

O CONHECIMENTO SEMÂNTICO REPRESENTADO EM ONTOLOGIAS APLICADAS À BUSCA E EXTRAÇÃO DE INFORMAÇÃO NA WEB

Rove CHISHMAN¹ (UNISINOS)

Isa Mara da Rosa ALVES² (UNISINOS e UNESP-Ar)

Anderson BERTOLDI³ (UNISINOS)

RESUMO: Este artigo tem como objetivo central refletir sobre a tarefa de codificação da descrição semântica no desenvolvimento de ontologias. Para tanto, discute a indefinição do conceito de ontologia e destaca seu papel em sistemas de busca e extração de informação na *web*. Através do relato de três propostas de estruturação ontológica – uma enfocando os nominais do domínio *universidade* e duas enfocando os verbos e os adjetivos do domínio *jurídico*, destaca também que o desempenho destas aplicações computacionais requer descrições do conhecimento teoricamente bem fundamentadas. A noção de *semântica ontológica* serve de sustentação para a perspectiva metodológica assumida nos experimentos relatados.

ABSTRACT: The main purpose of this article is to reflect upon the task of semantic description codification to structure ontologies. In order to do so, we discussed de proplems about the definition of ontololgy highlighting contribution of ontologies for the improvement of web information retrieval systems. Based on three ontological structuring proposals – one focusing on nominals belonging to the *university* domain and two focusing on verbs and adjectives belonging to the juridical domain -, we also highlight that the performance of these computational applications require theoretically well founded knowledge descriptions. The notion of ontological semantics is the rationale the methodology adopted in the described here experiments.

1. Introdução

Este artigo, tomando como base três propostas de estruturação ontológica – uma enfocando os nominais do domínio *universidade* e duas enfocando os verbos e os adjetivos do domínio *jurídico* -, tem como objetivo central ilustrar o papel que a Linguística desempenha, mais precisamente os estudos semânticos, na tarefa genuinamente interdisciplinar de desenvolvimento de ontologias. Através do relato desses três estudos, pretendemos demonstrar que o sucesso do desempenho de sistemas de recuperação e extração de informações na *web*, aplicação sobre a qual temos nos debruçado, está em grande medida na dependência de escolhas teóricas definidas tendo em vista a boa cobertura do domínio em questão. Para tanto, organizamos o artigo em dois blocos: o primeiro, compreendendo as seções 2 e 3, traz um reflexão sobre o lugar que as ontologias ocupam em sistemas de extração de informação e apresenta a perspectiva da semântica ontológica, termo cunhado por Nirenbrug & Raskin (1999, 2005) que expressa um complexo integrado de teorias, metodologias, descrições e implementações. O segundo bloco, compreendendo as demais seções, relata a descrição semântica, baseada em *corpus*, de nomes, verbos e adjetivos, seguindo abordagens semânticas computacionalmente representáveis. Destacamos, nessa segunda parte, as escolhas teóricas feitas e os principais desafios em se tratando de representação do conhecimento semântico em uma ontologia.

2. Ontologias e sistemas de busca e extração de informação na *web*

No âmbito da Filosofia, ontologia é o estudo do que existe e do que devemos assumir como existentes a fim de se obter uma descrição da realidade. Já no contexto da Engenharia do Conhecimento, não se distanciando totalmente da acepção filosófica, o termo *ontologia* tem recebido muitas definições. Uma das definições mais citadas na literatura é a proposta por Gruber (1993): “Uma ontologia é uma especificação formal explícita de uma conceitualização compartilhada.” Uma conceitualização ou modelo conceitual, nessa perspectiva, remete a uma visão abstrata do mundo que se objetiva representar em um sistema

¹ E-mail: rove@unisinis.br

² E-mail: isamralves@gmail.com

³ E-mail: andersonbertoldi@yahoo.com.br

computacional. Uschold (1999), por sua vez, associa as ontologias a um vocabulário de termos e à especificação de seus significados. Em linhas gerais, pode-se assumir também que uma ontologia é um documento ou arquivo que define formalmente as relações entre os termos, apresentando uma taxonomia e um conjunto de regras de inferência.

Nas seções subseqüentes, contudo, poderemos perceber, através dos relatos dos experimentos, que a modelagem ontológica não se dá de forma tão previsível, conforme sugerem as definições. Há uma série de abordagens, além da semântica de relações taxonômicas, também conhecidas como relações do tipo É-UM ou inclusão de classes, que se prestam à representação semântica nas diferentes propostas de ontologia. Entre as decisões metodológicas que afetam essas escolhas teóricas, enfatizamos a captura dos termos e conceitos e a própria formalização dos conceitos. Isso significa, por exemplo, que o *design* de uma ontologia de um sistema de tradução tem particularidades não encontradas em uma ontologia de um sistema de busca na internet, aplicação de interesse dos estudos relatos neste artigo. Há níveis distintos de granularidade do sistema, haja vista as diferentes necessidades de cobertura descritiva. Questões como essas nos permitem ressaltar que, no contexto computacional, o papel assumido pelas ontologias é relativamente distinto do papel exercido na filosofia. As ontologias dos sistemas de informação têm caráter funcional, característica não almejada nos estudos filosóficos.

Feitas essas considerações iniciais acerca da ambigüidade conceitual do termo *ontologia*, convém também chamar a atenção para a ampla gama de aplicações que têm se beneficiado do desenvolvimento de ontologias. Segundo Benjamins & U. (1999), a análise ontológica, ao explicitar a estrutura do conhecimento, representa o componente central de qualquer sistema, de onde se depreende que análises inconsistentes e superficiais compõem bases de conhecimentos incoerentes também. Sistemas de recuperação de informação, bibliotecas digitais, integração de fontes heterogêneas de informação e motores de busca na internet necessitam de ontologias de domínio para organizar a informação e orientar os processos de busca.

O crescimento em progressão geométrica de coleções de textos eletrônicos tem dificultado a localização de documentos relevantes. A maioria dos motores de busca adota sistemas de indexação baseados em palavras-chave, estratégia que nem sempre atende às expectativas dos usuários em termos de busca bem sucedida. A anotação dos documentos armazenados na *web* com informação semântica estruturada tem se mostrado uma solução promissora para fornecer automaticamente seções ou páginas conceitualmente similares. Como bem enfatiza Berners-Lee (2001), em seu clássico artigo introduzindo a tecnologia da *web semântica*, as ontologias podem contribuir para a concretização dessa idealização. Elas podem ser usadas para melhorar a acurácia das busca à medida que o programa acessa somente as páginas relevantes.

