

UNIVERSIDADE DO VALE DO RIO DOS SINOS – UNISINOS  
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA APLICADA  
NÍVEL MESTRADO

Carolina Müller

M\_ONTO:

Proposta de Modelagem Semântica para uma Ontologia do Domínio EAD

São Leopoldo

2011

Carolina Müller

M\_ONTO:

Proposta de Modelagem Semântica para uma Ontologia do Domínio EAD

Dissertação apresentada como requisito parcial para a obtenção do título de Mestre em Linguística Aplicada pela Universidade do Vale do Rio dos Sinos – UNISINOS.

Orientadora:

Profª Drª Rove Luiza de Oliveira Chishman

São Leopoldo

2011

*Ao meu pai (in memoriam) e minha mãe por sempre me incentivarem e valorizarem meu estudo.*

*Aos meus amores, Denis e Daniel, pela compreensão, paciência e amor.*

## AGRADECIMENTOS

É tempo de agradecer...

...agradecer a todos que estiveram ao meu lado e que, de uma forma ou outra, contribuíram para a realização deste trabalho.

De modo especial, agradeço:

À professora Rove Chishman por ter acreditado em mim e me permitido fazer parte do seu grupo de pesquisa; pela paciência e incentivo; pela compreensão às minhas limitações e dificuldades; por ser uma orientadora sempre atuante, mesmo à distância e por me apresentar de forma tão apaixonada a Linguística Computacional e a Semântica.

Ao meu marido Denis, pelo incentivo, ajuda e, sobretudo, paciência e compreensão.

Ao meu amado Daniel pela compreensão nos momentos em que estive afastada para estudar.

À professora Isa Mara Alves, por ter me concedido a oportunidade de realização do estágio de docência; por toda a sua paciência e compreensão e pelo espaço de trocas.

Às demais professoras do PPG pelo incentivo e ensinamentos, me permitindo conhecer a área da Linguística Aplicada.

Ao Programa de Pós-Graduação em Linguística Aplicada da Universidade do Vale do Rio dos Sinos, pela oportunidade de realizar esta pesquisa.

Ao Programa de Bolsas Fundo Pe. Milton Valente de Apoio Acadêmico à Pós-Graduação, por ter me concedido a bolsa de estudos.

Aos meus queridos colegas do mestrado pelo apoio, pelas conversas, pela ajuda, pelas trocas.

À minha amiga Sabrina por ter me incentivado a buscar esta qualificação e pelo apoio durante todo o mestrado.

Aos demais amigos por compreenderem as minhas angústias e me apoiarem sempre que necessitei.

Aos meus queridos familiares que acompanharam e incentivaram esta minha caminhada.

E, sobretudo, a Deus por me dar saúde e condições para alcançar este objetivo.

*“Não é o desafio que define quem somos  
nem o que somos capazes de ser, mas  
como enfrentamos esse desafio. Podemos  
incendiar as ruínas ou construir através  
delas, passo a passo, um caminho que nos  
leve à liberdade.”*

*Richard Bach*

## RESUMO

Este trabalho propõe a modelagem Semântica para a construção de uma ontologia para o domínio da Educação a Distância. Busca na Semântica uma abordagem para representar a descrição e conceituação do domínio EAD, partindo do pressuposto de que é necessário modelar os relacionamentos entre conceitos com base em uma abordagem capaz de ampliar a rede de relações. Fundamenta-se na Semântica Lexical (CRUSE, 1986, 2000) e na Semântica de Frames (FILLMORE, 1982), baseando-se nas relações paradigmáticas e sintagmáticas (SAEED, 2003; CALZOLARI et. al, 2010), levando em conta as relações de hiponímia, sinonímia e meronímia, para estabelecer a taxonomia (hierarquia de classes) e papéis temáticos e *frames* para outros tipos de relações. Tomando como *corpus* manuais de uso do ambiente Moodle, este estudo propõe a construção de uma ontologia com vistas a contribuir com a melhoria dos Ambientes Virtuais de Aprendizagem. O resultado desta investigação aponta para a Semântica de *Frames* como uma possível abordagem para descrever o significado em ontologias.

Palavras chave: Ontologia, Educação a Distância, Semântica, Frames

## **ABSTRACT**

This work proposes the Semantics Modelling for the construction of an ontology for the Remote Education dominion. It quests in Semantics an approach to represent the description and conceptualization of the EAD dominion, having as a starting point that it is necessary to model relationships among concepts with base in a capable approach of extending the relations network. Based on Lexical Semantics (CRUSE, 1986, 2000) and on the Frames' Semantics (FILLMORE, 1982), based on the Paradigmatic and Sintagmatic relations (SAEED, 2003; CALZOLARI et. al, 2010), considering relations of hiponomy, synomny, meromny, to establish taxonomy (class hierarchy) and thematic roles and frames for other kind of relations. Takins as corpus Moodle environment usage manuals, this study porposes the construction of and onthology having in sight a contribution on improving Learning Virtual Environments. The result of this investigaton points towards Frame Semantics as a possible approach on describing meanings in onthology.

Keywords: Onthology, Remote Education, Semantics, Frames.

## LISTA DE FIGURAS

Figura 1: Complexidade dos vocabulários controlados .....	28
Figura 2: Taxonomia lineana dos seres vivos – classificação dos humanos ..	30
Figura 3: Tipos de Ontologias e seus relacionamentos .....	36
Figura 4: Herança de classes .....	41
Figura 5: Temas relacionados à Web semântica.....	45
Figura 6: Ontologia como chave para o entendimento comum.....	46
Figura 7: Contribuições das ontologias para os AVA's .....	47
Figura 8: Continuum de ações entre ambiente, ontologia e usuário.....	49
Figura 9: Relação entre agente e ontologia para melhoria do AVA.....	50
Figura 10: Mapa conceitual – síntese do capítulo.....	51
Figura 11: Exemplos de relações paradigmáticas .....	63
Figura 12: Diferentes tipos de relações paradigmáticas .....	64
Figura 13: Relação de hiponímia .....	68
Figura 14: Classes semânticas dos verbos (Borba, 1996).....	77
Figura 15: Elementos de Frame - <i>EDUCATION_TEACHING</i> (visão parcial)..	83
Figura 16: Elementos de Frame - <i>STUDYING</i> (visão parcial).....	84
Figura 17: Categorias da M_ONTO .....	87
Figura 18: Mapa conceitual – síntese do capítulo 2.....	88

Figura 19: Ambiente Moodle – categorias de cursos e comunidades.....	96
Figura 20: Interface do ambiente Moodle .....	97
Figura 21: Lista parcial de termos – extração <i>e-Termos</i> .....	105
Figura 22: Consulta TUTOR – concordanceador e-Termos .....	106
Figura 23: Lista de termos do ambiente Moodle .....	106
Figura 24: Relações paradigmáticas – classe RECURSOS .....	109
Figura 25: Relações paradigmáticas – classe EVENTOS .....	110
Figura 26: Estrutura hierárquica da M_ONTO .....	111
Figura 27: Frame <i>EAD_TEACHING</i> e subframes – estrutura da FN.....	120
Figura 28: Definição da classe CURSO.....	124
Figura 29: Hierarquia – classes disjuntas e relação de hiponímia .....	126
Figura 30: Representação gráfica – classe RECURSO.....	127
Figura 31: Representação da relação de Meronímia.....	128
Figura 32: Representação da relação de Equivalência .....	130
Figura 33: Representação gráfica da relação de Equivalência.....	130
Figura 34: Classe <b>Usuario</b> .....	132
Figura 35: Propriedades restritivas – Classe <b>Professor</b> .....	133
Figura 36: Verbos – instâncias classe AcaoProcesso .....	134
Figura 37: Hierarquia de classes – modelagem dos frames .....	135
Figura 38: Classe <b>EventosPossiveis</b> .....	136
Figura 39: Descrição da classe <b>PostarForum</b> .....	137

Figura 40: Descrição da classe <b>EvocaFrameForum</b> .....	138
Figura 41: Representação gráfica da M_ONTO (visão parcial) .....	139

## LISTA DE TABELAS

Tabela 1: Relacionamentos semânticos de um tesouro .....	34
Tabela 2: Áreas de aplicação de ontologias .....	39
Tabela 3: Relações sintagmáticas e paradigmáticas.....	62
Tabela 4: Casos de papéis semânticos .....	78
Tabela 5: Modelo da Estrutura Ontológica.....	87

## SUMÁRIO

<b>1</b>	<b>CONSIDERAÇÕES INICIAIS.....</b>	<b>14</b>
<b>2</b>	<b>SITUANDO ONTOLOGIAS E RECURSOS LEXICAIS.....</b>	<b>18</b>
2.1	Ontologias, tesouros e taxonomias: organizando conceitos.....	19
2.1.1	Ontologias: definições e conceitos básicos.....	20
2.1.2	Vocabulários controlados e ontologias: principais diferenças .....	27
2.2	Classificando os tipos de ontologias.....	34
2.3	Áreas que fazem uso de ontologias e seus diferentes fins.....	37
2.4	O que compõe uma ontologia .....	40
2.5	Ontologias aplicadas aos AVA's: uma abordagem situada na <i>Web Semântica</i> .....	42
2.6	Ontologia + agente = uma aplicação.....	48
<b>3</b>	<b>O CONTEÚDO DA M_ONTO.....</b>	<b>52</b>
3.1	Semântica: da teoria à aplicação.....	54
3.2	Nossas escolhas teóricas.....	58
3.3	Sobre as relações semânticas .....	61
3.3.1	Relações paradigmáticas .....	62
3.3.2	Relações sintagmáticas .....	74
3.4	Categorias da M_ONTO.....	85
<b>4</b>	<b>CONSTRUÇÃO DA M_ONTO .....</b>	<b>89</b>
4.1	Caracterização do domínio: Ambiente Moodle.....	90
4.2	Caracterização da aplicação: o agente .....	98
4.3	O percurso metodológico .....	100
4.3.1	Preparação e análise do <i>corpus</i> .....	103
4.3.2	Etapa linguística .....	107
4.3.3	Etapa computacional.....	121
4.4	Resultados da pesquisa .....	139
<b>5</b>	<b>CONSIDERAÇÕES FINAIS .....</b>	<b>141</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>145</b>

## 1 CONSIDERAÇÕES INICIAIS

A interface entre linguagem e computação tem sido foco de diversos estudos, tanto na área da Linguística quanto na da Informática. Muitas são as possibilidades de interação entre linguagem e computação, sendo as ontologias uma das mais profícuas.

Desde os anos noventa, as ontologias vêm sendo estudadas na área de Inteligência Artificial como forma de representar e organizar o conhecimento. Recentemente, pesquisas sobre ontologias têm-se expandido para outras áreas como forma de integração de sistemas de informação aplicáveis em vários campos, criando uma relação de interdisciplinaridade. Através das ontologias, o domínio do conhecimento pode ser representado computacionalmente, viabilizando a comunicação entre pessoas e computadores, automaticamente, de forma inteligente.

Atualmente, as ontologias têm sido amplamente utilizadas por informatas para refinamento de sistemas de busca na *Web* e criação de agentes que visam ao aprimoramento de *softwares* através do Processamento da Linguagem Natural (doravante PLN).

Uma das aplicações das ontologias é o aprimoramento da *Web* atual, também chamada de *Web Sintática*<sup>1</sup>, para uma *Web Semântica*, passando a considerar não só a estrutura, mas também as relações de significados.

Os Ambientes Virtuais de Aprendizagem (doravante AVA's) têm sua estrutura centrada na *Web Sintática* e, conforme Stojanovic et.al. (2001), alguns de seus

---

<sup>1</sup> Termo apresentado por Breitman (2005): "A internet atual pode ser definida como *Web Sintática*. Nela os computadores fazem apenas a apresentação da informação, porém a interpretação fica a cabo dos seres humanos mesmo." (p. 2)

aspectos podem ser aperfeiçoados seguindo os padrões da Web Semântica com base em ontologias que descrevam o conteúdo semântico.

Além disso, as ontologias podem ser muito úteis para o aperfeiçoamento de agentes inteligentes ou pedagógicos que tenham por finalidade complementar aplicações computacionais para os AVA's, ampliando seu grau de interatividade com o sistema e com os usuários.

A construção de ontologias interessa tanto a linguistas quanto a informatas, pois o aperfeiçoamento dos sistemas passa por um aprofundado estudo semântico, capaz de contribuir e ampliar as possibilidades em PLN.

O estudo aqui proposto buscará subsídios nas áreas da Ciência da Informação (doravante CI), Ciência da Computação (doravante CC) e Linguística, mais precisamente a Semântica e a Linguística Computacional, a fim de possibilitar a construção de uma ontologia do ambiente Moodle.

Para a realização deste estudo contaremos com a colaboração do escritório de Educação a Distância (EAD) da Unisinos, que fornecerá os dados necessários para a composição do *corpus* de pesquisa e realizará a efetiva aplicação da ontologia criando um agente inteligente que atuará diretamente no ambiente Moodle.

Justificamos a escolha deste tema pela relevância dos aspectos supracitados, por existirem poucos relatos de trabalhos deste tipo relacionados a EAD na literatura e pela necessidade advinda do Escritório de EAD da Unisinos de descrever o conhecimento relacionado ao ambiente Moodle com vistas à criação de um agente que venha a aperfeiçoar a interação dos usuários com o sistema.

Esta pesquisa busca na Semântica uma abordagem para descrever e conceituar o domínio EAD, de modo a possibilitar a construção de uma ontologia para este domínio. Além disso, objetivamos definir os relacionamentos mais adequados para a descrição do conhecimento relacionado à EAD, bem como buscar soluções para a formalização dos dados em um editor de ontologias.

Para alcançar nosso objetivo, elaboramos as seguintes questões norteadoras:

- Qual a importância das ontologias para o aperfeiçoamento das tecnologias da informação, mais precisamente os AVA's?
- Que abordagens se mostram mais adequadas à descrição semântico-conceitual do domínio em questão a partir da Semântica?
- Como modelar os dados de forma consistente em um editor próprio para a construção de ontologias?

Estas questões balizarão nossas reflexões nos capítulos seguintes e, com base nelas, estruturamos nosso trabalho em cinco capítulos.

O primeiro capítulo corresponde às considerações iniciais e apresenta nossa justificativa e objetivos para o desenvolvimento desta pesquisa.

No capítulo dois apresentamos aspectos gerais relacionados à conceituação de ontologia, suas características e áreas de aplicação. Salientamos a diferenciação entre ontologias e vocabulários controlados, dando destaque às taxonomias e tesouros. Neste capítulo buscamos subsídios para responder o primeiro questionamento e destacamos pontos que relacionam as ontologias aos AVA's.

O capítulo três tem como finalidade esclarecer qual será o conteúdo da nossa ontologia, ou seja, qual será a abordagem semântica a ser seguida para sua construção. Enfatizamos neste capítulo questões referentes à Semântica buscando estabelecer um paralelo entre os diferentes enfoques dados pela Ciência da Computação, Ciência da Informação e Linguística. Apresentamos e justificamos nossas escolhas teóricas pela Semântica Lexical e pela Semântica Cognitiva, mais especificamente a Semântica de Frames. As relações semânticas e a forma como são vistas pela abordagem lexical e cognitiva, bem como o paralelo com as demais áreas já citadas, também são foco de estudo neste capítulo. Outro aspecto discutido diz respeito às categorias que devem compor uma ontologia e, acerca deste assunto, buscamos subsídios nas áreas da Ciência da Informação e Computação.

No capítulo quatro dedicamos uma seção ao detalhamento do ambiente Moodle, foco do nosso estudo, e outra a questões relacionadas aos agentes inteligentes, aplicação à qual nossa ontologia se vincula. Além disso, apresentamos

nosso percurso metodológico, organizado em duas etapas: a linguística e a computacional. Na etapa linguística ocupamo-nos de descrever como serão representadas as relações na ontologia e a etapa computacional visa a detalhar a proposta de formalização do conteúdo em um editor próprio de ontologias.

Por fim, tecemos algumas considerações finais destacando nossas conclusões e resultados de pesquisa, além de apresentarmos as limitações, dificuldades e perspectivas futuras de ampliação da M\_ONTO (*Moodle Ontology*), nome com o qual batizamos nossa ontologia.

## 2 SITUANDO ONTOLOGIAS E RECURSOS LEXICAIS

O crescente aumento de informações armazenadas na *Web* tem conferido significativa importância às técnicas de organização de informação. Diferentes estruturas têm sido utilizadas para essa tarefa: estruturas que utilizam termos para sua organização, tais como *glossários* e *dicionários*; estruturas que criam categorias e classificações, tais como as *taxonomias*; ou ainda, as estruturas que se organizam a partir de conceitos e seus relacionamentos, conhecidas por *ontologias*.

As ontologias, como estruturas para organização do conhecimento, têm sido amplamente estudadas nos últimos anos por diversas áreas do conhecimento, entre elas a Ciência da Informação (CI), a Inteligência Artificial (IA), com destaque para a área de Processamento da Linguagem Natural (PLN), e a Linguística, em especial a área da Semântica e da Linguística Computacional.

O termo *ontologia* deriva do grego “*onto*”, ser, e “*logos*”, palavra. Este termo tem sido empregado ao longo da história pela Filosofia e estuda as teorias sobre a natureza da existência. Diversas áreas utilizam-se das ontologias para estruturar e descrever conceitos, dentre elas, a Ciência da Computação (CC), que descobriu este tipo de recurso e tem feito grandes progressos em termos de Inteligência Artificial e Processamento da Linguagem Natural utilizando esta forma de representação do conhecimento.

Diversas aplicações computacionais fazem uso de ontologias referentes a um determinado domínio para a realização do processamento da linguagem natural, sendo que tal recurso pode contribuir significativamente para a melhoria dos sistemas. Um exemplo é a criação de agentes inteligentes para incremento dos Ambientes Virtuais de Aprendizagem (doravante AVA's).

Na literatura podem-se encontrar diversas definições para *ontologia*. Tal diversidade dificulta a escolha e a utilização das técnicas disponíveis para a construção deste tipo de estrutura. Outro aspecto importante é a semelhança entre léxicos computacionais, ou vocabulários controlados, e ontologias, pois, por serem estruturas semelhantes, sua diferenciação é um tanto sutil e passível de esclarecimentos.

Os aspectos supracitados nos levam a dedicar este capítulo ao estudo sobre as ontologias e os recursos lexicais, a fim de discutirmos pontos importantes sobre os mesmos e também situarmos nosso trabalho no âmbito de tais estudos.

Organizamos o capítulo em seis seções. Dedicamos a seção 2.1 para tecer esclarecimentos sobre a conceituação de alguns termos importantes e correlatos ao termo *ontologia*. Os diferentes tipos de ontologias e suas possíveis aplicações são apresentados de forma sucinta na seção 2.2. As diferentes áreas que fazem uso de ontologias e suas aplicações são apresentadas na seção 2.3. Na seção 2.4, apresentamos a composição da ontologia. Na seção 2.5, apresentamos a aplicação das ontologias aos AVA's, discutindo as principais vantagens. Por fim, destacamos a importância da ontologia para o desenvolvimento de agentes inteligentes seção 2.6.

## **2.1 Ontologias, tesauros e taxonomias: organizando conceitos**

De acordo com a introdução deste capítulo, muitos são os conceitos para ontologias, sendo que estes diferem conforme a área de aplicação. Dedicamos esta seção a fim de elucidar diferentes conceitos e definições acerca do termo *ontologia*, uma vez que este é um ponto importante para a escolha de métodos e para a construção deste tipo de estrutura. Além disso, consideramos adequado apresentar a distinção entre ontologias, taxonomias e tesauros destacando suas semelhanças e diferenças com o intuito de trazer esclarecimentos sobre as diferentes formas de organizar o conhecimento.

Organizamos esta seção de modo a apresentar as definições e conceitos básicos referentes às ontologias em 2.1.1. Os vocabulários controlados, mais

especificamente as taxonomias e os tesouros, bem como as principais diferenças entre eles e as ontologias, são apresentados em 2.1.2.

### 2.1.1 Ontologias: definições e conceitos básicos

Conforme referimos na introdução do capítulo, historicamente o termo *ontologia* – “*onto*” e “*logos*” – tem sua origem no grego. Suas bases são fundamentadas na Filosofia, mais precisamente nas categorias criadas por Aristóteles. Categorias estas que servem como alicerce para classificar qualquer entidade, atribuindo propriedades que distinguem diferentes espécies do mesmo gênero. O filósofo desenvolve dez categorias básicas que, segundo o autor, servem para classificar qualquer objeto, a saber: *substância, quantidade, qualidade, relação, lugar, tempo, posição, estado, atividade e passividade*. A partir dessa noção inicial apresentada por Aristóteles, diversos filósofos fundamentam seus estudos sobre ontologias.

A Ciência da Informação (CI), a Inteligência Artificial (IA), a Ciência da Computação (CC), a Linguística e uma variedade de outras áreas tomaram o termo emprestado da Filosofia com o intuito de organizar o conhecimento, mas passam a empregá-lo com um sentido distinto. Por se tratar de diferentes áreas, os conceitos de ontologia variam de acordo com a aplicação e o domínio ao qual estão relacionados.

Nesta seção procuramos apresentar as diferentes perspectivas para a definição de ontologias, buscando referencial nas áreas da CI, CC, IA, PLN e Ciências da Linguagem.

Na literatura encontramos uma distinção para “Ontologia” e “ontologia”, que consideramos aqui importante destacar. Para Guarino (1998), *Ontologia* (com inicial maiúscula) se refere aos estudos ancorados na Filosofia, isto é, aos estudos referentes ao ramo da filosofia que trata da natureza e da organização da realidade. Para o autor, *Ontologia* é um sistema de categorias que corresponde a uma determinada visão do mundo. Em contrapartida, *ontologia* (com inicial minúscula)

compreende um estudo de um vocabulário específico que tem por objetivo descrever uma determinada realidade.

Breitman (2005), em seu livro sobre a Web Semântica, também se refere a esta distinção. Segundo a autora, na Filosofia, *Ontologia* (com letra maiúscula) “é a ciência do que é, dos tipos de estruturas dos objetos, propriedades, eventos, processos e relacionamentos em todas as áreas da realidade” (p. 30). Dessa forma, Este vocábulo foi introduzido na Filosofia com o objetivo de distinguir o estudo do ser do estudo de diferentes seres existentes no mundo, sendo que a principal finalidade é a organização de sistemas de categorização que organizem a realidade.

Moreira et.al (2004) ressaltam que o uso do termo *ontologia* (com letra minúscula) tornou-se frequente na Computação por volta dos anos 90, em projetos que buscavam a organização de grandes bases do conhecimento, tornando-as interoperáveis e melhor estruturadas. Estes projetos, desenvolvidos no contexto da *Web* e da IA, contribuíram para uma adaptação para o significado da palavra *ontologia*.

Para os profissionais da Ciência da Informação (SOERGEL, 1997; VICKERY, 1997; GILCHRIST, 2003), ontologia constitui-se em um documento (ou arquivo) que define as relações entre termos e conceitos de maneira formal, apresentando similaridades com os vocabulários controlados (muito difundidos na área da Biblioteconomia), tais como os tesouros e as taxonomias. Segundo os autores, as similaridades encontram-se principalmente na forma de elaboração da estrutura destes instrumentos, uma vez que se faz necessária a organização dos conceitos em processos de categorização, classificação e definição das relações entre estes conceitos.

Breitman (2005) apresenta uma definição de ontologia voltada à área da computação.

Ontologias são modelos conceituais que capturam e explicitam o vocabulário utilizado nas aplicações semânticas. Servem como base para garantir uma comunicação livre de ambiguidades. Ontologias serão a *língua franca* da Web Semântica (grifo da autora) (p. 7).

De acordo com essa definição, as ontologias buscam fornecer o vocabulário necessário para a comunicação entre os agentes de *software*<sup>2</sup> e as páginas da *Web*, definindo as relações entre os conceitos e trazendo informações sobre os documentos. Essas informações são chamadas de *metadados*, isto é, “dados sobre os dados. O termo se refere a qualquer informação utilizada para a identificação, descrição e localização de recursos”, conforme definição apontada pela *International Federation of Library Associations (IFLA)*<sup>3</sup>.

Conforme o *World Wide Web Consortium (W3C)*<sup>4</sup>, os metadados são definidos como “informações para a *Web* que podem ser compreendidas por máquinas.” E, de maneira sucinta, aponta que as ontologias, como metadados, devem prover descrições para os tipos de conceitos, a saber: (a) classes nos domínios de interesse; (b) relacionamentos entre estas classes e (c) propriedades ou atributos que estas classes devem possuir (BREITMAN, 2005). Para o W3C, ontologia é “a definição dos termos utilizados na descrição e na representação de uma área do conhecimento”.<sup>5</sup>

Uma das definições mais frequentes é apresentada por Gruber (1993, p1), que conceitua ontologia como “uma *especificação formal e explícita* de uma *conceitualização*, o que existe é aquilo que pode ser representado” (grifo nosso). Para o autor, uma ontologia é, formalmente, uma afirmação da lógica teórica.

A definição de Gruber (1993) é a mais conhecida e a mais citada na literatura sobre ontologias; porém, cabem alguns esclarecimentos acerca dos termos utilizados pelo teórico. Por *formal*, Gruber compreende que as ontologias devem ser processáveis por máquina, isto é, os computadores devem ser capazes de compreender o que está descrito, através de uma representação semântica, lógica e

---

<sup>2</sup> Na seção 1.6 apresentaremos considerações sobre os agentes de software e a importância das ontologias para este tipo de aplicação.

<sup>3</sup> Disponível em: <<http://www.ifla.org/>>. Acesso em: 25/05/2010

<sup>4</sup> Disponível em: <<http://www.w3c.br/> e <http://www.w3.org/>>. Acesso em: 20/03/2010

<sup>5</sup> Informações advindas do site do W3C. Disponível em <<http://www.w3c.br/>>. Acesso em: 12/11/2010.

formal, favorecendo o raciocínio automático. Complementando *formal*, Gruber utiliza os termos *especificação* e *explícita*, que, para o autor, referem-se aos conceitos, propriedades, relações, funções, restrições, axiomas que são definidos de forma explícita na modelagem dos dados.

Para Gruber (1993), o termo *conceitualização* é utilizado para referir-se à descrição do significado de cada termo/conceito incluindo suas relações semânticas com outros termos/conceitos. Uma *conceitualização* pode ser considerada uma visão abstrata e simplificada daquilo que se pretende representar. Segundo o autor, todo o tipo de agente de software (caso da aplicação da ontologia aqui proposta) está comprometido com uma forma de *conceitualização* explícita ou implícita, sendo que compreende uma relação extensional, ou seja, um inventário da composição do domínio, uma listagem de todos os aspectos pertinentes àquele domínio, ainda que em nível de abstração.

Na área de IA, o termo *ontologia* foi tomado emprestado da Filosofia e, para os pesquisadores da *Web* e estudiosos destas áreas, tem outro sentido. “É um documento que define as relações entre termos e conceitos” (BERNERS-LEE; HENDLER; LASSILA, 2001).

Borst (1997) define ontologia de maneira muito semelhante a Gruber (1993). Para o autor, uma ontologia é

uma especificação formal e explícita de uma conceitualização compartilhada. Nessa definição, “formal” significa legível para computadores; “especificação explícita” diz respeito a conceitos, propriedades, relações, funções, restrições e axiomas explicitamente definidos e manipulados por computadores; “compartilhado” quer dizer conhecimento consensual e “conceitualização” diz respeito a um modelo abstrato de algum fenômeno do mundo real (Borst, 1997, p. 12).

Assim, uma ontologia é uma descrição formal dos conceitos e relacionamentos que existem dentro de um domínio, o que significa que uma ontologia se relaciona com um vocabulário específico e com uma linguagem específica (DAUM, 2002 apud ARAUJO, 2003).

Alguns autores discordam de tais afirmações, principalmente no que diz respeito à definição de “conceitualização”. Guarino e Giaretta (1995, p. 5) apresentam “*conceitualização*” como “uma estrutura semântica intensional que codifica as regras implícitas que determinam a estrutura de uma porção de realidade”. De acordo com os autores supracitados, uma ontologia deve ser definida em sentido restrito como “uma teoria lógica que fornece uma proposta explícita e parcial de uma conceitualização” (p. 5)

Diante desta questão sobre a “conceitualização”, consideramos importante apresentar a visão da chamada Ontolinguística<sup>6</sup>, abordada por Schalley e Zaefferer (2007), estando inserida no âmbito da Linguística Cognitiva. Segundo os autores, o conhecimento linguístico é um tipo de conhecimento ontológico, sendo que a Ontolinguística assenta-se no pressuposto de que existem universais mentais<sup>7</sup>. Dessa forma, compreendem ontologia como um “sistema de conceituações”, isto é, “uma rede de conceituações interconectadas dos fenômenos que constituem o mundo” (SCHALLEY e ZAEFFERER, 2007, p. 3). Pode-se perceber que a visão da Ontolinguística não difere muito da definição apresentada por Gruber (1993), a qual referimos anteriormente.

Smith (2002) propõe uma discussão entre Ciência e Ontologia<sup>8</sup>, afirmando que a Ciência tem a função de “explicar” a natureza e a Ontologia vem para descrever, organizar e sistematizar este conhecimento obtido através dos experimentos científicos. A CI busca referencial nesta discussão apresentada pelo autor, sendo responsável por uma vasta tradição no que se refere à organização dos domínios do conhecimento, aplicadas originalmente aos vocabulários controlados, visando à organização documental. Esta área constitui um amplo arcabouço teórico

---

<sup>6</sup> Termo utilizado como título do livro editado por SCHALLEY e ZAEFFERER em 2007. Este livro reúne diversos estudos referentes às contribuições do conhecimento linguístico para o conhecimento ontológico. Gruber (1993) também utilizou o termo “ontolinguá” em suas discussões.

<sup>7</sup> Para maiores informações sobre universais mentais, ver GEERAERTS, 2006.

<sup>8</sup> Ontologia enquanto domínio de conhecimento preocupado com a natureza dos seres.

e metodológico potencialmente aplicável ao desenvolvimento de ontologias (OTLET, 1934; RANGANATHAN, 1967; DAHLBERG, 1978; HJØRLAND<sup>9</sup>, 2002).

Outro ponto importante e de destaque na literatura relacionada ao PLN é a diferenciação entre ontologias linguísticas e conceituais ou formais (VOSSEN, 1998<sup>a</sup>; PALMER, 2001; FARRAR e BATEMAN, 2005). As ontologias linguísticas têm como principal característica registrar conceitos lexicalizados, isto é, as palavras que compõem o léxico de uma determinada língua. Por esta ótica, “uma ontologia é um inventário dos sentidos de uma dada língua, ou seja, é um inventário somente daqueles conceitos compartilhados por uma comunidade linguística” (Di FELLIPO, 2008, p. 470).

Conforme o viés pelo qual é construída, a ontologia pode ser *formal* ou *linguística*, isto é, uma ontologia pode ser construída sob o viés da Engenharia da Linguagem (área da computação ligada ao PLN) ou sob o viés da Linguística. A principal diferença entre as duas está no tamanho e no grau de formalização.

Segundo Di Felippo (2008), as ontologias linguísticas mais conhecidas no PLN são Mikrokosmos (VIEGAS et.al., 1996), SENSUS (HOVY, 1998) e WordNet<sup>10</sup> (FELLBAUM, 1998)<sup>11</sup>. Para a autora, o uso deste tipo de ontologia em PLN ocorre porque as referidas ontologias são construídas a partir do léxico e da semântica de uma língua específica, isto é, fazem uso de teorias linguísticas para a descrição das unidades lexicais.

---

<sup>9</sup> Birger Hjørland, PhD em Supervisão de ensino da Organização do Conhecimento e conceitos básicos. Professor de IVA – Informations Videnskabelige Akademi, em Copenhagen. Suas principais pesquisas estão relacionadas à organização do conhecimento e conceitos básicos; diferentes abordagens teóricas e tecnológicas para a organização do conhecimento. Disponível em: <[http://forskning.iva.dk/da/persons/birger-hjoerland\(3713cca6-7e15-4a2f-8938-680a776029d2\).html](http://forskning.iva.dk/da/persons/birger-hjoerland(3713cca6-7e15-4a2f-8938-680a776029d2).html)>. Acesso em: 18/01/2011

<sup>10</sup> É importante atentar para o fato de que Di Felippo (2008) considera a WordNet (WN) uma ontologia linguística, também podendo ser caracterizada como um tipo especial de léxico computacional. Segundo a autora, a WordNet (WN) foi elaborada de modo que cada synset é, por definição, “construído de modo a codificar ou representar um único conceito lexicalizado por suas unidades constituintes. Dessa forma, WN armazena apenas os conceitos lexicalizados na variante Americana da língua inglesa, o que a caracteriza como uma ‘ontologia linguística’” (p. 471).

<sup>11</sup> Apud Di FELLIPO, 2008, p. 470.

Em contrapartida, as ontologias conceituais ou formais têm como característica a organização entre conceitos. Há a inserção de níveis artificiais para que se alcance uma estrutura controlada dos conceitos referentes a um determinado domínio. Segundo Palmer (2001 apud Di FELLIPO, 2008), a ontologia conceitual mais difundida no PLN é a CYC<sup>12</sup>.

As ontologias formais aplicam técnicas puramente lógicas para a seleção e a descrição do conhecimento, sendo necessário pouco envolvimento humano, o que as torna maiores. Já as ontologias linguísticas, como a proposta neste trabalho, descrevem a semântica das unidades lexicais com base em diferentes perspectivas de organização conceitual.

Tanto as ontologias formais quanto as linguísticas contribuem muito para o trabalho cooperativo entre homens e máquinas, exigindo a integração entre profissionais da computação e da linguística e, muitas vezes, sendo importante a participação de um especialista no domínio sobre o qual se está tratando.

Avaliamos que, apesar desta diferenciação entre ontologias linguísticas e conceituais apontada pelos estudiosos de PLN, ambas possuem um aspecto em comum, isto é, permitem o trabalho colaborativo entre máquinas e seres humanos através de uma equipe interdisciplinar composta por linguistas, informatas e especialistas do domínio ao qual a ontologia visa a descrever (ALVES, 2005).

Na nossa visão, e de acordo com as necessidades do trabalho aqui proposto, a definição de ontologia deve reunir as ideias expostas por DAUM (2002 apud ARAUJO, 2003), BERNERS-LEE; HENDLER; LASSILA (2001) e SCHALLEY e ZAEFFERER (2007). Consideramos que uma ontologia descreve formalmente os conceitos que estruturam um determinado domínio, bem como as relações semânticas existentes entre eles, isto é, ontologia é um sistema que define e classifica o conhecimento, estruturando seus conceitos de forma clara e precisa e formalizando-os de modo a serem tratados por computadores. Em suma, uma

---

<sup>12</sup> CYC – “redução de *enCYClopædia*, ou enciclopédia. Este projeto objetivou a criação de uma enorme base de conhecimento que, segundo seus projetistas, contém os ‘termos mais gerais da realidade consensual dos humanos’” (BREITMAN, 2005, p. 68)

ontologia deve ser construída com base em conceitos e nas relações existentes entre eles, levando em consideração o contexto do domínio no qual está inserida.

Consideramos importante esclarecer alguns aspectos referentes aos vocabulários controlados e às ontologias e, para isso, dedicamos a seção 2.1.2, na qual apresentamos as principais diferenças entre ontologias, taxonomias e tesouros.

### **2.1.2 Vocabulários controlados e ontologias: principais diferenças**

Diante do exposto anteriormente acerca da imprecisão que envolve a definição do termo *ontologia*, julgamos importante salientar alguns aspectos no sentido de elucidar conceitos relacionados ao assunto. Consideramos fundamental tal distinção a fim de minimizar possíveis equívocos em relação aos conceitos apresentados.

Nesta seção abordaremos com maior ênfase os vocabulários controlados, mais especificamente apontando diferenças e semelhanças entre tesouros, taxonomias e ontologias.

A *National Information Standards Organization* (NISO) publicou a norma ANSI/NISO Z39.19 intitulada *Guidelines for the construction, format and management of monolingual controlled vocabularies* em 2005. Tal publicação prevê diretrizes para a construção, formatação e gestão de vocabulários controlados.

De acordo com a ANSI/NISO Z39.19-2005, vocabulários controlados correspondem a uma listagem de termos enumerados de forma explícita, sendo que todos os termos pertencentes ao vocabulário controlado devem ter uma definição livre de redundâncias e inequívoca; porém, a norma evidencia que nem sempre é possível pôr em prática tais restrições.

Proporcionar um meio para organizar as informações é o objetivo dos vocabulários controlados. Segundo o documento, eles servem para cinco propósitos, a saber:

- (a) tradução – para possibilitar a conversão da linguagem natural em um vocabulário para ser usado na indexação da informação;
- (b) consistência – para promover a uniformidade no formato dos termos;
- (c) indicação de relacionamentos – para explicitar as relações semânticas entres os termos;
- (d) etiquetas para navegação – para fornecer uma hierarquia para ajudar o sistema de navegação;
- (e) recuperação – para auxiliar na busca e localização de conteúdo.

