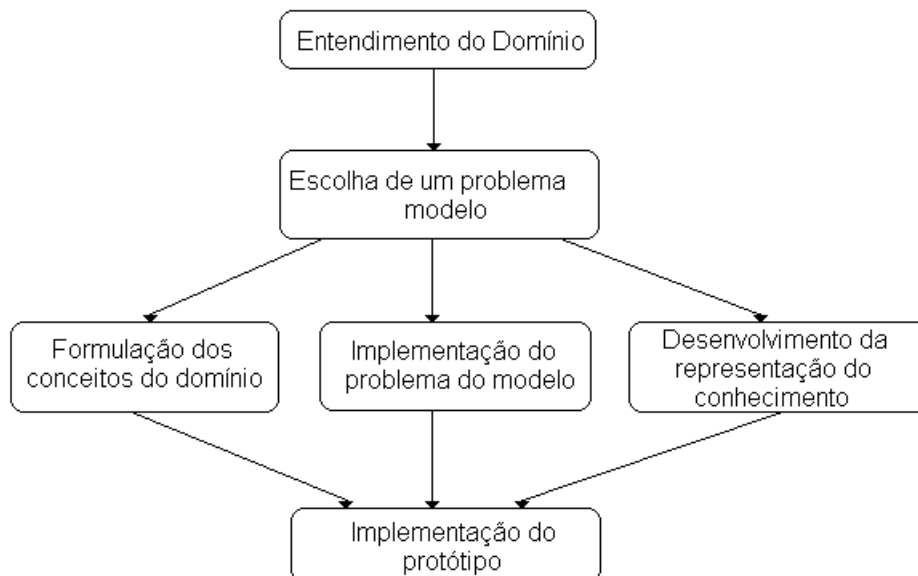


Figura 4: Processos de Aquisição do Conhecimento

¹Especialistas, também chamados de peritos ou *experts*, são pessoas que possuem um alto grau de conhecimento em dado domínio e habilidade para transmitir esse conhecimento. Em muitos casos eles são a fonte de conhecimento para a funcionalidade de um sistema especialista [17].

- **Entendimento do Domínio:** É o primeiro passo na aquisição do conhecimento; é um período de familiarização do domínio, uma visão de muito alto nível deste por parte do engenheiro do conhecimento. Nessa fase, obtém-se uma descrição geral do problema (domínio), uma relação de referências bibliográficas importantes e um glossário que descreve os termos próprios do domínio, símbolos e siglas;
- **Identificação do problema modelo:** Essa fase é a identificação de exemplos de problemas a serem usados no protótipo inicial, através de uma análise em amplitude que deve identificar uma lista de possíveis problemas existentes[8];
- **Formulação dos conceitos do domínio:** Identificar as subtarefas que formam o domínio[8];
- **Representação do conhecimento:** O Engenheiro do Conhecimento deve selecionar uma forma de representação do conhecimento procurando especificamente mantê-lo em um nível explícito[17];
- **Implementação de um problema modelo:** A forma mais comum é a simulação de cenários progressivos. O Engenheiro do Conhecimento deve manter uma visão global e extrair o conhecimento que torna-se disponível somente durante o processo de resolução do problema. A investigação completa do problema deve ser em profundidade, pois limita a complexidade e permite o acesso a conhecimento de baixo nível seqüencialmente[17];
- **Implementação do protótipo:** Essa é a fase final. Quando o protótipo já passou por todas as fases anteriores e já está pronto para adquirir conhecimento necessário e adicioná-lo à Base de Conhecimento[17].

Métodos ou técnicas para aquisição de conhecimento variam, pois não seguem um protocolo formal. Lira e Fantinato[17] consideram que as seis melhores classes de técnicas comumente usadas são:

- **Entrevistas:** é a interação entre pessoas onde o elicitor é o entrevistador e a fonte de conhecimento humano é o entrevistado. Baseia-se em uma estratégia de perguntas e respostas.
- **Protocolos:** tratando-se de elicitação de conhecimento, protocolos devem ser entendidos como registros externos, usualmente em vídeo ou áudio, em tempo real ou retrospectivamente.
- **Programação Neurolingüística:** método para capturar informações a partir de sinais não verbais, que representam atividades internalizadas (visão, audição, etc.) do provedor de conhecimento.
- **Traits:** a força da maioria dos sistemas baseado em conhecimento está na habilidade de selecionar opções, baseando-se em características das informações. A maioria dos especialistas tomam suas decisões com base em atributos de objetos, em eventos capazes de diferenciar os mesmos de outros objetos ou eventos. *Traits* é a técnica em que o engenheiro do conhecimento e o especialista estabelecem fatores e atributos característicos de objetos, que são refinados e então incorporados à aplicação.