Os experimentos apresentados neste artigo são aplicações em sistemas de extração de informações. A proposta de estruturação ontológica que apresentamos para os nominais tem como meta melhorar um sistema de busca na internet relacionado ao domínio *universidade*. Já a descrição dos verbos e dos adjetivos tem como meta o desenvolvimento de uma ontologia de domínio *jurídico* e visa contribuir para a melhoria do desempenho de sistemas de busca e extração de informações na *web*, em especial, o sistema da Procuradoria Geral da República de Portugal (PGR). Este estudo reúne pesquisadores da Linguística do Brasil e da Informática de Portugal que colaboram no âmbito de projetos maiores unindo alguns países da União Européia no intuito de construir uma *wordnet* multilíngüe especializada para o domínio jurídico. Espera-se, com isso, que a busca de informações desse domínio, em sites jurídicos oficiais desses países em suas próprias línguas (português, italiano, inglês, checo e holandês), seja facilitada.

3. Justificando a metodologia

Considerando que o objetivo principal deste trabalho é ilustrar o papel dos estudos semânticos no desenvolvimento de ontologias, reservamos esta seção para refletir sobre algumas escolhas metodológicas. Interessa-nos, em especial, discorrer sobre um estágio central da tarefa de construir uma ontologia, que é a fase de codificação, ou seja, fase em que se define o formato das descrições semânticas que comporão o trabalho.

Antes, porém, de tratar desse ponto, convém fazer menção a aspectos mais gerais relacionados à metodologia, aos quais também estamos atentos. McGuinness & Noy (1999), embora enfatizem que não há uma única forma correta de desenvolver ontologias, propõem quatro questões para nortear a fase inicial do processo: (1) Qual o domínio que a ontologia pretende cobrir?, (2) Para que a ontologia será usada?, (3) Para que tipo de questões a informação na ontologia deve fornecer respostas? E (4) Quem vai usar e manter a ontologia?

King & Uschold (1995), por sua vez, preocupados com o todo da tarefa, propõem quatro estágios: (1) identificação da proposta; (2) construção da ontologia: captura da ontologia, codificação da ontologia e integração de ontologias existentes; (3) avaliação e (4) documentação.

A seção anterior, ao fazer referência aos sistemas de busca e extração de informação na *web* e aos domínios abordados, já indica a preocupação do grupo em atender as etapas apontadas pelos autores. Os relatos apresentados a seguir trazem maiores detalhes sobre os demais estágios, em especial, a descrição da codificação ontológica e a formalização no editor de ontologias Protégé.

Central para a fase de codificação da ontologia é a tarefa de escolher as abordagens que melhor se ajustam à aplicação. Para explicar melhor essa questão, remetemo-nos a Niremburg & Raskin (2005), que empregam a denominação *semântica ontológica* para se referirem às abordagens semânticas que se prestam ao tratamento do significado computacionalmente. O conceito de semântica ontológica está relacionado a uma sociedade de microteorias, cujas descrições levam em conta não apenas as teorias, mas as exigências das aplicações em si. Em outras palavras, a semântica ontológica, tendo um caráter híbrido e versátil, é guiada pela necessidade de certas tarefas de processamento da linguagem, como manipulação, análise e geração de textos.

Assumindo essa perspectiva comprometida com a demanda do PLN, afastamo-nos, em certo sentido, da visão adotada pelas teorias genuinamente lingüísticas, ou seja, teorias cujo principal critério de validação é a adequação explanatória. Em se tratando de teorias comprometidas com a computação, o critério último de validação é o sucesso das aplicações pretendidas. Isso significa, por exemplo, que uma abordagem semântica para PLN deve se ajustar às necessidades tanto de sistemas de geração como de sistemas de compreensão da linguagem.

Niremburg & Raskin (2005), atentando para essa tensão entre teorias comprometidas com a generalização e com ênfase na relação léxico/sintaxe e teorias comprometidas com o conteúdo semântico com vistas à computação, empregam a denominação *method-driven theory* ou *supply-side theory*, para se referirem às primeiras, e *ontology-driven lexical semantics* ou *demand-side theory*, para se referirem às que têm o compromisso com as aplicações.

As primeiras são seletivas e primam pela generalização e elegância teórica, à medida em que descrevem apenas os fenômenos que se ajustam à metodologia. O exemplo clássico que a literatura traz é a proposta de Katz & Fodor (1963) de restringir o escopo da semântica em função da dificuldade de tratar de questões contextuais, por exemplo. As teorias que exploram a relação da semântica com a sintaxe também ilustram essa perspectiva. O problema, para fins de aplicação computacional, é que a dita semântica gramatical descreve apenas uma parcela do conhecimento semântico, aquela vinculada à relação entre predicados e argumentos e sua grade temática.

As teorias escolhidas tendo em vista a aplicação, ao contrário das primeiras, incluem como parâmetro metodológico a eficiência e a computabilidade, o que significa priorizar a ampla cobertura de descrição e o conteúdo semântico dessas descrições.

Considerando a tarefa de construção de ontologias, a proposta de Niremburg & Raskin serve de justificativa para as opções assumidas nos trabalhos a seguir expostos. O que objetivamos destacar, nas próximas seções, é que a estrutura ontológica que definimos para cada uma das categorias – nominais, verbos e adjetivos – tem flexibilidade teórica porque a tarefa assim exige. Nesse sentido, assumindo a perspectiva que prioriza a tarefa, o problema, temos a flexibilidade de abarcar abordagens lexicais, semântico-gramaticais e semântico-pragmáticas, sem parecer uma atitude incoerente do ponto de vista teórico. É esse o enfoque que permeia os três relatos que seguem.

4. Ontologia de nominais

A descrição semântica dos nominais do domínio *universidade* partiu do levantamento de itens lexicais do domínio estudado, extraídos de um *corpus* constituído de ementas e documentos da universidade. Ementas são documentos contendo informações sobre conteúdos, critérios de avaliação, metodologia e bibliografia adotada em determinada disciplina dos cursos de graduação. Vale ressaltar que muitos dos nomes constantes nesses documentos são os consultados com maior frequência no sistema de busca do *site* da universidade⁴, constatação que vem confirmar a escolha dos documentos que comporiam o *corpus*. Unidades lexicais como *professor*, *curso*, *disciplina*, *universidade*, *centro*, *aluno* e *estágio* estão entre os mais frequentes. À medida que procedemos ao estudo semântico, em especial ao estudo das relações lexicais,

⁴ Tomamos como referência o sistema de busca da Universidade do Vale do Rio dos Sinos.

outros termos também relacionados ao domínio, tais como *coordenador, diretor, reitor*, entre outros, passaram a fazer parte da ontologia.