A norma ANSI/NISO (2005) assim descreve os vocabulários controlados:

Os vocabulários controlados são geralmente usados para descrever o conteúdo através da atribuição de termos para representar metadados associados aos objetos de conteúdo, a organização de conteúdos em sites, e afins. Para os fins desta Norma, o termo é definido como sendo uma ou mais palavras usadas para representar um conceito. Termos são selecionados a partir da linguagem natural para a inclusão em um vocabulário controlado (ANSI/NISO Z39.19-2005, p. 11, tradução nossa).

Os vocabulários controlados permitem a visualização de diferentes tipos de relações que os termos contêm através da forma como são estruturados. Segundo a norma ANSI/NISO, existem quatro tipos de vocabulários controlados, diferenciados basicamente por sua estrutura mais ou menos complexa: lista, anel de sinônimos, taxonomia e tesouro.

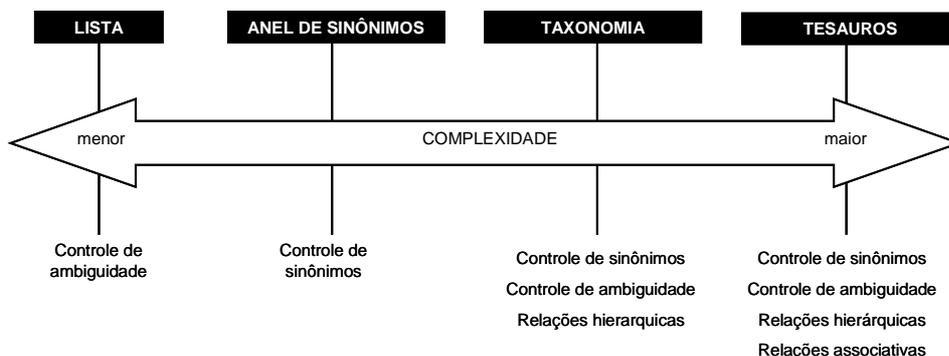


Figura 1: Complexidade dos vocabulários controlados

Na figura 1<sup>13</sup> é possível visualizar o aumento do grau de complexidade entre o nível mais simples de vocabulários controlados, as listas, ao mais complexo, os tesouros, de acordo com os tipos de relações que cada um deve acomodar. Na figura também se pode perceber que as estruturas mais complexas incluem os requisitos das estruturas mais simples.

No caso do nosso trabalho, julgamos que apenas os tesouros e as taxonomias podem gerar algum tipo de ambiguidade, uma vez que apresentam em suas estruturas as relações, tanto hierárquicas quanto associativas, que também compreendem o âmbito das ontologias. Por esta razão, apresentamos a seguir características das taxonomias e tesouros, destacando suas diferenças com as ontologias.

### **a)Taxonomias x Ontologias**

Segundo o dicionário Aulete Digital<sup>14</sup>, taxonomia é a “ciência da classificação”. Já o dicionário digital Merriam Webster<sup>15</sup> define como “o estudo dos princípios gerais de classificação científica: classificação sistemática; em particular, classificação ordenada de plantas e animais segundo relacionamentos naturais”.

Uma definição mais adequada à área da Ciência da Informação é elaborada por Daconta (2003). Segundo o autor, uma taxonomia é a “classificação de entidades de informação no formato de uma hierarquia, de acordo com relacionamentos que estabelecem com entidades do mundo real que representam” (p. 34).

---

<sup>13</sup> Figura adaptada de ANSI/NISO Z39.19-2005, p. 17 – tradução nossa.

<sup>14</sup> Disponível em <[http://aulete.uol.com.br/site.php?mdl=aulete\\_digital](http://aulete.uol.com.br/site.php?mdl=aulete_digital)>. Acesso em 12/10/2010.

<sup>15</sup> Disponível em <<http://www.merriam-webster.com/>>. Acesso em: 12/10/2010.

O padrão ANSI/NISO Z39.19-2005 define taxonomia como “um vocabulário controlado que consiste em condições preferenciais, as quais estão conectadas em uma hierarquia” (p. 18, tradução nossa).

Assim, pode-se entender taxonomia como uma **hierarquia** que serve para **classificar** informações através de um **relacionamento** pai-filho (relacionamento de generalização, ou tipo-de).

Breitman (2005) apresenta três exemplos clássicos de taxonomias: (a) a classificação de humanos, segundo a taxonomia lineana, ilustrada na figura 2<sup>16</sup>; (b) o sistema Dewey Decimal de classificação de assuntos, muito utilizado em bibliotecas para a indexação de livros; e (c) a estrutura de diretórios, utilizada para organizar documentos nos computadores (ainda que este seja um exemplo um tanto simples e sua hierarquia seja variada, de acordo com cada usuário).

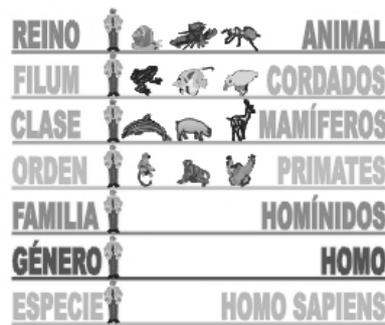


Figura 2: Taxonomia lineana dos seres vivos – classificação dos humanos

Nas taxonomias, os itens são relacionados através de relacionamentos de especialização, isto é, *pai-filho* ou *classe-subclasse*. Neste tipo de estrutura não podem ser atribuídas características ou propriedades aos termos (atributos) e nem exprimir outros tipos de relacionamentos, tais como parte-de, causa-efeito, localização, associação, entre outros. Para isso, é necessário construir uma ontologia.

<sup>16</sup> Imagem disponível em <<http://lacomunidad.elpais.com/manurentzat/2008/8/16/clasificacion-los-seres-vivos>>. Acesso em: 05/01/2011

Contudo, vale salientar que **toda ontologia é também composta por uma taxonomia**, isto é, a estrutura de classes e subclasses compreende uma taxonomia.

## **b) Tesouros x Ontologias**

O padrão ANSI/NISO (2005) define tesouro como:

Um tesouro é um vocabulário controlado organizado em uma ordem conhecida e estruturada de modo que as diversas relações entre os termos são claramente exibidas e identificadas por indicadores de relacionamento padronizados. Indicadores de relacionamento devem ser empregados mutuamente (p. 18, tradução nossa).

Assim, um tesouro tem o objetivo de garantir que os conceitos sejam descritos de forma consistente permitindo o refinamento de buscas e facilitando a localização da informação.

Os tesouros são formas encontradas para estruturar o léxico e, segundo Breitman (2005), reúnem um conjunto de relacionamentos entre termos organizados em uma taxonomia. Conforme a autora, uma taxonomia define os termos, sendo que a única relação existente entre eles é a generalização, ou seja, os termos são associados em uma hierarquia do geral para o específico (é-um).

Segundo Gusmão (1985), os tesouros possibilitam um número finito de possibilidades de relacionamento, a saber: relações de equivalência ou sinonímia; relação de hierarquia e relação de associação. Para Breitman (2005), este tipo de estrutura visa a garantir que conceitos sejam descritos de maneira consistente, permitindo que os usuários possam refinar suas buscas e localizar a informação de que necessitam.

Os tesouros contam com tipos de relacionamentos bem definidos e finitos entre seus termos, sendo que este conjunto de relacionamentos é útil para a criação de vocabulários, mas não suficiente para a organização de conceitos do “mundo real”. Em alguns casos é necessário utilizar outros tipos de relacionamentos, como parte-de, membro-conjunto, fase-processo, lugar-região, entre outros. Tais

relacionamentos não podem ser incluídos em um tesouro e esta pode ser considerada a principal diferença dos tesouros em relação às ontologias.

Atualmente, o léxico computacional mais conhecido e utilizado é o WordNet<sup>17</sup> (FELLBAUM, 1998) e, justamente por sua semelhança com as ontologias, é citado em diversas obras como ontologia linguística, conforme referimos na seção anterior. No nosso trabalho não é relevante discorrer profundamente sobre tais léxicos, mas consideramos importante salientar que os léxicos computacionais se aproximam muito das ontologias.

Noy e McGuinness (2001) afirmam que existem três propriedades fundamentais que diferem uma ontologia de taxonomias e tesouros, a saber:

(a) Estrita hierarquia de subconceitos (hierarquia do tipo formal), isto é, toda a instância de uma subclasse tem que ser também uma instância da classe-pai e a organização dos termos se dá através do relacionamento “tipo-de” (generalização), sendo este o fio condutor da ontologia;

(b) Interpretação livre de ambiguidades para os significados e relacionamentos, isto é, as propriedades para cada classe são definidas pelo usuário de acordo com o domínio da ontologia;

(c) Utilização de um vocabulário controlado restrito e finito, mas que pode ser estendido conforme a aplicação e a necessidade da ontologia.

Café e Sales (2009) reforçam os aspectos citados para diferenciar ontologias e tesouros, afirmando que as ontologias transcendem a representação de conteúdos formais, de um modo geral atribuída aos tesouros, pois assumem um papel de ferramenta fundamental para os sistemas automatizados de recuperação da informação, para acesso às bases de conhecimento e para os serviços ofertados pela Web, sobretudo ao que se refere ao PLN, à Web Semântica e à melhoria dos AVA's.

---

<sup>17</sup> Disponível em: <<http://wordnet.princeton.edu/>>. Acesso em: 20/04/2009

Segundo Moreira et.al. (2004), o emprego do termo *ontologia* para denominar uma estrutura de termos e as relações entre eles em um determinado domínio é mais comum na área da Ciência da Computação e, mais particularmente, na Inteligência Artificial. Esta definição favorece uma imprecisão entre o emprego dos termos *tesauro* e *ontologia*; contudo, Breitman (2005) esclarece a fundamental diferença:

Muitas vezes é necessário relacionar conceitos utilizando relacionamentos do tipo parte-de, membro-conjunto, fase-processo, lugar-região, material-objeto, causa-efeito, entre muitos outros. Um tesauro não permite a seus usuários a criação destes e novos tipos de relacionamento, para tal é necessário utilizar uma ontologia (p. 37).

Dessa forma, podemos argumentar que as ontologias são mais flexíveis e complexas do que os tesouros, porque permitem que novos relacionamentos sejam estabelecidos entre os termos, de acordo com a necessidade.

Outro aspecto relevante na diferenciação entre tesauro e ontologia é referente ao número de relacionamentos. Segundo Daconta (2003), em um tesauro, os tipos de relacionamentos entre os termos são finitos e bem definidos, sendo que são suficientes para a criação de um vocabulário, mas pouco eficazes para modelar aspectos do mundo real, pois para este tipo de aplicação são necessários outros tipos de relacionamentos, conforme já discutido anteriormente.

Na tabela 1 apresentamos características dos tesouros evidenciando seus relacionamentos finitos e definidos descritos por Daconta (2003)<sup>18</sup>, fato considerado o principal diferenciador entre este tipo de estrutura e as ontologias.

---

<sup>18</sup> Apud BREITMAN (2005), p. 36

Relacionamento semântico	Definição	Exemplo
Sinônimo Similar a Equivalente Usado para	Um termo X tem quase o mesmo significado que o termo Y.	Aluno é um sinônimo para estudante.
Homônimo Mesma grafia	O termo X tem a mesma grafia que o termo Y, porém tem significado diferente.	Tanque – veículo militar Tanque – recipiente para guardar líquidos Tanque – lugar onde se lava a roupa
Mais amplo do que (hierarquia – pai de superclasse)	Um termo X tem significado mais amplo que o termo Y.	Organização tem significado mais amplo que instituição financeira.
Mais restrito (hierarquia filho de subclasse)	Um termo X tem significado mais restrito que o termo Y.	Instituição financeira tem significado mais limitado que organização.
Associado Associativo Relacionado a	O termo X está associado a um termo Y, isto é, existe um relacionamento não especificado entre os termos.	Um prego está associado a um martelo.

Tabela 1: Relacionamentos semânticos de um tesouro

Tais relacionamentos certamente são úteis também na construção de ontologias, mas existem outros que permitem uma melhor representação do conhecimento relacionado ao domínio da ontologia. Dedicaremos uma seção do capítulo seguinte para esclarecermos e apresentarmos os outros tipos de relacionamentos que irão compor a M\_ONTO.

Concluído este sobrevôo aos termos correlatos à ontologia e suas definições, passamos às seções seguintes nas quais discutiremos sobre as formas de classificação de ontologias, as áreas nas quais vem sendo aplicadas e os itens que compõem as ontologias.

## 2.2 Classificando os tipos de ontologias

Diferentes autores propõem formas de classificar as ontologias, dentre eles podemos citar Guarino (1998) e Gómez-Perez et.al. (2004). O primeiro sugere o desenvolvimento de diferentes tipos de ontologias de acordo com o nível de generalidade necessária, enquanto que os demais propõem outro tipo de classificação concentrada no tipo de informação a ser modelada.

Segundo Gómez-Perez et.al. (2004), as ontologias, organizadas de acordo com o tipo de informação a ser modelada, seriam divididas em (a) ontologias para a representação do conhecimento, (b) ontologias gerais e de uso comum, (c) ontologias de topo ou de nível superior (*upper ontologies*), (d) ontologias de domínio, (e) ontologias de tarefas, (f) ontologias de domínio-tarefa, (g) ontologia de métodos e (h) ontologias de aplicação.

Guarino (1998) classifica as ontologias em quatro tipos: (a) ontologias de alto nível, (b) ontologias de domínio, (c) ontologias de tarefa e (d) ontologias de aplicação. No âmbito do nosso trabalho, seguiremos a divisão proposta por Guarino (1998), pois entendemos que os quatro tipos propostos pelo autor englobam em suas características todos os tipos propostos por Gómez-Perez et.al. (2004), tornando a classificação mais abrangente, ainda que alguns tipos sejam semelhantes. Diante deste posicionamento, teceremos esclarecimentos acerca dos tipos propostos por Guarino (1998).

As ontologias de alto nível ou *top-level* (também pode ser encontrado o termo *fundacionais* para definir este tipo de ontologia) são as que descrevem conceitos gerais, tais como espaço, tempo, assunto, objeto, evento, ação, sendo que são totalmente independentes de um problema particular ou domínio. Tem como principais características a abrangência e a reusabilidade, estando, algumas vezes, vinculadas a ontologias menores (de domínio, por exemplo) usadas por grandes comunidades (GUARINO, 1998).

Podemos citar como exemplos deste tipo de ontologia as ontologias DOLCE (*Descriptive Ontology for Linguistics and Cognitive Engineering*)<sup>19</sup> e SUMO (*Suggested Upper Merged Ontology*)<sup>20</sup>, resultado de um trabalho cooperativo e interdisciplinar entre linguistas, filósofos, estudiosos das ciências cognitivas, entre outros. O objetivo das ontologias supracitadas é servir de base de conhecimento

---

<sup>19</sup> DOLCE é uma ontologia “fundacional” criada pelo *Instituto di Ontologia Applicada (Roma-Trento)*.

<sup>20</sup> SUMO - Ontologia formal usada para mapear o WordNet. Mais informações em: <http://www.ontologyportal.org/>

para ontologias mais simples, sendo compartilhadas por grandes comunidades de usuários.

As ontologias de domínio têm como característica a descrição de um vocabulário relacionado através de um domínio genérico, buscando especializar os conceitos introduzidos nas ontologias de alto nível. Um bom exemplo é a ontologia BFO (*Basic Formal Ontologie*)<sup>21</sup> muito usada na área da saúde e da biologia.

As ontologias de tarefa se caracterizam pela descrição de um vocabulário relacionado a uma tarefa ou atividade genérica, buscando especializar os conceitos introduzidos nas ontologias de alto nível. E as ontologias de aplicação são as mais específicas, pois são utilizadas em aplicações. É o tipo de ontologia que especializa os conceitos, tanto das ontologias de domínio quanto das de tarefa. Um exemplo é a ontologia que propomos neste trabalho.

Na figura 3 apresentamos, com base no modelo de Guarino (1998), os tipos de ontologias e seus relacionamentos.

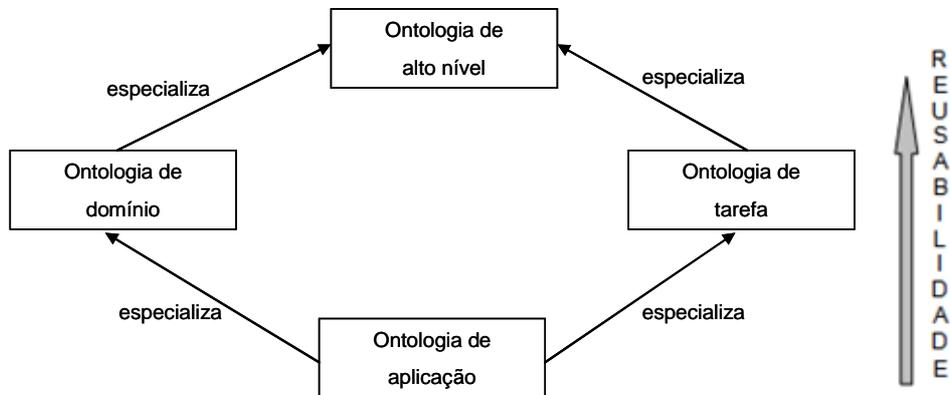


Figura 3: Tipos de Ontologias e seus relacionamentos

Pode-se observar que as ontologias de alto nível são as que têm mais capacidade de reuso, pois definem conceitos genéricos. Já as ontologias de

<sup>21</sup> Ontologia de Domínio que apresenta distinções entre objetos e processos e podem ser unidos utilizando relações básicas.

aplicação, por definirem conceitos relativos a uma aplicação específica, possuem uma menor capacidade de reutilização.

As ontologias podem ser utilizadas para vários fins, de acordo com a área à qual estão relacionadas. Por esta razão, na seção seguinte, realizamos uma sinopse acerca de diferentes áreas que se valem das ontologias como estruturas para a organização dos conhecimentos e suas principais aplicações.

### **2.3 Áreas que fazem uso de ontologias e seus diferentes fins**

Ontologias são amplamente utilizadas por diversas áreas, entre elas: processamento da linguagem natural, gestão do conhecimento, *web semântica* e comércio eletrônico.

O PLN tem por objetivo a implementação de sistemas computacionais que sejam capazes de processar automaticamente a linguagem natural, sendo um campo de natureza teórica, aplicada e interdisciplinar. Assim, agrega diversas áreas do conhecimento, como a Linguística, a Filosofia, a Psicologia, a Lógica, a Matemática, a Linguística Computacional, a Ciência da Computação e a Inteligência Artificial.

Os sistemas de PLN que visam a emular conhecimentos linguísticos, mesmo os mais simples, exigem um grande conhecimento linguístico para a sua aplicação. Pria (2008) ressalta que o PLN exige léxicos estruturados que forneçam informações morfológicas, sintáticas e semânticas das palavras para que os sistemas possam ser implementados de forma eficiente<sup>22</sup>.

Segundo Jasper e Uschold (1999 apud BREITMAN 2005), as ontologias têm grande aplicação e trazem muitas vantagens para a construção de um sistema de PLN. Uma delas é a facilidade de documentação, manutenção e a confiabilidade dos dados no sistema, pois a definição dos termos relevantes não se dá de forma

---

<sup>22</sup> Cf. Sanfilippo e Handke, 1995.

alfabética. Outra vantagem apresentada pelos autores é a possibilidade de reutilização do conhecimento definido pela ontologia.

Conforme referimos anteriormente, as ontologias têm sido utilizadas por diversas áreas para organizar o conhecimento, uma delas é a Computação, que faz uso das ontologias para fornecer dados para sistemas. No caso do nosso trabalho, analisaremos os itens lexicais sob a perspectiva linguística, verificando os termos e seus relacionamentos através da análise semântica.

Na área da gestão do conhecimento, as ontologias apresentam-se como estruturas nas quais são construídas as bases do conhecimento, sendo possível a anotação semântica sobre as informações e facilitando a recuperação de informação.

Usadas no âmbito da *Web Semântica*, as ontologias permitem uma estrutura para anotação das páginas da *Web*, permitindo que os agentes tenham maior capacidade de recuperação da informação buscando informações mais precisas. Uma vez que a M\_ONTO tem um compromisso direto com o agente inteligente que integrará o ambiente *Moodle*, permitindo que ele realize inferências no sistema, contribuindo para uma aplicação para a *Web Semântica*, destacaremos este ponto na seção 2.5.

Ainda ao que faz referência à *Web*, as ontologias têm importante papel na recuperação da informação, pois atualmente a *web* possui um ilimitado acervo de documentos dos mais variados (*sites*, *blogs*, vídeos, imagens, bancos de dados, etc), os quais são criados livremente sem qualquer tipo de censura ou supervisão, sendo que os próprios usuários podem estabelecer as relações entre os variados conteúdos. Essa liberdade gera um conteúdo incalculável, tornando cada vez mais difícil a localização de informações precisas. Através das ontologias, este problema de recuperação da informação pode ser consideravelmente minimizado, uma vez que ontologias possuem uma estrutura com anotação semântica.

Em relação ao comércio eletrônico, as ontologias tornam-se úteis para a descrição e organização dos diferentes produtos ofertados nos sites de vendas, ou lojas virtuais.

De acordo com Lopes et.al. (2009), as áreas supracitadas fazem uso de ontologias porque estas permitem a conceituação, a estruturação e a representação dos seus conteúdos em um documento permitindo que o conhecimento relativo ao domínio seja mais facilmente compartilhado. Na tabela 2 apresentamos uma síntese das áreas que usam ontologias para estruturar o conhecimento e a aplicabilidade das mesmas.

<b>Área</b>	<b>Aplicabilidade</b>
Processamento de Linguagem Natural	- redução dos problemas de ambiguidade através do uso de uma ontologia sobre o domínio do discurso do texto. - a ontologia funciona como um dicionário de conceitos dentro do domínio do texto.
Gestão do Conhecimento	- as ontologias fornecem a estrutura básica sobre a qual se constroem bases de conhecimento. - ao usar ontologias é possível anotar informações semânticas.
Web Semântica	- as ontologias fornecem uma estrutura semântica para anotação de páginas da Web possibilitando buscas mais precisas e dando maior capacidade para os agentes de software que utilizam conteúdo da Web.
Comércio Eletrônico	- integrar sistemas de vendas pela internet, através de uma ontologia de descrição de produtos.

Tabela 2: Áreas de aplicação de ontologias

Para o desenvolvimento de tais atividades, muitas pesquisas seguem os modelos psicolinguísticos de Collins e Quillian (1972) e de Miller e Johnson-Laird (1976)<sup>23</sup>. Estes modelos se apóiam na criação de redes semânticas, que têm o objetivo de descrever e explicar como o cérebro humano armazena as informações em forma de redes, isto é, “as pesquisas psicolinguísticas podem explicar como se organizam conceitos na memória humana e como essa informação é acessada no léxico mental” (ALVES, 2005, p. 30).

As ontologias, sendo construídas sob esta perspectiva, descrevem como o léxico está organizado na memória para que o falante tenha acesso e uso imediatos.

No contexto do trabalho aqui proposto, a ontologia será a representação de um conjunto de termos (referentes ao domínio EAD), organizados em classes e subclasses conforme seu conteúdo semântico formalmente definido. A aplicação

---

<sup>23</sup> Apud ALVES (2005).

desta ontologia será alimentar e qualificar um agente inteligente integrado ao ambiente Moodle possibilitando maior interação entre usuário e ambiente.

Ao findarmos esta sinopse sobre as áreas de aplicação das ontologias, reiteramos a grande importância deste recurso nos mais variados campos e salientamos que alguns aspectos em relação à estrutura das ontologias são comuns em todos os âmbitos. Assim, trataremos sobre a estrutura geral da ontologia e os itens que a compõem na seção que segue.

## **2.4 O que compõe uma ontologia**

Segundo Gruber (1996), os componentes básicos de uma ontologia são as classes e as subclasses, organizadas em uma taxonomia; as relações, que representam o tipo de interação entre os conceitos de um domínio; as propriedades ou atributos referentes a cada entidade; os axiomas, usados para modelar sentenças sempre verdadeiras, e as instâncias, utilizadas para representar elementos específicos, ou seja, os próprios dados.

Para o autor, as classes são descrições de conceitos referentes ao domínio, sendo que tais conceitos representam um conjunto de objetos abstratos. Já as subclasses são ocorrências particulares do objeto em relação à classe considerada. Uma subclasse também descreve conceitos de forma única, individualizada e concreta, fazendo referência a um objeto real. Por exemplo, no domínio EAD, temos a classe *usuário*, que tem como subclasses *tutor* e *aluno*.

As classes possuem características que são definidas pelos atributos ou propriedades que têm valor na subclasse, dessa forma se estabelece uma relação hierárquica, compondo uma taxonomia. Assim, as classes filho (subclasses) herdam atributos das classes pai.

Vejamos o exemplo citado acima: a classe *usuário*.

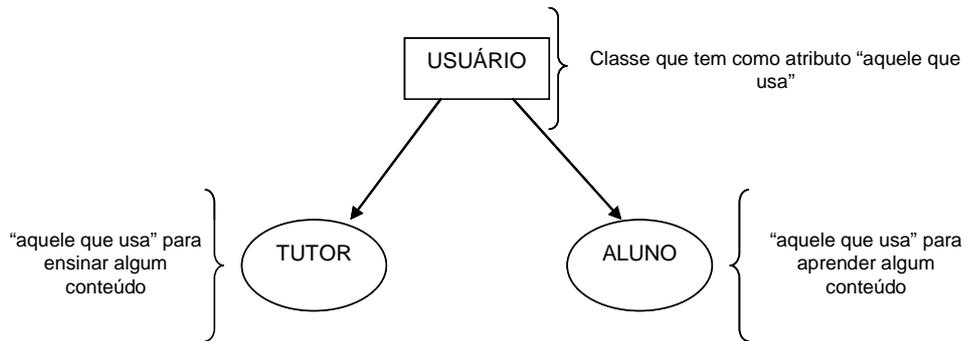


Figura 4: Herança de classes

As subclasses TUTOR e ALUNO são *classes filho* que herdaram o atributo da *classe pai* USUÁRIO, estabelecendo uma relação hierárquica.

As propriedades ou atributos são muito importantes na composição da ontologia, pois se referem à descrição de outros tipos de relações, sendo que estas correspondem a informações adicionais ligadas à estrutura hierárquica dada pelas classes e subclasses. É através das propriedades que ocorre a caracterização e a qualificação da categoria com a qual estão relacionadas.

Tais aspectos, por serem essenciais para a construção de ontologias, têm sido amplamente discutidos pelos autores das áreas da computação, da IA e da Ciência da Informação. Porém, não só a organização taxonômica e as propriedades/atributos são importantes para o desenvolvimento de sistemas capazes de depurar a linguagem natural, é fundamental levar-se em conta as questões referentes às relações existentes entre tais conceitos. Uma discussão mais detalhada acerca da estrutura e do conteúdo da M\_ONTO, enfatizando a semântica e os diferentes relacionamentos, será apresentada no capítulo 3.

Na seção seguinte discorreremos sobre as contribuições das ontologias para a melhoria e o aperfeiçoamento dos AVA's, apontando as questões relacionadas à Web Semântica.

## 2.5 Ontologias aplicadas aos AVA's: uma abordagem situada na *Web Semântica*

A internet atingiu grandes proporções nos últimos tempos e evoluiu muito rapidamente, tornando-se um poderoso meio de aprendizado e interação. Um exemplo disso são os ambientes digitais ou virtuais de aprendizagem.

Uma abordagem para os ambientes digitais é trazida por Santos (2003, p. 2), que afirma “que um ambiente virtual é um espaço fecundo de significação onde seres humanos e objetos técnicos interagem potencializando assim, a construção de conhecimentos, logo a aprendizagem”.

Dessa forma, os espaços digitais criados na *World Wide Web*<sup>24</sup> estão relacionados ao desenvolvimento de condições, estratégias e intervenções de aprendizagem em um ambiente virtual, organizados de modo que propiciem a interação entre alunos, professores e o objeto de conhecimento.

Os primeiros AVA's que foram criados com objetivo de utilização na educação foram modelados com base em quatro estratégias relacionadas às suas funcionalidades: incorporar elementos já existentes na *Web*, como correio eletrônico e grupo de discussão; agregar elementos para atividades específicas de informática, como gerenciar arquivos e cópias de segurança; criar elementos específicos para a atividade educacional, como módulos para o conteúdo e a avaliação; adicionar elementos de administração acadêmica sobre curso, alunos, avaliações e relatórios (ARAÚJO, 2003).

---

<sup>24</sup> A **World Wide Web** (que em português significa, "**Rede de alcance mundial**"; também conhecida como **Web** e **WWW**) é um sistema de documentos em hipermídia que são interligados e executados na Internet.

Informação disponível em: <<http://sites.google.com/site/historiasobreositesdebusca/www-world-wide-web>>. Acesso em: 17/11/2010

Foram criados na internet diversos ambientes virtuais, alguns com fins educacionais e outros somente de entretenimento. *Softwares* como o TelEduc<sup>25</sup>, Solar<sup>26</sup>, Sócrates<sup>27</sup>, Moodle, AulaNet<sup>28</sup>, entre outros, ganham espaço no cotidiano dos educadores virtuais pelo fato de possibilitarem fácil manuseio e controle de aulas, discussões, apresentações; enfim, atividades educacionais de forma virtual.

Neste contexto virtual, em que podemos encontrar variados *softwares* disponíveis, a Educação a Distância (EAD) ganhou espaço nos meios acadêmicos, sendo cada vez mais difundida e aperfeiçoada.

Para Moran (2002), educação a distância é o processo de ensino-aprendizagem, mediado por tecnologias, no qual professores e alunos estão separados espacial e/ou temporalmente. Esta modalidade de educação se dá através de AVA's que disponibilizam diferentes ferramentas e um ambiente interativo para o desenvolvimento do processo educativo.

Conforme apresentado anteriormente, existem diversos sistemas e ambientes virtuais que se prestam ao ensino a distância. O Moodle é um deles e vem sendo usado como ferramenta de apoio à EAD por diversas universidades que oferecem esta modalidade de ensino. Nesta pesquisa, o Moodle será o objeto de análise, pois a ontologia proposta estará baseada nas páginas e tutoriais referentes a esse ambiente.

A Educação a Distância está centrada na Web, que, apesar de ser bem estruturada e com recursos variados, ainda não possui uma estrutura que possibilite

---

<sup>25</sup> O TelEduc é um ambiente para a criação, participação e administração de cursos na Web. Ele foi concebido tendo como alvo o processo de formação de professores para informática educativa, baseado na metodologia de formação contextualizada desenvolvida por pesquisadores do NIED (Núcleo de Informática Aplicada à Educação) da Unicamp. É um software livre. Informação disponível em: <http://www.teleduc.org.br/>. Acesso em: 03/01/2010.

<sup>26</sup> Software livre desenvolvido pelo Instituto UFC Virtual, da Universidade Federal do Ceará. Informação disponível em: <http://www.solar.virtual.ufc.br/>. Acesso em 03/01/2010.

<sup>27</sup> Sistema On-Line para Criação de Projetos e Comunidades, desenvolvido Núcleo de Estudos e Pesquisas em Educação Continuada da Universidade Federal do Ceará. Informação disponível em: <http://www.solar.virtual.ufc.br/>. Acesso em 03/01/2010.

<sup>28</sup> Desenvolvido pela Universidade Federal do Rio de Janeiro. Informação disponível em: <http://www.les.inf.puc-rio/aulanet/index.html>. Acesso em: 03/01/2010.

uma busca semântica através das relações de significados. Sua estrutura é baseada na linguagem *HTML*<sup>29</sup>, cujo propósito é determinar como ter acesso e como apresentar a informação; porém, não considerando nesse propósito significado nem conteúdo. Dessa forma, os sistemas são capazes de reconhecer corpo do texto e imagem, mas não de relacionar a foto ao texto que se encontra ao seu lado, por exemplo (BAUM, 2006).

Com a intenção de melhorar tais sistemas, linguistas computacionais e cientistas da computação têm se dedicado à construção de uma nova *Web*, que tenha sistemas de busca mais completos e consistentes. Com este intuito, a W3C<sup>30</sup> criou um novo padrão: a Web Semântica (do inglês *semantic web*).

A Web Semântica é uma tecnologia que vem sendo desenvolvida pela comunidade científica com o objetivo de aperfeiçoar os sistemas de busca, possibilitando uma consulta mais ágil e precisa para o usuário.

Para melhor explicar esta nova tecnologia, Breitman (2005) faz uma comparação com a padronização da classificação dos animais:

De modo a organizar a informação na Internet, pesquisadores de Inteligência artificial vêm propondo uma série de modelos. A ideia central é categorizar a informação de maneira padronizada, facilitando seu acesso. Essa ideia é semelhante à solução utilizada para a classificação dos seres vivos. Os biólogos utilizam-se de uma taxonomia bem definida, adotada e compartilhada pela maior parte dos pesquisadores do mundo. Existem vários esforços no sentido de criar um modelo estruturado para a informação da Internet (p. 5).

Esse novo padrão de internet propõe que as bases textuais sejam anotadas com informações semânticas através de ontologias. Dessa forma, o conteúdo semântico dos documentos disponibilizados na rede, atualmente legíveis apenas por humanos, serão também lidos pelas máquinas, tornando o sistema apto a realizar

---

<sup>29</sup> *Hyper Text Markup Language* (Linguagem de Marcação de Hipertexto).

<sup>30</sup> O *World Wide Web Consortium* (W3C) cria padrões para a internet.

inferências, estabelecendo relações de sentido entre os bancos de dados e as informações fornecidas no momento da busca (BREITMAN, 2005).

Segundo Breitman (2005), alguns temas estão presentes em todas as discussões sobre a Web Semântica, conforme pode ser visualizado na figura 5<sup>31</sup>.

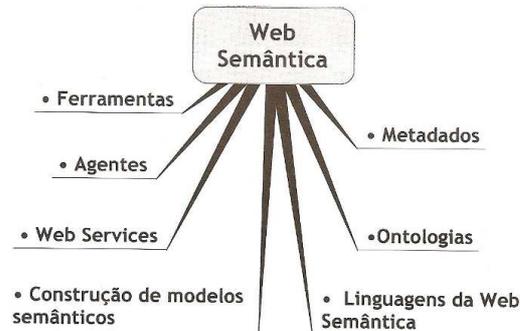


Figura 5: Temas relacionados à Web semântica

Neste trabalho damos destaque às ontologias, uma vez que elas podem representar a semântica dos documentos e permitir o uso em aplicações Web e por agentes de *software*. Ao se usarem ontologias em aplicações Web ou para habilitar agentes de *software* a entendê-las e processá-las, abre-se caminho para que as aplicações que venham a surgir no futuro sejam mais inteligentes, no sentido de considerarmos uma capacidade maior de execução de tarefas num nível conceitual mais próximo do humano.

Assim, ontologias possibilitam o preenchimento do "vazio" semântico entre a representação sintática da informação e sua conceitualização, como esquematizado na figura 6:

---

<sup>31</sup> BREITMAN, 2005, p. 6.

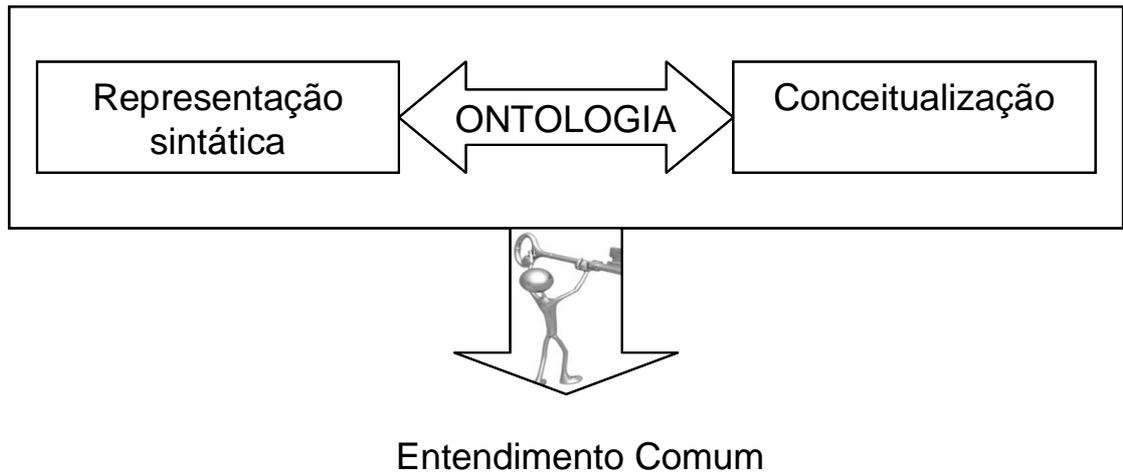


Figura 6: Ontologia como chave para o entendimento comum

Esse novo padrão para a *Web* propõe a construção de ontologias que contenham detalhamento semântico e informações estruturadas para tornar o conteúdo acessível aos sistemas de PLN. Assim, a linguagem de marcação de texto *HTML* será complementada pela *Ontology Web Language (OWL)*. Esta linguagem possui semântica e sintaxe próprias para a modelação de conteúdo semântico, deixando de ser apenas uma linguagem de marcação de texto. Com esse novo modelo, o sistema será capaz de “realizar inferências, estabelecendo relações de sentido entre as bases eletrônicas e as informações fornecidas pelo usuário na solicitação da busca” (ALVES, 2005, p. 20).