- **Análise:** deve-se analisar e organizar toda fonte de dados explorada, além de compará-la com os objetivos do domínio. O objetivo da análise de material que fornece conhecimento é construir representações de conhecimento para validação e construção do sistema. É usada para esclarecer situações problemáticas, considerações alternativas e conjuntos de soluções apropriadas.
- **Aquisição Automatizada:** a aquisição manual de conhecimento é um processo caro e demorado, porém, a utilização de ferramentas para aquisição automatizada de conhecimento reduz estes pontos negativos. Uma série de módulos integrados realizam alguns aspectos dos processos de aquisição e representação de conhecimento, e geração de código. Outro tipo de aquisição automatizada é o aprendizado de máquina, onde o sistema consegue buscar novos conhecimentos a partir de bases de dados, textos, periódicos que serão digitalizados, analisados, sintetizados e incorporados a uma base de conhecimento inicial. Este processo emula o processo humano de remontar modelos através da captação de informações recentes.

3.1.2 Representação do Conhecimento

A manifestação inteligente pressupõe aquisição, armazenamento e inferência de conhecimento. Para que o conhecimento possa ser armazenado é essencial que se possa representá-lo. A IA tem se esforçado para aperfeiçoar o formalismo para a representação do conhecimento.

A representação do conhecimento é uma forma mais adequada para implementação de soluções dos problemas mais complexos na IA pois, para isso, é necessário uma grande quantidade de conhecimento e certos mecanismos para manipulá-lo[20].

A representação do conhecimento pode ser definida como um conjunto de convenções sintáticas e semânticas que torna possível descrever coisas. Consiste na utilização de linguagens específicas, frases ou números que correspondem à descrição ou condição do mundo[10]. Há uma enorme quantidade de técnicas que são utilizadas para poder representar conhecimento. Caberá ao engenheiro do conhecimento escolher aquela que melhor se adequar ao problema em questão.

A seguir, são apresentadas três perspectivas no campo de representação do conhecimento[17]:

1. Representação do conhecimento como epistemologia aplicada: utiliza como princípio a complexidade do conhecimento, ou seja, envolve-se com o problema do conhecimento pressuposto na atividade inteligente e a forma de representá-lo em uma base como estruturas e programas.
2. Representação do conhecimento como um módulo de pergunta resposta: é a mais baixa expectativa de um módulo de representação do conhecimento. Provém, no mínimo, duas operações:

Tell (K, F) = Dada uma base de conhecimento K , o fato F é somado a ela, resultando em uma nova base K' .

Ask (K, F) = A base de conhecimento K é examinada sobre um fato F .

A resposta, dependendo do paradigma de Representação do Conhecimento usado, pode ser sim, não ou desconhecido...

3. Representação do conhecimento como incorporação de sistemas de IA: esta é a visão conexionista, para a qual há várias unidades interconectadas idênticas que são coletivamente responsáveis por representar vários conceitos. Um conceito é representado num senso distribuído e é indicado por um envolvimento em atividades sobre uma coleção de unidades

Lira e Fantinato[17] citam algumas das características das formas de representação do conhecimento:

- escopo e granulosidade: partes do domínio consideradas, detalhamento.
- indeterminância e definição das noções primitivas de representação: alternativas de modelagem.
- modularidade/compreensibilidade: agrupamento (*clusterização*) do conhecimento, legibilidade.
- conhecimento explícito e flexibilidade: toda a informação necessária à solução do problema na Base de Conhecimento, e não embutida em outro componente.

3.2 Paradigmas de Representação do Conhecimento

Para se representar o conhecimento é necessária a escolha de métodos adequados. A linguagem associada ao método escolhido deve ser suficientemente expressiva (mas não mais do que o suficiente) para permitir a representação do conhecimento a respeito do domínio escolhido de maneira completa e eficiente.