Com base nesse primeiro estágio de estudo do vocabulário do domínio, passamos a nos envolver com a descrição explícita do conhecimento semântico. Duas concepções teóricas serviram de principal referencial para esse propósito: a semântica relacional (Evens, 1988; Miller & Fellbaum, 1991) e a Teoria do Léxico Gerativo (Pustejovsky, 1995). A escolha dessas abordagens, aliada às particularidades do domínio em questão, determinou que tipo de informação deveria ser codificado. A forma como as palavras são associadas no WordNet (Miller, 1986) serviu de ponto de partida para o estudo de duas relações semânticas: a relação de meronímia (parte-todo) e a relação de hiponímia (superordenação e subordinação). A Teoria do Léxico Gerativo, por sua vez, ao postular a *estrutura qualia*, nível de representação responsável pela semântica dos nominais, parece dispor de um aparato descritivo capaz de capturar não apenas as relações de meronímia e hiponímia como também outras informações importantes para organizar a ontologia. Destacam-se as informações contidas nos papéis AGENTIVO e TELICO, assim como o tratamento proposto para os casos de polissemia lógica ou sistemática.

A relação parte-todo, conhecida pelos semanticistas como meronímia/holonímia, serviu para organizar um número considerável de termos. Seguindo proposta de Evens et alli (1988), que examinam diferentes formas de manifestação dessa relação, aplicamos o termo geral *parte de* a noções mais específicas, como *membro, seção* ou *segmento*. Diz-se, então, que reitoria *é parte* da universidade, professor *é membro* do colegiado do Curso de Letras. Reitoria e professor são *merônimos* e universidade e colegiado, *holônimos*. Aplicação semelhante *é* feita para ementa e suas partes (objetivos, metodologia, bibliografia), reitoria e suas divisões administrativas (administração superior, administração geral e administração de centros) e assim por diante. Uma questão controversa relacionada a essa relação diz respeito à propriedade da transitividade, que significa supor que se nós sabemos que A *é* uma parte de B e B *é* uma parte de C, nós não podemos afirmar que A *é* parte de C.

A relação de hipônimo/hiperônimo, mais conhecida por *é-um (is-a)*, ainda que seja considerada como a base de uma taxonomia, no caso do domínio *universidade*, não foi aplicada de forma tão produtiva como a relação de meronímia. Há, contudo, amplos conjuntos de termos que se prestam a essa forma de organização. *Sintaxe II é uma disciplina, Rove é uma professora, Unisinos é uma universidade, Curso de extensão é um tipo de curso e Universidade é um tipo de instituição de ensino* são exemplos de hiponímia. Nos três primeiros exemplos, a expressão *é um(a)* focaliza instâncias (Sintaxe II, Rove e Unisinos). A expressão *é um tipo de*, empregada nos dois últimos exemplos, não indica os termos acomodados no nível mais baixo da hierarquia; tem-se subclasses apenas (curso de extensão e universidade).

Distinguir a relação *é-um* da relação *parte de* não é tarefa fácil, até porque um subconjunto de um conjunto *é* uma parte deste. É o que se pode perceber na relação entre Cursos e Cursos de Letras, Curso de Jornalismo, por exemplo.

Investigando o problema à luz da teoria do Léxico Gerativo, mais precisamente a partir das quatro dimensões fornecidas pela estrutura qualia – papéis CONSTITUTIVO, FORMAL, TÉLICO e AGENTIVO –, *é* possível obter um tipo de informação mais refinada. As relações de meronímia e hiponímia podem ser equiparadas às informações previstas pelos papéis CONSTITUTIVO e FORMAL. Os papéis TÉLICO e AGENTIVO também permitem a codificação de muitas informações semânticas ligadas ao domínio *universidade*. A faceta télica, que define a proposta ou finalidade associada a um item lexical, tem uma aplicação especial: os nominais com ocupações definidas, como *professor, reitor, pesquisador*. Nesses casos, a faceta télica diz respeito à informação relacionada às atividades típicas que exercem. A representação lexical de *professor*, por exemplo, inclui o evento de *lecionar* ou *ministrar*, acompanhado das informações dos argumentos. *Rove leciona Sintaxe na sala 353 no horário 31* teria a seguinte representação: *lecionar (e, Rove, Sintaxe, sala 353, horário 31)*.

A relação AGENTIVA, que se refere originalmente à forma como algo surge ou *é* criado, delineada por Pustejovsky para distinguir tipos naturais e artefatos, pode aplicar-se à forma como disciplinas, cursos e estágios ocorrem.

Ainda explorando a capacidade descritiva e explanatória da Teoria do Léxico Gerativo, procuramos aplicar o tratamento que Pustejovsky propõe para dar conta do fenômeno da polissemia para organizar os itens lexicais da ontologia. Consideram-se como sistematicamente polissêmicos sentidos que são sistematicamente relacionados e predizíveis para classes inteiras de palavras. O nome *escola*, por exemplo, pode referir-se no mesmo discurso a *grupo de pessoas, processo de aprendizagem* ou *construção*. *Janela e porta* também são termos que expressam o mesmo tipo de polissemia, por exprimir idéia de *abertura* ou *objeto físico*.

Além de identificar as relações semânticas que melhor organizam os termos, interessa-nos também analisar em que medida há classes de termos sistematicamente polissêmicos no domínio e avaliar a melhor forma de representá-los na ontologia.

Guiados pela iniciativa de Buitellar (1998), que vem desenvolvendo um banco de dados semânticos - Corelex - com base no fenômeno da polissemia, identificamos duas grandes classes de termos polissêmicos: termos referentes a publicações, tais como *livro*, *artigo*, *tese*, e termos como *universidade*, *mestrado*, *curso*, que podem exprimir idéia de instituição de ensino, edificação ou pessoas da universidade.

A relevância de incluir esse tipo de fenômeno ou relação para organizar a ontologia pôde ser percebida em dois momentos: consultando o WordNet, em especial as informações para o termo *universidade*, constatou-se não haver distinção entre ambigüidade contrastiva (homonímia) e ambigüidade complementar (polissemia sistemática). Ao se ampliar a busca, consultando os resultados para hiponímia ou meronímia, evidenciou-se a formação de redes completamente distintas, como se os diferentes sentidos fossem totalmente excludentes, o que não é o caso. Ao construir manualmente a ontologia, pudemos experimentar a mesma dificuldade. Sem um mecanismo para indicar que os diferentes sentidos dos termos polissêmicos estão inter-relacionados, obrigatoriamente serão construídas redes separadas, já que teremos informações sobre merônimos e hipônimos completamente distintas.

5. Ontologia de verbos

A primeira tarefa para a construção da ontologia dos verbos foi o estudo do *corpus* composto por 6 Acórdãos Judiciais sobre o tema “acidentes rodoviários”, homologados entre 2002-2003, disponibilizados eletronicamente⁵ nas bases de dados jurídicas do Instituto das Tecnologias de Informação na Justiça de Portugal. Feito o reconhecimento do *corpus*, foram selecionados os 10 verbos mais relevantes para o domínio, e analisados em 101 concordâncias, levando sempre em consideração seu papel para a situação de comunicação como um todo. A escolha do tema deve-se a integração da pesquisa aos projetos ONTOVERB⁶ e LOIS⁷.