Conforme Stojanovic et.al. (2001), alguns aspectos da EAD podem ser melhorados com base na estrutura da Web Semântica, mais precisamente com a criação de ontologias, possibilitando novas formas de navegação e acesso ao conteúdo. Dentre eles, destacam-se:

- (a) *Entrega*: possibilita a consulta semântica aos tópicos de interesse, uma vez que os materiais de aprendizagem distribuídos na Web são conectados através de uma ontologia;
- (b) *Acesso*: possibilita a realização de pesquisas semânticas nos materiais disponíveis. Por exemplo: consultas do tipo “Quais materiais de aprendizagem sobre linguística estão disponíveis no Unisinos Virtual?”

devolveriam como resposta não só os materiais disponíveis no curso em que o aluno está matriculado, mas também o que está disponível em outros cursos, permitindo ao aluno uma complementação de estudos;

- (c) *Integração*: permite a integração entre variados cursos e seu conteúdo;
- (d) *Distribuição*: possibilita a maior distribuição de materiais de aprendizagem através da anotação semântica;
- (e) *Personalização*: possibilita a criação um material de aprendizagem personalizado, de acordo com as necessidades de cada estudante. Dessa forma, a ontologia estabelece a conexão entre as necessidades e o material de aprendizagem.

Sintetizamos as contribuições das ontologias para os ambientes EAD através da figura 7.

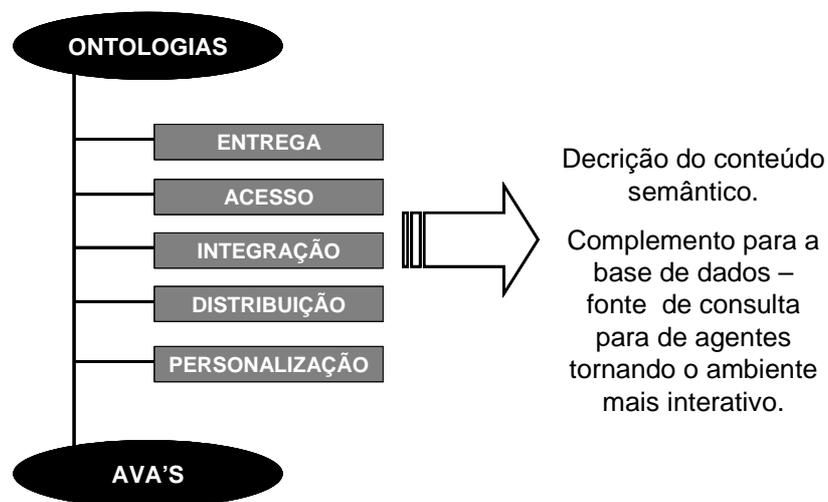


Figura 7: Contribuições das ontologias para os AVA's

As ontologias, como um dos temas relacionados à Web Semântica, têm um importante papel na melhoria dos sistemas de EAD, uma vez que permitem a descrição semântica do conteúdo, podendo servir de base para a construção de agentes e tornando os ambientes mais interativos (STOJANOVIC et.al., 2001).

A representação semântica dos dados, possível através das ontologias, permite aos agentes maior precisão e qualidade na recuperação de informações, beneficiando diretamente o desenvolvimento de aplicações para EAD. Assim, as ontologias são fundamentais para a normalização de conceitos e as relações semânticas estabelecidas entre eles. Por essa razão, apresentamos na seção 2.6 a aproximação da ontologia que desenvolvemos para o ambiente Moodle com o agente a ser implementado futuramente.

## **2.6 Ontologia + agente = uma aplicação**

As ontologias possibilitam que os documentos relacionados ao ambiente *Web*, no nosso caso o ambiente virtual de aprendizagem Moodle, sejam acessados de forma inteligente, através da dedução e da inferência dos dados que foram declarados explicitamente na sua organização hierárquica e semântica (STAAB & MAEDCHE, 2002). A dedução e a inferência em relação aos dados se dão através do agente, que será capaz de “buscar” as informações na ontologia e torná-las úteis para seu propósito.

Muitas são as concepções de agentes inteligentes apresentadas na literatura e, por esta razão, consideramos importante situar nossa visão em relação a este conceito. Baseamo-nos na proposta de Wooldridge & Jennings (1995), que explicam *agente* como sendo um sistema computacional que recebe como entrada estímulos do meio (no nosso caso, a interação do usuário com o ambiente) e, com base neles, realiza ações que poderão modificar este meio de alguma forma, sendo uma tarefa realizada sem a inferência direta do usuário.

Os agentes têm sido usados para diversas aplicações, até mesmo no campo pedagógico, sendo ferramentas de apoio ao processo de ensino e aprendizagem. No caso do ambiente virtual, maiores são os ganhos com o uso de agentes, uma vez que eles são capazes de interagir com os usuários aumentando o desempenho e a motivação dos estudantes.

Conforme Bogo (2003), um sistema educacional – no nosso caso o Moodle – pode ter a inserção de um agente capaz de “perceber” a ação do usuário e “reagir” executando algum processo. Por exemplo: o agente pode exibir alertas e notificações sobre quaisquer novidades que surjam na comunidade da disciplina/curso que o aluno esteja cursando. Além disso, o agente pode ter a função de verificar a conectividade, garantido que a conexão do aluno esteja ativa ao realizar alguma interação com o ambiente e, caso não esteja, manter um histórico das atividades não completas, de modo a completá-las posteriormente sem a necessidade de o aluno interagir com o *software*. Como outro exemplo de ação do agente, podemos citar que ele será capaz de fazer tratamento dos dados digitados pelo aluno em tempo real, fornecendo *feedback* quando alguma operação não puder ser concluída e também informando ao professor/tutor de modo que o aluno não seja prejudicado.

Assim, o agente, com sua capacidade de perceber as ações ocorridas no ambiente educacional, poderá auxiliar o processo de aprendizagem através de sucessivas interações com os usuários, promovendo um *continuum* entre ações e reações.

Na figura 8 esquematizamos o *continuum* de ações e reações estabelecendo a relação entre os usuários e o ambiente Moodle e destacando a participação da ontologia.

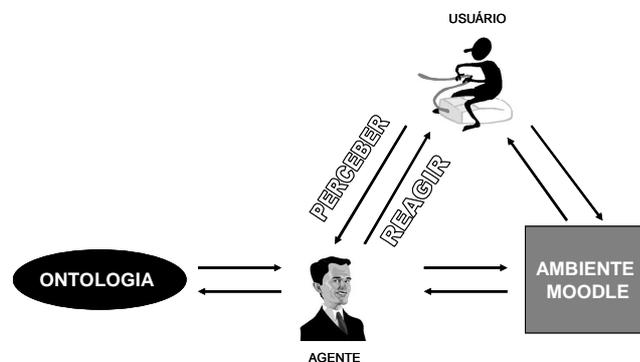


Figura 8: Continuum de ações entre ambiente, ontologia e usuário

Essa “inteligência” do agente tem como base a descrição semântica e hierárquica delineadas na ontologia. Através dos dados consultados na ontologia, o agente será capaz de compreender os relacionamentos entre as entidades – alunos/atividades e tutor/atividades, por exemplo – e fazer as inferências necessárias.

Na figura 9, ilustramos a posição da ontologia em relação ao agente e sua inserção na Web Semântica através do PLN.

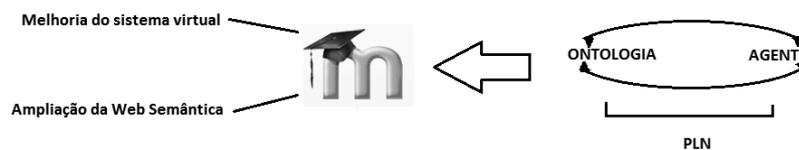


Figura 9: Relação entre agente e ontologia para melhoria do AVA

O agente alimenta-se de informações contidas na ontologia, rica em informações semânticas, e amplia as potencialidades do Ambiente Moodle, tornando-o mais interativo e favorecendo as aplicações para a ampliação da Web Semântica.

Tanto para a ontologia quanto para o agente, a descrição das relações semânticas e hierárquicas são fundamentais e, por esta razão, no capítulo seguinte apresentaremos o que consideramos como o conteúdo da ontologia, que norteará nossa descrição e formalização.

Para finalizar, apresentamos um mapa conceitual<sup>32</sup> (figura 10) que sintetiza os principais conceitos discutidos neste capítulo.

---

<sup>32</sup> Mapa conceitual (MC) é uma ferramenta que serve para estruturar o conhecimento, representando ideias e conceitos em forma de um diagrama hierárquico, explicitando relações entre estes conceitos com o objetivo de refletir sobre a estrutura cognitiva de um determinado assunto. Segundo a literatura, apresentam semelhanças com ontologias; porém, o fato de que os difere é em relação à reusabilidade, isto é, os MC são específicos de uma determinada área, não podendo ser reutilizados, já as ontologias permitem o reuso de sua estrutura (LIMA, 2004).

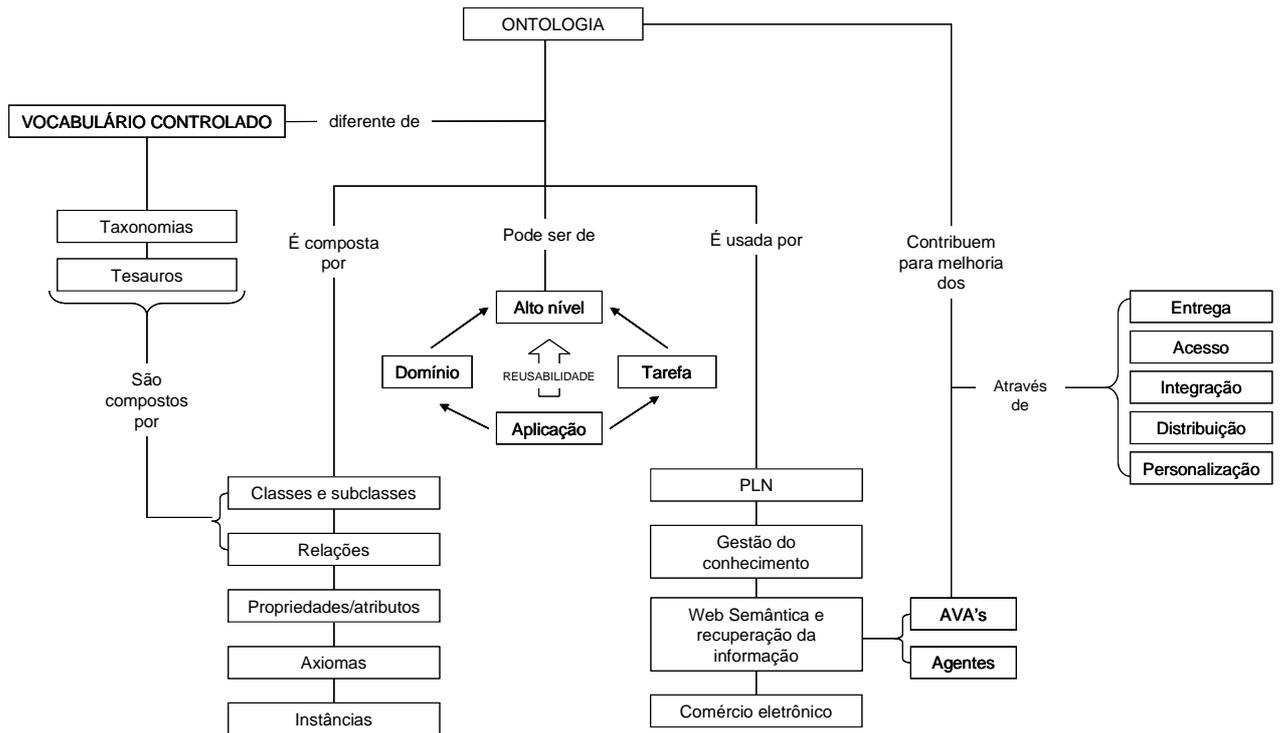


Figura 10: Mapa conceitual – síntese do capítulo

No capítulo seguinte, enfatizamos o conteúdo da M\_ONTO, isto é, as relações e categorias que farão parte da sua modelagem.

### 3 O CONTEÚDO DA M\_ONTO

As ontologias são compostas de conceitos integrados através de relações semânticas, cuja natureza e número têm sido alvo de inúmeros estudos nas áreas da Computação, da Ciência da Informação e da Linguística. De acordo com Green (2001, p. 5-6), “o inventário das relações semânticas inclui um conjunto fechado de relações (incluindo relações essencialmente hierárquicas e de equivalência) e um conjunto aberto de relações”, referindo-se especificamente aos diferentes tipos de relações que podem se estabelecer entre os conceitos de acordo com cada domínio.

Estudos sobre as relações que devem compor uma ontologia são realizados nas diferentes áreas que as utilizam como forma de descrever o conhecimento e apresentam perspectivas distintas. Dependendo dos aspectos da língua a serem focalizados, os estudiosos da Linguística, da CI e da CC propõem diferentes formas de descrição semântica para organizar os conceitos. Para os estudiosos da Computação, o que é relevante é a conveniência computacional e, de acordo com White (1988)<sup>33</sup>, todos os problemas semânticos devem ser resolvidos usando modelos relacionais. O linguista Sowa (1988)<sup>34</sup>, pesquisador envolvido com as questões computacionais, considera que é possível combinar as relações semânticas livremente com outros modelos de descrição.

Segundo Hjørland (2007a), estudos que envolvem as relações semânticas são abordados de forma fragmentada no âmbito da CI. Para o autor, a área da CI deveria promover um debate teórico mais aprofundado acerca das questões semânticas e os pesquisadores desta área, mais precisamente os que trabalham

---

<sup>33</sup> Apud EVENS, 1988.

<sup>34</sup> Apud EVENS, 1988.

com a representação do conhecimento, deveriam fundamentar seus trabalhos em teorias semânticas, favorecendo uma maior aproximação entre as áreas. Analisando diferentes trabalhos da área da CI, percebemos que há um grande interesse em sistematizar, isto é, criar padrões metodológicos para a criação de ontologias; porém, tais trabalhos tornam-se semanticamente fracos, sendo que poderiam se valer mais fortemente das teorias semânticas desenvolvidas na Linguística, compromisso adotado neste trabalho.

Levando em consideração os pontos citados acima, objetivamos, neste capítulo, refletir sobre a modelagem semântica da ontologia, isto é, discorrer sobre os principais aspectos sobre significado e relações semânticas, buscando identificar quais são os relacionamentos mais adequados ao nosso propósito, considerando que ontologias são constituídas de classes e subclasses de conceitos interligados fortemente através de relações semânticas.

Consideramos importante apresentar *qual a Semântica* que vai compor a nossa ontologia e de que forma o domínio do Ambiente Moodle vai ser descrito levando em conta abordagens teóricas que a Semântica apresenta para a descrição do significado. Saliemos que, neste capítulo, seguiremos um viés semântico para a construção desta modelagem, destacando seu potencial e diferencial na construção da ontologia proposta. No entanto, chamamos a atenção para o fato de que não deixaremos de apresentar as significativas contribuições das áreas da CI e da CC buscando estabelecer um paralelo entre as diferentes abordagens.

A fim de esclarecermos os aspectos citados inicialmente, organizamos o capítulo em quatro seções. Na seção 3.1, apresentamos, em linhas gerais, os aspectos relacionados às diferentes abordagens semânticas que são essenciais para nosso trabalho. A seção 3.2 complementa a anterior e destaca nossa escolha pela Semântica Lexical e pela Semântica Cognitiva, mais especificamente a Semântica de Frames. Na seção 3.3 enfatizamos as relações semânticas organizando-as em paradigmáticas e sintagmáticas e oferecendo detalhamentos sobre as mesmas. Por fim, a seção 3.4 tem como finalidade apresentar as diferentes categorias que irão compor a nossa ontologia, sendo tal aspecto considerado o espaço de transição entre o conteúdo da M\_ONTO e as etapas metodológicas a serem seguidas.

### 3.1 Semântica: da teoria à aplicação

A semântica é uma área da Linguística que se apresenta através de estudos complexos e variados, uma vez que decorre de diversas escolas e linhas de pensamento desenvolvidas ao longo da história, sendo algumas até mesmo consideradas obsoletas. O estudo do significado é o foco da semântica levando em conta o conhecimento do falante, isto é, considerando suas habilidades linguísticas (SAEED, 2003).

Saeed (2003) aponta que, de acordo com os estudos da Linguística moderna, o falante de uma língua tem diferentes tipos de conhecimento linguístico, incluindo a pronúncia das palavras, a forma adequada de construir as frases e o significado das palavras e sentenças, fazendo-se necessário um estudo linguístico que contemple diferentes níveis de análise<sup>35</sup>. Segundo o autor, nos diferentes níveis de análise teríamos o léxico, a fonologia, a morfologia, a sintaxe e a semântica. Dessa forma, “conhecer uma palavra une diferentes tipos de conhecimento” (SAEED, 2003, p. 4, tradução nossa).

Na presente pesquisa, a fonologia, a morfologia e a sintaxe não apresentam papel relevante, uma vez que não são abordagens capazes de representar interpretações das diferentes entidades do mundo; elas podem, porém, auxiliar na análise da estrutura da ontologia com informações complementares, caso haja necessidade. No caso deste estudo, a ontologia é uma solução para a representação do conhecimento linguístico relacionado ao ambiente Moodle, cujas informações serão descritas no âmbito semântico ou semântico-pragmático.

Há um consenso entre os semanticistas de que a semântica é a ciência que estuda o significado das línguas (ALLAN, 2001); porém, não há unanimidade em relação às formas de descrição deste significado, pois se trata de uma definição um tanto abstrata.

---

<sup>35</sup> Cançado (2005), em seu Manual de Semântica, também discute os mesmos aspectos sobre a semântica corroborando com as ideias apresentadas por Saeed (2003).

Segundo Saeed (2003), a questão em relação à definição do significado na semântica reúne, pelo menos, três grandes desafios. Como primeiro desafio, está exatamente a definição das palavras; a este problema o autor chama de *circularidade*. Por exemplo, ao buscarmos um significado para professor teremos aquele que ministra aulas e conseqüentemente teremos que ter um significado para ministra e aulas e, para cada uma destas palavras, teremos um novo significado e novas palavras para definir, ou seja, para compreender o significado de professor é necessário compreender o significado de diversas outras palavras; portanto, “se a definição do significado das palavras é dada em palavras, o processo nunca pode acabar” (p. 6, tradução nossa).

Um segundo problema apontado pelo autor é como podemos ter certeza de que as definições do significado das palavras são exatas, pois, para Saeed (2003), o significado é um tipo de conhecimento na mente dos falantes nativos de uma língua, isto é, uma palavra poder ter um significado para a pessoa X que é diferente para a pessoa Y, e isso vai estar estritamente ligado a sua vivência e à forma como organiza as informações em sua mente.

E o terceiro problema apontado por Saeed (2003) é referente ao contexto. As palavras possuem diferentes significados de acordo com o contexto no qual estão inseridas. Por exemplo: tutor, no contexto EAD, significa *professor*, já no contexto jurídico significa aquele que detém a tutela sobre alguém.

O semanticista Alan Cruse também aborda as questões referentes ao significado e seus pressupostos diferem parcialmente da concepção apresentada por Saeed. De acordo com Cruse (2000), o significado de uma palavra, apesar de ser imensamente mutável e sensível ao contexto, possui “regiões de ‘alta densidade significativa’ formando ‘blocos de significações’” (ALVES, 2005, p. 74), gerando uma maior ou menor estabilidade em relação a mudanças contextuais. Podemos trazer como exemplo o caso do item lexical professor. O significado desta palavra, independentemente do contexto (desde que não usada de forma metafórica), sempre envolverá um mesmo “bloco de significação”, isto é, aquele que ministra aulas; porém, no nosso caso, por se tratar de um contexto EAD, este item lexical contará com uma propriedade diferenciada, uma vez que as aulas não são presenciais e a forma de ministrá-las é diferente. Assim, de acordo com Cruse

(2000), nem toda mudança de contexto produz uma expressiva mudança no significado das palavras, conforme proposto por Saeed (2003).

Tal complexidade, apresentando diferentes concepções acerca do significado e as implicações trazidas pelo contexto, tem influência direta na modelagem da ontologia. Diante disso, consideramos que um aspecto fundamental para a nossa ontologia é a delimitação clara e objetiva do domínio ao qual ela se refere, eliminando, assim, as dificuldades relacionadas ao problema de contexto. No que se refere à caracterização do domínio da nossa ontologia, consideramos importante aprofundar em uma seção específica e, portanto, não detalharemos este aspecto neste capítulo, destacando para ele uma seção no capítulo 4.

Salientamos outro ponto importante no que tange às relações que podem compor uma ontologia. Neste tipo de recurso, na representação da informação, ao contrário de recursos tradicionais, não há uma estrutura pré-definida e limitada de relações semânticas, mas há a possibilidade de que sejam refinadas e incorporadas à estrutura ontológica de acordo com a necessidade de representação e com as características do próprio domínio.

Conforme apresentado anteriormente, muitos são os estudos que abordam a semântica e suas relações para a construção de ontologias, principalmente na área da CI e da CC.

A área da CI busca soluções prontas para tratar as questões sobre semântica, criando modelos detalhados de relacionamentos semânticos para o desenvolvimento de seus trabalhos, procurando desenvolver um inventário de relacionamentos semânticos incluindo distinções entre tipologias (GREEN, 2001).

Já a área da CC trata das questões de relacionamentos através das redes semânticas, o que podemos considerar muito próximo da abordagem Linguística. Para Quilliam e Collins (1969), os conhecimentos na memória humana e a forma como estão organizados podem ser expressos através de uma rede semântica.

Neste trabalho, buscamos uma aproximação com as áreas da CI e da CC, mas não temos a pretensão de criar um modelo definitivo, pois a sistematização dos modelos semânticos deve ser feita de acordo com a necessidade do estudo em

questão, no nosso caso, a ontologia que será utilizada pelo agente para a interação entre usuário e ambiente Moodle.

Em nosso estudo enfatizamos as abordagens semânticas que se prestam à descrição do significado dos nominais e de entidades verbais compatíveis com a elaboração da ontologia. Assim, o significado será investigado sob dois diferentes enfoques: o enfoque semântico-lexical, mais precisamente a Semântica Lexical ou Relacional<sup>36</sup>, e o enfoque semântico-cognitivo, mais precisamente a Semântica Cognitiva, com destaque para a semântica de frames.

Procurando elucidar nosso posicionamento, buscamos subsídios em Niremburg & Raskin (2005), que denominam de *Semântica Ontológica* os estudos referentes às abordagens que buscam tratar do significado através de recursos computacionais.

Semântica Ontológica é uma integração complexa de teorias, metodologias, descrições e implementações. Na Semântica Ontológica, uma teoria é vista como um conjunto de afirmações que determinam o formato das descrições dos fenômenos dos quais a teoria trata. Uma teoria é associada com uma metodologia usada para obter as descrições. As implementações são sistemas computadorizados que usam as descrições para resolver problemas específicos em processamento de textos. As implementações da Semântica Ontológica são combinadas com outros sistemas de processamento para produzir aplicações, tais como a extração da informação ou a tradução por máquina (NIREMBURG & RASKIN, 2005, p. 6 - tradução nossa).

Diante da definição proposta pelos autores, percebe-se que este conceito está relacionado a um universo de microteorias que levam em conta, não só a teoria em si, como também as necessidades das aplicações nas quais as teorias serão aplicadas, isto é, a Semântica Ontológica é norteada pela necessidade da tarefa de processamento da linguagem natural.

---

<sup>36</sup> Uma Semântica Lexical, embora não siga uma abordagem cognitivista, pode ser também considerada uma Semântica Relacional, com base na Linguística Cognitiva (MURPHY, 2003). Tal aspecto torna-se importante uma vez que as relações taxonômicas são parte da ontologia.

Dessa forma, ao assumirmos que nossa demanda está comprometida com a demanda do PLN, podemos optar pela teoria semântica que melhor se adapta ao nosso propósito, uma vez que uma abordagem semântica para PLN deve se ajustar às necessidades do sistema, tanto de geração quanto de compreensão da linguagem.

Perante o exposto e de acordo com a necessidade de aplicação da nossa ontologia, assumimos a Semântica Lexical (análise do léxico) como principal perspectiva semântica para o desenvolvimento do nosso trabalho. Faremos uso, também, da Semântica Cognitiva, mais especificamente a Semântica de Frames (FILLMORE, 1982), como apoio para o estabelecimento das relações na ontologia e para a definição do significado dos termos e conceitos. Salientamos, porém, o fato de que uma Semântica Lexical também pode ser vista como uma abordagem cognitiva, mas optamos por fazer esta distinção entre Lexical e Cognitivo a fim de melhor organizar nosso texto.

Julgamos importante tecer esclarecimentos acerca das teorias que elegemos como fundamentais para a elaboração da M\_ONTO.

### **3.2 Nossas escolhas teóricas**

A Semântica Lexical considera as propriedades referentes a cada unidade lexical para analisar o significado das palavras e estabelecer as relações necessárias para a construção da ontologia proposta. Outro aspecto que nos leva à Semântica Lexical é o fato de esta estar associada às categorias de palavras conhecidas como “palavras de classe aberta”<sup>37</sup>, isto é, os verbos, os substantivos e os adjetivos, uma vez que nossa investigação se dará em relação aos nomes e às ações realizadas pelos participantes no ambiente virtual de aprendizagem. Além

---

<sup>37</sup> As chamadas “palavras de classe aberta” são assim denominadas em função de terem um número ilimitado e pelo fato de o sistema permitir a criação ou inclusão de novos membros – os neologismos e os estrangeirismos. Além disso, seus membros estão sujeitos a uma reviravolta relativamente rápida nas classes a que pertencem (CRUSE, 1986).

disso, a Semântica Lexical se preocupa em estudar as propriedades dos significados das palavras, o que a torna relevante para a construção de uma ontologia.

Cançado (2005) apresenta uma breve revisão sobre os estudos de Semântica Lexical, organizando-os em duas correntes ou linhas. Uma primeira linha, que tem como foco a relação entre a estrutura sintático-semântica do léxico e da sentença, explora a noção de papel temático, sendo estas teorias compatíveis com a gramática gerativa. Ainda nesta linha, a autora inclui os estudos do Léxico Gerativo de Pustejovsky<sup>38</sup>, teoria que “propõe estruturar o léxico de uma forma organizada, composicional e geradora de novas estruturas. É um trabalho que busca uma viabilidade de aplicação da teoria linguística à prática computacional” (p. 145). Os trabalhos realizados por Cruse, 1986, 2000; Evens, 1988; Lehrer e Feder, 1992<sup>39</sup>, são apresentados por Cançado como uma segunda linha para os estudos em Semântica Lexical, sendo que tais teorias se ocupam das palavras e das relações entre elas, ou seja, investigando somente as relações de sinonímia, antonímia, hiponímia e meronímia.

As informações semânticas estudadas pela Semântica Lexical, conforme definição de Cruse (1986, 2000) – *sinonímia*, *antonímia*, *hiponímia* e *meronímia* – parecem centrais na construção de ontologias, uma vez que estas são compostas principalmente por uma taxonomia, ou seja, uma hierarquia de classes e subclasses, conforme descrito no capítulo anterior. Destacamos as relações semânticas supracitadas na seção seguinte, na qual enfatizamos a descrição e a importância das mesmas para a nossa ontologia, apresentando alguns exemplos com base na nossa análise.

A Linguística Cognitiva (LC), por sua vez, surge no final dos anos 70 a partir da crítica aos paradigmas anteriores (estruturalismo e gerativismo), que concebiam a linguagem como um sistema autônomo. Segundo esse novo paradigma, há uma relação entre o desempenho linguístico e a mente, desta forma o foco do estudo é o sistema complexo das correlações entre a linguagem e as estruturas mentais. No

---

<sup>38</sup> 1995.

<sup>39</sup> Apud CANÇADO (2005).

âmbito da LC, o homem compreende o mundo através de diferentes modelos de categorização. Surge daí a corrente semântica derivada desta linha e designada Semântica Cognitiva (Talmy, 1988; Langacker, 1995 e Fillmore, 1982, estão entre os principais expoentes desta visão teórica).

Langacker (1995)<sup>40</sup> compreende que existem duas habilidades cognitivas básicas: a abstração e a categorização. De acordo com o autor, existem apenas três elementos que podem ser atribuídos ao sistema linguístico:

(a) estruturas semânticas, fonológicas e simbólicas, através das expressões usadas usualmente na situação de comunicação;

(b) abstrações ou esquematizações;

(c) relações entre os esquemas e suas estruturas.

Segundo Feltes (2007), o que fica evidente nas considerações de Langacker é o fundamental papel da categorização na Linguística Cognitiva; logo, “a categorização humana é o coração do programa global da Semântica Cognitiva” (p. 80).

Complementando tais afirmativas, lembramos Silva (1997), que apregoa que uma das formas de aquisição do conhecimento decorre da categorização, como uma capacidade cognitiva fundamental. Para o autor, categorização significa “o processo mental de identificação, classificação e nomeação de diferentes entidades como membros de uma mesma categoria” (p. 7).

Para o estabelecimento do significado das palavras e para a construção da ontologia a qual nos propomos desenvolver neste trabalho, consideramos que o ponto central encontra-se na análise das relações semânticas, sejam elas as fechadas - que já são preestabelecidas, ou as abertas - que serão determinadas através do domínio (GREEN, 2001). Na seção seguinte, teceremos considerações sobre as relações semânticas, destacando-as na abordagem lexical e na abordagem cognitiva e situando-as no que tange à nossa aplicação.

---

<sup>40</sup> Apud FELTES (2007, p. 80).

### 3.3 Sobre as relações semânticas

Tendo discorrido na seção 3.2 acerca de aspectos gerais da semântica, mais precisamente a Semântica Lexical e a Semântica Cognitiva, sem, no entanto, ter detalhado aspectos relativos aos relacionamentos semânticos para uma ontologia, reservamos esta seção para aprofundar as questões referentes às relações, respeitando as escolhas semânticas descritas na seção anterior.

Há consenso entre semanticistas e informatas de que as relações semânticas são fundamentais na construção de ontologias. Segundo Saint Dizier e Viegas (1995), as relações semânticas exercem um papel central na Semântica Lexical e interferem no plano do processamento da linguagem natural, uma vez que as ontologias são consideradas ferramentas para a representação do conhecimento, pois tais estruturas se caracterizam por um sistema classificatório bem descrito e marcado, permitindo que sejam passíveis de entendimento para a máquina.

Além disso, de acordo com Oltramari (2010, p. 12), “em ontologias os conceitos são integrados em um todo coerente com as relações” (tradução nossa) e as relações são conceitualmente orientadas e assumem conceitos como argumentos. O autor também argumenta que há diferenças na proposição de relações para a criação de ontologias formais e ontologias linguísticas, uma vez que as mesmas relações podem significar diferentes organizações nestes dois níveis.

De acordo com Gomes et.al. (2008), as abordagens para o estabelecimento das relações diferem entre as variadas áreas que apresentam interesse sobre este tipo de estudo, uma vez que tratam de diferentes objetos de pesquisa. Diante disso, as autoras prescrevem que é necessário definir claramente os objetivos a serem atingidos, pois a partir destes objetivos é que serão determinadas as abordagens a serem seguidas.

Khoo e Na (2006) ressaltam a importância dos tipos, da maneira como são representadas e das especificidades das relações para a construção de ontologias, sendo aspectos mais profundamente discutidos pela Linguística. De acordo com os autores, as relações semânticas são o caminho para o aperfeiçoamento do PLN e da recuperação da informação.

Diante de nossa escolha teórica pela Semântica Lexical e pela Semântica Cognitiva, optamos por dividir as relações semânticas em dois grupos. Consideramos as relações que se referem a uma abordagem lexical como relações paradigmáticas e destacamos a subseção 3.3.1 para tratar sobre elas. Já as relações do âmbito cognitivo estão relacionadas a abordagens sintagmáticas e são apresentadas na subseção 3.3.2.

### 3.3.1 Relações paradigmáticas

As relações semânticas em uma abordagem lexical podem ser vistas como paradigmáticas ou sintagmáticas. As relações paradigmáticas ocorrem entre as palavras de uma mesma classe gramatical, ou categoria sintagmática, conforme Cruse (2000), e refletem as escolhas semânticas realizadas para proferir a sentença, sendo possível realizar substituições sem prejuízo à estrutura correta da frase. Já as relações sintagmáticas ocorrem em sentido linear, dentro de uma mesma sentença, através das regras gramaticais e sintáticas.

A tabela 3<sup>41</sup> ilustra a forma como ocorrem as relações sintagmáticas e paradigmáticas. Os itens que têm relação horizontal (nas linhas) representam a visão sintagmática, já os itens que têm relação vertical (nas colunas) representam a visão paradigmática e podem ser substituídos um pelo outro.

RELAÇÕES	←..... Sintagmáticas .....→				
 Paradigmáticas 	Este	cachorro	dormiu	nesta	cadeira
	O	gato	sentou	no	chão
	Aquela	mulher	comeu	com	chapéu

Tabela 3: Relações sintagmáticas e paradigmáticas

<sup>41</sup> Adaptação da tabela disponível em:

<[http://changingminds.org/explanations/critical\\_theory/concepts/syntagm\\_paradigm.htm](http://changingminds.org/explanations/critical_theory/concepts/syntagm_paradigm.htm)>. Acesso em: 12/01/2011.

Segundo Saeed (2003), nas relações paradigmáticas podemos nos limitar à substituição de termos em um determinado contexto, conforme pode ser visto na figura 11.

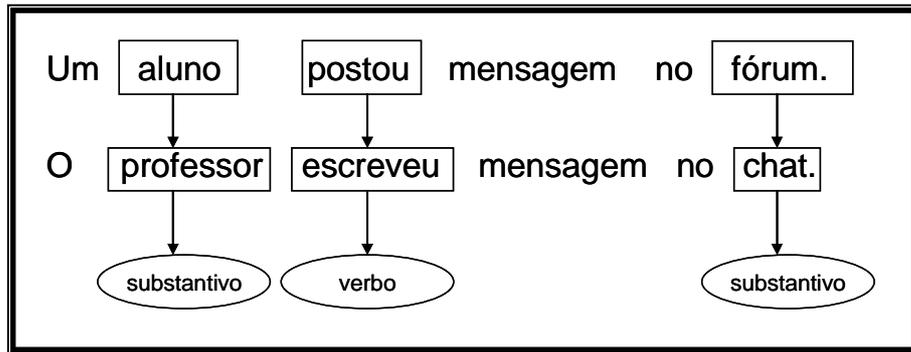


Figura 11: Exemplos de relações paradigmáticas

Di Felippo (2004) apresenta uma subdivisão para as relações paradigmáticas, as quais são apresentadas detalhadamente por Alves (2005, p. 79). A figura 12<sup>42</sup> traz uma aproximação entre as diferentes relações paradigmáticas e teorias semânticas.

---

<sup>42</sup> Adaptado de ALVES (2005).

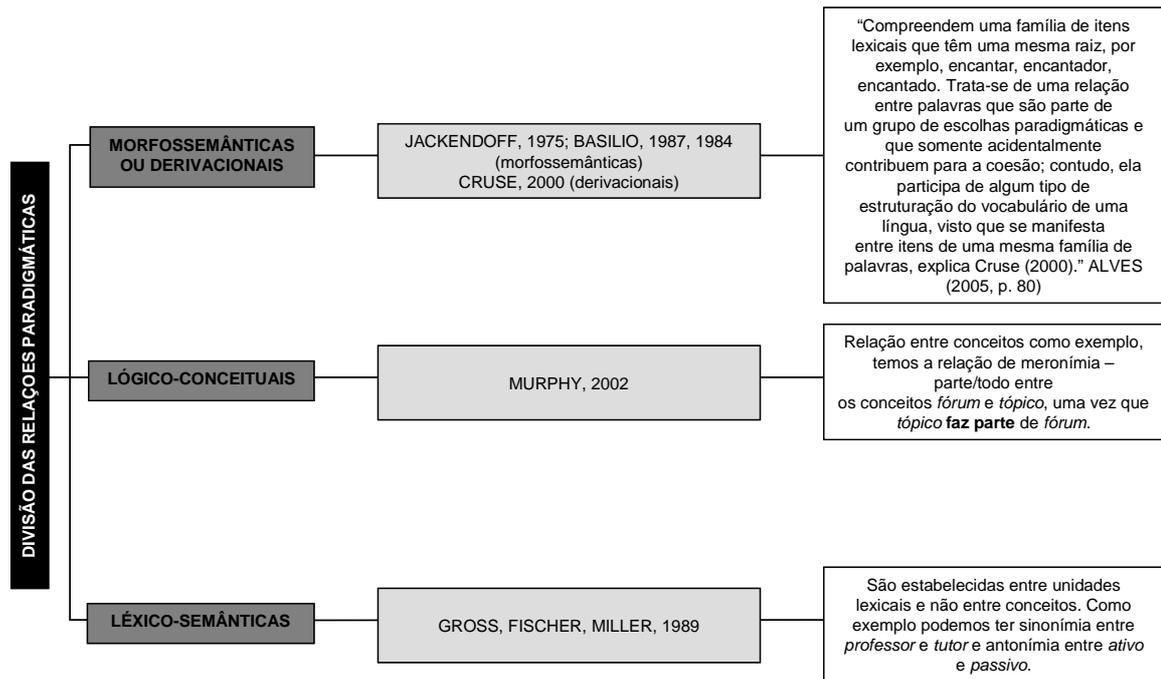


Figura 12: Diferentes tipos de relações paradigmáticas

Consideramos importante elucidar um ponto acerca da divisão proposta por Di Felippo (2004), uma vez que optamos por organizar nosso referencial semântico separando em Semântica Lexical e Cognitiva. Nesta subdivisão temos abordagens cognitivistas, tais como as de Jackendoff e Murphy, o que nos mostra que uma abordagem lexical também pode ser uma abordagem cognitiva. Ao realizarmos uma separação entre Semântica Lexical e Cognitiva, não desconsideramos este fato nem vemos as duas abordagens como totalmente antagônicas, apenas buscamos organizar o conteúdo visando a uma forma mais clara e objetiva de aprofundar tais teorias.