Em tese, uma representação geral como a lógica seria suficientemente expressiva para representar qualquer tipo de conhecimento. No entanto, problemas de eficiência, facilidade de uso e a necessidade de expressar conhecimento incerto e incompleto levaram ao desenvolvimento de diversos tipos de paradigmas (ou técnicas de representação) e o desenvolvedor de um sistema deve avaliar qual destes se adequa ao problema que ele tem para solucionar. São descritos a seguir alguns paradigmas chave, observando-se que a quantidade existente é muito grande[2].

- **Conhecimento procedimental:** O conhecimento é representado em forma de funções/procedimentos. Estabelece como fazer algo, métodos de questionamento; critérios para utilização de habilidades, algoritmos, técnicas e métodos.
- **Árvores de decisão:** Conceitos são organizados em forma de árvores.

A Árvore de Decisão é um dos modelos mais práticos e mais usados em inferência indutiva. Este método representa funções como árvores de decisão. Estas árvores são treinadas de acordo com um conjunto de treino (exemplos previamente classificados) e, posteriormente, outros exemplos são classificados de acordo com essa mesma árvore.

As árvores de decisão classificam instâncias partindo da raiz da árvore para algum nodo folha que fornece a classe da instância. Cada nodo da árvore especifica o teste de algum

atributo da instância, e cada arco alternativo que desce daquele nodo corresponde a um dos possíveis valores deste atributo. Uma instância é classificada começando no nodo raiz da árvore, testa o atributo relacionado a este nodo, segue o arco que corresponde ao valor do atributo na instância em questão. Este processo é repetido então para a sub-árvore abaixo até chegar a um nodo folha [15].

Na Figura 5 é apresentada uma árvore de decisão típica. Esta árvore de decisão classifica os dias para jogar tênis, conforme eles são satisfatórios ou não.

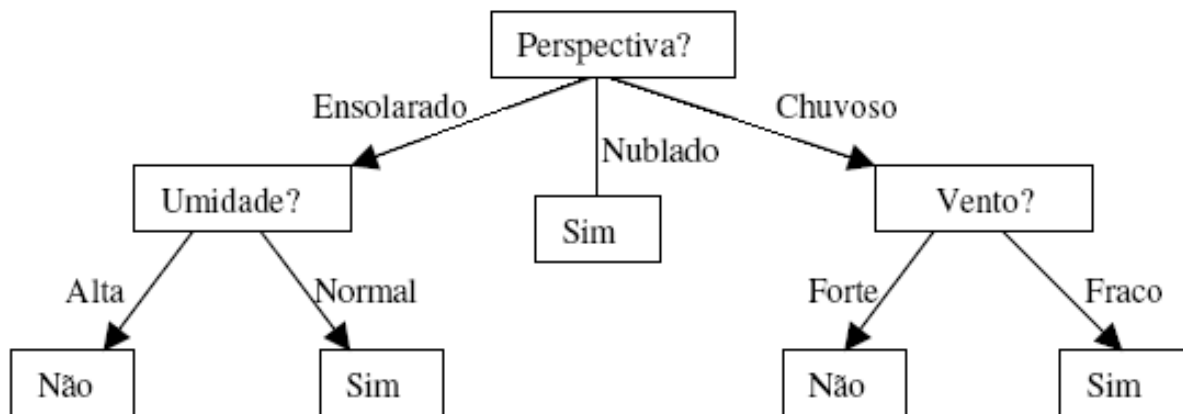


Figura 5: Árvore de decisão para jogar tênis

- **Processamento paralelo distribuído:** Utiliza-se de modelos conexionistas. Consiste em dividir uma tarefa em partes independentes e na execução de cada uma destas partes em diferentes processadores [4].

Basicamente é necessário:

- paralelizar os algoritmos;
 - um mecanismo para distribuição do processamento pelos diversos processadores disponíveis;
 - um mecanismo para troca de mensagens (informações e dados) entre os diferentes processos.
- **Esquemas híbridos:** Qualquer representação do formalismo que emprega a combinação de esquemas de representação do conhecimento.
 - **Regras:** Regra é a forma mais conhecida de representar conhecimento, usada atualmente apenas em sistemas de pequeno porte[17]. Usam sistemas de produção para codificar regras de condição-ação.