A ontologia de verbos contou com a sistematização dos aspectos da semântica verbal em quatro níveis: (i) definição; (ii) relações lógico-semânticas; papéis semânticos; e elementos *frame*. A definição é um nível de representação importante para uma melhor interação da ontologia com humanos, mas, para o sistema, esse nível não é compreensível.

Sistematizar informações sobre as relações lógico-semânticas possibilita que o sistema reconheça o significado dos conceitos definidos na ontologia através das relações expressas entre eles. A identificação de conceitos sinônimos, por exemplo, possibilitará que a consulta ao sistema de busca seja feita através de palavras equivalentes, facilitando o acesso do usuário à base de dados. Informações de ordem sintático-semântica com base em tipos de situação e papéis semânticos auxiliam o sistema a interpretar e a construir frases coerentemente. As relações lógico-conceituais representadas na ontologia foram as seguintes: (a) acarretamento; (b) antonímia; (c) causa; (d) hiponímia; (e) sinonímia. A codificação desse nível da semântica verbal possibilitou uma grande ampliação da ontologia, de 10, passou a 107 entidades verbais. A hiponímia e a sinonímia foram a relação mais produtiva do domínio. A hiponímia facilitará a realização de inferências por parte do sistema, possibilitando, por exemplo, a interpretação de perguntas como: *que veículos automotores mais se acidentam?* O sistema, para responder a esse tipo de pergunta, deve ser capaz de associar *carro*, *motocicleta*, *motociclo*, *autocarro* (“ônibus” em PE) com *veículos automotores*. A relevância da sinonímia na ontologia aqui proposta deve-se à possibilidade de auxiliar o usuário no momento da busca, na medida em que ele não precisará se limitar a empregar um termo jurídico específico, podendo realizar suas consultas através de termos equivalentes.

Inserir informações sobre a estrutura argumental de um verbo em uma ontologia possibilitará que o sistema de busca reconheça que determinado verbo exige a presença de um conjunto X de argumentos e que esses argumentos precisam ser representados por conceitos de determinado tipo (ex.: humano, animado, abstrato, etc). Para classificar os argumentos dos verbos, fez-se uma sistematização dos papéis propostos por diferentes autores (ex.: Fillmore, 1968; Chafe, 1970; Jackendoff, 1975; Dowty, 1979; Frawley, 1992; Borba,

⁵ <http://www.dgsi.pt>

⁶ O projeto ONTOVERB propõe-se a estudar a semântica verbos a fim de incluí-lo em uma ontologia de domínio *universidade* com o intuito de aperfeiçoar sistema de busca de informação na web. Projeto financiado pela UNISINOS, FAPERGS e CNPq.

⁷ *Lexical Ontologies for Legal Information Sharing – LOIS* - é um projeto que se propõe a construir uma *WordNet* multilíngüe especializada para o domínio Jurídico nas seguintes línguas européias: italiano, inglês, alemão, checo, português e holandês.

1996; Saeed, 1997; Kearns, 2000). Os papéis semânticos identificados no *corpus* foram: (a) agente; (d) instrumento; (b) beneficiário; (e) paciente; (f) objetivo; (c) origem; (d) locativo; (e) propósito; (f) razão.

Entretanto, há informações que auxiliam na representação dos aspectos semânticos das entidades verbais que não se encontram nos limites da Semântica Lexical nem na Sintaxe; são informações situacionais que uma abordagem baseada em *frames* permite codificar. A noção de *frame* recebe interpretações variadas tanto no escopo da Linguística quanto da Computação. Algumas abordagens mantêm-se mais ligadas à estrutura da sentença, como as *wordnets*), outras menos, como o FrameNet, abordagem aqui adotada. Os elementos *frame* identificados no *corpus* foram: (a) avaliador (juiz, magistrado, tribunal); (b) avaliado (réu, representante, autor); (c) argumentador (réu, representante, autor); (d) reconhecedor (réu, representante, autor // juiz, magistrado, tribunal) (e) meio; (f) base legal; (g) razão; (h) propósito; (i) evidência; (j) tópico; (k) conteúdo; (l) mensagem; (m) pedido. Além desses, incluímos elementos *frame* contidos no cabeçalho dos documentos: (n) tempo; (o) condição; (p) local; e (q) maneira.

Com o auxílio de uma abordagem baseada em papéis semânticos, descobrimos como se estrutura a predicação, percebendo que *condenar*, por exemplo, exige argumentos como agente (argumento externo) e paciente (um dos argumentos internos), sendo ambos do tipo *humano*. Essa informação, no entanto, não é suficiente para que o sistema diferencie se é uma entidade como *réu* que deve ser o agente ou uma entidade como *juiz*. É necessário fornecermos subsídios ao sistema para que ele reconheça que *juiz* se presta para ser o agente de verbos como *condenar* e *julgar*, mas não para ser agente de verbos como *recorrer* e *provar*. Já a entidade *réu* somente pode ser paciente de *condenar* e *julgar*, mas pode ser agente de verbos como *recorrer* e *provar*.

Os níveis de análise dedicados aos papéis semânticos e aos elementos *frame* organizam informações semelhantes na semântica verbal. Tanto uma abordagem de papéis semânticos quanto de *frames* explicitam papéis para os elementos envolvidos. A diferença fundamental é que uma abordagem de papéis semânticos focaliza a *predicação* e classifica os participantes pertencentes da estrutura argumental e outras entidades acessórias e um abordagem baseada em *frames*, por sua vez, identifica participantes da *situação de comunicação* expressos ou não lexicalmente, previstos ou não pela estrutura argumental. Por localizar-se na interface entre sintaxe e semântica, uma abordagem baseada em papéis semânticos mostra-se mais rígida do que uma abordagem baseada em *frames*, a qual amplia as possibilidades de identificação dos participantes da situação de comunicação em que o verbo se insere. Os elementos *frame* propostos pelo FrameNet codificam essas informações contextuais; portanto, possibilitam diferenciar o tipo de agente ou de paciente de determinado verbo, como nos casos descritos acima. Uma abordagem baseada em *frames* possibilitará ainda incluímos na ontologia papéis para elementos da situação que não estão presentes na estrutura das sentenças de maneira explícita, mas que se referem a informações que podem ser recuperadas em porções anteriores do texto, em elementos do co-texto (como o cabeçalho) ou até a partir do conhecimento de mundo do falante. Os papéis semânticos e os *frames* incluíram na ontologia 74 entidades entre nomes e adjetivos, ampliando assim a abrangência da ontologia.