As relações paradigmáticas são associadas por Cruse (2000) à coerência entre as classes, estando ligadas à identidade, inclusão, sobreposição e disjunção. As relações mais conhecidas no domínio lexical como sendo do eixo paradigmático são a hiponímia/hipernímia, a meronímia/holonímia e a sinonímia/antonímia (informações detalhadas sobre cada tipo de relação serão apresentadas no decorrer deste capítulo). Em ontologias este tipo de relacionamento está ligado às relações conceituais de inclusão do tipo *é-um* ou *parte-de*.

Tais relações também são conhecidas como taxonômicas, principalmente quando estão relacionadas à criação de taxonomias, glossários e tesouros, estando mais diretamente ligadas às áreas de Biblioteconomia, Ciência da Computação, Terminologia e Ciência da Informação; porém, conforme já afirmamos anteriormente, todas as ontologias são compostas de uma taxonomia e, portanto, tais relações são fundamentais para o estabelecimento da hierarquia de classes que compõe a ontologia.

A CI também trata das relações paradigmáticas, incluindo em seus estudos as relações hierárquicas (hiponímia), partitivas e associativas (meronímia) – no plano do conceito e de equivalência (sinonímia) – no plano da língua.

Guarino (1995) trata as relações taxonômicas – para a Semântica, relações de hiponímia e meronímia – como relações “estruturantes”, e as demais como “não estruturantes”, referindo-se àquelas que oferecem informação adicional aos conceitos – as relações associativas.

Smith et.al. (2005) discorrem sobre a dificuldade em enumerar todos os tipos de relações existentes e buscam sistematizar através de três tipos de relações binárias<sup>43</sup>, a saber:

- (a) relação *classe-classe* (é-um);
- (b) relação *instância-classe* (tipo-de);
- (c) relação *instância-instância* (parte-de).

Com esta classificação, as relações podem ocorrer entre classes, entre classes e instâncias e entre as instâncias. Tal modelo estabelece muito bem as relações taxonômicas; porém, desconsidera as relações sintagmáticas para o desenvolvimento de ontologias para o PLN.

---

<sup>43</sup> Almeida (2006) explica que relação unária é a relação entre um e outro conceito (que é seu atributo, uma característica do conceito) e relação binária é a relação entre dois conceitos.

Na Computação, as redes semânticas também produzem este tipo de relação taxonômica, que podem ser representadas através do relacionamento “é-um” (*is a*). Através destas relações, é introduzido o conceito de herança<sup>44</sup>, isto é, as propriedades e características são herdadas através do relacionamento. Por exemplo: <peessoa> tem <endereço><sup>45</sup>, <aluno> é uma <peessoa>, logo <aluno> tem <endereço> – característica herdada através do relacionamento. Tais relacionamentos equivalem à hiponímia e à meronímia, respectivamente (RUSSEL e NORVIG, 2004).

Diante do exposto, percebe-se que as relações de hiponímia/hipernímia e meronímia/holonímia aparecem como fundamentais para a construção de ontologias, visão compartilhada pelas áreas da CI, da Linguística e da Computação. Passaremos, a seguir, a um detalhamento de cada uma destas relações, buscando apresentar exemplos baseados em nossa análise.

### a) Hiponímia/hipernímia

A relação de hiponímia é conhecida como uma relação de inclusão, ou seja, a relação entre um termo mais específico em um mais geral (LYONS, 1987). Por exemplo: *banana* e *maçã* são **hipônimos** de *fruta*, logo *fruta* é **hiperônimo**, ou seja, inclui *banana* e *maçã* na categoria das *frutas*.

Em relação a esta estrutura hierárquica entre classes e subclasses, a visão cognitivista traz os conceitos de superordenado, subordinado (hiponímia e hipernímia) e nível básico, a exemplo dos estudos realizados por Rosch e Mervis

---

<sup>44</sup> Herança é um princípio de orientação a objetos, que permite que classes compartilhem atributos e métodos, através de "heranças". Ela é usada na intenção de reaproveitar código ou comportamento generalizado ou especializar operações ou atributos. O conceito de herança de várias classes é conhecido como herança múltipla. Informação disponível em: <http://pt.wikipedia.org/wiki/Heranca>. Acesso em: 09/01/2011.

<sup>45</sup> Seguiremos este padrão para destacar os exemplos, marcando as classes entre < > e os relacionamentos com a fonte Courier New - <classes> relacionamentos <classes>. Este padrão será adotado no decorrer da parte teórica e na etapa linguística de análise. Na etapa computacional faremos um ajuste para adequar as classes e relações ao Protegé.

(1976)<sup>46</sup>. Conforme Silva (2003), os estudos da Psicolinguística sobre as hierarquias lexicais constataam um nível cognitivamente mais saliente, chamado “*nível básico*”, sendo caracterizado como o mais genérico em uma taxonomia. Além disso, identificam o nível *superordenado*, para destacar atributos mais gerais e *subordinados*, com atributos mais específicos e relacionados aos níveis superiores (herança).

Conforme Fellbaum (1998), a hiponímia é a relação mais importante para a construção de uma ontologia, uma vez que é a mais frequente entre os nominais. Pode-se dizer que é uma relação lexical que corresponde à inclusão de uma classe em outra; tal inclusão depende do ponto de vista com que se analisará o significado. Este significado pode ser analisado (a) sob um ponto de vista extensional, isto é, tratando do significado da palavra como um todo (a relação entre as entidades) ou (b) sob um ponto de vista intensional<sup>47</sup>, isto é, tratando da representação do conteúdo interno do significado da palavra.

No caso do nosso trabalho, a relação de hiponímia será abordada sob o ponto de vista extensional, pois para o desenvolvimento de uma ontologia são representadas as relações entre as entidades. Também podemos dizer que a construção de uma ontologia segue o ponto de vista onomasiológico, ou seja, o conceito é o ponto de partida.

A construção da ontologia proposta terá como base os nominais e os verbos (ações dos participantes e eventos) que compõem o léxico do ambiente Moodle. Para identificarmos as relações de hiponímia nos nominais, podemos fazer uso do esquema *X é um Y*, ou *X é um tipo de Y*. Por exemplo: *Chat é uma atividade*. *Chat é*

---

<sup>46</sup> Conforme LYONS (1987).

<sup>47</sup> Extensão e intensão são termos clássicos empregados no âmbito da Semântica. Para melhor explicar tais termos, apresentamos a definição do dicionário Oxford “A extensão de um predicado é a classe dos objetos que ele descreve: a extensão de ‘vermelho’ é a classe das coisas vermelhas. A intensão é o princípio pelo qual os objetos são escolhidos ou, em outras palavras, a condição que um objeto tem de satisfazer para ser corretamente descrito pelo predicado. Dois predicados (‘... é um animal racional’, ‘...é um bípede sem penas’) podem identificar uma mesma classe, mas o fazem por meio de uma condição diferente. [...] Um predicado ou qualquer outro termo ocorre num contexto extensional, numa frase, se puder ser substituído por outro predicado ou termo com a mesma extensão sem que o valor de verdade da frase sofra alteração: se João é um animal racional, e substituímos ‘é um animal racional’ pelo predicado co-extensivo ‘é um bípede sem penas’, então João é um bípede sem penas. Outros contextos, como ‘Maria acredita que João é um animal racional’, podem não permitir a substituição, sendo denominados contextos intencionais”.

um tipo de atividade. Porém, para os verbos há uma importante discussão a ser apresentada. Miller e Fellbaum (1991) acreditam que não é possível aplicar esta regra lógica quando buscamos relações entre os verbos e justificam seu posicionamento afirmando que as características que diferenciam dois verbos superordenados são diferentes das que diferenciam dois nomes; por esta razão denominam esta relação entre verbos de *troponímia*. No entanto, outros autores, como Cruse (2000) e Vossen (1997), não fazem tal distinção, apesar de considerarem as diferenças existentes entre uma taxonomia verbal e uma nominal, e mantêm a mesma nomenclatura. No caso do nosso estudo, seguiremos a posição de Cruse (2000) e Vossen (1997) quando se mostrar necessário estabelecer este tipo de relação hierárquica entre os verbos.

Como um exemplo para a relação de hiponímia, podemos apresentar o seguinte esquema:

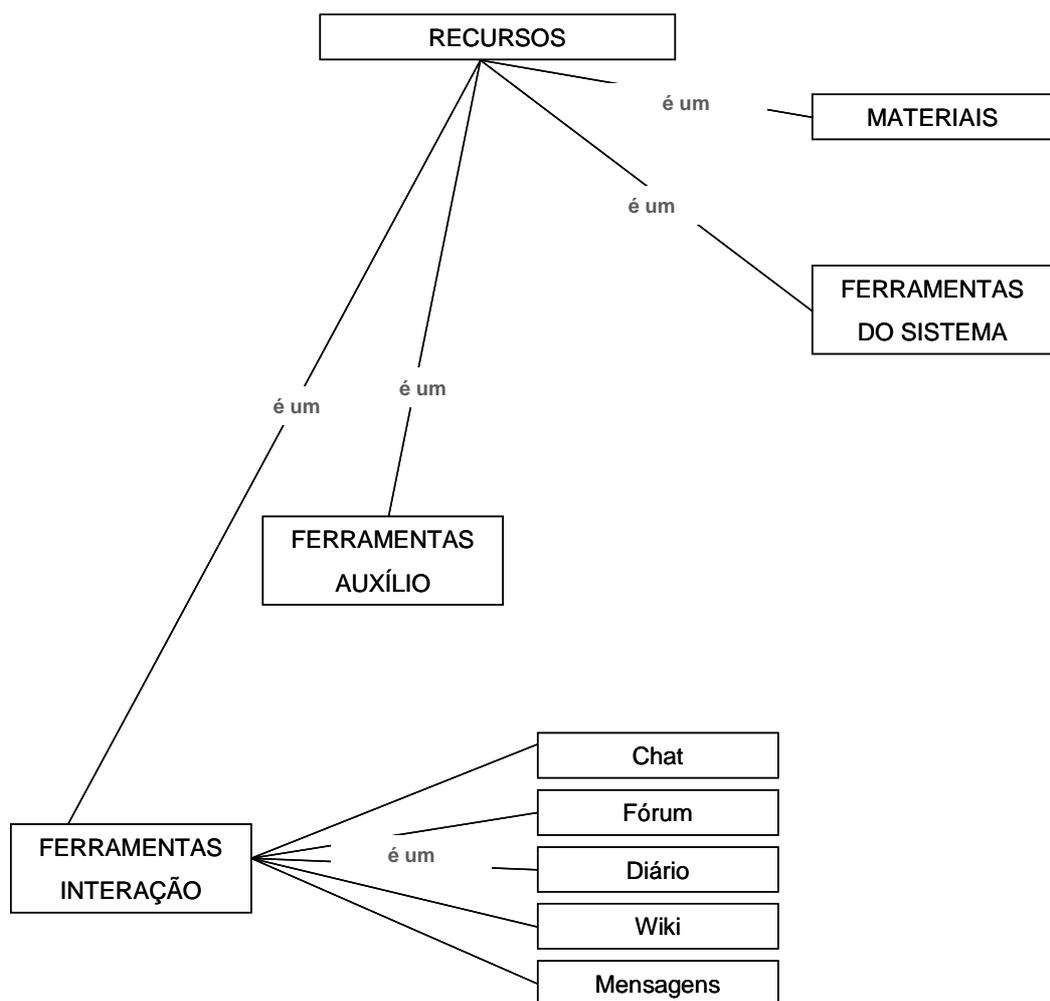


Figura 13: Relação de hiponímia

Na figura 13 temos a representação da classe <recursos>, da qual derivam as subclasses <ferramentas\_interação>, <ferramentas\_auxílio>, <ferramentas\_do\_sistema> e <materiais>, sendo que a primeira apresenta mais outras subclasses.

Neste caso, temos a relação de hiponímia ao estabelecermos que <chat> é uma <ferramenta\_interação> e <ferramenta\_interação> é um <recurso>, por exemplo.

### **b) Meronímia/holonímia**

A meronímia é igualmente uma relação lexical caracterizada pelo traço de inclusão. No entanto, difere da hiponímia, pois estabelece uma relação hierárquica em que Y é parte de um outro objeto A. Dessa forma, trata-se de uma relação parte-todo, sendo caracterizada, dentro da hierarquia, pelo verbo *ter*. Para identificar a meronímia, pode-se fazer uso do esquema *X é uma parte de Y; logo, Y tem um X*<sup>48</sup>. Por exemplo: Um tópico *é parte de* um fórum. Um fórum *tem* tópicos.

Já a holonímia caracteriza-se pelas relações contrárias, isto é, do todo para as partes.

Winston et.al. (1987) consideram a relação de meronímia como de ordenação parcial e, por esta razão, propõem uma tipologia de seis categorias de merônimos: componente-objeto integral, membro-coleção, porção-massa, material-objeto, elemento-atividade e lugar-área, baseando a distinção em elementos de natureza relacional. A seguir, apresentamos uma breve descrição para cada categoria e alguns exemplos:

---

<sup>48</sup> Cf. CRUSE (1986).

- (a) *Componente-objeto integral*: Há uma clara relação estrutural e funcional entre o todo e suas partes. Exemplos: <curso> contém <módulos>. <questionário> contém <perguntas>;
- (b) *Membro-coleção*: As partes não necessariamente têm uma relação estrutural ou funcional em relação ao todo e são distintas umas das outras. Exemplo: <exercício> é-*parte-de* <tarefa>;
- (c) *Material-objeto*: Esse tipo de relação descreve o material com o qual um objeto é construído, criado ou elementos constitutivos e um objeto. Exemplo: <fórum> contém <tópicos>;
- (d) *Porção-massa*: Há completa similaridade entre as partes e seus todos; o limite entre as partes são arbitrários, e as partes não têm qualquer função específica *a priori* no que se refere a seus todos. Exemplo: <perguntas> são-*parte-de* <questionário>;
- (e) *Lugar-área*: A parte não contribui de fato para o todo de maneira funcional; o lugar existe sem a divisão em localidades. Exemplo: <Unisinos Virtual> possui <tipos de cursos>, onde <tipos de cursos> não interfere de maneira funcional no domínio <Unisinos Virtual>;
- (f) *Elemento-atividade*: Descreve as diferentes subatividades que formam uma atividade de maneira estruturada, por exemplo, de maneira temporalmente organizada. Exemplo: Planejar aulas é uma *subatividade* de lecionar.

Esta caracterização permite afirmar que a meronímia não implica necessariamente um processo de inclusão entre classes, mas a conexão entre dois elementos que estão reciprocamente implicados, assim X está implicado no sentido de Y. Por exemplo: O fórum *tem* tópicos, o estudante *postou* um tópico no fórum (relação de elemento-atividade). Esse aspecto fica mais claro com as palavras de Silva (2003, p. 6):

Há, por este fato, uma espécie de inclusão entre a entidade que sofre a divisão e o resultado que dela decorre, não estando, no entanto, as propriedades do todo obrigatoriamente incluídas nas suas partes, dado que uma parte não é semanticamente idêntica ao sentido do todo. Por este motivo, a meronímia é assimétrica e frequentemente intransitiva.

Lyons (1977) já tratava da meronímia e apontava problemas de transitividade, conforme a citação acima. De acordo com o autor, este problema pode ser devido ao fato de existirem muitos tipos de relação parte/todo. Fato este que pode ser comprovado nos estudos realizados por Gomes et.al. (2008). As autoras apresentam em seu artigo uma vasta revisão bibliográfica sobre as relações conceituais que podem ser encontradas em trabalhos sobre ontologias na área da CI, sendo grande parte delas relações de parte/todo, isto é, relações meronímicas.

Marrafa (2001) também apresenta diferentes maneiras de representar a meronímia. Para a autora, responsável pela construção da WordNetPT, existem cinco tipos de relações parte-todo, a saber:

- (a) Mero-parte: corresponde à relação parte-todo típica. Por exemplo: o tópico é *mero-parte* de fórum;
- (b) Mero-membro: este subtipo de meronímia expressa a relação típica de indivíduo e um grupo. Por exemplo: Paula é *mero-membro* de comunidade;
- (c) Mero-porção: uma relação de meronímia atípica, “na medida em que o todo é sempre existente à porção, as fronteiras da porção não são definidas e, a um nível muito produtivo, não é lexicalizada” (MARRAFA, 2001, p. 40). Por exemplo: gota é *mero-porção* de líquido;
- (d) Mero-matéria: as entidades podem ser relacionadas àquilo de que são constituídas. Por exemplo: madeira é *mero-porção* de cadeira;
- (e) Mero-localização: permite o estabelecimento de relações topológicas, ou seja, denotando os nomes em relação ao espaço que ocupam. Por exemplo: centro é *mero-localização* de cidade.

A meronímia torna-se necessária na estrutura relacional da ontologia, uma vez que é fundamental descrever as relações que ocorrem entre as partes e o todo.

### **c) Sinonímia/antonímia**

É muito comum a sinonímia estar relacionada a uma relação de igualdade; porém, este pressuposto já foi contestado por diversos estudos. Para justificar tal afirmativa, apoiamo-nos nos estudos realizados por Lyons (1977) e Cruse (1986, 2000), que apregoam a existência de diferentes condições (ou graus) de sinonímia. De acordo com os estudiosos, poderíamos classificar a sinonímia em:

- sinonímia absoluta (ponto questionado por outros autores, entre eles Lyons, 1977);
- sinonímia proposicional;
- quase sinonímia.

Cruse (2000) traz maiores esclarecimentos sobre cada grau de sinonímia apresentado acima. Segundo o autor, a sinonímia absoluta está relacionada à total igualdade de significado entre duas unidades lexicais, independentemente do contexto. Lyons (1977) compreende que este tipo de sinonímia é raro, ou quase inexistente, uma vez que seria preciso atender às severas exigências abaixo:

- todos os significados dos itens lexicais envolvidos devem ser idênticos;
- os itens lexicais devem ser sinônimos em todos os contextos;
- as expressões têm que ser semanticamente equivalentes em todas as dimensões do significado.

Quando dois itens lexicais podem ser substituídos em qualquer conjectura, sem que haja alteração na sentença, ocorre a sinonímia proposicional; porém, é possível que ocorram situações de diferenciação na expressividade do significado,

ou diferentes formas de expressão (coloquial ou formal), ou, ainda, diferenças de campo pressuposto de discurso.

Por fim, Cruse (2000) explica a quase-sinonímia como um grau de semelhança difícil de ser identificado e sugere que sejam utilizadas as seguintes estratégias para identificar tal situação:

- consideração da percepção do falante;
- não há correspondência entre proximidade semântica e grau de sinonímia; portanto, não se pode basear a análise somente pela distância semântica do significado;

Assim sendo, Cruse (2000) descreve os sinônimos como palavras cujas características comuns são mais proeminentes que as diferenças.

O problema da sinonímia absoluta é discutido também por Miller e Fellbaum (1991), que baseiam suas suposições na ideia de que a sinonímia precisa ser relativa ao contexto; assim, dois itens lexicais ou expressões podem ser sinônimas se, ao serem substituídos, não ocorrer alteração na condição de verdade da sentença.

A norma ANSI/NISO (2005, p. 43) trata da sinonímia como relações de equivalência.

Quando o mesmo conceito pode ser expresso por dois ou mais termos, um deles é selecionado como termo preferido. O relacionamento entre termos preferidos e não preferidos é uma relação de equivalência, em que cada termo é considerado como uma referência ao mesmo conceito (tradução nossa).

A sinonímia mostra-se útil para a nossa ontologia, pois o agente inteligente deverá ser capaz de reconhecer diferentes itens lexicais e associá-los a uma mesma classe, como, por exemplo, o item lexical *professor*, que, no nosso contexto, é sinônimo de *docente*. Trataremos esse tipo de relação usando o termo “equivalência”, a exemplo da norma ANSI/NISO (2005).

A antonímia é uma relação de exclusão, envolvendo conceitos opostos, como, por exemplo, *aumentar* e *diminuir*, *quente* e *frio*, *alto* e *baixo*. Limitamo-nos a esta breve definição, pois esta relação não será relevante na modelagem da M\_ONTO. Vale acrescentar que a antonímia não é citada na norma ANSI/NISO (2005) como um tipo de relação a ser incluída nos vocabulários controlados.

As relações acima apresentadas são caracterizadas como paradigmáticas. Na nossa ontologia também serão incluídas as relações sintagmáticas, sendo que tais relações estão ligadas à abordagem cognitiva, sobre a qual trataremos na seção seguinte.

### 3.3.2 Relações sintagmáticas

Na seção anterior detalhamos as relações paradigmáticas e as destacamos como fundamentais na construção da M\_ONTO. No entanto, apesar deste tipo de relação representar-se como o “corpo” da ontologia, pois designa a estrutura hierárquica da mesma, sabemos, com base nos estudos da Computação, da Linguística e da CI, que é insuficiente para a descrição completa dos conceitos envolvidos no domínio EAD, mais precisamente do ambiente Moodle. Por esta razão, dedicamos a atual seção para tecermos considerações e apresentarmos de forma exemplificada as relações sintagmáticas.

Conforme salienta Oltramari (2010), os itens inventariados através de relações sintagmáticas ocorrem com frequência, mas não há possibilidade de realizar substituições entre eles, até porque eles podem ser lexicalizados a partir de palavras com categorias sintáticas diferentes.

Na Semântica Lexical, as relações sintagmáticas estão ligadas ao estudo da semântica gramatical com enfoque para os papéis temáticos e a predicação, sendo que, com a crescente evolução dos estudos em PLN, este tipo de relação passou a ser de fundamental importância para a implementação de sistemas inteligentes, uma vez que a parte formal envolvendo as ligações entre classes e subclasses não é suficiente.

Neste estudo seguiremos uma abordagem semântica que enfatiza a descrição das situações e seus participantes; porém, vale esclarecer que existem diferentes formas para descrever tais relações que estão relacionadas diretamente aos recursos computacionais. Chishman e Alves (2005) destacam três recursos como os mais representativos:

- os modelos relacionais;
- os modelos baseados na noção de predicação e papéis temáticos;
- os modelos baseados em frames.

De acordo com as autoras, os modelos relacionais são importantes pontos de partida para a organização de grandes bases de dados, tais como a WordNet<sup>49</sup> (Princeton) e a EuroWordNet<sup>50</sup>. Tais bases de dados, ao organizarem os itens lexicais a partir de relações semânticas conceituais, não considerando a ordem alfabética, apresentam uma organização que poderia ser comparada à de um léxico mental. As relações paradigmáticas que citamos anteriormente (sinonímia, antonímia, hiponímia e meronímia) também fazem parte do acervo das Wordnets.

Para as ontologias, as informações do tipo relacional são empregadas quando se trata dos nominais; no entanto, para a análise dos verbos são insuficientes. Neste caso, outros modelos tornam-se mais eficazes, tais como o modelo de frames e dos papéis temáticos. No caso do nosso estudo, o modelo de frames apresenta-se mais adequado, mas também faremos menção ao modelo de papéis temáticos, uma vez que o mesmo será útil na identificação das relações referentes aos verbos em se tratando da organização em classes semânticas.

---

<sup>49</sup> Disponível em: <<http://wordnet.princeton.edu/>>. Acesso em: 02/04/2009.

<sup>50</sup> Disponível em: <<http://www.illc.uva.nl/EuroWordNet/>>. Acesso em: 02/04/2009.

### a) Papéis Temáticos

Os modelos de papéis temáticos baseiam-se nos participantes e nas situações nas quais eles podem estar envolvidos, explorando a ligação entre a sintaxe e a semântica em um nível sentencial. Em se tratando de verbos, entidades que funcionam como organizadoras da sentença, este tipo de descrição torna-se especialmente rico, uma vez que possibilita a descrição do papel que cada participante pode ter em situações específicas.

Saeed (1997)<sup>51</sup> utiliza o termo *situação* ao referir-se à relação sintático-semântica formada entre um verbo (elemento predicador) e seus argumentos externos e internos (participantes).

Chishman e Alves (2005) apresentam uma revisão da literatura desta área e defendem que existem diferentes formas de descrição das *situações*. São citados os estudos realizados por Vendler (1969) sobre a classificação aspectual – verbos de atividade, estado, accomplishment e achievement<sup>52</sup> - que serviram de inspiração para vários trabalhos posteriores, como é o caso dos estudos de Dowty (1979), Van Valin (1997) e Pustejovsky (1995)<sup>53</sup>. Segundo as autoras,

ainda que tenhamos à disposição diferentes abordagens para o estudo dos verbos sob essa ótica, havendo notável distinção terminológica, todas compartilham de um mesmo pressuposto teórico: a centralidade no verbo. Eles sustentam que é o verbo que dita a presença e a natureza do nome, e não o contrário (Chishman e Alves, 2005, p. 49).

---

<sup>51</sup> Apud Chishman e Alves (2005).

<sup>52</sup> Optamos pela não tradução destes verbos, seguindo a linha de raciocínio adotada por Chishman e Alves (2005). As autoras não realizam a tradução devido à falta de consenso entre os demais autores brasileiros sobre a tradução destes termos e, por esta razão, são referidos os termos originais apresentados por Vendler (1967).

<sup>53</sup> Apud Chishman e Alves (2005)

As autoras citam Borba (1996, 2002) e a forma de classificação proposta na obra *Uma gramática de valências para o português* e seu *Dicionário de usos do português do Brasil* para a análise dos verbos no corpus de sua pesquisa.

Nesta pesquisa, consideramos que os estudos de Borba trazem uma boa visão sobre a teoria dos papéis semânticos e seguiremos esta mesma linha, adotando, na medida do possível e com os ajustes necessários, a classificação proposta pelo autor. Para Borba (1996), o resultado da associação entre um verbo e um nome é uma classe e um caso/papel, respectivamente.

Borba (1996) organiza os verbos em quatro classes semânticas: verbos de ação, de processo, de ação-processo e de estado. Criamos um esquema para apresentar as principais características de cada classe semântica dos verbos proposta pelo autor, exemplificando-as com sentenças referentes ao domínio de nossa ontologia.

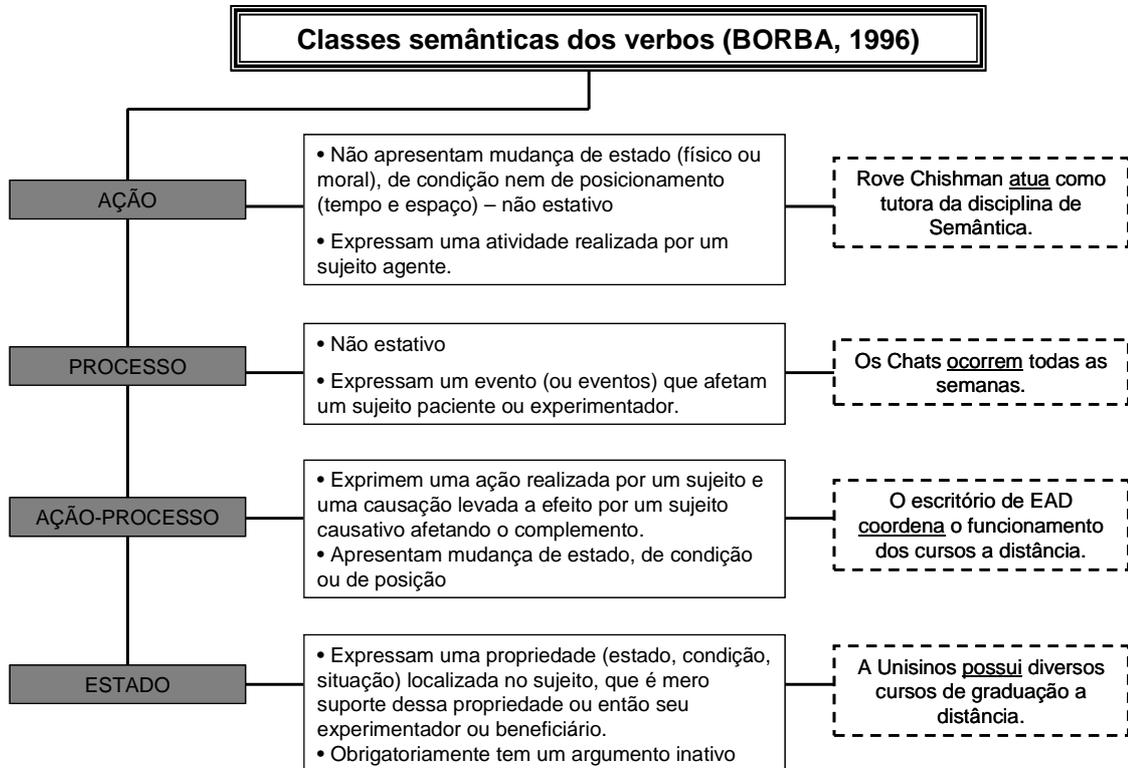


Figura 14: Classes semânticas dos verbos (Borba, 1996)

A definição de papéis semânticos segundo a visão de Borba (1996) é apresentada por Chishman e Alves (2005, p. 50):

O autor define papéis semânticos como noções relacionais que se apresentam como configurações estruturais, com estatuto comparável aos das noções de sujeito e objeto em muitas teorias gramaticais. Como noções relacionais, as relações temáticas (ou semânticas) representam um sistema de casos ou gramática de casos, sendo caso definido como a atuação dos argumentos na predicação.

A partir dessa visão e seguindo ideias apresentadas por Fillmore (1971), Borba (1996) lista doze casos de papéis semânticos e apresenta sua caracterização com base em traços semânticos. Na tabela 4, listamos os casos adotados pelo autor e suas caracterizações<sup>54</sup>.

<b>Papéis semânticos</b>	<b>Características</b>
Agentivo	O que age ou faz; desencadeia uma atividade, sendo origem dela e seu controlador.
Beneficiário	O que se beneficia de
Locativo	O que localiza
Experimentador	Traduz uma experiência ou uma disposição mental.
Objetivo	Mais neutro, é o afetado por aquilo que o verbo indica.
Instrumental	Exprime uma causa direta
Causativo	Provoca um efeito ou desencadeia algo; expressa uma atividade ligada a um estímulo.
Meta	Contém os traços "afetado" e "transição" e expressa o ponto de partida.
Origem	Contém os traços "afetado" e "transição" e expressa o ponto de chegada.
Resultativo	É um efetuado, liga-se a verbos de existência.
Temporal	Indica localização no tempo
Comitativo	Indica associação, é sempre um afetado

Tabela 4: Casos de papéis semânticos

Diante do apresentado, parece-nos que há uma lista única de papéis semânticos e que estes são imutáveis; porém, esta teoria tem sido constantemente

---

<sup>54</sup> Apud Chishman e Alves (2005).

reformulada, sendo que não há um conjunto finito de papéis semânticos. Kearns (2000) afirma que há variação nos papéis semânticos, sendo que alguns têm sido reformulados, subdivididos em outros tipos ou até mesmo abandonados.

Consideramos importante salientar que não há um conjunto de papéis semânticos fechado e estático ideal para a descrição das relações de sentido estabelecidas entre predicador e argumentos. Este aspecto dificulta a aplicação da proposta de Borba na sua totalidade, além de trazer certa insegurança na identificação adequada dos casos em nosso corpus. Por esta razão, buscamos subsídios nos *frames* a fim de definirmos os papéis de forma mais refinada e precisa.

Nossa discussão teórica apresentada até aqui concentrou-se no âmbito lógico-conceitual e gramatical. Conforme referido anteriormente, este nível é fundamental para a ontologia; porém, não é suficiente, pois há necessidade de uma avaliação de aspectos contextuais. Para isso, apresentaremos a Semântica de Frames, enfatizando a noção de *frame*, que cumprirá o papel de abarcar os aspectos contextuais em nossa ontologia. Ao recorrermos aos *frames*, buscamos a possibilidade de incluir, de forma estruturada, informações de cunho enciclopédico, ultrapassando, assim, os limites da Semântica Lexical e adentrando no terreno cognitivo. Tal estratégia nos propiciará uma investigação mais ampla, o que não seria possível adotando os papéis temáticos.

## **b) Semântica de Frames**

A Semântica de *Frames* tem origem nos estudos de Fillmore (1982) e considera fatores culturais e situacionais para descrever a estrutura cognitiva de um evento, ou seja, considera o chamado conhecimento enciclopédico, avaliando como o conhecimento geral do falante reflete na forma como ele interpreta o mundo e como compreende o significado das palavras. Tem como base a multiplicidade de significados atribuídos a uma palavra, de acordo com a experiência humana. Esta abordagem para a construção dos significados e do conhecimento insere-se no âmbito da LC, conforme Croft e Cruse (2004). Segundo os autores, “Fillmore utiliza-

se de uma ampla gama de exemplos para demonstrar que existem fenômenos significativos da semântica, os quais não podem ser facilmente capturados em um modelo de semântica estrutural e traços semânticos” (CROFT e CRUSE, 2004, p. 8).

Conforme Geeraerts (2003), a Semântica de Frames baseia-se no pressuposto de que o conhecimento humano é organizado em um todo, não através de conceitos isolados. Para o autor, os conceitos se estruturam internamente em conjuntos, por ele chamados de *cenias*, que compreendem crenças, experiências ou imaginações.

O termo *cenias* equivale ao que Minsky (1974), na área da IA, e Fillmore (1977), na área de Linguística, chamaram de *frame*. Para Fillmore (1982), um *frame*, sob o ponto de vista linguístico-pragmático, considera o conhecimento de mundo em sua estrutura de representação, tornando-se uma visão mais ampla. Diante disso, estruturar o conhecimento através de *frames* é um aspecto fundamental para o entendimento do significado de uma palavra na Semântica de Frames.

Trabalhos em Linguística Computacional seguem a mesma linha de Beaugrande e Dressler (1998), que consideram que os *frames* são padrões que englobam conhecimento do senso comum com base em algum conceito central (por exemplo, uma aula) e que as entidades aparecem juntas, mas não é definida a ordem em que tais situações são mencionadas.

Na área da Computação, mais especificamente na IA, o termo *frame* também é utilizado para referir-se à forma como dados se estruturam de modo a representar uma determinada situação, ou seja, compreendem um conjunto de informações sobre uma situação, que pode ser organizada através de propriedades (*slots*) que caracterizam cada circunstância (MINSKY, 1974). Considerando os aspectos formais, a IA compreende *frame* como uma estrutura formada por *atributos*, *valores*, *relações* e *restrições* sobre os elementos envolvidos em cada situação. Para a Computação, em seus trabalhos voltados para a IA, os *frames* consistem em um conjunto de nós ou relações que representam características prototípicas de objetos ou situações, podendo ser alteradas de acordo com mudanças no mundo real, uma vez que eles contêm informações universais e particulares de uma comunidade (MINSKY, 1974).

Para Minsky (1974), que discorda da posição de representação do conhecimento através de fórmulas lógicas e defende que haja uma estrutura de dados para armazenar as informações relevantes sobre cada entidade, o significado deve ser organizado pelo meio de um conjunto de dados estruturados que representem situações típicas. Tais estruturas nada mais são que os *frames*. Os pressupostos do autor servem como base para diversos trabalhos na área da organização do conhecimento, sendo que o cerne de seus estudos concentra-se na hipótese de que a memória humana possui muitas propriedades típicas de situações ou objetos (estereótipos). Por exemplo: quando alguém fala em *estudar*, pode-se diretamente ligar este item lexical a *aula, escola, professor* e até mesmo *livros*.

Minsky (1974) apresenta uma ligação dos *frames* ao PLN, na medida em que eles podem trazer mais informações sobre como deve ser utilizado (compreendido) certo conceito, prevendo uma série de ações ou expectativas, isto é, em um *frame* pode-se ter muitas informações que vão além da situação, advindas das representações do conhecimento do falante. Neste contexto, as palavras são componentes que compartilham *frames*, mas não estão necessariamente inter-relacionadas como na abordagem relacional.

Sob uma abordagem linguística, Fillmore (1982) considera *frame* uma construção intuitiva que possibilita a formalização de relações entre semântica e sintaxe como decorrência de uma análise lexical. Também poderíamos dizer que um *frame* é uma estrutura de conhecimento que ocorre a partir das interações refletidas pela linguagem.