Regras de produção: Consiste em representar o domínio do conhecimento através de um conjunto de regras. Suas principais características são: modularidade, facilidade de implantação e também a grande quantidade de pacotes desta técnica existentes para o desenvolvimento de Sistemas Especialistas (SE)².

²SEs são sistemas computacionais que usam as técnicas de conhecimento e raciocínio para resolver problemas que normalmente requeriam perícia humana [19].

A representação do conhecimento baseada em regras pode ser identificada nas linguagens de computador, que permitem o armazenamento de conhecimentos na forma de um “SE” «Condição» e “ENTÃO” «Agir ou concluir alguma coisa». Exemplificando[12]:

SE um líquido inflamável foi derramado nas instalações

ENTÃO chamar os bombeiros

SE o “PH” do líquido derramado é menor do que 6

ENTÃO o material do líquido é ácido.

- **Lógica:** Uma das mais primitivas formas de representação do raciocínio ou conhecimento humano.

Lógica proposicional: é considerada a forma mais comum da lógica. Usam-se proposições e relações lógicas entre proposições. As proposições podem ser verdadeiras ou falsas, assim como demonstrado na Tabela 1 [10][22].

Tabela 1: Ilustração da irrelevância do valor-verdade

Valor-Verdade condição e conclusão	Argumento Válido	Argumento Inválido
(Verdadeiro, Verdadeiro)	Todos os homens são mortais Sócrates é homem ∴ Sócrates é mortal	Todos os números pares são múltiplos de 2 ∴ todos os múltiplos de 2 são números pares
(Verdadeiro, Falso)	-	Todos os cachorros são animais. ∴ todos os animais são cachorros
(Falso, Verdadeiro)	Todos os pássaros são humanos Todos os humanos têm pena ∴ todos os pássaros têm pena	todos os animais são cachorros ∴ todos os cachorros são animais
(Falso, Falso)	Todos os gatos são caninos todos os caninos têm pena ∴ todos os gatos têm pena	Todos cachorros são gatos ∴ todos os gatos são cachorros

Em linguagem natural, a premissa e a conclusão são normalmente separadas por palavras tais como “então”, “portanto”, “assim”, e assim por diante. Nos argumentos utilizados na tabela acima, o símbolo “∴” está sendo usado para separar a premissa da conclusão.

Considerando os argumentos apresentados na Tabela 1, para dizer se são válidos ou inválidos, a forma dos mesmos pode ser estudada independentemente do assunto das proposições. De fato, um argumento é válido ou inválido devido à sua forma e não por causa de seu assunto. Isto significa que todos os argumentos com a mesma forma poderão ser válidos ou inválidos [10].

Lógica de predicados: é considerada como uma extensão da lógica proposicional. Na lógica de predicados os elementos fundamentais são, além do objeto, também os seus predicados[2].

- **Inferências:** Examinam as regras existentes e deduzem novos procedimentos quando possível. Existem várias estratégias para inferir novos fatos. Uma delas é a “modus ponens”, a mais comum, na qual se inferem novos fatos na base de dados, sendo que se o antecedente da regra é verdadeiro então o conseqüente é aceito como verdade[10].
- **Redes:** Conhecimento é representado por um rótulo de grafos direcionados cujos nós representam conceitos e entidades, enquanto os arcos representam a relação entre entidades e conceitos.

Redes Semânticas: São primitivas de representação. É a forma de representação mais adequada para domínios onde problemas podem ser descritos como taxonomias, classificações complexas. Redes semânticas são representadas como um conjunto de nós ou nodos que são ligados por meio de arcos, onde cada nodo representa um objeto, uma entidade conceitual ou um evento e cada arco representa o relacionamento existente entre cada par de nodos, sendo que cada par de nodos representa um determinado fato [20].

As redes semânticas foram originalmente projetadas como um meio de representar os significados de palavras de língua inglesa. A representação em redes inspirou-se no modelo psicológico da memória humana associativa[19].

As redes semânticas também podem ser definidas como um conjunto heterogêneo de sistemas. Em última análise, a única característica comum a todos estes sistemas é a notação utilizada - uma rede semântica consiste em um conjunto de nodos conectados por um conjunto de arcos. Os nodos em geral representam objetos e os arcos, relações binárias entre esses objetos. Mas, os nodos podem também ser utilizados para representar predicados, classes, palavras de uma linguagem, entre outras possíveis interpretações, dependendo do sistema de redes semânticas em questão[2].