6. Ontologia de adjetivos

Ao contrário dos nominais e dos verbos, os adjetivos têm despertado pouco interesse por parte dos estudiosos. Quando falamos em formalização do conhecimento em ontologias, os adjetivos se tornam muito menos atrativos. Nominais e eventos são as entidades que, por excelência, se destacam nesse tipo de aplicação. A inclusão da informação veiculada pelos adjetivos é tema controverso.

Enquanto os substantivos designam os seres e os verbos as suas ações, os adjetivos representam as propriedades dos seres. Assim sendo, como organizá-los em uma ontologia? Se os adjetivos estão modificando um nome, formando um sintagma nominal, eles devem ter o mesmo *status* dos substantivos e dos verbos, entrando como uma terceira categoria dentro da ontologia? Além dessas dificuldades de natureza organizacional, há outras de natureza semântica.

Nosso objetivo é avaliar a problemática da formalização da semântica dos adjetivos em uma ontologia jurídica, pois, tradicionalmente, uma ontologia se presta bem à formalização dos nomes e dos eventos. Este trabalho dá continuidade ao estudo anterior, apresentando um estudo preliminar dos adjetivos do domínio

jurídico⁸. A escolha desse tema está diretamente ligada à sua inserção nos projetos SEMANTEC⁹ e LOIS, cujo objetivo é a criação de uma rede semântica multilíngüe para o Direito. É através dessa relação com o LOIS que objetivamos, para o futuro, uma ligação multilíngüe entre o nosso trabalho e outras ontologias e léxicos semânticos do domínio jurídico. Para desenvolvermos a parte aplicada, utilizamos os mesmos 6 acórdãos que compuseram o trabalho de descrição dos verbos (Alves, 2005). O nosso *corpus* apresenta 66 diferentes adjetivos, totalizando 132 ocorrências. A descrição semântica desses 66 *types* foi baseada em Borba (1996).

Para a realização desse experimento, escolhemos 5 adjetivos, segundo a frequência e a riqueza de informações a formalizar. Nosso propósito é descrever a semântica dos adjetivos que ocorrem nesses acórdãos e formalizá-los no editor de ontologias Protégé, criando uma ontologia de amostra do domínio jurídico. Para este exercício, os adjetivos serão classificados em *qualificadores* e *classificadores*, segundo Borba (1996). O motivo dessa escolha está ligado à natureza dos adjetivos do domínio jurídico. Através de uma pré-análise, verificamos que a maioria dos adjetivos caracteriza-se por criar classes para os substantivos que o acompanham.

Os adjetivos *qualificadores* se incorporam à natureza do nome, representando uma avaliação subjetiva, como em *interpretação inadmissível*. Os *classificadores* apenas colocam o nome em uma classe, classificam as entidades do mundo, como em *código civil*.

Conforme Borba (1996), as principais características dos adjetivos *qualificadores* que os distinguem dos *classificadores* são: (i) a possibilidade de ocupar tanto a posição atributiva como a predicativa (1); (ii) a nominalização (2); (iii) a gradação (3); (iv) a combinação com verbos de avaliação (4); e a construção exclamativa (5). Todas essas possibilidades são vetadas aos adjetivos classificadores, conforme os exemplos a seguir:

- (1) a. interpretação *inadmissível* / a interpretação é *inadmissível*
b. código *civil* / *o código é *civil*
- (2) a. interpretação *inadmissível* / a *inadmissibilidade* da interpretação
b. código *civil* / *a *civilidade* do código
- (3) a. A interpretação é extremamente *inadmissível*. / A sua interpretação da lei é a mais *inadmissível* de todas.
b. *O ataque foi *muito* cardíaco. / *Esse ataque foi *mais* cardíaco que o outro.
- (4) a. Eu *acho* esta música muito bonita. / Eu *considero* este exercício difícil.
b. *Eu *acho* este ataque cardíaco. / *Eu *considero* este código civil.
- (5) a. Que interpretação *inadmissível!*
b. *Que código *civil!*

Assumimos aqui que adjetivos relacionais e classificadores correspondem, a grosso modo, à mesma classe de adjetivos, porém com nomes diferentes. Assim sendo, acreditamos que a descrição semântica empregada para esses adjetivos pelo WordNet (Miller, 1999), pelo SIMPLE (Peters & Peters, 2000) e pelo Mikrokosmos (Raskin & Nirenburg, 1998) pode ser ampliada. Segundo essas propostas, a principal informação codificada para os adjetivos relacionais é: “relativo a”, “pertencente a”.

Para ilustrar, podemos citar como exemplo o adjetivo *cardíaco*. Devido a sua natureza denominal, um adjetivo desse tipo relaciona-se ao nome do qual deriva. Ou seja, a principal definição semântica para *cardíaco* descreve a sua relação com o nome *coração*: *cardíaco* – relativo, pertencente a *coração*.

Essa definição pode ser ampliada se considerarmos informações adverbiais que alguns classificadores portam consigo. Borba (1996: 186) inclusive cria uma subclasse entre os adjetivos classificadores chamada de *circunstanciais*. Os adjetivos circunstanciais são aqueles que provêm de uma função adverbial subjacente. Eles podem ser *locativos* (6), *temporais* (7), *instrumentais* (8), *causativos* (9) etc.

- (6) traumatismo *craniano* (<que ocorreu no crânio)
- (7) publicação *mensal* (= feita por mês)
- (8) conversa *telefônica* (= por telefone)
- (9) doença *tropical* (causada pelo clima tropical)

⁸ Essa versão preliminar que apresentamos aqui está sendo desenvolvida de forma mais aprofundada em dissertação de mestrado de Bertoldi

⁹ Projeto coordenado pela Profa. Dr^a. Rove Chishman.

Essas informações adverbiais podem ser formalizadas através de relações semânticas. A relação semântica que liga *craniano* a *traumatismo* é *ocorre em*. A relação *feita por* liga *mensal* a *publicação*. A relação *por meio de* liga *telefônica* a *conversa* e *causada por* liga *tropical* a *doença*. A formalização dessas informações amplia as relações “relativo a”, “pertencente a”, já adotadas pelos diferentes léxicos apresentados anteriormente.

Com relação aos papéis temáticos, Borba afirma que “se o adjetivo funciona como um predicado, o nome evidentemente terá uma função casual” (1996: 190). Assim, quando o adjetivo estiver em posição predicativa, ele atribuirá aos nomes que o acompanha na sentença uma relação temática (12). Conforme vimos até agora, são os adjetivos qualificadores que podem assumir a posição predicativa. Isso significaria dizer que somente qualificadores podem atribuir papéis temáticos aos componentes de uma predicação?