Considerando as ideias apresentadas anteriormente, podemos dizer que um *frame* é uma estrutura que leva em consideração o conhecimento enciclopédico (de mundo) e que permite observar as relações (sintáticas e semânticas) manifestadas na língua.

Estudos sobre Semântica de Frames inspiraram a criação do projeto FrameNet (FN), sediado no Instituto Internacional da Ciência da Computação de Berkeley. A FN consiste em um repositório com mais de dez mil unidades lexicais, organizadas em cerca de oitocentos frames e mais de cento e vinte mil exemplos para o inglês. Trata-se de um recurso lexicográfico que identifica e descreve *frames*

*semânticos*, analisando o significado das palavras e estabelecendo ligações com outros *frames* cujos significados estejam implícitos através de propriedades sintáticas e semânticas (FILLMORE, 1982).

Para Fillmore (1982), *frame semântico* é uma representação em forma de esquema de uma situação que envolve vários participantes e papéis conceituais onde cada uma das noções representa um *elemento de frame*, este, por sua vez, corresponde a uma categoria da FrameNet (versão computacional a partir da Semântica de Frames). Assim, podemos concluir que cada argumento semântico corresponde a um *elemento de frame* do *frame semântico* ao qual a palavra está associada (FILLMORE, WOOTERS e BAKER, 2001; JOHNSON e FILLMORE, 2000; PETRUCK, 1996)<sup>55</sup>.

A FrameNet tem como objetivo principal identificar padrões de valência para verbos, substantivos, adjetivos, advérbios e preposições (somente na língua inglesa), através de anotações e da exemplificação com sentenças reais.

Na FrameNet um *frame* descreve uma situação típica de uma determinada língua, levando em consideração os aspectos culturais a ela relacionados e incluindo os participantes e suas condições. Cada *frame*, como uma categoria cognitiva, manifesta-se na língua por meio de palavras que o introduzem, isto é, *evocam* o *frame*.

Neste recurso lexical, os *elementos de frame* das situações podem ser organizados em nucleares (*Core*) e não nucleares (*Non-Core*), noções estas que estão fortemente ligadas à visão dos papéis semânticos que ocupam posições argumentais. Os nucleares são os *elementos frame* fundamentais para a caracterização do *frame* e se manifestam até mesmo na estrutura argumental traduzida pelo verbo em questão (predicador). Já os não nucleares não são tão importantes na conceituação, uma vez que não caracterizam o *frame* sozinhos.

---

<sup>55</sup> Disponível em: <<http://www.icsi.berkeley.edu/framenet/>>. Acesso em: 2010

Vejamos o exemplo do frame *EDUCATION\_TEACHING* (ensino)<sup>56</sup>, evocado pelo verbo *studying* (estudar). Na figura 15, aparecem os elementos nucleares (core): curso, fato, instituição, material, preceito, qualificação, papel, ação, estudante, assunto, professores – tais elementos são fundamentais para a caracterização da cena de ensino. O frame apresenta também os elementos não nucleares: grau, descrição, duração, nível, maneira, meios, local, efeito, resultado e tempo – elementos que contribuem para a caracterização da situação, mas não são relevantes; todos apresentando uma explicação e/ou exemplo. Através destes elementos, especialmente os nucleares com seus exemplos, podemos obter maiores subsídios para a caracterização dos papéis semânticos de cada participante e o estabelecimento de suas relações.

Core:	
<b>Course [cou]</b>	FN: a program of lectures or other matter dealing with a subject
<b>Fact [fac]</b>	A piece of information that the <b>Student</b> is informed of by the <b>Teacher</b> .
Excludes: Precept	When I was two she <b>TAUGHT</b> me <b>that a quarter was 25 cents and a dime was 10 cents</b> .
<b>Institution [insti]</b>	An educational establishment, such as a school or college.
<b>Material [mat]</b>	Educational <b>Material</b> , such as books, tapes, or videos, used by a <b>Teacher</b> or a <b>Student</b> to acquire skills or knowledge.
	All employees have received <b>Books</b> <b>INSTRUCTING</b> them on how to report: anything from unwanted touching and uncalled-for blue jokes to overt demands for sexual favours. *
<b>Precept [pre]</b>	A guideline for correct behavior. In most cases, this concerns morally or socially desirable actions.
Excludes: Skill	My brother <b>TAUGHT</b> me <b>not to be mean to guys when I'm not interested in them</b> .
<b>Qualification [Qual]</b>	A formal Qualification such as an academic degree or a certificate for which a student is aiming.
Excludes: Fact	She's <b>STUDYING</b> for a <b>Master's degree</b> .
	She's an <b>MR STUDENT</b> .
<b>Role [rol]</b>	A role, typically professional or vocational, that the <b>Student</b> is meant to be able to fill as a result of their training.
	When his father died he went to live with an uncle in Kau who <b>TRAINED</b> him <b>as a scientist</b> during his teenage years.
<b>Skill [ski]</b>	An action which the <b>Student</b> is able to perform as a result of instruction.
Excludes: Subject	Taut with worry, Dad <b>TAUGHT</b> us <b>how to drive</b> .
<b>Student [stu]</b>	One who is instructed by a <b>Teacher</b> in skills or knowledge.
Semantic Type: Sentient	
<b>Subject [sub]</b>	The area of knowledge or skill which is taught by a <b>Teacher</b> or to a <b>Student</b> .
Excludes: Precept	She's a <b>French TEACHER</b> .
<b>Teacher [tea]</b>	One who instructs a <b>Student</b> in some area of knowledge or skill.
Semantic Type: Sentient	

Figura 15: Elementos de Frame - *EDUCATION\_TEACHING* (visão parcial)

Uma abordagem baseada em *frames*, que identifica os *elementos de frame* que participam da situação evocada pelo item lexical, permite a inclusão de importantes informações na nossa ontologia, ampliando sua capacidade de

<sup>56</sup> Este frame será apresentado com maiores detalhes no capítulo 3, mais especificamente na seção 4.3.2 quando descrevermos nosso percurso metodológico para a descrição dos relacionamentos elencados como importantes para a M\_ONTO.

representação relacionada ao domínio do ambiente Moodle. Estas propriedades ou *slots* são compatíveis com os papéis semânticos que apresentamos anteriormente.

Na figura 16, podemos ver parte da caracterização do *frame STUDYING* que está relacionado ao *frame EDUCATION\_TEACHING*. Através da descrição do *frame*, podemos contextualizar os participantes da cena de educação e a forma como se relacionam na situação descrita. Mesmo que este *frame* não apresente situações específicas do domínio EAD, como, por exemplo, que as atividades podem ser síncronas<sup>57</sup> e assíncronas<sup>58</sup>, ele é um *frame* útil para nossa descrição.

## Studying

[Lexical Unit Index](#)

### Definition:

A **Student** enrolls and then remains at an **Institution** for the purposes of education within a **Subject**. They may receive instruction from a particular **Teacher** at the **Institution**. The **Place**, **Time**, and **Duration** of the study may also be specified.  
 He spent ten years **STUDYING** physics at the **tyccc** before he flunked out.

**Semantic Type:** Non-perspectivalized\_frame

### FEs:

#### Core:

<b>Institution</b> [Instn]	An educational establishment, such as a school or college.
<b>Student</b> [stu] Semantic Type: Sentient	One who receives instruction from a <b>Teacher</b> or <b>Institution</b> .
<b>Subject</b> [sub]	The area of knowledge or skill which is taught to a <b>Student</b> . She <b>STUDIES</b> astrophysics.
<b>Teacher</b> [tea] Semantic Type: Sentient	One who instructs a <b>Student</b> in some area of knowledge or skill. She <b>STUDIED</b> from him for years.

Figura 16: Elementos de Frame - *STUDYING* (visão parcial)

Diante dos aspectos apresentados nesta seção, destacamos os *frames* como uma abordagem possível para a modelagem da nossa ontologia, sendo que se mostram eficazes no que diz respeito à contextualização e caracterização de cada conceito, ampliando as possibilidades de relacionamentos.

<sup>57</sup> As atividades ocorrem em tempo real, isto é, alunos e professor estão interagindo ao mesmo tempo, como no caso de um chat (bate-papo) através da internet.

<sup>58</sup> As atividades ocorrem em tempos distintos, isto é, não há necessidade de todos estarem conectados ao mesmo tempo, como no caso do fórum, onde as mensagens podem ser postadas em diferentes momentos.

### 3.4 Categorias da M\_ONTO

Na área da CI, as abordagens para a criação e organização de categorias em ontologias são bastante variadas. Na nossa visão, as questões apresentadas pelos teóricos da CI, mais precisamente Dahlberg (1978b) e Ranganathan (1967), podem ser úteis, ainda que mescladas entre si e com a visão linguística, para a categorização dos conceitos em nossa ontologia. Por esta razão, teceremos alguns comentários acerca das visões de dois teóricos da CI, Dahlberg (1978) e Ranganathan (1967), que, apesar de serem referências um tanto antigas, ainda são amplamente citadas em obras sobre ontologias. Salientamos que não temos a pretensão de discorrer profundamente sobre tais teorias nem mesmo analisá-las sob um olhar crítico, uma vez que as mesmas terão uma função secundária na organização das categorias neste trabalho.

A Teoria do Conceito, proposta na área da CI por Dahlberg (1978b)<sup>59</sup>, apresenta um “método de raciocínio analítico-sintético para organizar os conceitos de um domínio em grandes categorias” (CAMPOS et.al, 2010, p. 5). De acordo com esta teoria, a organização das categorias deve partir da análise da definição (etapa analítica) e da busca por características comuns para a criação de grupos (etapa sintética). Para a autora, tais características podem ser divididas em simples e complexas, sendo que as simples referem-se a uma única propriedade, por exemplo, “redondo”, “colorido”, e as complexas a mais de uma característica, por exemplo, “moldado em argila”, “pintado de verde”. O resultado é uma propriedade, pois em ambos os casos trata-se de um material combinado com um processo. A teórica toma como base as categorias Aristotélicas para a criação de categorias simples, afirmando que não há possibilidade de combinação entre elas (DAHLBERG, 1978b).

Ranganathan (1967)<sup>60</sup>, outro importante teórico da área da CI, defende que existe “uma série de princípios classificatórios para a organização de conceitos de

---

<sup>59</sup> Apud Campos, 2010.

<sup>60</sup> Apud Campos, 2010.

um domínio dentro de classes, as quais são estruturadas de acordo com categorias utilizadas para pensar sua organização” (CAMPOS, 2010, p. 5). As categorias propostas por Ranganathan são *personalidade, matéria, energia, espaço e tempo*, sendo referidas pelos estudiosos da CI pelo acrônimo PMEST (*personality, mater, energy, space, time*).

Dahlberg (1978) e Ranganathan (1967) divergem no que se refere às propriedades das categorias. De acordo com Ranganathan, as propriedades são consideradas manifestações de uma categoria, não fazendo parte da mesma. Já Dahlberg (1978) considera as propriedades como sendo o “caráter” da categoria. No entanto, ambos concordam que a organização das categorias se dá independentemente do domínio.

Seguindo a linha de pensamento de Dahlberg (1978) e Ranganathan (1967), podemos citar as ontologias de topo BFO e DOLCE (já referidas no capítulo 1), que definem suas categorias independentemente do domínio.

A BFO organiza suas categorias em *continuanes (continuant)* e *ocorrentes (occurent)*, sendo que estas ainda possuem outras subdivisões. Enquanto que a DOLCE trata como *endurantes (endurants)* e *perdurantes (perdurants)*, criando categorias com as mesmas características da ontologia BFO. Comparando com a proposta de Ranganathan (1967), poderíamos dizer que os endurantes correspondem à categoria de *personalidade* e os perdurantes, à categoria *energia*. Oltramari (2010) complementa afirmando que os endurantes se referem aos objetos/agentes e os perdurantes, aos eventos/processos.

De acordo com os aspectos acima citados, elaboramos nossa estratégia para a criação das categorias que irão compor a M\_ONTO. Tomaremos como base as categorias da ontologia de topo DOLCE e seguiremos o pressuposto de Dahlberg (1978b), organizando as categorias a partir de sua definição e verificando características comuns.

A M\_ONTO será composta por quatro categorias: atores, recursos, eventos e processos, desmembradas das categorias da ontologia fundacional DOLCE.

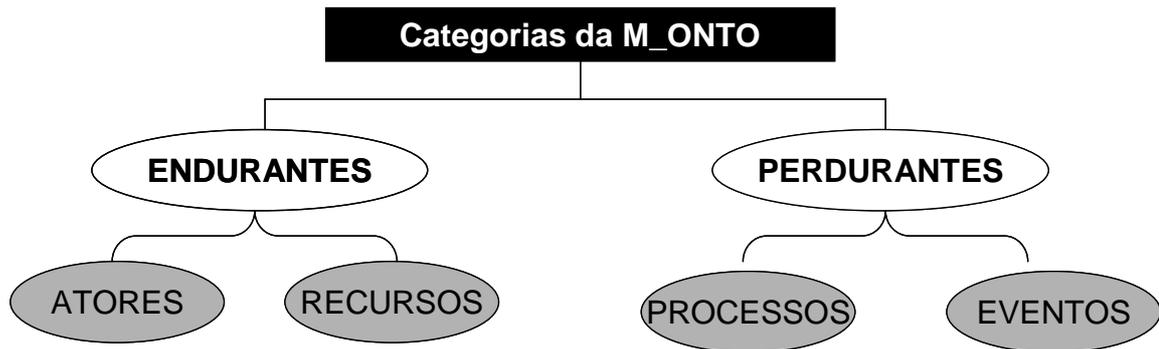


Figura 17: Categorias da M\_ONTO

Ao concluirmos nossas reflexões acerca das relações e das categorias que irão compor a M\_ONTO sintetizamos em uma tabela os aspectos semânticos norteadores de nossa análise.

ENTIDADE – nome da classe ou subclasse	
<b>Definição</b>	Informação importante para maior detalhamento da classe a ser descrita, porém, para o sistema este nível não é compreensível.
<b>Relações de hiponímia</b>	Definem a hierarquia de classes através da relação <i>é-um</i>
<b>Relações de meronímia</b>	Definem a hierarquia de classes através da relação <i>parte-de</i>
<b>Relações de equivalência</b>	Definem termos e classes que podem ser considerados equivalentes.
<b>Frames</b>	Nesse nível, serão incluídas descrições do papel contextual que cada entidade desempenha na situação

Tabela 5: Modelo da Estrutura Ontológica

Encerramos, apresentando um mapa conceitual que tem por objetivo sintetizar as principais ideias aprofundadas neste capítulo.

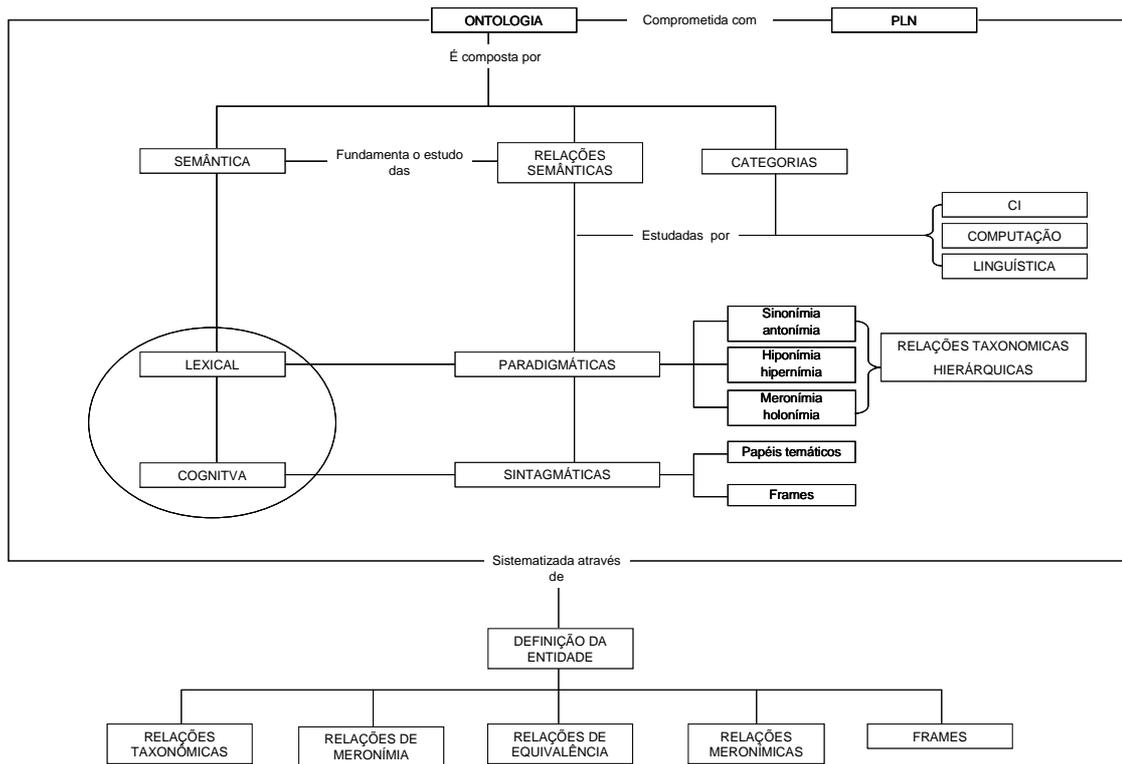


Figura 18: Mapa conceitual – síntese do capítulo 2

No capítulo seguinte, apresentaremos nosso percurso metodológico para a modelagem e implementação da M\_ONTO, através de duas etapas – Linguística e Computacional, enfatizando a caracterização do domínio e da aplicação.

## 4 CONSTRUÇÃO DA M\_ONTO

Nos capítulos anteriores, exploramos amplamente aspectos teóricos sobre a construção de ontologias e sobre o conteúdo das mesmas. Tais discussões ora convergem para a descrição semântica dos itens lexicais selecionados no *corpus* de pesquisa, dados estes que serão descritos semanticamente para a composição da M\_ONTO.

Este capítulo tem o objetivo de descrever o que explicitamos no capítulo 3 acerca das relações paradigmáticas e sintagmáticas e de promover, através deste estudo semântico, a implementação da ontologia por meio de uma ferramenta computacional específica para este fim.

Julgamos importante situar pontos referentes ao domínio e à aplicação da M\_ONTO, uma vez que isso é fundamental para a modelagem consistente da ontologia. Por esta razão dedicamos o início do capítulo para tecer esclarecimentos sobre o Ambiente Moodle e os agentes inteligentes.

Organizamos nossa metodologia para implementação da M\_ONTO em duas fases distintas: uma linguística e outra computacional. A fase linguística tem como objetivo o detalhamento semântico dos itens lexicais com base na Semântica Lexical e Cognitiva, conforme referimos no capítulo 3. E a fase computacional refere-se à modelagem formal da M\_ONTO através do editor de ontologias Protégé 4.1.

Para melhor organizarmos nossa análise, o capítulo foi dividido em quatro seções. Na seção 4.1 dedicamo-nos à caracterização do domínio ao qual nossa ontologia se refere. Aspectos sobre a aplicação da M\_ONTO, mais especificamente os agentes inteligentes, são apresentados na seção 4.2. A seção 4.3 corresponde ao percurso metodológico adotado para a implementação da ontologia, sendo

dividida nas fases linguística e computacional. Por fim, na seção 4.4 apresentamos M\_ONTO como resultado da nossa investigação.

#### 4.1 Caracterização do domínio: Ambiente Moodle

A M\_ONTO é uma ontologia inserida no domínio da EAD, estando especificamente ligada ao ambiente Moodle, uma vez que tem como propósito descrever suas categorias, conceitos e relações. Consideramos que EAD compreende o nível mais amplo do nosso domínio e que o ambiente Moodle, especificamente, é o nosso objeto de análise. Sobre EAD já discorreremos no capítulo 2 e consideramos importante aprofundar nossa discussão sobre o ambiente virtual em questão – Moodle.

O Moodle – *Modular Object-Oriented Dynamic Learning Environmet* – é um sistema de administração de atividades educacionais destinado à criação de comunidades *on-line*<sup>61</sup>, em ambiente virtual voltado para a aprendizagem. Foi desenvolvido pelo australiano Martin Dougiamas, em 1999. De acordo com a informação que consta no site do ambiente<sup>62</sup>:

A palavra Moodle referia-se originalmente ao acróstico: “Modular Object-Oriented Dynamic Learning Environment”, que é especialmente significativo para os programadores e acadêmicos da educação. É também um verbo que descreve o processo de navegar despreziosamente por algo, enquanto se faz outras coisas ao mesmo tempo, num desenvolvimento agradável e conduzido frequentemente pela perspicácia e pela criatividade. Assim, o nome Moodle aplica-se tanto à forma como foi feito, como a uma sugestiva maneira pela qual um estudante ou um professor poderia integrar-se estudando ou ensinando num curso on-line.

Segundo o idealizador e mantenedor do software, Martin Dougiamas, o Moodle tem uma proposta diferenciada, que visa a proporcionar o aprendizado em

---

<sup>61</sup> Significa algo ou alguém que está conectado em tempo real a uma rede de computadores, recebendo e enviando informação.

<sup>62</sup> Disponível em: <<http://moodle.org>>. Acesso em: 21/01/2010.

colaboração em um ambiente *on-line*. O *software* está baseado na pedagogia sócioconstrutivista, que “não só trata a aprendizagem como uma atividade social, mas focaliza a atenção na aprendizagem que acontece enquanto construímos ativamente artefatos (como textos, por exemplo), para que outros vejam ou utilizem”<sup>63</sup>.

Trata-se de um projeto “*open source*” (sob as condições GNU - “*General Public License*”), ou seja, é aberto, de distribuição livre e gratuita, podendo ser baixado, utilizado e modificado pelos usuários. Dessa forma, os usuários podem contribuir melhorando a programação do ambiente e criando *plug-ins*<sup>64</sup> ou agentes para aprimorá-lo. Esta possibilidade permite que o ambiente esteja em constante desenvolvimento e aperfeiçoamento, possibilitando que periodicamente sejam criados novos módulos com funções que atendem aos mais variados tipos de usuários. Além disso, cada vez mais se ampliam as possibilidades de aplicação em diferentes práticas pedagógicas.

É um *software* de fácil utilização e já está disponível em vários idiomas, incluindo o Português. Tecnicamente funciona em diferentes sistemas operacionais<sup>65</sup>, tais como *Unix*, *Linux*, *Windows*, *MacOS* ou em qualquer outro sistema que suporte a linguagem *PHP*<sup>66</sup>. Isso torna possível a sua hospedagem na maioria dos servidores. Necessita de um único banco de dados<sup>67</sup> que pode ser *MySQL*, *PostgreSQL*, *Oracle*, *Access*, *Interbase* ou *ODBC*.

O Moodle é considerado um *Learning Management System* (LMS), ou, traduzindo, um sistema de gestão da aprendizagem, sendo que dispõe de um

---

<sup>63</sup> Palavras de Martin Dougiamas no site de apresentação do Ambiente Moodle. Disponível em: < <http://www.moodle.org.br/>>. Acesso em: 04/12/2009.

<sup>64</sup> Na informática, um **plugin** (também conhecido por *plug-in*, *add-in*, *add-on*) é um programa de computador usado para adicionar funções a outros programas maiores, provendo alguma funcionalidade especial ou muito específica. Geralmente pequeno e leve, é usado somente sob demanda.

<sup>65</sup> Softwares que determinam o funcionamento de um equipamento computadorizado.

<sup>66</sup> PHP (um acrônimo para Hypertext Preprocessor) é uma linguagem de programação de computadores interpretada, livre e muito utilizada para gerar conteúdo dinâmico na Web.

<sup>67</sup> Banco de dados (ou base de dados) é um conjunto de registros dispostos em estrutura regular que possibilita a reorganização dos mesmos e produção de informação. Um banco de dados normalmente agrupa registros utilizáveis para um mesmo fim.

grande conjunto de ferramentas que tem por finalidade o apoio pedagógico, podendo ser selecionadas pelo professor conforme suas necessidades.

A plataforma<sup>68</sup> Moodle, quando utilizada como ferramenta de ensino, apresenta alguns pontos fortes, dentre eles o aumento da motivação dos alunos, a maior facilidade na produção e distribuição dos conteúdos, a possibilidade de compartilhar conteúdos entre instituições, a gestão total do ambiente virtual, a realização de avaliações dos alunos, o suporte tecnológico para disponibilização de conteúdos de acordo com um modelo pedagógico e o design institucional e o controle de acessos. Permite a transmissão e a organização de conteúdos através da criação de cursos e páginas Web, facilitando a comunicação (síncrona<sup>69</sup> ou assíncrona<sup>70</sup>) e contribuindo para melhoria nos níveis de ensino, tanto presencial quanto à distância. Para permitir toda esta interatividade, ambiente disponibiliza diversas ferramentas de apoio à aprendizagem, sendo as mesmas gerenciadas pelo professor. São elas:

- (a) Materiais: consistem em um conjunto de recursos disponíveis como ferramentas de apoio à aprendizagem, sendo postados pelo professor de acordo com as necessidades do conteúdo e do módulo. São eles:
- Textos em diferentes formatos (doc, pdf, páginas web);
  - *Links* para sites ou arquivos com materiais complementares;
  - Recursos de vídeo: vídeo-aulas, videoconferências, animações, simulações ou arquivos de rádio;

---

<sup>68</sup> No contexto da Informática, **plataforma** é o padrão de um processo operacional ou de um computador. É uma expressão utilizada para denominar a tecnologia empregada em determinada infraestrutura de Tecnologia da Informação (TI) ou telecomunicações, garantindo facilidade de integração dos diversos elementos dessa infraestrutura.

<sup>69</sup> Processo de comunicação no qual as mensagens emitidas por uma pessoa são imediatamente recebidas e respondidas por outras pessoas, ocorre em tempo real.

<sup>70</sup> Processo de comunicação no qual mensagem emitida por uma pessoa é recebida e respondida mais tarde pelas outras, não há necessidade de estarem todos conectados na mesma hora.

(b) Ferramentas de interação: são ferramentas que permitem a interação entre os alunos e professores, possibilitando atividades síncronas e assíncronas. São elas:

- *Chat*: para atividades síncronas, através de um bate-papo sobre determinado assunto;
- Fórum: para atividades assíncronas, onde os alunos podem trocar mensagens, links ou documentos anexados, relacionados ao tópico proposto pela atividade;
- Diário: também uma atividade assíncrona, que tem por objetivo o registro de percepções, dificuldades, aprendizagens dos alunos, sendo acompanhado e comentado pelo professor;
- *Wiki*: uma ferramenta de escrita colaborativa, onde todos possuem acesso a um mesmo texto e podem colaborar com novas informações ou com ajustes;
- Mensagens: através das quais os alunos podem se comunicar com os demais por meio de correio eletrônico;

(c) Ferramentas para auxílio da aprendizagem: permitem ao professor criar recursos para complementar suas aulas, tornando-as mais interessantes aos alunos e possibilitando avaliação.

- Glossário: permite a criação de uma lista de definições, como um dicionário, a qual pode ser acessada pelos alunos, tanto para consulta quanto para a criação de novas entradas, sendo que é possível criar *links* entre as acepções dos glossários e outros documentos;
- Lição: com este recurso o conteúdo é apresentado de modo flexível e interessante. Uma lição é composta por um número de páginas, sendo que ao final de cada página encontra-se uma pergunta, de acordo com a resposta o aluno avança ou não na lição;

- Pesquisas de opinião: o professor elabora uma pergunta com diversas opções de respostas. Serve para fazer pesquisas rápidas sobre a opinião sobre algum assunto, a solução para algum tipo de problema, etc;
  - Questionário: ferramenta que permite a construção de diferentes tipos de questionários; múltipla escolha, verdadeiro ou falso, resposta breve, etc. As questões são armazenadas por categorias e podem ser reutilizadas em outros questionários, cursos ou comunidades. Ao criar um questionário é possível configurar o período em que ficará disponível, a apresentação do feedback para o aluno, formas de avaliação e as possibilidades de tentativas de resposta;
- (d) Tarefas: consistem na descrição de uma atividade que deverá ser desenvolvida pelos alunos, por exemplo: projetos, relatórios, apresentações, imagens, produções de textos, etc. as tarefas envolvem criação de texto on-line, envio de arquivo único ou realização de atividade off-line.

Muitas universidades e escolas já utilizam o Moodle, não só para cursos totalmente virtuais, como também para apoio aos presenciais. Também é indicado para outros tipos de atividades que envolvem formação de grupos de estudo, treinamento de professores e até desenvolvimento de projetos. Existem outros setores, não ligados diretamente à educação, que utilizam o Moodle, como, por exemplo, empresas privadas, ONGs e grupos independentes que interagem na Internet.

Os cursos no Moodle podem ser configurados em três formatos, escolhidos de acordo com a atividade educacional a ser desenvolvida. São eles:

- (a) Formato Social, no qual o tema é articulado em torno de um fórum publicado na página principal;
- (b) Formato Semanal, no qual o curso é organizado em semanas, com datas de início e fim;

- (c) Formato em Tópicos, no qual cada assunto a ser discutido representa um tópico que não tem limite de tempo pré-definido.

A UNISINOS, através da Unisinos Virtual, oferece cursos de Graduação, Pós-Graduação e Extensão, além de algumas disciplinas de cursos fundamentalmente presenciais que utilizam o Moodle como plataforma de interação e aprendizagem.

Segundo o Guia do Aluno (Unisinos Virtual. Guia do Aluno, 2008)<sup>71</sup>, disponibilizado pela Unisinos e elaborado pela Unisinos Virtual, a organização do ambiente Moodle da Unisinos se dá através da criação de diferentes comunidades, sendo que geralmente cada turma corresponde a uma comunidade. As comunidades são organizadas dentro dos diferentes cursos oferecidos pela universidade. Os alunos são identificados através do sistema acadêmico da Universidade e não há necessidade de um cadastro prévio, pois o sistema disponibiliza o acesso aos alunos de acordo com as comunidades referentes a sua matrícula, isto é, aquelas às quais o aluno pode ter acesso. A identificação do aluno é feita através do sistema de autenticação do ambiente *Minha Unisinos*, sendo que os alunos devem usar o mesmo nome de usuário e senha previamente cadastrados. Na figura 19, é possível visualizar a forma como as diferentes comunidades aparecem para o acesso aos alunos.

---

<sup>71</sup> Disponível em [http://www.moodle.unisinos.br/file.php/1/Tutoriais/Guia\\_do\\_Aluno\\_Moodle.pdf](http://www.moodle.unisinos.br/file.php/1/Tutoriais/Guia_do_Aluno_Moodle.pdf), acesso em 12/01/2011

Este é o Ambiente Virtual de Aprendizagem onde você tem acesso às disciplinas cursadas através da Mediação Digital Virtual.

Selecione o item "Ajuda" para obter as instruções necessárias e para entrar em contato sempre que desejar.

O item "Acesso" permite a sua identificação junto ao sistema acadêmico da Unisinos.

**Menu Principal**

- Novidades
- Site Unisinos
- Unisinos EAD
- Ajuda
- Tutoriais Moodle
- Vídeo de apresentação

**Meus cursos**

EAD - GRAD - Introdução a Educação a Distância - T2 - 1º Bim - 2010/2	Introdução a Educação a Distância - T2 - 1º Bim - 2010/2
EAD - GRAD - Introdução a Educação a Distância - T4 - 1º Bim - 2010/2	Introdução a Educação a Distância - T4 - 1º Bim - 2010/2
EAD - GRAD - Introdução a Educação a Distância - T1 - 1º Bim - 2010/2	Introdução a Educação a Distância - T1 - 1º Bim - 2010/2
EAD - GRAD - Introdução a Educação a Distância - T3 - 1º Bim - 2010/2	Introdução a Educação a Distância - T3 - 1º Bim - 2010/2

Buscar cursos:

**Categorias de Cursos**

Administrativo - Unisinos Virtual	103
Capacitação	22
GRUPOS DE TRABALHO TEMPORÁRIOS	27
EDUCAÇÃO CONTINUADA	
Unilinguas	21
Ext 2008	3

Figura 19: Ambiente Moodle – categorias de cursos e comunidades

Conforme consta no referido manual, o Moodle conta com *atores* e *recursos*. Os *atores* são aqueles que interagem com o ambiente, isto é, aluno, professor, tutor e monitor, sendo que cada um tem níveis diferentes de acesso e interação. Os **alunos** têm acesso aos materiais publicados nas comunidades às quais estão ligados e interagem com os colegas realizando as atividades previstas em cada módulo. O **professor** é o responsável pela metodologia de trabalho, organização dos conteúdos e atividades e pela avaliação. Normalmente alunos e professor interagem, sendo que o professor tem a função de sanar dúvidas e atender às dificuldades dos alunos. Em alguns casos, dependendo da necessidade da comunidade, ela poderá contar com a presença de um **tutor** ou de um **monitor**, ambos têm a função auxiliar os alunos em caso de dúvidas e auxiliar os alunos, sendo que o primeiro também interage com o grupo e o segundo atende dificuldades de forma mais individual (Unisinos Virtual. Guia do Aluno, 2008).

Figura 20: Interface do ambiente Moodle

O Moodle apresenta uma interface<sup>72</sup> simples e de fácil acesso aos usuários, seguindo o conceito de portal<sup>73</sup>, como pode ser visualizado na figura 20. As páginas dos cursos, ou comunidades, caso da UNISINOS, podem ser divididas em três colunas, permitindo que o professor as personalize de acordo com seus interesses e necessidades do curso, inserindo calendário, usuários on-line, lista de atividades e outros recursos. Na coluna central, são colocadas as atividades e os conteúdos do curso, através de uma sequência de tópicos, sendo que este espaço também pode ser personalizado pelo professor. Este é o espaço onde o curso efetivamente é realizado.

Conforme referimos anteriormente, o Moodle é uma ferramenta livre, isto é, permite que programadores tenham acesso ao código de programação e façam alterações para melhorias no sistema. Assim, é possível criar novos módulos ou agregar agentes inteligentes ao ambiente. Na seção seguinte faremos uma breve explanação com o intuito de apresentar e caracterizar o agente que fará uso da M\_ONTO.

<sup>72</sup> Podemos entender interface neste contexto como a forma de comunicação entre o usuário e o aplicativo, também relativo ao visual ou layout do sistema.

<sup>73</sup> Site que pretende ser uma experiência completa para o usuário, oferecendo vários tipos de conteúdo e serviços. O UOL e o Terra são exemplos de portais.

## 4.2 Caracterização da aplicação: o agente

Nesta seção, apresentamos brevemente questões relativas aos agentes de software, sem, no entanto, pretendermos realizar uma discussão detalhada acerca do assunto, uma vez que não é o foco principal desta pesquisa. O agente de software que fará uso da M\_ONTO é um trabalho a ser desenvolvido futuramente; por esta razão, apresentamos questões gerais sobre este tipo de aplicação.

Os ambientes virtuais de aprendizagem, em função do distanciamento físico que ocorre entre seus participantes, devem buscar formas de minimizar tal dificuldade procurando maneiras diferenciadas de interação, através de mecanismos mais eficientes de adaptabilidade capazes de auxiliar o aluno na solução de problemas. Uma das soluções encontradas é a criação de agentes de *software* ou agentes pedagógicos que visam a auxiliar os alunos no processo de aprendizagem e possibilitam uma maior interação tornando o ambiente menos impessoal.

Um agente de *software* pode ser definido como uma entidade competente para a execução de uma determinada tarefa fazendo uso de informações colhidas a partir do próprio ambiente no qual está inserido (JAQUES, 2001). Este tipo de agente apresenta as seguintes qualidades: autonomia, habilidade social, proatividade, persistência, reatividade, continuidade temporal, aprendizagem e flexibilidade.

Outro ponto que julgamos importante apresentar são os agentes pedagógicos animados. Segundo Jaques e Vicari (2005), este tipo de agente, que pode se caracterizar por um personagem animado na tela do computador é capaz de auxiliar o aluno na realização de tarefas, apresentando dicas e respostas para diferentes situações de aprendizagem. Em breves palavras, trata-se de um agente de software especificamente desenvolvido para os AVA's, que possui uma representação gráfica animada, cuja finalidade é facilitar e aperfeiçoar a aprendizagem.

Os agentes pedagógicos animados<sup>74</sup>, além de propiciarem uma maior interação entre alunos e ambiente, também podem produzir um efeito positivo na experiência educacional, aumentando a motivação do aluno e melhorando seu comprometimento com as tarefas. Jaques e Vicari (2005) listam alguns benefícios do uso de agentes pedagógicos animados para a aprendizagem:

- Um agente pedagógico que acompanha o processo do aluno e causa a impressão de estar preocupado com seu progresso dá a impressão de “estar junto”, sempre encorajando o aluno a ir adiante e superar dificuldades;
- Um agente pedagógico que seja sensível aos resultados obtidos pelo aluno pode intervir quando ocorrer alguma frustração, recuperando o interesse do aluno;
- Um agente pedagógico pode mostrar entusiasmo em relação a um conteúdo ou atividade, incentivando o aluno a realizar a tarefa com mais disposição;
- Um agente com uma personalidade interessante pode auxiliar o aluno a tornar sua aprendizagem divertida fazendo com que o mesmo dedique-se mais.