Os nós podem representar objetos, conceitos ou eventos e os arcos definem as relações hierárquicas existentes entre os nós(é-um, tem-um, tipo-de, maior-que) ou definem novas entidades(altura, cor). Eles podem estar em diferentes níveis: *tokens* estão nos níveis hierárquicos mais baixos e denotam indivíduos ou instâncias, e são conectados por arcos é-um; *types* são nós de níveis hierárquicos mais altos, representam classes ou categorias de indivíduos e são abstrações definidas sobre *tokens* que captam generalizações úteis daqueles.

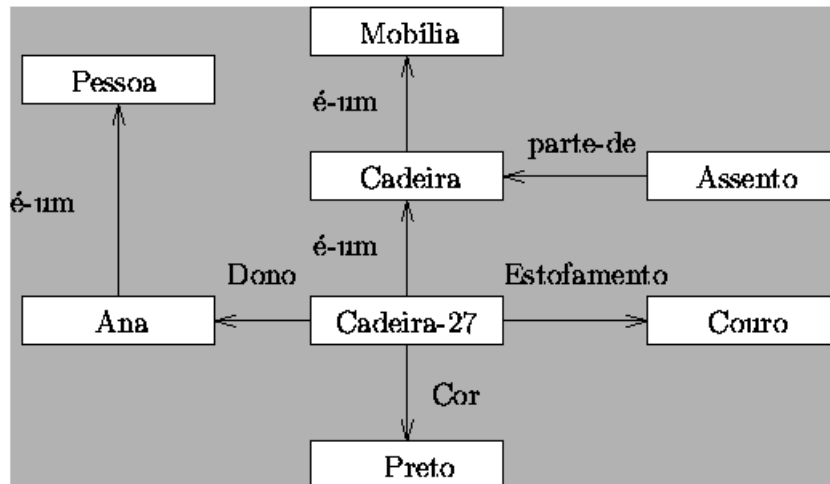


Figura 6: Rede Semântica

Exemplo: Considere-se a rede semântica da Figura 6, que representa conceitos relacionados com mobiliário. Os arcos **é-um** e **parte-de** estão determinando a herança de propriedades. Os demais arcos (**Dono**, **Cor**, **Estofamento**) são específicos do domínio e, como já foi citado acima, representam propriedades de conceitos.

Os nós **Pessoa**, **Mobília**, **Couro**, **Preto** são exemplos de *types*, os nós **Cadeira-27** são exemplos de *tokens* e o nó **Cadeira** é um *type* e um *token* ao mesmo tempo, pois está em dois níveis.

Sendo uma simples coleção de nós e arcos, essa rede é uma estrutura semântica menos complexa que os quadros, empregada com um conjunto de regras de inferência, especialmente projetadas para tratar de modo correto os tipos específicos de arcos presentes na rede[19].

Conforme Rover[19] as vantagens das redes semânticas são que as associações podem ser feitas de modo explícito e sucinto, isto é, os fatos relevantes acerca de um objeto ou conceito podem ser deduzidos dos nós a que eles estão ligados diretamente, sem uma procura através de uma extensa base de dados, à existência de metodologias de desenvolvimento e a repressão natural. As desvantagens são que, por ser uma simples coleção de nós e arcos, cujas informações estão armazenadas neles de forma direta, torna-se um fator de fragilidade do formalismo quando objetos bem mais complexos fizerem parte do domínio de conhecimento a ser implementado. Sendo assim, verifica-se que os quadros são estruturas mais poderosas.

- **Frames ou Quadros:** Representação do conhecimento baseada em estruturas, que categorizam o conhecimento em objeto, atributo e valor. Conhecimento em objeto é a descrição baseada em seus atributos[12].

Quadros são muito parecidos com as redes semânticas, exceto que cada nó representa conceitos e/ou situações. Cada nó tem várias propriedades que podem ser especificadas ou herdadas por padrão. É a forma de se representar o conhecimento de um objeto através da “observação visual”, ou seja, tendo-se uma idéia do objeto pré-definida na memória, faz-se a comparação da mesma com aquelas propriedades que se pode observar visualmente[17].