(10) O perfume da rosa é *suave*. (O perfume – objetivo)

Pensando em termos ontológicos, classificar os adjetivos em classificadores e qualificadores tem a vantagem de explicitar as relações de hiponímia que permeiam uma ontologia. Como os adjetivos classificadores colocam o nome que acompanham em uma classe, reconhecer os adjetivos classificadores nos ajuda a reconhecer os diversos níveis de classes possíveis dentro de uma ontologia.

O ponto forte para a formalização dos adjetivos qualificadores é que, como eles assumem a posição predicativa, são predicadores. Assim sendo, eles podem ter o mesmo *status* que os verbos dentro da ontologia. Mas todos os adjetivos em posição predicativa são qualificadores?

Conforme sugere a análise, a distinção entre classificadores e qualificadores é muito simples e as possibilidades sintáticas de ambos os adjetivos estão muito bem definidas. Mas o que podemos dizer de um adjetivo como *ilegal*, que Borba (2002) considera como classificador mas que ocorre na posição predicativa em nosso *corpus*?

Os dados analisados, contudo, parecem contrariar a divisão sintática dos qualificadores como predicativos e classificadores como atributivos. Dos 6 adjetivos que ocorrem em posição atributiva em nosso *corpus*, apenas 2 eram qualificadores, segundo Borba (2002). Isso significa que 66,67% dos adjetivos em posição predicativa eram classificadores, contra apenas 33,33% de qualificadores. Afinal, fica uma pergunta: o que diferencia adjetivos qualificadores de classificadores? Definitivamente, o critério sintático, conforme apresentado por Borba (1996), não é o principal deles.

A delimitação entre classificadores e qualificadores é um ponto muito frágil. Conforme Borba (1996: 177), os adjetivos qualificadores “se incorporam (de forma acidental ou essencial) à natureza do nome, como se constituíssem um traço dele (...)”. Borba (2002) considera *suposto* como um adjetivo qualificador. Isso é o mesmo que dizer que *suposto*, em *um suposto assassino*, constituiria um traço intrínseco ao nome *assassino*. Na verdade, um *suposto assassino* não é nem *suposto* nem *assassino*, portanto considerá-lo como um qualificador não é uma escolha indubitável.

A delicadeza dos limites entre qualificadores e classificadores fica ainda mais evidente se compararmos *suposto* a *falso*. Como *suposto*, não podemos dizer que *uma arma falsa* seja uma *arma*. Ou seja, tanto *falso*, como *suposto*, têm o mesmo comportamento: o significado é composto juntamente com o nome. Apesar da aparente similaridade, Borba (2002) considera *falso* como um classificador, enquanto *suposto* é qualificador. Essas questões demonstram que a aparente facilidade em se distinguir adjetivos classificadores e qualificadores não corresponde à realidade da língua em uso. Além do mais, o benefício de se adotar essa classificação, visando a sua formalização em uma ontologia, esbarra na dificuldade que alguns adjetivos possuem para se encaixar nessas duas únicas classes.

Analisando o *corpus*, encontramos um total de 66 adjetivos, considerando-se apenas os *types*. Esses adjetivos estão distribuídos da seguinte maneira: 61 em posição atributiva e 6 em posição predicativa. Apenas um adjetivo se repete em ambas as posições. Para a análise, utilizamos a classificação semântica proposta por Borba (1996), que divide os adjetivos em qualificadores e classificadores. Para uma classificação mais rigorosa, não baseada na intuição, seguimos o *Dicionário de Usos do Português do Brasil*, de Borba (2002).

Segundo Borba (1996), os adjetivos qualificadores são predicadores, enquanto os adjetivos classificadores são apenas atributivos. Na verdade, essa equação não corresponde aos dados encontrados no *corpus*. Podemos ver que os classificadores são a maioria no domínio jurídico: 59,09%, contra 40,91% de qualificadores. Mas o inesperado é que os classificadores ocorrem em percentual muito maior na posição predicativa, contradizendo a teoria de que classificadores ocorrem apenas na posição atributiva.

No estudo que estamos relatando, optamos por analisar os 5 adjetivos mais frequentes. São eles: *civil*, *legal*, *supremo*, *patrimonial* e *sumário*. Dentre esses cinco, 3 são classificadores e 2 são qualificadores, todos em posição atributiva.

O adjetivo *civil* ocorreu 27 vezes em nosso *corpus* em quatro contextos diferentes: Código *civil* (17 ocorrências); Código processual *civil* (05 ocorrências); Responsabilidade *civil* (04 ocorrências); e Construção *civil* (01 ocorrência). Cada um desses contextos apresenta informações semânticas distintas.

O adjetivo *legal* aparece em segundo lugar. Com 08 ocorrências, sua distribuição dentro do *corpus* se dá em 4 diferentes contextos: Taxa *legal* (03 ocorrências); Juros *legais* (02 ocorrências); Presunção *legal* (02 ocorrências); e Regime *legal* (01 ocorrência).

O adjetivo *supremo* ocorreu 7 vezes, em apenas um contexto, que também é um nome próprio: Supremo Tribunal de Justiça. O adjetivo *patrimonial*, por sua vez, ocorreu 6 vezes no *corpus*, em dois contextos diferentes: danos *não patrimoniais*, com 4 ocorrências; e danos *patrimoniais* com 2 ocorrências. O adjetivo *sumário* teve 4 ocorrências no *corpus*, aparecendo em 3 contextos diferentes, mas com o mesmo sentido para o Direito, são eles: Ação *sumária* (01 ocorrência); Forma *sumária* (01 ocorrência); e Processo *sumário* (01 ocorrência).

7. Possibilidades de formalização das ontologias usando a ferramenta Protégé

Tendo explicitado a fase de descrição semântica de nominais, verbos e adjetivos, apresentamos, por fim, a solução que encontramos para formalizar as representações propostas. Com isso, ilustramos a fase que, nos termos de King & Uschold (1999), compreende a codificação e integração de ontologias. Optou-se por utilizar a ferramenta *Protégé*¹⁰, que permite que a ontologia seja exportada em diversos formatos, tais como *rdf schema*, *html*, *texto*, entre outros, facilitando sua implementação. Desta forma, estes dados formalizados poderão ser acessados pelo programa, de modo a melhorar o sistema de busca e extração na *web*. Outra vantagem de escolhermos o Protégé é a integração das ontologias de domínio em uma ontologia e a integração desta, em um trabalho futuro, a ontologias do tipo *top-level* do mesmo domínio.

Vale ressaltar que o Protégé é uma ferramenta Java de código aberto, que proporciona uma arquitetura extensível para a criação de ferramentas de bases de conhecimento personalizadas. Para o Protégé, ontologia é um modelo de um campo específico de conhecimento, os conceitos e seus atributos, bem como as relações entre estes conceitos. Neste editor, uma ontologia é representada como um conjunto de *classes* e seus *slots*.