Ressaltamos que, além das vantagens citadas acima, é importante que eles tenham habilidades pedagógicas semelhantes às dos sistemas tutores inteligentes, isto é, que tenham condições para responder dúvidas, gerar explicações e realizar questionamentos.

Um sistema tutor inteligente é um sistema computacional composto de modelos que descrevem o que ensinar a partir de estratégias adequadas. Este tipo

---

<sup>74</sup> Podemos citar como exemplo de agente pedagógico animado o R-PAM desenvolvido por Edson Macedo Magalhães em trabalho de conclusão de curso de Bacharelado em Sistema de Informação (2010). O R-PAM (Remote Personal Assistant for Moodle – assistente pessoal remoto para o Moodle) possui uma interface de comunicação com os usuários e tem a proposta de assessorar o aluno na realização das atividades assíncronas, através do acompanhamento e do monitoramento das atividades da(s) comunidade(s) da(s) qual(is) o aluno participa.

de sistema tem a capacidade de interagir com o usuário e de atualizar sua própria base de conhecimento. Assim, os agentes pedagógicos representam um componente essencial para o sucesso dos sistemas tutores inteligentes junto aos usuários. Uma vez que tais tutores sejam modelados com base em sistemas multiagentes, o agente pedagógico representa a interface entre o tutor e o aluno.

Conforme referimos anteriormente, a M\_ONTO será utilizada por um agente pedagógico inteligente a ser modelado em um trabalho futuro de um colega da computação.

Ao usar uma ontologia, o agente poderá potencializar sua ação no ambiente, verificando a conectividade, mantendo um histórico das atividades mesmo que o aluno esteja *off-line* para que não haja necessidade de o aluno interagir diretamente com o ambiente. Além disso, o agente pode ser capaz de fornecer retornos tanto ao aluno quanto ao professor quando alguma atividade não foi concluída por algum problema técnico, por exemplo.

No caso deste trabalho, o agente a ser implementado fará uso da M\_ONTO para a busca de informações que sejam importantes na sua aplicação. Em especial, as descrições dos modelos de domínio e de interação poderão receber substanciais melhorias na qualidade.

Na seção seguinte, detalharemos nossos procedimentos metodológicos, explicitando a análise do *corpus* de pesquisa e desenvolvendo as duas etapas previstas para este trabalho: a linguística e a computacional.

### **4.3 O percurso metodológico**

Esta seção apresenta o nosso percurso metodológico para a análise semântica dos itens lexicais relacionados ao ambiente Moodle, compreendendo a etapa linguística, na qual prevemos a descrição dos aspectos teóricos apresentados no capítulo 3, e a etapa computacional, compreendendo a modelagem da ontologia no editor Protégé 4.1.

Sobre metodologias para a construção de ontologias pode-se encontrar diversos estudos na literatura. Almeida e Bax (2003a) realizaram uma ampla pesquisa sobre o assunto e sintetizaram os diferentes métodos. Segundo os autores (apud LÓPEZ, 1999), várias metodologias têm sido desenvolvidas com a finalidade de sistematizar a manipulação e a construção de ontologias; porém, complementam, ratificando a ideia de que as abordagens são muito distintas e variadas, que é improvável que haja unificação das propostas em uma metodologia única.

Nas metodologias apresentadas na literatura, há um consenso de que o processo de construção de uma ontologia envolve:

- (a) Definição das classes;
- (b) Organização das classes em uma hierarquia taxonômica (subclasses e superclasses);
- (c) Definição das propriedades (atributos) e valores;
- (d) Preenchimento dos valores das propriedades para cada instância.

Além disso, existem algumas estratégias para a definição de uma hierarquia, são elas:

- (a) *Top-down* (topo-para-baixo) – considerada a mais comum, pois remete à maneira cartesiana com que resolvemos problemas. Dessa forma, inicia-se definindo os conceitos mais gerais e segue-se através de um processo de decomposição, onde são colocados os termos mais abrangentes (superclasse) e abaixo os mais específicos (subclasse) através de relacionamentos;
- (b) *Bottom-up* (baixo-para-cima) – define-se primeiramente o conjunto de termos mais específicos para depois identificar os possíveis agrupamentos;
- (c) Combinação – utiliza um misto das duas estratégias descritas anteriormente. Os conceitos mais salientes são identificados e

escolhidos, assim o processo de generalização ou decomposição é guiado por esse conjunto de termos.

Para nossa pesquisa, utilizamos a estratégia chamada “combinação”, pois verificamos os conceitos mais salientes para depois estabelecer as relações e generalizar ou decompor cada um deles.

No nosso caso, ao definirmos as classes e subclasses estaremos adentrando na etapa linguística, estabelecendo as relações paradigmáticas – hiponímia, meronímia e sinonímia. A chamada *definição de propriedades* (atributos e valores) corresponde, na nossa análise, ao estabelecimento das relações sintagmáticas.

Por fim, é necessário que sejam criadas instâncias individuais para as classes e as subclasses a fim de verificar a consistência da ontologia e de seus relacionamentos. Nesta etapa, serão analisados os documentos dos diferentes cursos a distância promovidos pela UNISINOS, com vistas a verificar a consistência da ontologia.

Diante do exposto, salientamos que o diferencial desta pesquisa é a reflexão sobre a modelagem conceitual, fato este que nos leva a não seguir automaticamente os protocolos determinados por um tipo específico de metodologia. Assim, em linhas gerais, a metodologia adotada nesta pesquisa segue os seguintes pontos:

(a) determinação do escopo da ontologia, isto é, o domínio ao qual está relacionada e a sua aplicação (pontos já apresentados nas seções 4.1 e 4.2 deste capítulo);

(b) fase de preparação do *corpus*;

(c) etapa linguística;

(d) etapa computacional, ou seja, a modelagem da ontologia.

Passamos agora à seção 4.3.1, que corresponde à preparação e análise do corpus de pesquisa.

### 4.3.1 Preparação e análise do corpus

Esta etapa tem como principal objetivo preparar o *corpus* para a análise semântica a ser realizada posteriormente. Para efetivar esta etapa, realizamos procedimentos a fim de selecionar, definir e organizar o *corpus*.

#### a) Seleção do *corpus*

O *corpus* de pesquisa é constituído de manuais do Ambiente Moodle disponibilizados eletronicamente. Para a escolha destes documentos, realizamos uma busca na Web procurando por manuais completos sobre o ambiente e selecionamos aqueles que julgamos conter maiores informações sobre a plataforma e suas funcionalidades – um manual completo sobre o Moodle<sup>75</sup>, um manual para professores<sup>76</sup> e um guia para alunos<sup>77</sup>. A escolha por adotar os manuais como *corpus* de pesquisa e não as páginas *Web* das comunidades ocorreu pelo fato de necessitarmos de itens lexicais gerais sobre o Moodle e, caso nossa opção fosse pelas páginas, teríamos uma variedade de termos referentes ao conteúdo de suas disciplinas, fato este que dificultaria a extração e seleção dos termos para nossa ontologia.

---

<sup>75</sup> Manual completo do Ambiente Moodle.

Disponível em:

<[http://www.google.com.br/url?sa=t&source=web&cd=1&ved=0CBwQFjAA&url=http://www.pdadigital.com.br/downloads/doc\\_download/2-manual-completo-do-moodle&rct=j&q=manual%20%20do%20moodle&ei=q1c4TZDmLoK8lQewl7GtBw&usg=AFQjCNE8LtTuvAQmHowH6tFxuqBeovOXbQ&cad=rjt](http://www.google.com.br/url?sa=t&source=web&cd=1&ved=0CBwQFjAA&url=http://www.pdadigital.com.br/downloads/doc_download/2-manual-completo-do-moodle&rct=j&q=manual%20%20do%20moodle&ei=q1c4TZDmLoK8lQewl7GtBw&usg=AFQjCNE8LtTuvAQmHowH6tFxuqBeovOXbQ&cad=rjt)>. Acesso em: 21/01/2010.

<sup>76</sup> Manual do professor.

Disponível em: <[http://www.moodle.ufba.br/file.php/1/Manual\\_do\\_Moodle\\_para\\_professor\\_-\\_Vers\\_o\\_1.9.9.pdf](http://www.moodle.ufba.br/file.php/1/Manual_do_Moodle_para_professor_-_Vers_o_1.9.9.pdf)>. Acesso em: 21/01/2010.

<sup>77</sup> Guia do aluno.

Disponível em: <[http://www.moodle.unisinos.br/file.php/1/Tutoriais/Guia\\_do\\_Aluno\\_Moodle.pdf](http://www.moodle.unisinos.br/file.php/1/Tutoriais/Guia_do_Aluno_Moodle.pdf)>. Acesso em: 21/01/2010.

Delimitados os documentos que fariam parte do nosso *corpus*, realizamos a compilação dos mesmos utilizando a ferramenta *e-Termos*<sup>78</sup>, que gerou um documento único contendo todas as informações para a nossa análise.

#### b) Extração e seleção de termos

Com o *corpus* na base de dados do *e-Termos*, passamos para uma nova etapa e realizamos a extração de candidatos a termos. Em uma simples extração, o *e-Termos* trouxe todas as palavras e o número de ocorrências, sendo que isso não se mostrou muito produtivo para a nossa finalidade. Criamos, então, uma *stop-list*, ou seja, uma lista de palavras que não devem ser incluídas na análise do extrator, como, por exemplo, *de, a, uma, um, está*, etc. Para a geração da *stop-list*, analisamos a primeira lista fornecida pelo programa e verificamos quais seriam as palavras que deveriam compor a nossa lista de restrições.

Gerando uma nova extração, obtivemos um resultado mais satisfatório, que pode ser visualizado parcialmente na figura 21; porém, percebemos ainda uma grande ocorrência de termos duplos, diferenciados apenas por estarem no plural. Optamos por excluir manualmente esta variação por considerarmos desnecessário levar este ponto em conta na nossa análise, uma vez que o significado destes itens lexicais não se altera pela variação de número. A ferramenta organiza os termos extraídos pelo número de ocorrências; porém, este aspecto é irrelevante na nossa análise, pois temos o objetivo de buscar todos os termos relacionados ao Moodle, uma vez que, mesmo que eles sejam pouco expressivos nos manuais, podem ser importantes para a modelagem da ontologia.

---

<sup>78</sup> O *e-Termos*, acrônimo de Termos Eletrônicos, é um ambiente computacional colaborativo web de acesso livre e gratuito dedicado à gestão terminológica. Seu principal objetivo é viabilizar a criação de produtos terminológicos, sejam eles para os fins de pesquisa acadêmica ou de divulgação, por meio da (semi)automatização das etapas do trabalho terminológico. É um projeto acadêmico desenvolvido em parceria entre a Embrapa Informática Agropecuária (CNPTIA), Universidade de São Paulo (USP Campus de São Carlos, SP) e Universidade Federal de São Carlos (UFSCar), representados pelos laboratórios de pesquisa LabInfo (Laboratório de Organização e Tratamento da Informação Eletrônica), NILC (Núcleo Interinstitucional de Linguística Computacional) e o GETerm (Grupo de Estudos e Pesquisas em Terminológicos), localizados nas três instituições, respectivamente. Disponível em: <<http://www.etermos.cnptia.embrapa.br/index.php>>. Acesso em: 02/01/2011.

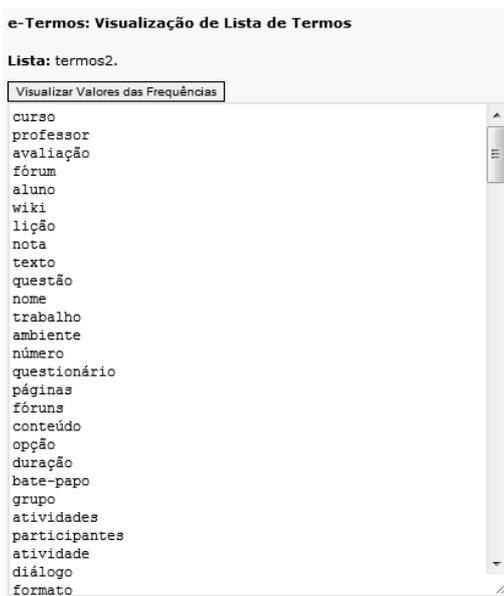


Figura 21: Lista parcial de termos – extração *e-Termos*

De posse da lista gerada pelo *software*, passamos a uma análise mais criteriosa dos termos. Nesta etapa, acessamos a página do ambiente Moodle da UNISINOS e buscamos pelos itens lexicais que haviam sido selecionados de forma automática. Além disso, fizemos uso da ferramenta *concordanceador*<sup>79</sup>, disponível no sistema *e-Termos*, para visualizar o termo no contexto do corpus. Na figura 22, é possível ver o retorno do *concordanceador* do *e-Termos* ao consultarmos o item lexical TUTOR.

---

<sup>79</sup> Ferramenta que busca as ocorrências do item lexical consultado no contexto em que ele está inserido no corpus.

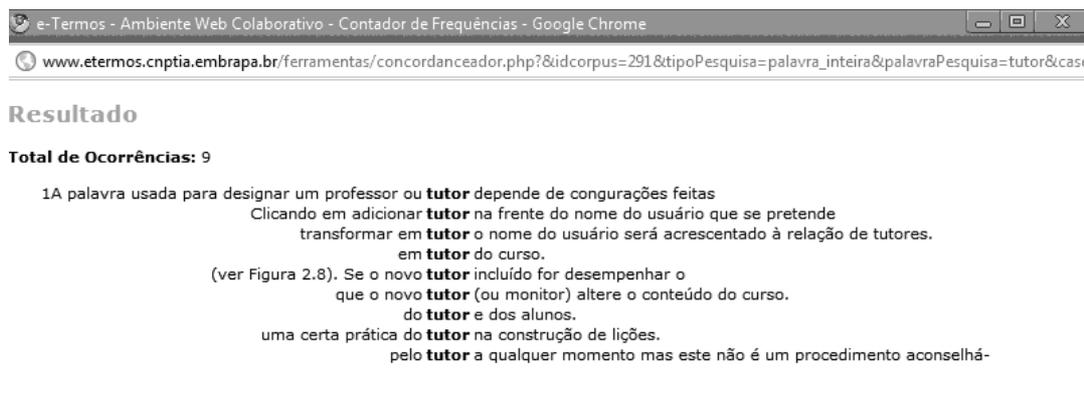


Figura 22: Consulta TUTOR – concordanceador e-Termos

Realizamos este procedimento para verificar o contexto de cada um dos itens lexicais e a importância (ou não) para a ontologia, fazendo uma nova seleção. Esta análise resultou em 87 itens lexicais que podem ser visualizados na figura 23.

1. agenda	26. disciplina	51. monitor	76. texto
2. aluno	27. documento	52. nome	77. tipo
3. arquivo	28. doutorado	53. nota	78. título
4. assíncrona	29. download	54. notícia	79. tópico
5. assunto	30. duração	55. objetivo	80. trabalho
6. atividade	31. e-mail	56. on-line	81. turma
7. auto-avaliação	32. endereço	57. participante	82. tutor
8. avaliação	33. escolha	58. período	83. universidade
9. bate-papo	34. estudante	59. peso	84. usuário
10. cadastro	35. exercício	60. pesquisa	85. visitante
11. caixa correio	36. feedback	61. pós-graduação	86. vocabulário
12. calendário	37. formulário	62. professor	87. wiki
13. chat	38. fórum	63. projeto	
14. código	39. glossário	64. prova	
15. comentário	40. graduação	65. questão	
16. comunidade	41. grupo	66. questionário	
17. conteúdo	42. hora	67. recurso	
18. critério	43. idade	68. semana	
19. curso	44. inscrição	69. semestre	
20. data	45. instrutor	70. senha	
21. descrição	46. lição	71. síncrona	
22. diálogo	47. matrícula	72. sobrenome	
23. diário	48. mensagem	73. tarefa	
24. dicionário	49. mestrado	74. tempo	
25. diretório	50. módulo	75. teste	

Figura 23: Lista de termos do ambiente Moodle

Através desta lista, é possível selecionarmos os termos mais representativos e agrupá-los de acordo com as categorias eleitas para constituição da M\_ONTO. Nesta etapa, seguimos uma abordagem *bottom-up*, conforme referido anteriormente.

Tendo definido a listagem de termos, passamos à análise semântica, levando em conta os aspectos teóricos discutidos no capítulo 3.

#### 4.3.2 Etapa linguística

Nesta subseção, temos o objetivo de apresentar de forma explícita as relações discutidas no capítulo 3, que elegemos como elementos estruturais para a nossa ontologia. Nesta análise, seguiremos a estrutura ontológica prevista no final do capítulo anterior, que compreende a definição da entidade; as relações paradigmáticas: relações de equivalência, hiponímia e meronímia; e as relações sintagmáticas: frames.

a) Seleção da definição: a definição do significado dos nominais deu-se a partir das definições apresentadas no manual do ambiente Moodle. Já os verbos foram definidos com base nos frames evocados por eles e pelas acepções consultadas no Dicionário Aulete Digital. Salientamos que a informação da definição não é relevante para a ontologia, uma vez que os significados não se expressam por meio de definições, mas a partir das relações entre os termos. Optamos por incluir este nível de representação a fim de documentar de forma mais detalhada os conceitos da ontologia e auxiliar linguistas e informatas que trabalharão com estes dados; além disso, diferentes manuais do Protégé referem-se à importância de incluir documentação nas ontologias, a fim de detalhar e explicar as formalizações descritas;

b) Identificação das relações paradigmáticas: tomamos a Semântica Lexical como base para a representação da significação nesta etapa. Os itens lexicais organizados hierarquicamente serão descritos através das relações de hiponímia, meronímia e equivalência. Ao selecionarmos os itens lexicais correspondentes às classes e, a partir deles, estabelecermos as subclasses e os relacionamentos, estamos seguindo uma abordagem *top-down*. Listamos, a seguir, as classes identificadas no domínio do ambiente Moodle: CATEGORIA\_CURSOS, CURSO,

ALUNO, PROFESSOR, VISITANTE, RECURSOS e EVENTOS, sendo as duas últimas subdivididas em outras subclasses:

- Classe RECURSOS, subdividida em 4 subclasses:
  - Materiais, composta por textos, *links* e vídeos;
  - Ferramentas\_de\_interação, composta por *chat*, fórum, diário, *wiki* e diálogo;
  - Ferramentas\_de\_auxílio, composta por glossário, lição, pesquisa, questionário, tarefa;
  - Ferramentas\_do\_sistema, composta por agenda, calendário, lista (atividades e participantes), *e-mail*.
- Classe EVENTOS, cujas subclasses serão organizadas de acordo com as classes semânticas dos verbos (Borba, 1996): ações, processos, ações-processos e estados.

A fim de melhor detalharmos a hierarquia definida, apresentamos cada classe através de um esquema representativo das relações de hiponímia, meronímia e equivalência (sinonímia). Para melhor visualização, optamos por apresentar cada classe em separado, fazendo, no final, um esquema geral da taxonomia. Salientamos que, para as relações de equivalência, foram usados itens lexicais retirados do *corpus*, uma vez que consideramos a aplicabilidade da ontologia em relação ao agente, isto é, quem vai realizar inferências é o agente e não um humano. Dentro desta perspectiva, é importante que a relação de sinonímia seja identificada dentro do *corpus*, por exemplo: podemos dizer que *mestre* é sinônimo de *professor*, porém, *mestre* não é um item lexical identificado no *corpus* e, portanto, não passível de influência na análise do agente. Na figura 24, apresentamos a hierarquia proposta para a classe RECURSOS.

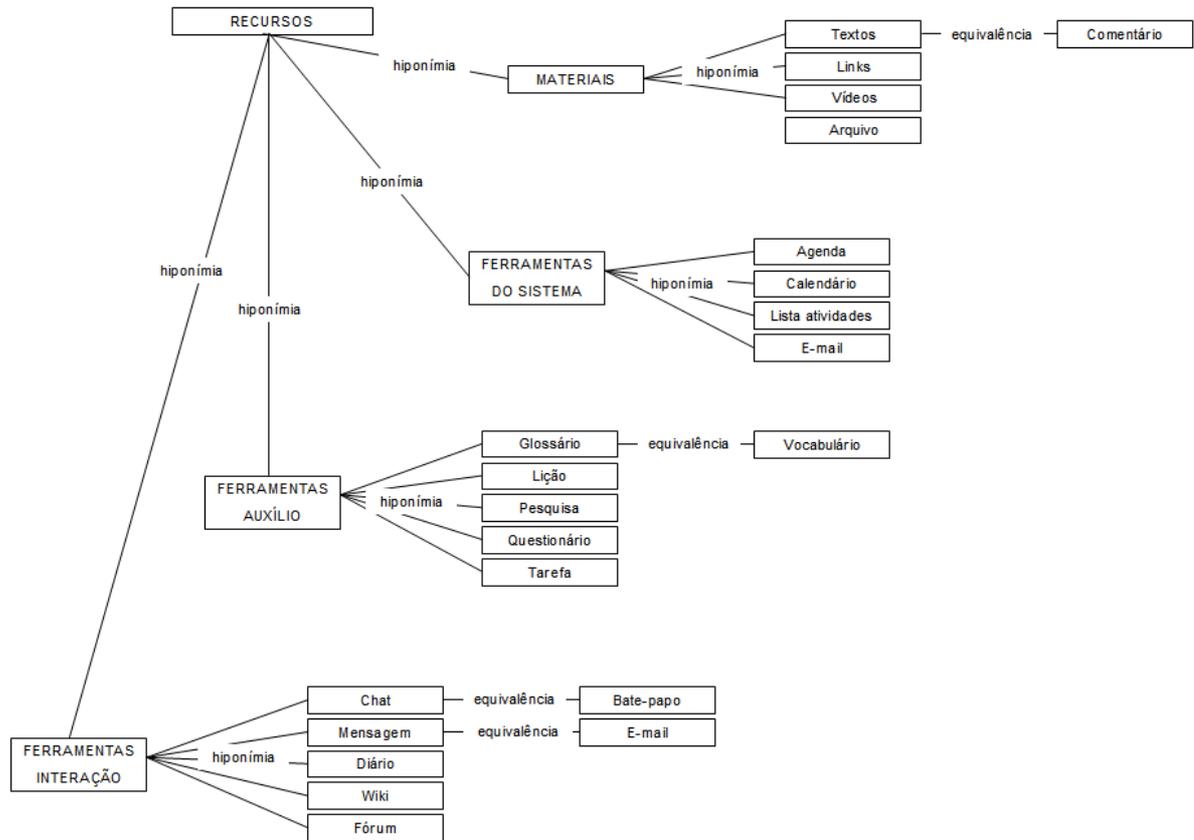


Figura 24: Relações paradigmáticas – classe RECURSOS

A classe EVENTOS, por ser composta unicamente por verbos, será passível de dois tipos de análise: uma no nível paradigmático e outra no nível sintagmático. No nível paradigmático, seguiremos a semântica verbal, classificando os verbos em ação, processo, ação-processo e estado (BORBA, 1996). Já o nível sintagmático envolverá a estrutura de *frames*, que será apresentada mais detalhadamente na sequência.

No *corpus* em estudo, destacamos os verbos mais representativos e correspondentes aos eventos de ação-processo. Salientamos que as demais classes de verbos podem ser exploradas em trabalhos futuros, através de um estudo dos verbos de estado por meio de mapeamento dos sentimentos dos usuários em suas interações no diário, por exemplo.

Na figura 25, apresentamos a estrutura hierárquica proposta para a classe EVENTOS, destacando que os verbos selecionados em nosso *corpus* correspondem à classe ação-processo.

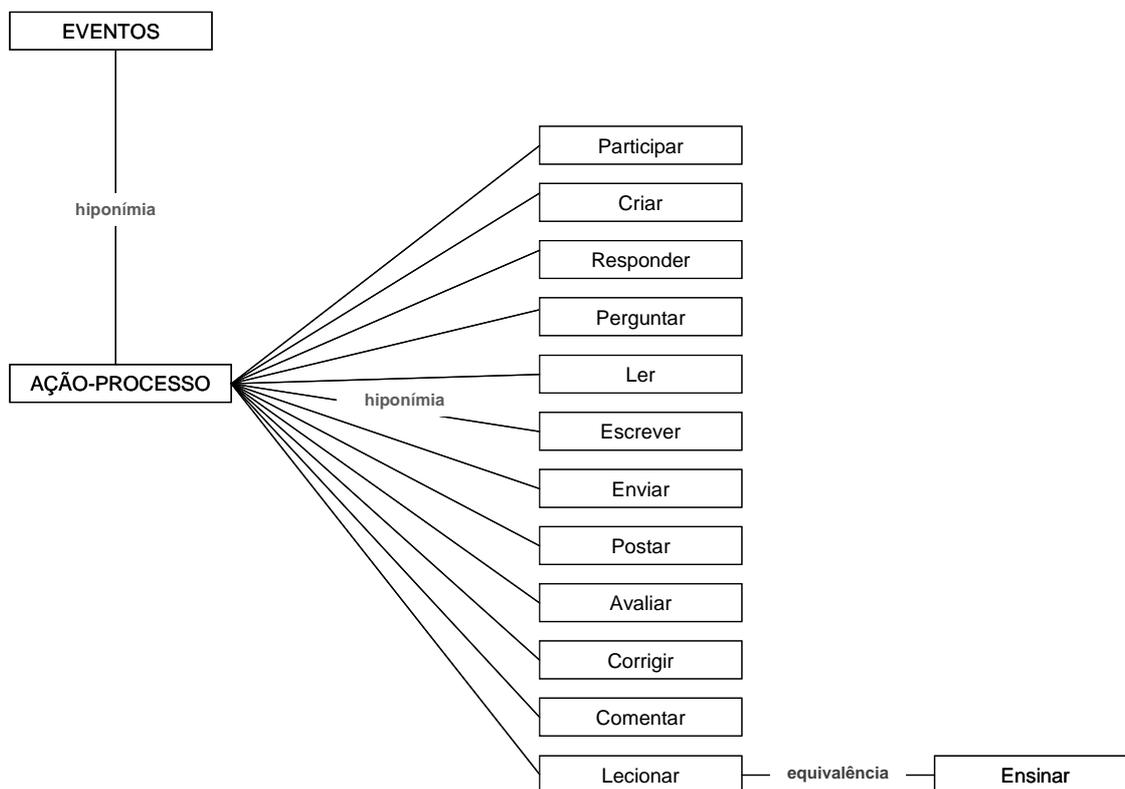


Figura 25: Relações paradigmáticas – classe EVENTOS

Na figura 26, apresentamos todas as classes elencadas para compor a M\_ONTO explicitando as relações existentes entre elas.

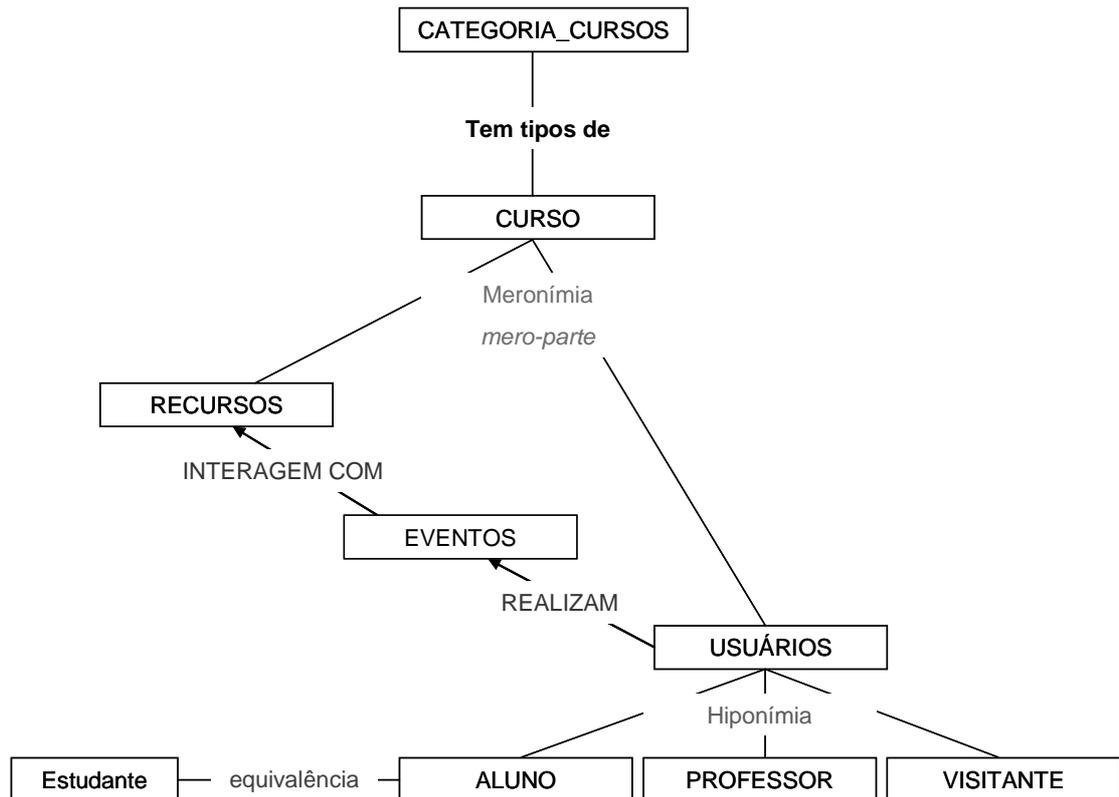


Figura 26: Estrutura hierárquica da M\_ONTA

Na etapa computacional, daremos seguimento a esta estrutura hierárquica incluindo as instâncias e, conseqüentemente, verificando a consistência da modelagem.

Como é possível visualizar na figura 26, a M\_ONTA apresenta outros tipos de relacionamentos (representados com letras maiúsculas), além dos chamados taxonômicos (hiponímia, meronímia e equivalência). Conforme referimos no capítulo 3, as áreas da CI e da CC tratam estes tipos de relacionamentos como relações associativas, tornando-os pouco consistentes para a modelagem da ontologia. No caso deste estudo, incluiremos os frames para maior detalhamento dos relacionamentos entre as classes. Assim, os verbos da classe EVENTOS passam a ser evocadores de *frames*, agregando valor à ontologia, à medida que os eventos passam a receber uma organização em forma de *frames*. A seguir, apresentamos maiores detalhes sobre os *frames* e sua contribuição na modelagem da M\_ONTA.

c) Identificação das relações sintagmáticas: no capítulo 3, apresentamos os papéis temáticos e os *frames* para a descrição dos relacionamentos sintagmáticos, justificando nossa escolha pelos *frames* por possibilitarem a inclusão de novas categorias, agregando valor à descrição de significados da ontologia.

Em uma abordagem baseada em *frames*, ao contrário do que ocorreria com os papéis temáticos, faz-se necessário partir de uma análise das situações nas quais os verbos estão envolvidos, isso porque se trata de uma abordagem que estabelece uma relação entre o léxico e o contexto extralingüístico.

Tomaremos como base os *frames* existentes da FN, realizando uma busca dos verbos do nosso domínio e verificando os elementos de frame relacionados a cada cena, a fim de verificar se corresponde ou não ao escopo de nossa ontologia. Encontramos nove *frames* na FN que são evocados pelos verbos retirados do nosso *corpus*, sendo muitos deles não relacionados diretamente ao domínio da educação, mas que poderão ser úteis na descrição semântica dos verbos selecionados no *corpus*. O *frame* mais diretamente relacionado à nossa aplicação é o *EDUCATION-TEACHING*, o qual descrevemos abaixo. Os demais *frames* encontrados serão apenas citados a fim de comprovar a grande rede de relacionamentos existente na FN e capaz de ampliar a descrição semântica dos verbos; são eles: *PARTICIPATION* (evocado pelo verbo *participar*), *INTENTIONALLY\_CREATE* (evocado pelo verbo *criar*), *COMMUNICATION\_RESPONSE* (evocado pelo verbo *responder*), *QUESTIONING* (evocado pelo verbo *perguntar*), *READING* (evocado pelo verbo *ler*), *TEXT\_CREATION* (evocado pelo verbo *escrever*), *EXAMINATION* (evocado pelos verbos *avaliar* e *corrigir*) e *SENT\_ITEMS* (evocado pelos verbos *postar* e *enviar*).

O *frame* mais significativo para o nosso estudo que encontramos na FN é o *frame EDUCATION-TEACHING*, pois se refere especificamente à situação de aprendizagem dos sujeitos. Consideramos importante detalhar este *frame*, traduzindo-o, para melhor compreensão dos relacionamentos entre os verbos e os elementos de *frame*. Passamos, então, à descrição do *frame* supracitado.

**a) Frame EDUCATION-TEACHING (evocado pelos verbos *lecionar, ensinar e estudar*)**

**Definição:** este frame contém as palavras referentes ao ensino e aos participantes no ensino. Um aluno começa a aprender sobre um assunto, uma habilidade, um preceito ou um fato como resultado da instrução de um professor.

**Elementos de frame:**

**\* nucleares**

Curso: um programa de palestras ou outras matérias que tratam de um assunto.

Fato: a informação que o estudante recebe do professor.

Instituição: estabelecimento de ensino, como uma escola ou faculdade.

Material: material didático, tais como vídeos, textos, livros, utilizados pelo professor ou pelo estudante para adquirir habilidades ou conhecimentos.

Preceito: a orientação para um comportamento correto. Trata-se, na maioria dos casos, de ações sociais e moralmente aceitáveis.

Qualificação: uma qualificação formal, um grau acadêmico ou certificado que o aluno busca ao estudar.

Papel: posição que o estudante será capaz de ocupar após a sua formação.

Habilidade: uma ação que o aluno seja capaz de executar como resultado de uma instrução.

Estudante (aluno): aquele que é instruído por um professor em habilidades e conhecimentos.

Assunto: área do conhecimento ou habilidade que é ensinada.

Professor: aquele que instrui um aluno em alguma área do conhecimento ou habilidade.

**\* não nucleares**

Grau: grau em que o evento ocorre.

Descrição: frase que descreve o ator de uma ação.

Duração: a quantidade de tempo que mantém um processo em andamento.

Nível: identifica o nível de um estudante em sua formação. Isso difere do objetivo ou qualificação para a qual o aluno está estudando.

Maneira: forma como é executada uma ação.

Meios: uma ação intencional realizada pelo ator (aluno ou professor) que lhes permite atingir a aprendizagem e o ensino ao qual se propõem.

Local: local onde o ensino ocorre.

Finalidade: o que o professor pretende promover através do ensino.

Resultado: resultado de um evento

Tempo: o tempo em que o ensino ocorre.

\* **relações entre frames**: Herda de *INTENTIONALLY\_AFFECT*, usado por *STUDYING*.

\* **unidades lexicais**: educar, educação, pós-graduação, instruir, instrução, aprender, professor, mestre, aluno, escola, professora, estudante, ensinar, treinar, treino, tutor.

Observando os detalhamentos apresentados acima, pode-se perceber que os *elementos de frame* organizam as informações de forma semelhante aos papéis temáticos, quando se trata de uma semântica verbal. Ambos explicitam diferentes papéis entre os envolvidos. A principal diferença é que a abordagem baseada em *frames* identifica os participantes de uma situação expressos ou não lexicalmente. A abordagem de *frames*, por estar ligada a uma situação de comunicação, possui um conjunto maior e mais detalhado de situações, permitindo uma análise mais ampla e completa, ao contrário dos papéis temáticos, que representam um grupo finito de papéis para as situações.

Conforme pode ser visto no *frame* destacado, a FN não se ocupa da descrição de *frames* especializados, mas daqueles que descrevem cenas gerais de situações de comunicação. Diante disso, propomos a criação de um *frame* específico para a EAD, o qual chamamos de *EAD-TEACHING*. O processo de descrição de um *frame* envolve uma definição, o estabelecimento dos elementos de *frame*, a listagem das unidades lexicais correspondentes e os relacionamentos entre *frames*. Este *frame* foi subdividido em outros quatro que correspondem a situações específicas de EAD, mais precisamente referindo-se às ferramentas de interação: *chat*, *fórum*, *wiki* e *diário*. A seguir, descrevemos cada *frame* criado seguindo parcialmente a estrutura da FN<sup>80</sup>, sendo que elencamos o subframe *fórum* para apresentar exemplos de sentenças e os papéis de cada elemento, tomando como base a representação na FN. Apenas a título de organização do texto, deixamos este *frame* para o final da descrição.

**a) Frame *EAD-TEACHING* (evocado pelos verbos *lecionar*, *ensinar* e *estudar*)**

**Definição:** este *frame* contém as palavras referentes ao ensino e aos participantes no ensino à distância. Um aluno começa a aprender sobre um assunto, uma habilidade, um preceito ou um fato como resultado da instrução de um professor. As atividades ocorrem através de um ambiente de ensino a distância, envolvendo atividades síncronas e assíncronas.

**Elementos de *frame*:**

Curso: programa de conteúdos específicos que serão desenvolvidos através de atividades.

Conteúdo: a informação que o estudante recebe do professor.

Área: área do conhecimento ou habilidade que é ensinada.

Instituição: estabelecimento de ensino, como uma escola ou faculdade.