O método de quadros também está na origem das idéias que levaram às linguagens de

programação orientadas a objetos e consistem em sistemas de quadros especializados na descrição de seqüências de eventos.

Em geral, um quadro consiste em um conjunto de atributos que, através de seus valores, descrevem as características do objeto representado por eles. Os valores atribuídos a estes atributos podem ser outros quadros, criando-se uma rede de dependências entre eles. Quadros são também organizados em uma hierarquia de especialização, criando uma outra dimensão de dependência entre si. Os atributos também apresentam propriedades, que dizem respeito ao tipo de valores e às restrições de número que podem ser associados a cada atributo. Essas propriedades são chamadas *facetas*[2].

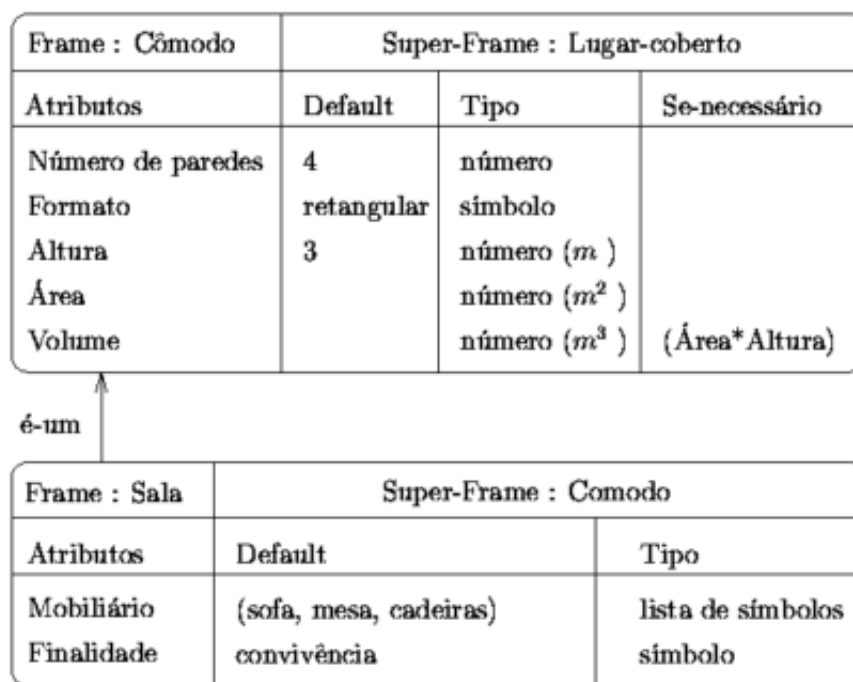


Figura 7: Quadros

Exemplo: Quadros descrevendo um cômodo e uma sala são mostrados na Figura 7. Por esta descrição, pode-se concluir que uma sala é um tipo de cômodo, normalmente com quatro paredes e de formato retangular, com um mobiliário específico. As facetas dos atributos especificam os tipos de valores esperados e, se for o caso, procedimentos adequados para calcular o valor do atributo [17].

Rover [19] diz que as vantagens na utilização de linguagens baseadas em quadros, como por exemplo, Common Lisp, são consideráveis: capturam a maneira que especialistas pensam a respeito de seu conhecimento, fornecem uma representação estrutural concisa de relações e dão suporte a uma técnica de definição por especialização que é fácil de ser utilizada pela maioria dos especialistas. As linguagens baseadas em quadros são particularmente poderosas, uma vez que as relações taxonômicas entre os mesmos capacitam as informações descritivas serem compartilhadas entre múltiplos quadros - via herança - e também devido ao fato que a estrutura interna dos mesmos capacita a manutenção automática de restrições semânticas.

Por outro lado, é consenso geral em tecnologia de sistemas de conhecimento que o domínio do conhecimento pode ser efetivamente utilizado e mais facilmente entendido pelos usuários, se for representado na forma declarativa, ao invés da forma procedimental. En-

tretanto, sistemas baseados em quadros não fornecem facilidades diretas para descrever, declarativamente, como o conhecimento armazenado nos quadros deve ser utilizado [19].

A Figura 8 mostra a representação abstrata da estrutura de um frame.