Em se tratando da representação dos nominais, partindo de redes construídas manualmente, foi possível explicitar as relações semânticas de hiponímia e meronímia. A relação de hiponímia pode ser explicitada no Protégé como classe e subclasse. A relação de meronímia é estabelecida através de *slots*, que aparecem como propriedades de cada classe. Em nosso *corpus* temos registrado (como *slots*) os seguintes tipos de meronímia: *contém*, *é membro de*, *é parte de*, *possui*. Os casos de polissemia lógica também podem ser descritos como *slots*. Para um item como *universidade*, por exemplo, podemos ter como *slot* as diferentes facetas do significado.

No caso dos verbos, partimos das relações léxico-conceituais, que são, muitas vezes, a base das ontologias. A arquitetura geral da ferramenta Protégé utiliza como critério primeiro a relação entre as classes, suas subclasses e instâncias. Dessa forma, o início da construção da ontologia tratou da inclusão de entidades relacionadas por meio de *hiponímia/hipernímia*. A partir da inclusão das classes, é possível inserirmos os outros tipos de relações lógico-conceituais. Sejam elas relações de *acarretamento*, *sinonímia*, *antonímia* ou *causa*, sua forma de representação é a mesma; definimos essas relações/propriedades no que a ferramenta denomina estrutura *frame* (do Inglês: *template slot*). Cada elemento da estrutura *frame* (*slot*) pode estabelecer relação entre uma *classe* e outras quantas forem necessárias, desde que elas já tenham sido inseridas na hierarquia.

É importante comentarmos que a inserção dos *slots* do tipo relações lógico-semânticas sofreu uma restrição devido a limitações da ferramenta. O Protégé não permite que sejam definidos *slots* iguais com valores diferentes, ou seja, não é possível inserirmos uma mesma propriedade (relação) que estabeleça ligação entre unidades lexicais diferentes. A solução encontrada foi inserirmos juntamente com o nome da relação o nome do verbo a que ela se refere. Assim temos: *sinônimo_condenar*, *sinônimo_absolver*, *sinônimo_julgar*, etc.

¹⁰ Disponível para download em <http://www.smi.stanford.edu/projects/protege/>

As relações lógico-semânticas se confirmaram como um recurso bastante produtivo e ágil para a inclusão de classes em uma ontologia. Foram inseridas inicialmente 10 classes verbais as quais geraram um total de 120, isso porque conseguimos expandir a análise relacional para além dos 10 verbos base.

A ferramenta Protégé também se prestou à representação dos *frames*. Conforme vimos anteriormente, os papéis semânticos e os elementos *frame* referem-se a entidades semelhantes em uma situação de forma mais ou menos independente da estrutura sintática. A diferença reside no fato de que os papéis semânticos representam os participantes da predicção e os elementos *frame* representam os participantes da situação evocada pela unidade lexical em questão (nesse caso, os verbos). O desafio, nessa etapa da construção da ontologia, foi encontrar uma forma de não inserir informações sobrepostas nem perder preciosas informações.

A figura a seguir apresenta uma visão geral de como a semântica verbal pode inserida na ferramenta.

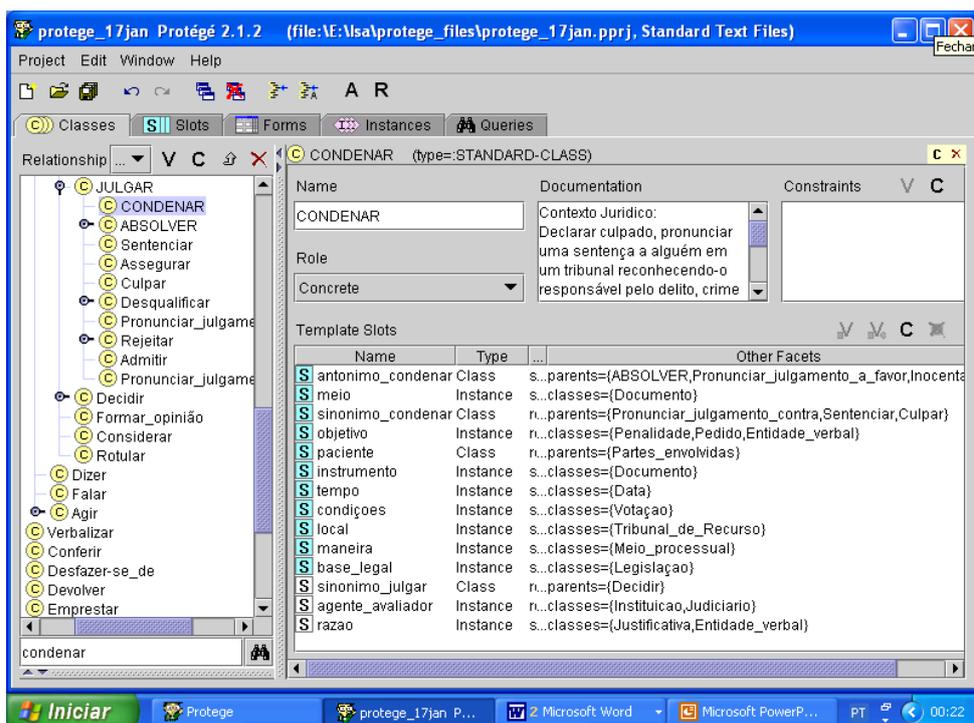


Figura 1: Visão Geral do Editor Protégé

No que tange à inclusão dos adjetivos, consideramos inicialmente que o principal critério estruturante na arquitetura do Protégé e na organização de ontologias é a relação de classe e subclasse. Criamos a superclasse *propriedades*, considerada pela ferramenta uma classe de papel abstrato, haja vista que não se aplica aos dados reais do *corpus*, servindo apenas para organizá-los na ontologia. O mesmo ocorreu com as categorias *classificadores* e *qualificadores*, consideradas *classes abstratas*. Outra estratégia que se aplica a todos os adjetivos é a inclusão de definições, recurso denominado pelo Protégé de *documentação*.

A relação hiponímica é mais uma vez contemplada com a inclusão dos adjetivos classificadores. Isso porque os nominais formados por N + Adj Classificador, como *código civil*, *responsabilidade civil*, aparecem, no Protégé, como subclasses de nominais. Além dos nominais hipônimos, os adjetivos classificadores podem gerar outras relações. Incluímos, por exemplo, como um SLOT, a relação *relaciona-se a* para explicitar o elo entre o adjetivo *civil* e o substantivo *cidadão*.