---

<sup>80</sup> A FN apresenta a descrição das cenas em um nível sintático-semântico, elencando os elementos nucleares e não nucleares, além de listar os itens lexicais relacionados. Consideramos que nossa descrição segue parcialmente o padrão da FN, pois não chegamos a esse nível de detalhamento.

Ambiente: o ambiente EAD no qual ocorre o curso.

Material: material didático, tais como vídeos, textos, e-books<sup>81</sup>, links utilizados pelo professor ou pelo estudante para adquirir habilidades ou conhecimentos.

Recurso: recursos disponíveis no sistema como ferramentas de aprendizagem. São criados pelo professor e destinados aos estudantes.

Estudante (aluno): aquele que é instruído por um professor em habilidades e conhecimentos. Acessa o ambiente e realiza as tarefas postadas pelo professor.

Professor: aquele que instrui um aluno em alguma área do conhecimento ou habilidade. Acessa o ambiente e cria as tarefas para os estudantes. Tem a função de avaliar as tarefas e organizar o ambiente.

Semana/módulo: formato como o curso é organizado.

Tutor: auxilia os alunos e professores no processo de aprendizagem. Em alguns casos assume o mesmo papel de professor.

Monitor: tem a função de esclarecer dúvidas e auxiliar os alunos no processo de aprendizagem.

Período: período em que o curso ocorre.

Avaliação: as atividades realizadas pelos estudantes através das ferramentas do ambiente podem ser avaliadas pelo professor, resultando em uma nota ou conceito.

**\* relações entre frames:**

Herda de *EDUCATION-TEACHING*,

---

<sup>81</sup> Termo em inglês que significa “*eletronic book*” ou livro eletrônico

Usa: *STUDYING, READING, PARTICIPATION, INTENTIONALLY-CREATE, SENT-ITENS, COMMUNICATION\_RESPONSE, QUESTIONING, TEXT-CREATION* e *EXAMINATION*

É herdado por: *CHAT, FÓRUM, WIKI E DIÁRIO*

#### **b) Frame *CHAT* (subframe de *EAD-TEACHING*)**

**Definição:** este frame descreve uma situação de uso da ferramenta chat de comunicação síncrona que permite que professor e alunos mantenham uma conversa em tempo real, ou seja, é necessário que todos estejam simultaneamente on-line para que haja a interação. Também conhecido como sala de bate-papo.

##### **Elementos de frame:**

Assunto: assunto que será discutido no chat. Pode ser sobre algum texto lido, sobre alguma dificuldade do grupo, etc.

Estudante: participam do chat em tempo real lendo e respondendo as mensagens digitadas sobre o assunto em questão.

Professor: participa do chat, criando e postando o assunto inicial. Gerencia o chat, definindo o horário e mediando as interações dos alunos. Lê e responde os comentários dos alunos, provocando discussões e levantando questionamentos.

Objetivo: explicita a finalidade da criação do chat.

Data: marca a data em que todos deverão estar logados aos sistema para a atividade.

Horário: marca o início e o término da seção de bate-papo.

#### **c) Frame *WIKI* (subframe de *EAD-TEACHING*)**

**Definição:** este frame descreve uma situação de uso da ferramenta wiki, ou seja, da criação de uma coleção de documentos criados de forma coletiva no ambiente da internet. Basicamente, uma página wiki é uma página web que

qualquer pessoa de um curso pode criar, diretamente no navegador, sem necessidade de conhecimento de uma linguagem específica. Uma wiki começa com uma página de rosto. Cada autor pode acrescentar outras páginas ao *wiki*, simplesmente criando um *link*. É uma ferramenta de trabalho colaborativo.

**Elementos de frame:**

Rosto: página inicial da *wiki*.

Páginas: todas as páginas produzidas pelos alunos e professor.

Link: ligação entre as diferentes páginas.

Grupos: define quem poderá criar e editar a *wiki*.

Mapa: visão hierárquica das páginas, começando pela primeira.

Índice: ordem alfabética do conteúdo criado na *wiki*.

Aplicação: define qual será o uso da *wiki*: apresentação de conteúdo, registro de impressões sobre a aula, etc.

**d) Frame *DIÁRIO* (subframe de *EAD-TEACHING*)**

**Definição**: este frame descreve uma situação de uso da ferramenta que permite ao professor conversar separadamente com o estudante e que o mesmo relate suas experiências e reflexões de maneira progressiva, podendo contar com o retorno do professor. É uma atividade que precisa ser feita constantemente, com intervalos de tempo de acordo com a necessidade da turma e do professor (dia, semana, quinzena).

**Elementos de frame:**

Anotação: recado ou tópico postado no diário pelo aluno.

Feedback: retorno do professor ao aluno.

Visualização: somente professor e aluno podem ver as postagens do diário.

Periodicidade: intervalo de tempo definido previamente para preenchimento do diário (diário, semanal, quinzenal)

### **e) Frame *FÓRUM* (subframe de *EAD-TEACHING*)**

**Definição**: este frame descreve uma situação de uso da ferramenta fórum como atividade de discussão e interação entre usuários. Os fóruns têm diversos tipos de estrutura e podem incluir avaliação recíproca em cada mensagem. As mensagens são visualizadas em diversos formatos e podem incluir anexos. Os participantes do fórum têm a opção de receber cópias das novas mensagens via e-mail e os professores, de enviar mensagens ao fórum com cópias via e-mail a todos os participantes. Não é necessário que todos os participantes estejam on-line para a interação no fórum.

#### **Elementos de frame:**

**Assunto**: assunto ao qual o fórum se refere

**Ana Guimarães** criou o fórum sobre **EAD**.

**Estudante**: participa do fórum postando comentários e enviando mensagens aos demais. Suas mensagens são avaliadas pelo professor. Pode também postar um tópico, caso esta opção seja liberada pelo professor.

**José Carlos** escreve **comentário**.

**Professor**: participa do fórum, criando e postando o tópico inicial. Gerencia o fórum, criando grupos, organizando os tópicos e avaliando as mensagens postadas pelos estudantes.

**Isa Mara Alves** avalia participação.

**Tópico**: texto que tem a função de dar início de uma discussão sobre um determinado assunto que deve nortear os demais comentários, delimitando o assunto da discussão. É inserido pelo criador (estudante ou professor).

**Rove** cria **tópico** sobre **Semântica**.

**Comentário**: mensagem postada no fórum relacionada ao assunto do tópico.

João comentou sua opinião sobre EAD.

**Objetivo:** explicita a finalidade da criação do fórum.

O fórum foi criado para avaliar os alunos.

**Período:** marca o período em que o fórum estará ativo para que os alunos e professores realizem sua interação.

Este fórum tem início em 28/02 e encerra-se em 03/03.

Diante dos dados dos *frames* expostos acima, tanto dos da FN quanto dos propostos por nós, é possível ampliar os relacionamentos baseados em verbos na nossa ontologia. Os *frames* revelam diferentes papéis dos participantes, por exemplo, <aluno> escreve <diário>, participa de <chat>, posta em <fórum>, escreve <texto>, lê <tarefas>, etc. A vinculação deste tipo de relacionamento torna-se muito mais fácil através dos *frames*, pois, caso tivéssemos optado pelos papéis temáticos, teríamos que criar uma grande lista de ações para cada participante; através dos *frames* e seus elementos estas ações ficam explícitas.

A figura 27 representa graficamente a proposição da inserção do frame *EAD-TEACHING* e seus subframes na FN.

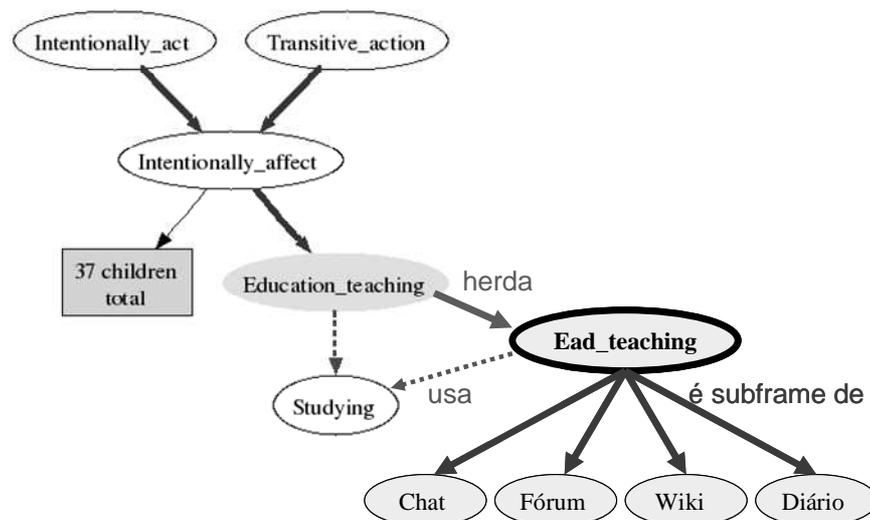


Figura 27: Frame *EAD\_TEACHING* e subframes – estrutura da FN

Concluída esta etapa, seguimos nosso percurso adentrando no terreno computacional, onde utilizaremos a ferramenta de construção de ontologias e estabeleceremos as formas de modelar o conteúdo da M\_ONTO

### **4.3.3 Etapa computacional**

Conforme referido anteriormente, organizamos a construção da M\_ONTO em duas etapas: a linguística e a computacional. A etapa linguística, apresentada na subseção 4.3.1, ofereceu detalhamentos sobre a hierarquia de classes e os relacionamentos da nossa ontologia, descrevendo como os conceitos serão representados linguisticamente através de relações paradigmáticas e sintagmáticas.

Nesta subseção, passamos aos procedimentos referentes à etapa computacional, sendo considerada assim por tratar da formalização dos dados linguísticos em uma ferramenta própria para a construção de ontologias. Esta etapa prevê a organização dos diferentes níveis de análise propostos na seção anterior através da ferramenta, de modo a possibilitar a conversão das informações linguísticas para a linguagem de implementação de ontologias, a *Ontology Web Language* (OWL).

Inicialmente, julgamos necessário apresentar a ferramenta computacional da qual nos valeremos para a modelagem dos dados da M\_ONTO em linguagem inteligível por computadores, o Protégé 4.1.

#### **a) Protégé 4.1**

O Protégé 4.1<sup>82</sup> é uma ferramenta que permite construir ontologias de domínio, personalizar formulários de entrada de dados, inserir e editar dados, possibilitando a criação de bases do conhecimento guiadas por uma ontologia. É um editor de ontologias de código aberto e foi desenvolvido pela *Stanford Medical Informatics na Stanford University School of Medicine*. Em seu projeto original, o *software* consistia em uma ferramenta de aquisição de conhecimento específica, limitada a um sistema especialista para a oncologia. Com o passar do tempo, foi sendo modernizado a fim de acompanhar os sistemas de organização do conhecimento, possibilitando a ampliação de seu uso para as demais áreas.

O Protégé possui duas formas para o desenvolvimento de ontologias, via Protégé-frames ou via Protégé-OWL, ambas apresentando características comuns, porém diferentes linguagens de formalização.

No Protégé-frames, é possível construir e preencher ontologias com o protocolo OKBC (Open Knowledge Base Connectivity). Nesta interface, a ontologia é desenvolvida através de quadros baseados no domínio, adaptando a forma de entrada de dados e inserindo diferentes dados de instanciação. O Protégé-Frames fundamenta-se nas ideias propostas por Minsky (1975), sendo que a noção de quadros/cenas deu origem às linguagens computacionais orientadas a objetos. Para o Protégé-Frames, um quadro (*frame*) consiste em um conjunto de atributos (*slots*) que descrevem as características dos objetos representados através de seus valores. Os valores atribuídos a cada quadro também podem servir como ligação com outros quadros.

O Protégé-OWL é uma extensão do *software* que suporta a *Web Ontology Language* (OWL), especificada pela W3C como linguagem oficial de construção de ontologias para a Web Semântica. Uma ontologia desenvolvida em OWL pode incluir descrições de classes, propriedades e suas instâncias.

---

<sup>82</sup> As informações sobre o Protégé 4.1 foram retiradas do Manual “*A Practical Guide To Building OWL Ontologies Using Protégée 4 and CO-ODE Tools*”, desenvolvido por Matthew Horridge, Simon Jupp, Georgina Moulton, Alan Rector, Robert Stevens, Chris Wroe, na University Of Manchester, 2007. Disponível em: <http://owl.cs.manchester.ac.uk/tutorials/protegeowltutorial/>. Acesso em: 10/01/2011

Verificando as diferentes abordagens do Protégé para a criação de ontologias, percebemos que ambas permitem a descrição detalhada das classes, subclasses e seus relacionamentos, mas que diferem em alguns aspectos.

Aparentemente, a ferramenta Protégé-frames poderia ser considerada mais adequada para nosso propósito, uma vez que definimos os *frames* como elementos fundamentais para a definição dos relacionamentos entre as classes. No entanto, optamos pela utilização do Protégé-OWL para a modelagem da nossa ontologia, por se tratar de um editor que retorna um arquivo no formato OWL, suportado por aplicações para a Web Semântica, sendo recomendado pela W3C. Temos ciência de que esta escolha nos remeterá a buscar uma forma original de representar as relações entre *frames* na nossa ontologia, além de não termos a opção de utilizar o mesmo nome para diferentes entidades, uma vez que o Protégé-OWL não permite que duas unidades de diferentes nomes se refiram a um mesmo ente e nem dois nomes diferentes se refiram a um mesmo indivíduo.

O Protégé sugere um padrão para a nomenclatura das classes e das propriedades. Na ferramenta, as classes devem ser nomeadas através do padrão **InicialMaiuscula**, ou seja, as palavras que compõem um nome de classe terão sempre a primeira letra maiúscula. E as propriedades inicial com minúscula e as palavras seguintes têm a inicial maiúscula como, por exemplo, em **propriedadeMaiuscula**. Além disso, os nomes das classes e propriedades não devem conter acentuação, sinais gráficos ou espaços.

Nesta seção, descreveremos como cada nível de descrição proposto na estrutura ontológica (apresentada no final do capítulo 3) foi modelado no editor Protégé. Para a inclusão dos dados propostos pela estrutura ontológica (definição, relações de hiponímia, relações de equivalência, relações de meronímia e *frames*), inicialmente fizemos um estudo das possibilidades e limitações do *software* e posteriormente estabelecemos critérios para a formalização do conteúdo linguístico na ferramenta através dos recursos disponíveis. A partir daí, passamos à inclusão dos dados no Protégé, dando forma à M\_ONTO.

Por estarmos trabalhando efetivamente na etapa computacional prevista na nossa metodologia, julgamos importante seguir o padrão para nomear classes e

propriedades apresentado pelo Protégé. Para fins de melhor visualização, continuaremos destacando os relacionamentos/propriedades com a fonte *Courier New* e padronizaremos as classes marcando-as em **negrito**.

Apresentamos, a seguir, cada elemento proposto na estrutura ontológica, descrevendo nossas escolhas para sua inclusão no Protégé, exemplificando-as.

### a) Representação da Definição

Reiteramos o fato já citado anteriormente de que a definição da entidade não é ponto fundamental na modelagem de uma ontologia, uma vez que o sistema buscará o significado das entidades realizando inferências a partir da análise de suas relações e propriedades. Porém, como também já mencionamos, o fato de explicitar o significado de um determinado termo pode ser útil na documentação da ontologia, facilitando a compreensão de lingüistas e informatas que farão uso das informações por ela formalizadas.

No Protégé 4.1, esta documentação pode ser realizada através do item *anotações (annotations)* inserindo um *comentário (comment)*. Na figura 28, exemplificamos a inserção das definições no Protégé, através da classe **Curso**.

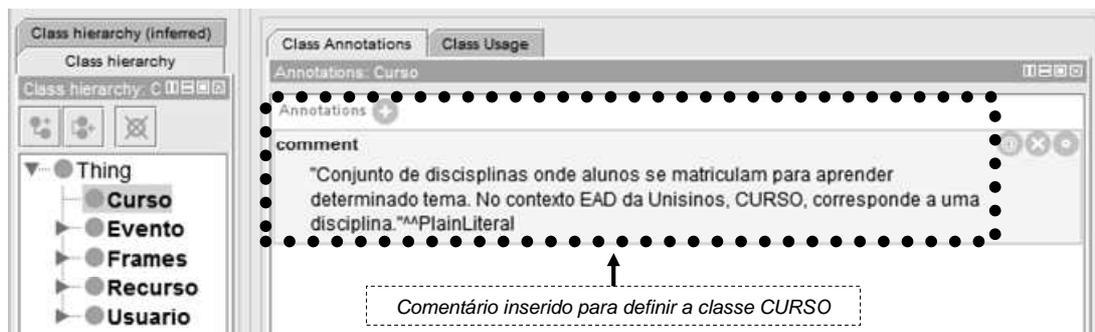


Figura 28: Definição da classe CURSO

### b) Representação das Relações de Hiponímia

Conforme referimos, a relação de hiponímia tem papel fundamental na modelagem de ontologias, pois representam a estrutura hierárquica proposta para a

organização das classes e subclasses, compreendendo a taxonomia que irá compor a ontologia.

No caso do Protégé, esta é a relação mais simples de ser modelada, uma vez que, ao criar uma classe e derivar subclasses, o próprio *software* já estabelece a relação *é-um* (hiponímia).

Para a inserção das classes, sempre se parte de uma classe inicial que compõe qualquer ontologia, mesmo quando está vazia; esta classe é chamada de **Thing**<sup>83</sup>. Para o Protégé, uma classe pode ser explicada como um conjunto de indivíduos (ou conjunto de objetos), sendo **Thing** a classe raiz da qual todas as demais derivam, isto é, todas as classes serão consideradas subclasses de **Thing**.

Outro aspecto importante a ser levado em consideração na criação da hierarquia de classes é o conceito de classes disjuntas (*Disjoint Classes*). As classes OWL podem se sobrepor e tornar a ontologia inconsistente, por isso é importante separar cada grupo de classes, criando-as como disjuntas, a fim de garantir que não ocorra a sobreposição. Isso garante que um indivíduo<sup>84</sup> instanciado como membro de uma classe não pode ser membro de nenhuma outra classe naquele grupo. Por exemplo, ao definir que **Curso**, **Evento**, **Frames** e **Recurso** são um grupo de classes disjuntas de **Usuario**, estamos determinando que os indivíduos que pertencem a esta classe não podem pertencer a nenhuma outra daquele grupo.

Na figura 29, apresentamos a hierarquia de classes estabelecida e a definição das classes disjuntas.

---

<sup>83</sup> **Thing** é uma palavra reservada da linguagem OWL, isto é, faz parte de seu vocabulário específico, conforme padrão da W3C e, por esta razão, não realizamos a tradução do termo.

<sup>84</sup> Indivíduo, membro e instância são considerados sinônimos no Protégé.

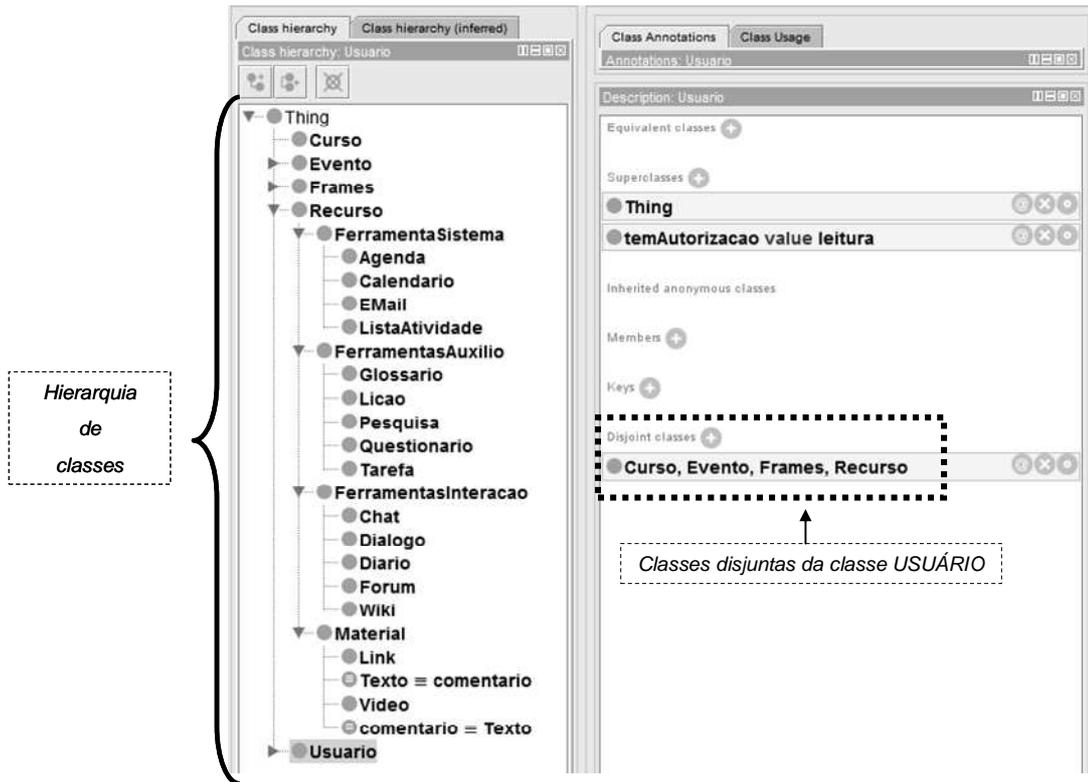


Figura 29: Hierarquia – classes disjuntas e relação de hiponímia

A hierarquia de classes também pode ser representada graficamente no Protégé, como pode ser visto na figura 30. Na figura, as classes são representadas em forma de elipses e os relacionamentos através da relação *is-a* (é-um).



Figura 30: Representação gráfica – classe RECURSO

Dando continuidade à representação das relações taxonômicas, descrevemos a seguir a forma como incluímos a relação de meronímia na ontologia.

### c) Representação das Relações de Meronímia

Sobre a relação de meronímia, apresentamos possibilidades de diferentes graus (WINSTON et.al., 1987) e diferentes formas de representá-la (MARRAFA, 2001) no capítulo 3; porém, ao analisar as classes que compõem a M\_ONTO, percebemos que não há necessidade deste tipo de detalhamento e que é possível seguirmos apenas o pressuposto *parte-de*, ou *mero-parte* (MARRAFA, 2001).

Para representar a relação de meronímia/holonímia na ferramenta, criamos uma propriedade de objeto<sup>85</sup> com nome de `meronimia` e suas subordinadas `parteDe` e `contem`, conforme pode ser visto na figura 31. O próximo passo foi adicionar os relacionamentos criados às nossas classes, para isso criamos propriedades restritivas<sup>86</sup> em cada uma das classes. Na figura 31, podemos visualizar as propriedades restritivas criadas para a classe **Usuario** e para a classe **Curso**.

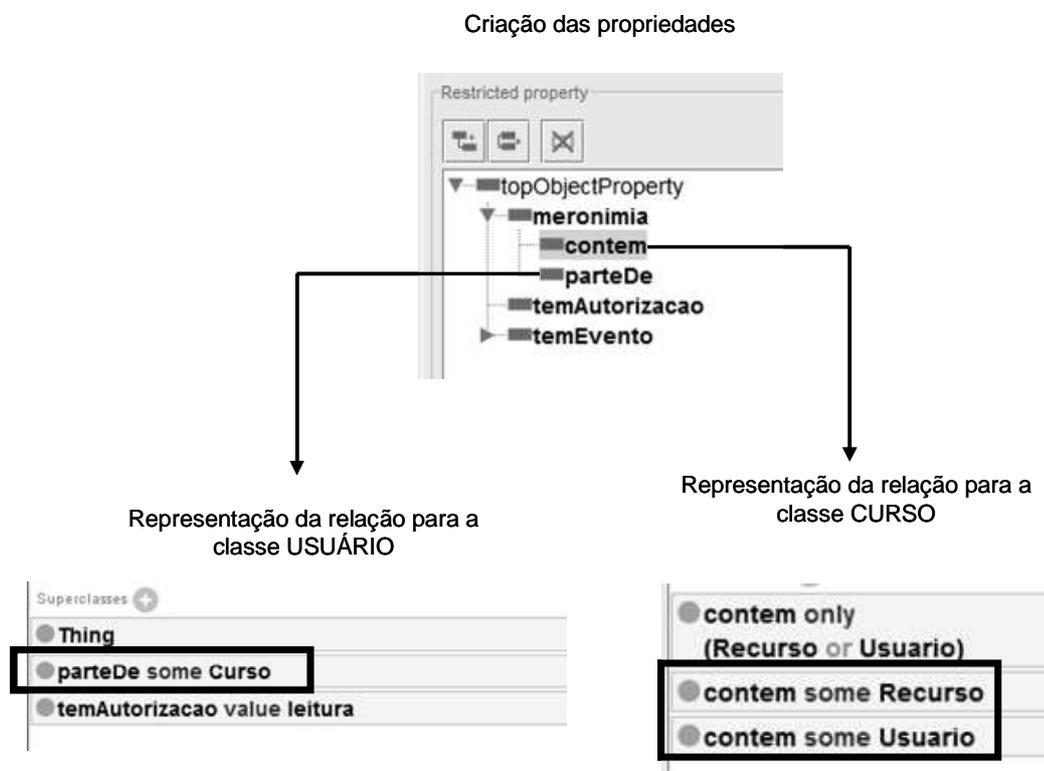


Figura 31: Representação da relação de Meronímia

<sup>85</sup> Propriedade de objeto é um termo da linguagem OWL. As propriedades OWL representam relacionamentos. Existem dois principais tipos de propriedades: propriedades de objetos e propriedades de tipo de dados. As propriedades do objeto são relações entre dois indivíduos.

<sup>86</sup> As propriedades restritivas são criadas para restringir os indivíduos a uma determinada classe.

Ao criar a propriedade restritiva <parteDe some Curso>, estamos definindo que os indivíduos da classe **Usuario** devem ter obrigatoriamente um relacionamento *parte-de* (mero-parte) com pelo menos um dos indivíduos da classe **Curso** (palavra reservada *some*), isto é, estamos garantindo que um aluno será *parte de* uma turma, por exemplo. Na classe **Curso**, as propriedades restritivas que foram criadas representam a relação inversa – holonímia.

#### d) Relações de Equivalência

Adotamos o termo “relações de equivalência” para representar a sinonímia, pois é desta forma que esta relação é apresentada na norma ANSI/NISO (2005), que define os padrões para a construção de vocabulários controlados. Outro aspecto já mencionado é que os termos sinônimos serão apenas aqueles encontrados no *corpus* em função da análise do agente.

Diante disso, apresentamos como exemplo a forma como foi definida a relação de equivalência entre **Texto** e **Comentario**, que são subclasses de **Material**. Para incluir o relacionamento que define que são equivalentes, há a opção “classes equivalentes” (*equivalent classes*), sendo que é necessário apenas clicar no botão de adicionar e escrever o nome da classe à qual ela está relacionada. Na figura 32, temos a visualização desta descrição, mostrando como a ferramenta identifica as classes equivalentes.

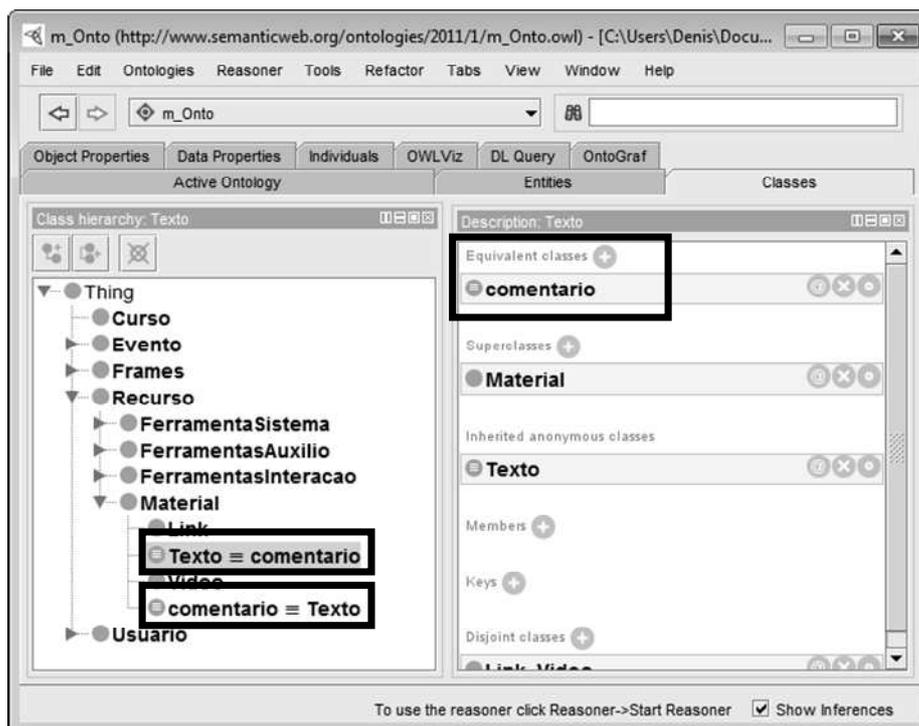


Figura 32: Representação da relação de Equivalência

Realizado este procedimento, as classes recebem uma identificação diferenciada (identificada na figura 33) e são representadas através da relação *is-a* na forma gráfica, como podemos ver abaixo destacado através de um retângulo pontilhado.

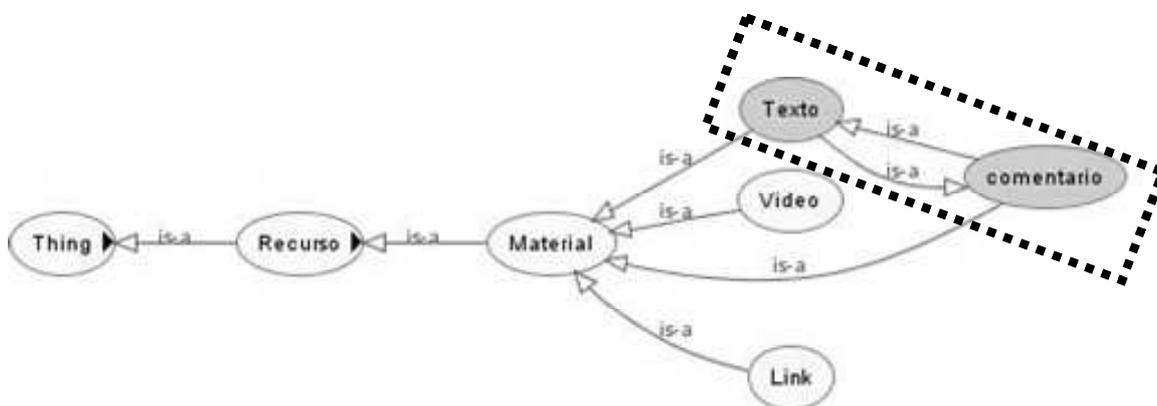


Figura 33: Representação gráfica da relação de Equivalência

Concluída a modelagem para as relações paradigmáticas, damos continuidade apresentando nossa proposta de modelagem dos *frames* na ontologia.

### e) *Frames*

Para desenvolvermos esta etapa do trabalho, foi necessário um estudo aprofundado da ferramenta a fim de verificar qual seria a melhor estratégia para inserir as informações sobre os *frames* na ontologia.

Partindo do pressuposto de que os *frames* são evocados pelos verbos e que relacionam um agente a um objeto através das ações, identificamos três componentes: atores (usuários), verbos (eventos) e objetos (recursos).

Em relação aos verbos, sabemos que o ambiente Moodle determina papéis específicos (permissões) para cada usuário. Em nossa análise, detectamos três tipos de permissões:

- I. Acesso aos conteúdos inseridos no ambiente – *leitura*. Por exemplo: o **Visitante** pode ler o conteúdo do **Fórum** ou o **Aluno** pode ler as mensagens colocadas no **Chat**;
- II. Interação com o conteúdo – *escrita*. Por exemplo: **Aluno** posta mensagem no **Fórum** ou **Professor** escreve comentário no **Diário**;
- III. Inserção de conteúdo – *publicação*. Por exemplo: **Professor** cria **Fórum** ou **Professor** elabora **Questionário**.

Identificados os componentes necessários para a inserção dos *frames* na ontologia, passamos a apresentar o modelo que projetamos para implementar as relações sintagmáticas da M\_ONTO no Protégé. A fim de melhor exemplificar, realizamos um recorte e apresentamos de que forma o *frame* **FÓRUM** será evocado pelo verbo POSTAR, além de demonstrarmos as classes, relacionamentos e instâncias envolvidas. Vale destacar que outros verbos também evocam o *frame* **FÓRUM**, mas exemplificaremos esta relação através do verbo POSTAR.

Para representar os atores, implementamos a hierarquia de classes apresentada na figura 34.

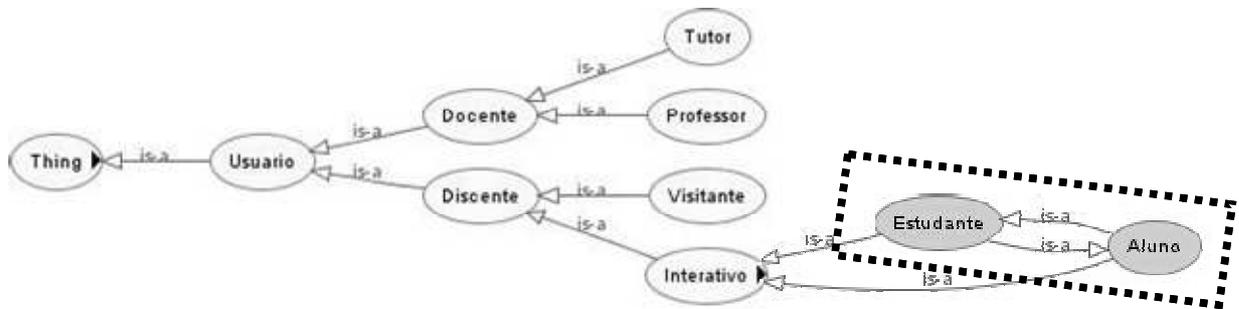


Figura 34: Classe **Usuário**

A hierarquia da classe **Usuário** foi organizada desta forma (dividindo entre **Docente** e **Discente**) em função das diferentes permissões que cada tipo de usuário tem no ambiente. A subclasse **Discente** foi subdividida em **Visitante** e **Interativo** para melhor acomodar a relação de equivalência entre as classes **Aluno** e **Estudante** e as permissões/restrições de cada tipo de usuário, pois esta organização permite possíveis alterações, caso a M\_ONTTO sofra atualizações. Por exemplo: caso, em algum trabalho futuro, seja necessário incluir uma classe **Academico** para representar que o estudante faz parte de uma Universidade, a estrutura já está adequada e permitirá que **Academico** seja incluído na subclasse **Interativo**, herdando as características da classe **Discente** e **Usuário**, respectivamente.

Dessa forma, podemos estabelecer as permissões, isto é, as propriedades restritivas de cada usuário. Para isso, criamos os membros (*members*) *leitura*, *escrita* e *publicação*, propriedades estas, que estão relacionadas à classe **Autorizacao** e atribuídas aos **Usuários**. Definimos que a classe **Usuário** tem *leitura* (propriedade “`temAutorizacao value leitura`” – figura 35), pela estrutura hierárquica, esta propriedade será herdada por todas as subclasses. Assim, fomos estabelecendo as propriedades restritivas de cada classe, exemplificado na figura 36 pelas restrições definidas para a classe **Professor**.