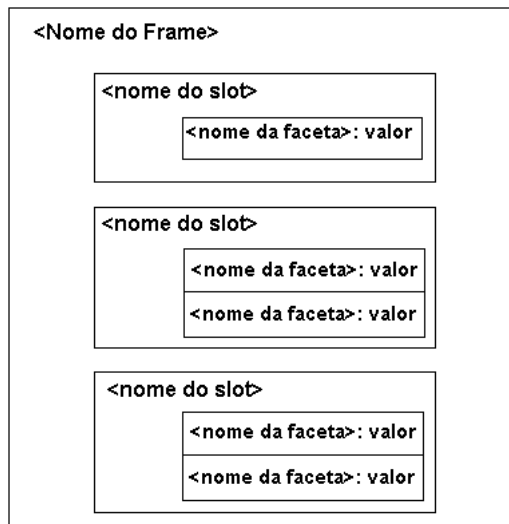


Figura 8: Representação abstrata da estrutura de um Frame

4 Considerações finais

Para não se perder o valor do conhecimento adquirido em uma organização é indispensável a utilização de sistemas capazes de validar esse conhecimento, armazená-lo e disponibilizá-lo para sua devida utilização e recuperação. É necessário fazer-se uma avaliação das concepções epistemológicas com relação às noções básicas que sustentam tais sistemas, como exemplificado aqui com a questão de conhecimento tácito e explícito.

Os sistemas computacionais capazes de representar o conhecimento utilizam técnicas ou métodos como regras, lógica, redes semânticas e quadros para ajudar na organização e aquisição do mesmo.

Conclui-se que a construção desses sistemas quase sempre exige a adoção de formas diferentes de representação de conhecimento, tendo-se em vista as deficiências que estas, isoladamente, possuem. Esses sistemas são produtos da combinação tanto de técnicas como de formalismos de representação diferentes.

Regras de produção e quadros são dois esquemas contrastantes e complementares por representar tipos diferentes de conhecimento. As árvores de decisão são mais simples e mais usados em inferências. O conhecimento procedimental difere das regras, pois utiliza funções/procedimentos para representar o conhecimento no lugar de condições. As regras são apropriadas para representar implicações lógicas nas quais é possível associar ações com condições. Os quadros são apropriados para definir termos e descrever objetos e as relações entre eles. As redes semânticas, como os quadros, são estruturas de finalidade geral, nos quais conjuntos particulares de conhecimentos específicos de domínio podem ser encaixados.

Quando o objetivo é combinar representações, é preciso focar a atenção tanto nas formas, quanto nos mecanismos, regras, procedimentos e objetos computacionais os quais dão conhecimento ao sistema.

Sendo assim, a IA está revolucionando o campo da gestão do conhecimento, sistematizando a realidade, facilitando e levando informações relevantes a todos os gestores que neces-

sitam do conhecimento para guiar e gerir uma organização, ou até mesmo, auxiliando pesquisadores e estudantes na evolução de novas pesquisas em vários domínios.

5 Agradecimento

Ao Prof. Dr. João Carlos da Silva, pela avaliação do presente texto e pelas sugestões feitas, as quais muito contribuíram para a melhoria do texto original.