No que tange à estruturação da semântica dos adjetivos qualificadores, vale enfatizar a possibilidade de incluir as típicas relações de sentido, como a sinonímia e a antonímia. Os adjetivos *supremo* e *sumário* exemplificam essas possibilidades.

Entre as limitações da ferramenta destacamos, primeiramente, a questão da herança obrigatória entre as classes. Não encontramos solução para isso. A ferramenta considera que uma subclasse da ontologia deve herdar todas as propriedades de sua superclasse, não há forma para restringir essa herança. Essa limitação afeta tanto a representação das relações lógico-semânticas quanto dos papéis semânticos e *frames*. Assim subordinados acabam herdando os mesmos *sinônimos* e *antônimos* do seu superordenado (na figura acima vemos *sinônimo_julgar*, entre os slots de *condenar*).

Outra limitação da ferramenta que afeta a inclusão das informações semânticas é o fato de não permitirem que determinado tipo de *slot* tenha diferentes valores, ou seja, remetam a diferentes entidades. Fato que contraria o que ocorre de fato entre as propriedades da língua. Um *slot paciente*, por exemplo, pode ser preenchido por uma entidade como *réu* em uma situação e em outra por uma entidade como um *documento* (ex.: acórdão). A solução para a questão foi, como já dissemos, a inserção de um atributo nos papéis e relações.

8. Considerações finais

Neste artigo, tomando como base três propostas de descrição semântica com vistas à construção de ontologias, objetivamos destacar o papel dos estudos semânticos nessa tarefa genuinamente interdisciplinar. Ao fazermos esse recorte, nosso propósito maior foi atentar para as estratégias que guiam o trabalho do semanticista. Partindo da perspectiva adotada por Niremburg & Raskin (1995), procuramos justificar as escolhas teóricas feitas.

O estudo sobre os nominais nos levou à escolha da semântica relacional e da Teoria do Léxico Gerativo. Fato surpreendente foi o número expressivo de casos de meronímia, comparados com os de hiponímia, considerada a relação estruturante de uma ontologia de nominais.

A ontologia verbal, em termos de descrição semântica, apresentou também resultados surpreendentes: a opção em incluir abordagens teoricamente contrastantes – semântica gramatical, semântica relacional e semântica de *frames* – surtiu resultados positivos em termos práticos.

O estudo dos adjetivos, por sua vez, trouxe-nos indicativos importantes no que tange às opções teóricas disponíveis. Vimos que não é tarefa fácil escolher, dentre as tantas abordagens e taxonomias disponíveis, a que melhor se ajusta aos diferentes propósitos de pesquisa. Exploramos aqui uma dentre tantas perspectivas, que é a que parte de um critério sintático e subclassifica os adjetivos em classificadores e qualificadores. Há, contudo, outras distinções a serem investigadas: a distinção em graduáveis e não-graduáveis e a própria polissemia dos adjetivos são temas que teremos que enfrentar quando ampliarmos este trabalho.

Em termos aplicados, chegamos a abordar a formalização no Protégé, ainda que não completa, principalmente no que se refere aos adjetivos. Como trabalho futuro, é nosso propósito incluir conceitos de uma ontologia de *top level*, como a DOLCE, a fim de que esta ontologia de domínio possa efetivamente interagir com aplicações similares.

9. Referências bibliográficas

ALVES, Isa Mara da Rosa. *O uso da semântica verbal em sistemas de extração de informação: a construção de uma ontologia de domínio jurídico*. Dissertação de Mestrado. São Leopoldo: UNISINOS, 2005.

BERNERS-LEE, Tim. *The web Semantic*. <http://www.sciam.com/2001/0501issue/0501berners-lee.html>

BORBA, Francisco da Silva. *Dicionário de usos do português do Brasil*. São Paulo: Ática, 2002.

_____. *Uma gramática de valências para o português*. São Paulo: Ática, 1996.

BUITELLAR, P. . *CoreLex: Systematic Polysemy and Underspecification*. PhD Thesis, Brandeis University, 1998.

BENJAMINS, V. et alli. What are ontologies, and why do we need them? In: *IEEE Experts*, Vol. 9, N. 6, 1999.

CHAFE, W. L. *Significado e estrutura lingüística*. Chicago: The University of Chicago Press, 1970

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

CRUSE, A. . *Meaning in Language: An Introduction to Semantics and Pragmatics*. New York: Oxford University Press, 1990.

DOWTY, D. *Word Meaning and Montague Grammar*. Dordrecht: D. Reidel, 1979.

- ERDMANN et alli. . From Manual to Semi-automatic Semantic Annotation: About Ontology-based Text Annotation Tools. In: *Proceedings of 18th International Conference on Computational Linguistics*, 2000.
- EVENS, M. . *Relational Models of the Lexicon*. Cambridge: Cambridge University Press, 1988.
- FILLMORE, C.J. The Case for Case. In.: Bach and Harms (Ed.): *Universals in Linguistic Theory*. Holt, Rinehart, and Winston, New York, 1968;
- FRAWLEY, W. *Linguistic Semantic*. London: Lawrence Erlbaum Associates, Publishers, 1992.
- JACKENDOFF, R. *Morphological and Semantics Regularities in the Lexicon*. *Language*, Baltimore: v.51, 3, 1975.
- KEARNS, K. *Semantics*. New York: St. Martin Press, 2000.
- KING, M. & USCHOLD, M. Towards a Methodology for Building Ontologies. In.: *Proceedings of Workshop on Basic Ontological Issues in Knowledge Sharing*, Edinburg, 1995.
- MILLER, G. Dictionaries in the mind. *Language and Cognitive Processes*, v. 1, 171-185, 1988.
- MILLER, G.; Fellbaum, Ch. Semantic networks of English. *Cognition*, v. 41, p 197-229, 1991.
- MILLER, Katherine J. Modifiers in WordNet. In: *WordNet: an electronic lexical database*. Christiane Fellbaum (ed.). Cambridge, MA: MIT Press, 1999.
- NcGUINNESS, D. & NOY, N. *Ontology Development 101: A Guide to Creating Your First Ontology*. <http://protege.stanford.edu/publications/ontology/~development/ontology101.pdf>
- NIRENBURG, Sergei & RASKIN, Victor. *An Applied Ontological Semantic Microtheory of Adjective Meaning for Natural Language Processing. Machine Translation*, Hingham, Kluwer Academic Publishers, 13(2-3):135-227, 1998.
- _____. *Ontological Semantic*. Cambridge University Press, 2005.
- PETERS, Ivonne & PETERS, Win . The treatment of adjectives in SIMPLE: theoretical observations. *Proceedings of LREC2000*, Athens, 2000.
- PUSTEJOVSKY, J. . *The Generative Lexicon*. Cambridge: The MIT Press, 1995.
- SAEED, J. *Semantics*. Oxford: Brasil Blackwell, 1997.