IME-USP

Instituto de Matemática e Estatística da Universidade de São
Paulo

Plano de Estudo para Monografia - Tópicos em Computação

ONTOLOGIAS, E WEB SEMÂNTICA, RDF, OWL,
E LINGUAGENS RELACIONADAS

São Paulo, Março de 2003

Aluno
Márcio Rodrigo de Freitas Carneiro

Orientadora
Prof. D. Renata Wassermann

1 Introdução e Motivação

O crescimento da teia mundial, ou mais conhecida como WWW, traz cada vez mais uma grande quantidade de informação muito diversificada para as pessoas dotadas de um simples computador e uma conexão com a Internet. O sucesso da WWW e de seus padrões de linguagens e protocolos se afirma no uso da mesma tecnologia dentro das empresas, por exemplo, numa Intranet. Os documentos em hipertexto (em geral escritos em HTML) são relativamente poderosos e de simples acesso, e portanto se tornam cada vez mais padrão no conhecimento básico de computadores.

Entretanto, o enorme volume da WWW, ou mesmo de uma Intranet, torna difícil e trabalhosa a organização da informação, e conseqüentemente a busca de informações específicas. Ainda hoje as ferramentas mais utilizadas na WWW para busca de sítios se baseiam na exaustão, procurando as palavras chaves no conteúdo do documento, sem nenhuma “consciência” do contexto dos termos e localidade do documento.

Na tentativa de melhorar não só as buscas, mas também a organização da informação, a idéia da *Web Semântica* (<http://www.w3.org/2001/sw/>) surge com intuito de adicionar semântica aos documentos da teia. Dentro dessa nova área de pesquisa, há diversas proposições de modelos, linguagens e ferramentas, que é nosso principal foco de estudo para a realização de uma monografia.

Duas linhas principais são desenvolvidas paralelamente. Linguagens para a definição de ontologias, e linguagens para a marcação de texto com referências às ontologias, para a adição de semântica a tais documentos. O consórcio W3C está no processo de definições de linguagens padrões para ambas proposições.

Uma aplicação direta dessas novas tecnologias, seria por exemplo melhores esquemas de buscas. O usuário poderia especificar a semântica de algumas palavras chaves, como por

exemplo, buscar documentos em que o autor seja a palavra chave requerida, ou que tenha uma relação específica (definida previamente) com outro documento ou recurso da teia. Com a premissa de que os documentos estão também descritos com metadados baseados numa ontologia definida, fica mais fácil a busca ser realizada nesses metadados, e bem mais frutífera que uma busca baseada em força bruta.

2 Plano de Estudo

Primeiramente, devemos estudar o que são ontologias e bases de conhecimento. Mesmo tendo como foco principal as soluções e ferramentas para a web semântica, temos que entender os conceitos e princípios básicos.

Um estudo geral sobre os principais modelos e soluções disponíveis deve ser feito, para principalmente termos base de seleção para nossa escolha.

Durante o semestre pretendemos estudar as alternativas existentes de ferramentas e linguagens, para que possamos selecionar algumas que serão mais tarde usadas em uma aplicação prática.

Para tal estudo, segue uma bibliografia selecionada.

Referências

- [BKvH02] Jeen Broekstra, Arjohn Kampman, and Frank van Harmelen. Sesame: A generic architecture for storing and querying RDF and RDF schema. In I. Horrocks and J. Hendler, editors, *Proceedings of the First International Semantic Web Conference*, number 2342 in Lecture Notes in Computer Science, pages 54–68. Springer Verlag, July 2002.
 - [DCvH⁺02] M. Dean, D. Connolly, F. van Harmelen, J. Hendler, I. Horrocks, and D. McGuinness. Web ontology language (OWL) reference version 1.0. *World Wide Web Consortium Working Draft*, November 2002.
-

-
- [FSvH02] Christiaan Fluit, Marta Sabou, and Frank van Harmelen. Ontology-based information visualisation. In Vladimir Geroimenko, editor, *Visualising the Semantic Web*. Springer Verlag, 2002.
- [GG95] N. Guarino and P. Giaretta. Ontologies and knowledge bases: Towards a terminological clarification. In N. Mars, editor, *Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing*, pages 25–32. IOS Press, 1995.
- [Hay03] Patrick Hayes. RDF semantics. *World Wide Web Consortium Working Draft*, January 2003.
- [PSHvH02] P. Patel-Schneider, I. Horrocks, and F. van Harmelen. Reviewing the design of DAML+OIL: An ontology language for the semantic web. In R. Dechter, M. Kearns, and R. Sutton, editors, *Proceedings of the Eighteenth National Conference on Artificial Intelligence*, July 2002.
- [vH02] F. van Harmelen. How the semantic web will change KR: challenges and opportunities for a new research agenda. *The Knowledge Engineering Review*, 17(1), 2002.
- [WSWS01] B. J. Wielinga, A. Th. Schreiber, J. Wielemaker, and J. A. C. Sandberg. From thesaurus to ontology. *International Conference on Knowledge Capture, Victoria, Canada*, October 2001.
-