Referências

- [1] BARR, A; FEIGENBAUM, E. **The Handbook of Artificial Intelligence**, volume v. I-II. William Kaufmann Inc., Los Altos, California, 1981.
- [2] BARRETO, J. M. **Inteligência Artificial no Limiar do Século XXI**. Florianópolis, Duplic, 3a. edição edition, 2001.
- [3] BURNHAM, T. F. **Aprendizagem Organizacional e Gestão do Conhecimento**. In: Informação, Conhecimento e Sociedade Digital, Bahia, Jun. 2007. In: Anais do VI CINFORM - Encontro Nacional de Ciência da Computação.
- [4] CÉLIO ODA MORETTI. **Análise de Estruturas Utilizando Técnicas de Processamento Paralelo Distribuído**. Master's thesis, Universidade de São Paulo, USP, 1997.
- [5] DE HOLANDA FERREIRA, A. B. **Dicionário Aurélio**. Editora Nova Fronteira, 4º edição edition, 2001.
- [6] DRUCKER, P. **Sociedade Pós-Capitalista**. Editora Thomson Learning, ed. 1 edition, 2001.
- [7] FILHO, J. T. **Gerenciando conhecimento: Como a Empresa Pode Usar a Memória Organizacional e a Inteligência Competitiva no Desenvolvimento dos Negócios**. Editora SENAC, Rio de Janeiro, 2000.
- [8] FURNIVAL, A. C. **Delineando as Limitações: Sistemas Especialistas e Conhecimento Tácito**. Ciência da Informação, 24(2), 1995.
- [9] GOLDSCHIMIDT, R; PASSOS, E. **Data Mining: Um Guia Prático**. Editora Campus, 2005.
- [10] JOHN, S. F. **Knowledge Representation: Logical, Philosophical, and Computational Foundations**. Brooks Cole Publishing Co., Pacific Grove, CA, actual publication date edition, Aug. 2000.
- [11] LÉVY, P. **As Tecnologias da Inteligência: O Futuro do Pensamento na Era da Informática**. Tradução Carlos Irineu da Costa. Editora 34, Rio de Janeiro, ed. 1 edition, 1993.
- [12] MAÑAS, A. V. **Administração de Sistemas de Informação : Como Otimizar a Empresa Por Meio dos Sistemas de Informação**. Editora Érica, quarta edição edition, 1999.

- [13] MENDES, R. D. **Inteligência Artificial: Sistemas Especialistas no Gerenciamento da Informação**. Ciência da Informação, 26(1), Jan./Abr. 1997.
- [14] MIRANDA, M. L. **Organização e Representação do Conhecimento: Fundamentos Teórico- Metodológicos na Busca e Recuperação da Informação em Ambientes Virtuais**. <http://www.portalppgci.marilia.unesp.br/enancib/viewabstract.php?id=274>, último acesso em Abril de 2007, 2006.
- [15] MITCHELL, T. **Machine Learning**. McGraw Hill, USA, 1997.
- [16] NONAKA, I; TAKEUCHI, H. **Criação de Conhecimento na Empresa: Como as Empresas Japonesas Geram a Dinâmica da Inovação**. Rio de Janeiro: Campus, 1997.
- [17] RICH, E; KNIGHT, K. **Artificial Intelligence**. McGraw Hill, USA, 2nd ed. edition, 1983.
- [18] RICH, E; KNIGHT, K. **Inteligência Artificial**. Makron Books, São Paulo, 2ª edição edition, 1994.
- [19] ROVER, A. J. **Representação do Conhecimento Legal em Sistemas Especialistas: O Uso da Técnica de Enquadramentos**. PhD thesis, Universidade Federal de Santa Catarina, UFSC, <http://www.infojur.ufsc.br/aires/princi.html>, último acesso em Junho de 2007, 1999.
- [20] RUSSELL, S; NORVIG, P. **Inteligência Artificial**. Editora Campus, tradução da segunda edição edition, 2004.
- [21] SCHWABE, D; CARVALHO, R. L. **Engenharia de Conhecimento e Sistemas Especialistas**. Editora Kapelusz - EBAI, 1987.
- [22] SOUZA, J. N. **Lógica para Ciência da Computação**. Campus, edição 1 edition, 2002.
- [23] SOUZA, T. F. C. **Documento, Informação e Conhecimento: Ainda uma Questão**. http://www.cinform.ufba.br/vi_anais/docs/TerezinhaSouza.pdf, último acesso em Abril de 2007, 2007.
- [24] TAIT, T. F. C. **Um Modelo de Arquitetura de Sistemas de Informação para o Setor Público: Estudo em Empresas Estatais Prestadoras de Serviços de Informática**. PhD thesis, Universidade Federal de Santa Catarina, UFSC, 2000.
- [25] TERRA, J. C. C. **Gestão do Conhecimento: Aspectos Conceituais e Estudo Exploratório Sobre as Práticas de Empresas Brasileiras**. PhD thesis, Escola Politécnica da USP, <http://www.terraforum.com.br/sites/terraforum/paginas/teses/teses.aspx>, último acesso em Junho de 2007, 1999.
- [26] WAH, L. **Muito Além de um Modismo**. HSM Management, Ano 4(22), Set./Out. 2000.
- [27] W.SETZER, V. **Dado, Informação, Conhecimento e Competência**. DataGramZero - Revista de Ciência da Informação, n. zero, Dez. 1999.