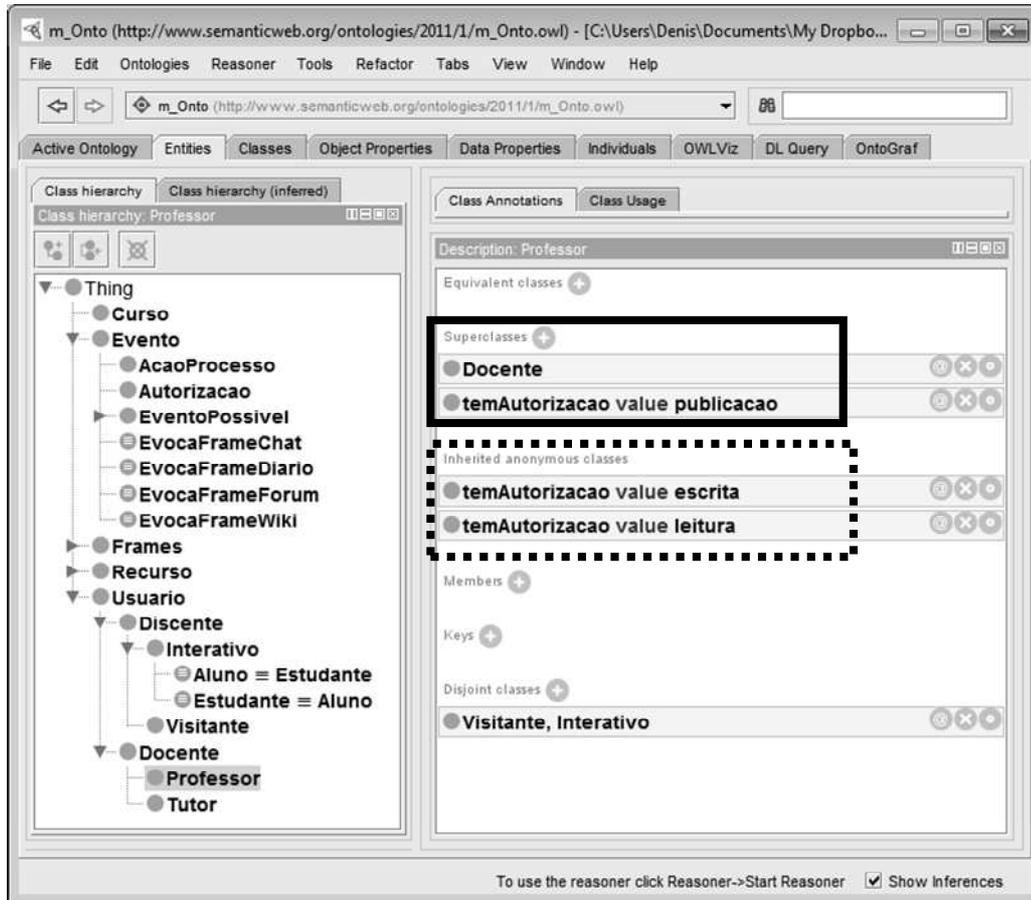


Figura 35: Propriedades restritivas – Classe **Professor**

As restrições de escrita e leitura (destacadas na figura pela caixa pontilhada) são herdadas da classe **Docente**. E a restrição publicacao foi associada somente à classe **Professor**, diferenciando-a das demais e permitindo que as relações entre os verbos e os frames possam ser formalizadas. Cabe aqui um esclarecimento acerca da classe **Tutor**. Já referimos anteriormente que, para o ambiente Moodle, o tutor é quem exerce a função de professor, mas no caso da UNISINOS temos atribuições diferentes para tutor e professor, sendo que o primeiro age como um auxiliar do segundo. Por esta ontologia estar diretamente relacionada ao ambiente EAD da UNISINOS, consideramos necessária a distinção entre estas duas classes – **Tutor** e **Professor**; porém, salientamos que nas descrições

encontradas para o ambiente Moodle (os manuais), tutor e professor desenvolvem as mesmas atividades.

Passamos agora à forma como realizamos a descrição dos verbos. De acordo com o que referimos na etapa linguística sobre a representação dos eventos na M\_ONTO, assumimos que eles podem ser descritos de duas formas diferentes, uma seguindo a abordagem paradigmática e outra a abordagem sintagmática. A primeira remete à criação de uma hierarquia simples de representar usando a relação *é-um* e a segunda estabelece a relação com os *frames*.

Optamos por representar os verbos através de membros (*members*), pois os mesmos não têm características de classes, ou seja, não conterão indivíduos (instâncias); eles próprios são os indivíduos da classe **AcaoProcesso** (figura 36).

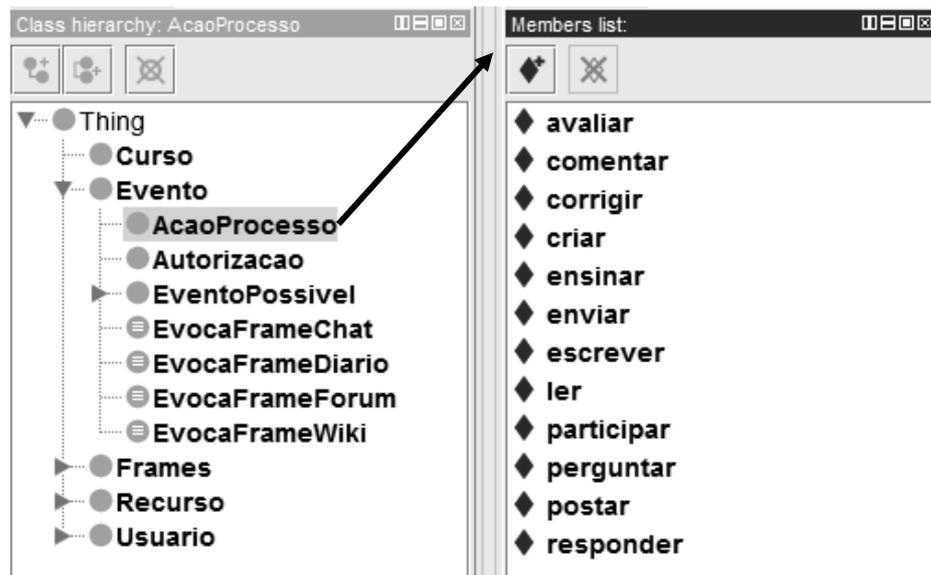


Figura 36: Verbos – instâncias classe AcaoProcesso

A atividade seguinte foi relacionar estes verbos (criados como membros) com a classe **Recursos**, pois definimos que para haver a interação é necessário que um ator relacione-se com um objeto. Para isso criamos novas classes que serviram para modelarmos estes relacionamentos. Escolhemos a representação do *frame FÓRUM* para exemplificar as soluções encontradas para que os verbos evoquem os *frames*.

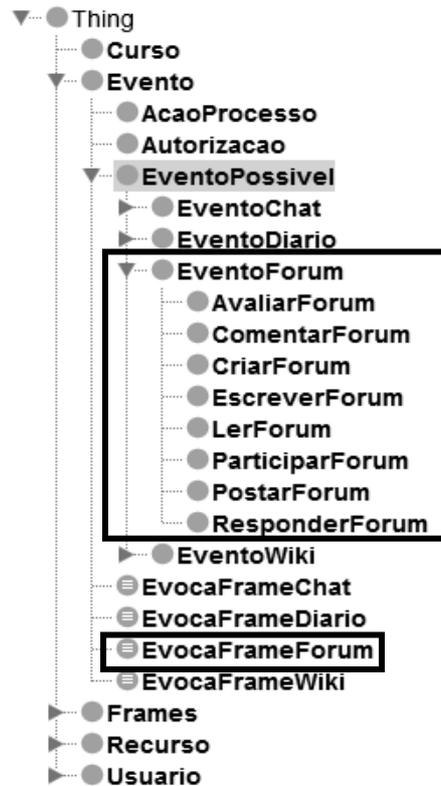


Figura 37: Hierarquia de classes – modelagem dos frames

Na figura 37, mostramos a hierarquia das novas classes que foram definidas para a modelagem dos *frames* na M\_ONTO, destacando as relacionadas ao *frame* que exemplificamos – *FÓRUM*.

Para o entendimento destas classes e suas funções, primeiramente precisamos esclarecer o conceito de *classe definida* e *classe primitiva*, de acordo com o Protégé 4.1.

As *classes definidas* são aquelas que têm pelo menos um conjunto de *condições necessárias e suficientes*, isto é, elas têm uma definição (condição) e qualquer indivíduo que satisfaça esta definição pertence à classe. As *classes primitivas* são aquelas que possuem apenas condições necessárias.

As classes que definimos até então possuem uma característica comum, elas são descritas apenas com *condições necessárias*, isto é, são classes primitivas.

No entanto, isso não é adequado às classes que criamos para modelar os *frames* na ontologia. Estas classes precisam ser descritas com *condições necessárias e suficientes*, para que o sistema possa realizar as inferências.

Apresentadas tais considerações, passaremos ao detalhamento da estratégia pensada para a modelagem dos *frames*.

Partimos do pressuposto de que os *frames* são evocados por verbos para criar a classe **EventosPossiveis**, que vai conter todas as ações possíveis de serem executadas no ambiente Moodle (figura 38). Os eventos a serem modelados estão relacionados às subclasses **EventoChat**, **EventoDiario**, **EventoForum** e **EventoWiki**, que contêm os eventos relativos a cada uma destas ferramentas, como pode ser visualizado no detalhamento de **EventoForum**. Organizamos os eventos desta forma (em grupos por ferramentas) apenas por uma questão de organização das informações, pois esta separação não é necessária e nem interfere na modelagem.

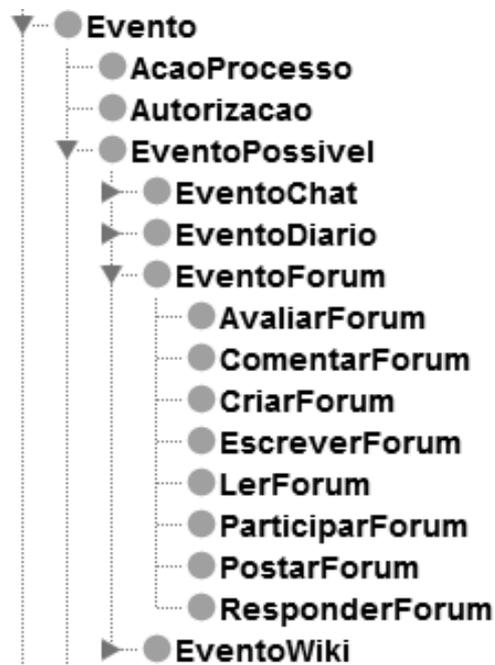


Figura 38: Classe **EventosPossiveis**

Na classe **PostarForum** descrevemos que ela tem que ter um ator pertencente à classe **Usuario**, e que tem que ter um objeto pertencente à classe **Forum** e que deve estar relacionada a uma ação de postar; e que este usuário deve ter autorização de escrita. A forma como estas regras foram introduzidas no Protégé pode ser visualizada na figura 39.

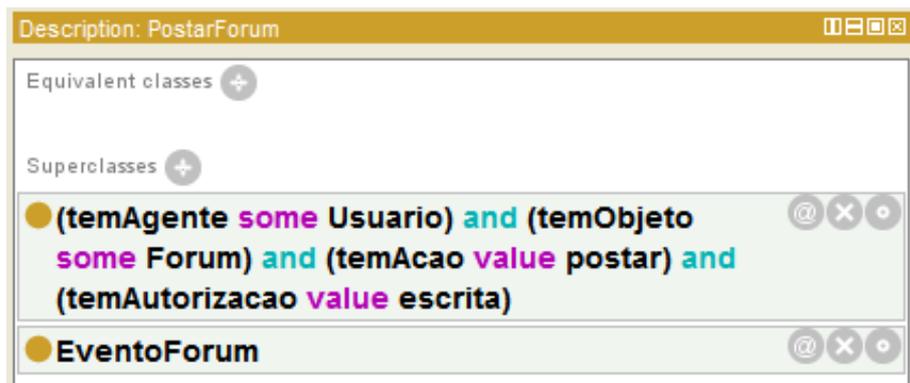


Figura 39: Descrição da classe **PostarForum**

A regra `temObjeto some Fórum`, que será inserida em todos os **EventosPossíveis**, é que permitirá a ligação entre os verbos e os frames.

Criamos as classes definidas do tipo **EvocaFrame** e todas elas têm a propriedade `temObjeto`. Isso significa que estas classes buscam a propriedade `temObjeto` nas demais que são filhas de **Evento**, tornando-se equivalentes a todas as classes que se encaixam nas suas próprias propriedades. Vejamos o exemplo de **EvocaFrameForum** na figura 40.

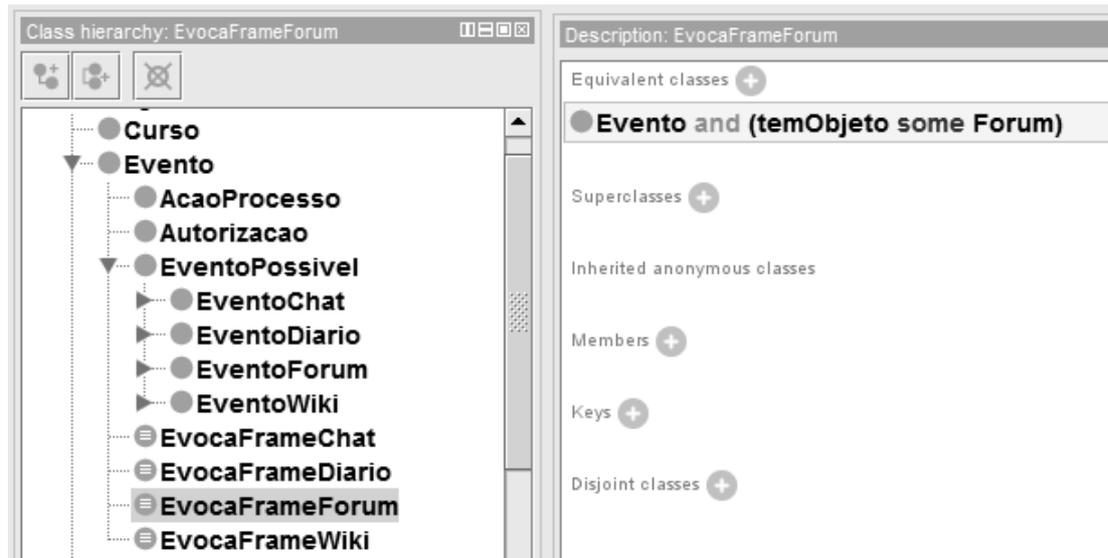


Figura 40: Descrição da classe **EvocaFrameForum**

Definimos que em **EventoForum** estarão todos os elementos do **Fórum**, isto é, os verbos, os elementos de *frame*<sup>87</sup> e as relações possíveis. Criamos a classe definida **EvocaFrameForum** para definir que sua classe equivalente deverá ser filha de **Evento** e temObjeto de **Fórum**.

Através desta estrutura que criamos, o sistema é capaz de inferir quem são os filhos de cada classe, possibilitando a ligação dos verbos aos frames.

Os relacionamentos com os demais *frames* foram modelados da mesma forma, de modo que o sistema consegue inferir sobre os verbos que evocam os frames, além de estabelecer, através dos **Eventos**, os relacionamentos existentes entre os **Usuarios** e os **Recursos**.

---

<sup>87</sup> Vale destacar que, neste estudo, apresentamos estratégias somente para a relação dos *frames* com o seu verbo evocador e que os aspectos relacionados aos elementos de *frame* são considerados possibilidades de estudos futuros.



paradigmáticas, criamos uma ontologia de base taxonômica e, ao estabelecermos os *frames* e as relações entre eles, criamos uma ontologia baseada em *frames*. A união destas duas resultou na M\_ONTO, sendo que a abordagem baseada em *frames* mostrou-se muito eficaz para o estabelecimento das relações existentes entre os atores, os verbos e os objetos.

Encerramos este capítulo confirmando nosso pressuposto de que os *frames* podem ser representados e formalizados através da linguagem OWL e tornarem-se importantes elementos para a representação dos relacionamentos em ontologias.

A seguir, apresentamos algumas considerações finais sobre esta pesquisa, destacando as possibilidades de trabalhos futuros.

## 5 CONSIDERAÇÕES FINAIS

Neste estudo, ocupamo-nos da construção de uma ontologia de domínio, com ênfase na investigação de sua modelagem semântica. Nossa pesquisa deu-se no âmbito da Linguística, da Computação e da Ciência da Informação, promovendo um debate entre estas áreas a fim de explorar os conceitos e verificar a aplicabilidade da Semântica Lexical e da Semântica de Frames ao nosso propósito.

Nosso objetivo com este estudo foi buscar na Semântica uma abordagem para representar a descrição e conceituação do domínio EAD, possibilitando a construção da ontologia. Além disso, buscamos soluções na Semântica Lexical e na Semântica de Frames para melhor definir e estabelecer os relacionamentos entre os conceitos referentes ao domínio, os quais foram modelados no Protégé.

Balizamos nossa pesquisa pelas três questões norteadoras conforme já apresentado nas considerações iniciais e organizamos o trabalho de modo a buscar respostas que satisfizessem nossos questionamentos e possibilitassem a construção da M\_ONTO.

Iniciamos tratando das questões gerais sobre ontologias a fim de situar o contexto no qual nossa pesquisa está inserida. No capítulo 2 buscamos subsídios para responder nossa primeira questão: “Qual a importância das ontologias para o aperfeiçoamento das tecnologias da informação, mais precisamente os AVA’s?”.

Com o intuito de melhor responder esta questão, procuramos definir o termo *ontologia* buscando a contribuição das áreas da CC, da CI e também da Linguística. Apresentamos termos correlatos, tais como taxonomia e tesouro, destacando as principais diferenças em relação às ontologias. Tais distinções mostraram-se importantes no sentido de elucidar a afirmativa de que **toda a ontologia é composta por uma taxonomia** e de demonstrar que os tesouros, apesar de serem

muito semelhantes às ontologias, não transcendem a representação dos conteúdos formais, ao contrário das ontologias, que assumem um importante papel para os sistemas informatizados de PLN.

Os aspectos teóricos apresentados no capítulo 2 foram relevantes para respondermos ao nosso questionamento inicial acerca da importância do uso de ontologias para aperfeiçoamento dos AVA's. Através dos agentes inteligentes ou pedagógicos, os AVA's tornam-se mais atrativos e aumentam a possibilidade de interação entre os usuários e o sistema, sendo que tais aspectos são ampliados através do uso de ontologias. Normalmente os agentes realizam consultas às bases de dados dos sistemas para inferirem suas ações; junto a estas bases de dados, podemos incluir as ontologias, que se tornam um precioso recurso para ampliar as relações existentes e aumentar a potencialidade do sistema.

Tendo esclarecido nosso posicionamento teórico em relação às ontologias, sua aplicabilidade e importância para o nosso domínio, passamos a buscar respostas para a segunda questão: "Que abordagens se mostram mais adequadas à descrição semântico-conceitual do domínio em questão a partir da Semântica?"

Dedicamos o capítulo três a responder a esse questionamento, trazendo reflexões teóricas acerca do conteúdo da M\_ONTO, ou seja, refletimos sobre diferentes aspectos da semântica apresentando como as áreas da CC e da CI tratam destas questões e estabelecendo um paralelo com a Linguística.

Levando em consideração as questões teóricas discutidas, percebemos que as diferentes áreas consideram a semântica como fundamental na construção de ontologias, enfatizando as relações paradigmáticas. As relações sintagmáticas, consideradas em nossa investigação como as mais representativas e adequadas para a descrição das entidades verbais, são pouco exploradas pelas áreas da CI e da CC, sendo representadas somente através de relações associativas.

No nosso estudo, as relações sintagmáticas tiveram uma grande influência na maneira como a ontologia foi construída e caracterizaram-se como uma forma diferenciada de estabelecer os relacionamentos entre as classes.

A Semântica Lexical muito contribuiu na nossa análise dos nominais elencados como classes da nossa ontologia, com especial destaque para as relações de hiponímia e equivalência. A meronímia mostrou-se útil apenas nos relacionamentos entre CURSO, USUÁRIOS e PROCESSOS, exemplo que detalhamos no capítulo 4.

Um dos aspectos de grande destaque em nossa análise foi a descrição semântica baseada em *frames*, recurso este que propiciou a ampliação da descrição de relacionamentos na M\_ONTO. Como pudemos comprovar, os verbos nos levam a diferentes *frames* que auxiliam na descrição das cenas de educação e, através dos *frames* que propusemos – *EAD-TEACHING*, *CHAT*, *FÓRUM*, *DIÁRIO* e *WIKI* –, ampliamos o escopo da ontologia, tornando-a muito mais completa e consistente. Concluímos nossa análise, apresentando uma forma de relacionar os *frames* à taxonomia através dos verbos evocadores de *frame*, demonstrando que é possível incluir este tipo de estrutura em ontologias. Diante disso, podemos destacar o valor dos *frames* não apenas como importantes elementos para modelar a M\_ONTO, mas para outras ontologias de diferentes domínios.

O fato de levarmos em conta a Semântica Lexical e a Semântica de Frames nos levou a fragmentar a M\_ONTO em duas ontologias menores, sendo uma totalmente taxonômica (o mais natural quando se fala em ontologias) e a outra capaz de cobrir uma gama de relações que levam a novos conceitos, promovendo a criação de uma grande rede de relacionamentos. Ambas, interligadas através dos verbos (evocadores de *frames*), constituem a M\_ONTO.

O quarto capítulo, através da análise do editor de ontologias Protégé e do estudo da melhor forma de representar as relações definidas como estruturantes da nossa ontologia, nos auxiliou a responder o questionamento: “Como modelar os dados de forma consistente em um editor próprio para a construção de ontologias?”.

O Protégé-OWL mostrou-se uma ferramenta adequada para a formalização das relações paradigmáticas, sendo estas de fácil representação. Para a descrição das relações sintagmáticas, isto é, as relações entre *frames*, foi necessário estabelecermos novas classes e buscar alternativas diferenciadas para representar os verbos e a forma como estes evocam cada um dos *frames* relacionados à

M\_ONTO. Estas diferentes formas de representação enriqueceram a nossa ontologia, estabelecendo as bases ou modelos para possíveis ampliações.

Por fim, no que tange às limitações desta pesquisa, consideramos o fato de não ter explorado a forma como os *frames* e as classes se inter-relacionam através dos elementos de *frame*. Outra limitação foi em relação aos demais *frames* citados, os quais não foram detalhados nesta pesquisa em função da grande abrangência dos mesmos. Tais limitações podem ser vistas como desafio para o desenvolvimento de trabalhos futuros, ampliando a potencialidade da descrição semântica através dos frames e seus elementos.

Concluimos acrescentando que uma das contribuições desta pesquisa é apresentar a Semântica de Frames como uma abordagem válida para a criação de ontologias e, diante disso, visualizamos a possibilidade de dar continuidade a esse estudo de como incluir os *frames* na modelagem de ontologias, buscando ampliar os estudos realizados até então. Esta perspectiva nos remete à possibilidade de ampliação da M\_ONTO incluindo novos frames e maior detalhamento destas estruturas.

## REFERÊNCIAS BIBLIOGRÁFICAS

ALLAN, K. *Natural Language Semantics*. Malden: Blackwell Publishers Inc, 2001.

ALMEIDA, M.B.; BAX, M.P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. *Ciência da Informação*. v. 26, n. 1. p. 39-45, 2003a.

ALVES, I. M. da R. *O uso da semântica verbal em sistemas de extração de informação: a construção de uma ontologia do domínio jurídico*. Dissertação de mestrado. São Leopoldo: UNISINOS, 2005.

ANSI Z39-19-2005. *Guidelines for the construction, format, and management of monolingual controlled vocabularies*. Bethesda: NISO Press, 2005.

ARAUJO, M. de. *Educação à distância e a Web Semântica: modelagem ontológica de materiais e objetos de aprendizagem para a plataforma COL*. 2003. Tese (Doutorado). Escola Politécnica da Universidade de São Paulo. São Paulo. Disponível em: < [www.teses.usp.br](http://www.teses.usp.br) > Acessado em: mai 2006.

ARAUJO FRANCO, M., CORDEIRO, L. M. e CASTILLO, R. A. F. O ambiente virtual de aprendizagem e suas incorporações na Unicamp. *Educação e Pesquisa*. v. 29 n. 2. São Paulo, FE/USP, jul-dez 2003.

BAUM, D. Semantic Break Through. *Oracle Magazine*, California, V. 20, n. 3, p. 42-46, mai./jun. 2006.

BEAUGRANDE R. A de e DRESSLER, W. U. *Introduction a La Lingüística dês Texto*. Barcelona: Ariel, 1998.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. *The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*. Scientific American. 2001. Disponível em: <<http://www.cs.nyu.edu/rgrimm/teaching/reading/semantic-web.pdf>> Acesso em: dez 2009.

BOGO, A. *Arquitetos de Sonhos*. São Paulo: Expressão Popular, 2003.

BORBA, F. S. *Dicionários de Usos do Português do Brasil*. São Paulo: Ed. Ática, 2002.

BORBA, F. S. *Uma Gramática de Valências para o Português*. São Paulo: Ed. Ática, 1996.

BORST, W. N. *Construction of engineering ontologies*. 1997. 243 f. Tese (Doutorado). – University of Twente, Enschede, 1997. Disponível em: <<http://doc.utwente.nl/fid/1392>> Acesso em: jun 2009.

BREITMAN, K. *Web Semântica: a Internet do Futuro*. Rio de Janeiro: LTC, 2005.

CANÇADO, M. *Manual de semântica: noções básicas e exercícios*. Belo Horizonte: Editora UFMG, 2005. 185.p.

CHAVES, E. O. *Ensino a distância: conceitos básicos*. 1999, p. 2-12. Disponível em: <[http://www.edutecnet.com.br/edconc.htm#Ensino a Distancia](http://www.edutecnet.com.br/edconc.htm#Ensino%20a%20Distancia)>. Acesso em: dezembro 2009.

CHISHMAN, R.; ALVES, I. M. da R.; BERTOLDI, A. *O Conhecimento Semântico Representado em Ontologias Aplicadas à Busca e Extração de Informação Na Web*. In: Anais do XI Simpósio Nacional & 1º Simpósio Internacional de Letras e Linguística. Uberlândia, 2006.

CHISHMAN, R. L. O. ; ALVES, I. M. da R. Extração de Informação e Web Semântica. *Calidoscópico*, São Leopoldo, v. 3, n. 1, p. 47-52, 2005.

COLLINS, A.M. E QUILLIAN, M.R.. How to Make a Language User. In: E. Tulving e W. Donaldson, editores, *Organisation of memory*. MIT Press, 1972.

CROFT, W.; CRUSE, D. A. *Cognitive Linguistics*. Cambridge Textbooks in Linguistics. 2004

CRUSE, D. A. *Lexical Semantics*. Cambridge: Cambridge University Press, 1986.

\_\_\_\_\_. *Meaning in Language: an Introduction to Semantics and Pragmatics*. New York: Oxford University Press, 2000.

DACONTA, M. C.; OBRST, L. J.; SMITH, K. T. *The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management*. Wiley Publishing: 2003.

DAHLBERG, I. A Referent-oriented analytical concept theory of interconcept. *International Classification*, Frankfurt, v.5, n.3, p.142-150, 1978a.

\_\_\_\_\_. *Ontical structures and universal classification*. Bangalore: Sarada Ranganthan Endowment, 1978b.

Di FELIPPO, A. *Representação Linguístico-Computacional dos Adjetivos Valenciais do Português*. Araraquara, 107p. Dissertação (Mestrado em Letras) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2004.

\_\_\_\_\_. Ontologias linguísticas aplicadas ao processamento automático das línguas naturais: o caso das redes wordnets. In: Magalhães, J. S.; Travaglia, L. C. (Orgs). *Múltiplas perspectivas em Linguística*. Uberlândia: Edufu, 2008.

- DOWTY, D. *Word Meaning and Montague Grammar*. Dordrecht: D. Reidel, 1979.
- EVENS, M. W. *Relational Models of the Lexicon*. Cambridge: Cambridge University Press, 1988.
- FELLBAUM, C. A Semantic Network of English Verbs. In: Fellbaum, Christiane. *WordNet: An Electronic Lexical Database*. Cambridge: MIT Press, 1998.
- FELTES, H. P. de M. *Semântica Cognitiva: ilhas, pontes e teias*. Porto Alegre: EdipucRS, 2007.
- FILLMORE, C. Em Favor do Caso. In: LOBATO, L. *A Semântica na Linguística Moderna: o Léxico*. Rio de Janeiro, Francisco Alves, 1977.
- \_\_\_\_\_. "Frame Semantics". The Linguistic Society of Korea. *Linguistic in the Morning Calm*, Seoul, Hansinh Publishing Co, 1982.
- FILLMORE C. et al. *FrameNet: Theory and Practice*. Universidade da Califórnia of California de Berkeley: Instituto Internacional de Ciências da Computação. Disponível em: <<http://www.icsi.berkeley.edu/~framenet/book/book.html>> 2003. Acesso ao longo do período de construção da dissertação.
- FILLMORE C.J.; BERYL T. e ATKINS. Toward a *Frame-Based* Lexicon: Semantics of Risk e its Neighbors. In Adrienne Lehrer and Eva Efer Kittay (eds). *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization*. Hillsdale: Lawrence Erlbaum. pp. 75-102, 1992.
- GEERAERTS, D. Cultural models of linguistic standardization. In: Dirven, R., Frank, R. e Pütz, M. (eds.), *Cognitive Models in Language and Thought. Ideology, Metaphors and Meanings*. Berlin/New York: Mouton de Gruyter, 25-68. 2003
- GOMES, H. E.; CAMPOS, M. L. de A.; SALES, L.. Ontologias de domínio: um estudo das relações conceituais. *Perspectivas em Ciência da Informação*, v. 13, n. 2, p. 62 – 76, maio/agosto 2008.
- GOMEZ-PEREZ, A; CORCHO, O; FERNANDEZ-LOPEZ, M. *Ontologicalm Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Heidelberg: Springer-Verlag. 2004.
- GREEN, R. Relationships in the organization of knowledge: an overview. In: Bean, A.; Green, R (Eds.) *Relationships in the organization of knowledge*. Dordrecht: Dordrecht: Kluwer Academic Publishers, 2001. p. 3-18
- GRUBER, T. R. A translation approach to portable ontology specifications. *Knowledge Acquisition*, v.5, 1993. p.199-220.
- GRUBER, T. *What is an ontology?* [S. l. : s. n.], 1996. Disponível em: <<http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>>. Acesso em: mar. 2009
- GRISHMAN, R. *Computational Linguistics: an Introduction*. Cambridge: Cambridge University Press, 1986.

GUARINO, N. Formal Ontology, Conceptual Analysis and Knowledge Representation. *International Journal of Human Computer Studies*, special issue on formal Ontology, conceptual analysis and knowledge representation. 1995.

GUARINO, N.; GIARETTA, P. Ontologies and knowledge bases: Towards a terminological clarification. In: MARS, N.J.I. (Ed.). *Towards very large knowledge bases*, IOS Press, 1995. p. 25-32. Disponível em: <<http://www.loacnr.it/Papers/KBKS95.pdf>>. Acesso em: out. 2010.

GUSMÃO, H. R. *Tesaurus: análise e utilização*. Niterói: UFF, 1985.

HJØRLAND, B. Domain analysis in information science: Eleven approaches - traditional as well as innovative. *Journal of Documentation*, v. 58, n. 4, p. 422- 462, 2002.

\_\_\_\_\_. Semantics and knowledge organization. *ARIST*, v.41, cap.8, p.367-405, 2007a.

\_\_\_\_\_. *Semantic relations* (meaning relations). 2007. Disponível em:<[http://www.db.dk/bh/Lifeboat\\_KO/CONCEPTS/semantic\\_relations.htm](http://www.db.dk/bh/Lifeboat_KO/CONCEPTS/semantic_relations.htm)>. Acesso em: 2008. 2007b.

JQUES, P; VICARI, R. (2005) “Estado da Arte em Ambientes Inteligentes de Aprendizagem que Consideram a Afetividade do Aluno”. In: *Informática na educação: Teoria e Prática*, Universidade Federal do Rio Grande do Sul, Porto Alegre, v. 8, n. 1, p. 15-38.

JASPER, R e USCHOLD, M. A *Framework* for Understanding and Classifying Ontology Applications. In.: *KAW'99 Twelfth Workshop on Knowledge Acquisition, Modeling and Management*. 1999. Voyager inn, Banff, Alberta, Canadá. Disponível em <<http://sern.ucalgary.ca/KSJ/KAW/KAW99/papers/Usehold2/final-ont-apnfmk>>

KEARNS, K. *Semantics*. New York, St. Martin Press. 2000

KLEIBER, G. *La Sémantique du Prototype. Catégories et sens lexical*, Paris, Presses Universitaires de France. 1990

KHOO, C.; NA, J. C. Semantic Relations in Information Science. *Annual Review of Information Science and Technology*, 2006.

LAKOFF, G. *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*, Chicago, The University of Chicago Press. 1987

LEVIN, B. Contribution of Linguistics. In: Bates, Madelaine e Weischedel, Ralph M. (eds). *Challenges in Natural Language Processing*. Cambridge: Cambridge niversity Press, p. 76-98, 1993

LOPES, L.; VIEIRA, R.; FINATTO, M. J.; ZANETTE, A.; MARTINS, D.; RIBEIRO Jr., Extração automática de termos compostos para construção de ontologias: um experimento na área da saúde. *Reciis- Revista Eletrônica de Comunicação Informação & Inovação em Saúde*, v.3, p.76 - 88, 2009.

LYONS, J. *Semântica*. Lisboa: Editorial Presença, 1977.

MARRAFA, Palmira. *WordNet do Português*. Instituto Camões, Lisboa, 2001.

MINSKY, M. A. *A Framework for Representing Knowledge*. Artificial Intelligence Memo 306, MIT AI Lab, 1974

MORAN, J. M. *O que é Educação a Distância?* Disponível em: <<http://www.eca.usp.br/prof/moran/dist.htm> 2002>. Acesso em: nov. 2008.

MOREIRA, A.; ALVARENGA, L.; OLIVEIRA, A. de P.. O nível do conhecimento e os instrumentos de representação: tesouros e ontologias. *Datagramazero: Revista de Ciência da Informação*, v. 5, n. 6, dez. 2004.

MURPHY, G. L. *The Big Book of Concepts*, MIT Press, A Bradford book, 2002.

NIRENBURG, S. & RASKIN, V. *Ontological Semantic*. Cambirdge University Press, 2005.

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

PRIA, A. D. A Relação entre o desenvolvimento de formalismos gramaticais de base lexicalista e as exigências do PLN. *Delta: Documentação de Estudos em Linguística Teórica e Aplicada*. Vol 24, nº 2, São Paulo, 2008.

RANGANATHAN, S. R. *Prolegomena to library classification*. New York: Asia Publishing House, 1967.

RECTOR, M.; YUNES, E. *Manual de semântica*. Rio de Janeiro: Ao Livro Técnico, 1980.

ROSCH, E. *Principles of categorization*. In: E. Rosch & B. Lloyd (eds.), pp. 27-48. 1978.

ROSCH, E. & MERVIS, C. B. Family resemblances: studies in the internal structure of categories. *Cognitive Psychology* 7, pp. 573-605. 1975.

\_\_\_\_\_. Basic objects in natural categories. *Cognitive Psychology* 8, pp. 382-439. 1976.

SAEED, J. *Semantics*. Oxford: Brasil Blackwell, 2003.

SAINT-DIZIER, P. e VIEGAS, E. *Computational Lexical Semantics*. Cambridge: Cambridge University Press, 1995.

SALES, R. de; CAFÉ, L.. *Tesouros e Ontologias sob o olhar da teoria comunicativa da terminologia*. Anais IX Congress ISKO-SPAIN. Valencia, 2009

SANTOS, E. O.. Ambientes virtuais de aprendizagem: por autorias livres, plurais e Gratuitas. *Revista FAEBA*, v.12, no. 18. 2003.

SCHALLEY, A. C; ZAEFFERER, D. *Ontolinguistics - An outline*. In: Schalley, Andrea C. and Dietmar Zaefferer. (ed.s), 2007.

SILVA, A. S. da. Metáfora, metonímia e léxico. *Diacrítica 7* (Revista do Centro de Estudos Humanísticos da Universidade do Minho), pp. 313-330. 1992

\_\_\_\_\_. A linguística cognitiva: uma breve introdução a um novo paradigma em lingüística. *Revista Portuguesa de Humanidades*, v.1, 1997. Disponível em: <<http://www.facfil.ucp.pt/lingcognit.htm>>. Acesso em: 10 set. 2008.

SILVA, D. S; SOUZA, R. R., ALMEIDA, M.B. *Princípios metodológicos para construção de ontologias: uma abordagem interdisciplinar*. Encontro Nacional De Pesquisas Em Ciência Da Informação - Enancib, 10, Anais, João Pessoa, 25 a 28 de outubro de 2009.

SMITH, B.; CEUSTERS, W.; KLAGGES, B.; KÖHLER, J.; KUMAR, A.; LOMAX, J.; MUGALL, C.; NEUHAUS, F.; RECTOR, A.; ROSSE, C. Relations in biomedical ontologies. *Genome Biology*, v. 6, n. 5. 2005.

STOJANOVIC, L.; STAABN, S.; STUDER, R. eLearning based on the Semantic Web. WebNet2001- World Conference on the WWW and Internet, 2001.

UNGERER, F. & SCHMID, H. Levels of Categorization. In: *An Introduction to Cognitive Linguistics*. Great Britain: Longman, 2007, 2ª edição.

VENDLER, Z. 1969. Adjectives and nominalizations. The Hague/ Paris, Mouton.

VIEIRA, R. E LIMA, V. L. S. Linguística computacional: princípios e aplicações. In: *IX Escola de Informática da SBC-Sul*. Luciana Nedel (Ed.) Passo Fundo, Maringá, São José. 2001.

VICKERY, B. C. Ontologies. *Journal of Information Science*, v.23, n.4, p.227-286, jan. 1997.

VOSSSEN, P. EuroWordNet: A Multilingual Database for Information Retrieval. In: *THIRD DELOS WORKSHOP Cross-Language Information Retrieval*, pp. 85-94. European Research Consortium For Informatics and Mathematics. 1997 Disponível em: <<http://www.ercim.org/publication/ws-proceedings/DELOS3/Vossen.pdf>>. Acesso em: 01 mar. 2003.

WINSTON, E.M.; CHAFFIN, R.; HERRMANN, D. A Taxonomy of Part-Whole Relations, in *Cognitive Science*, 11, 1987.

WITTGENSTEIN, L. *Philosophical Investigations*, Oxford, Basil Blackwell. 1953.

WOOLDRIDGE, M. & JENNINGS, N. R. Intelligent Agents: Theory and practice. *Knowledge Engineering Review*, 1995