

MAC5701 Tópicos em Ciência da Computação

Plano de Estudo: Web Semântica

Juliana Jabra Chahoud

Orientadora: Profa. Dra. Leliane Nunes de Barros

31 de março de 2003

1 Introdução

Este documento apresentará um plano de estudos a ser realizado durante o primeiro semestre de 2003, na área de Inteligência Artificial e Web Semântica. O objetivo deste estudo é realizar um levantamento bibliográfico, buscar recentes trabalhos realizados nessa área, encontrar as tecnologias utilizadas e possibilidades de pesquisa, a fim de elaborar uma proposta para dissertação de mestrado.

2 Motivação

O crescimento exponencial da Web tornou difícil a indexação de informações. As tecnologias utilizadas para construção da Web como, por exemplo, a linguagem HTML que é utilizada pela maioria dos sites, não possuem recursos que permitam atribuir significado à informação, apenas descrevem como a página deve ser exibida.

Para melhorar a indexação, é necessário que se crie padrões de conceitualização sobre o conteúdo da Web que permitam filtrar com eficiência, entre inúmeros dados disponíveis, aquilo que interessa ao usuário num contexto preciso. A diferença entre determinados assuntos é evidente para humanos mas não fazem sentido para os agentes de software, que não conseguem resolver as ambigüidades e compreender um texto ainda que simples.

A Web Semântica é uma tentativa de solução para esse problema. Ao invés de pensar na informação para os humanos, a idéia é pensar como ela poderia ser compreendida pelos agentes de software. Assim, serviços Web mais complexos poderão ser desenvolvidos.

O projeto da Web Semântica é iniciativa do consórcio W3C e tem como um dos líderes Tim Berners-Lee, considerado o pai da Web.

3 Plano de estudo

Além de estudar os conceitos básicos sobre Web semântica, serão estudados alguns componentes utilizados na sua implementação:

- Ontologias: especificação formal que define as relações entre conceitos e estabelece regras lógicas de raciocínio sobre eles. São descritas em linguagens formais e utilizadas pelo computador para interpretar o significado das informações.
- Linguagens da Web Semântica: necessárias para representação do conhecimento, algumas das linguagens mais utilizadas são SHOE, OML, OIL, DAML+OIL e RDF, esta última utilizada pelo consórcio W3C.
- Editores de ontologia: facilitam na construção de ontologias e podem gerá-las em várias linguagens diferentes. Alguns editores conhecidos são Protégé 2000, Ontomat SOEP, OilEdit e OntoEdit.
- Repositórios: armazenam e recuperam anotações semânticas. Um dos repositórios que será estudado é o Sesame, que permite o uso de qualquer banco de dados que armazene RDF e possui uma linguagem chamada RQL, que oferece suporte às consultas no repositório.
- Agentes Web Baseados em Conhecimento: programas que coletam informações como conteúdo e metadados na web.

Referências

- [1] Dave Beckett. *RDF XML Syntax Specification (Revised)*. W3C Working Draft.
- [2] Tim Berners-Lee. *Semantic Web Road map*. <http://www.w3.org/DesignIssues/Semantic.html>, Setembro 1998.
- [3] Graham Klyne; Jeremy Carroll. *Resource Description Framework (RDF): Concepts and Abstract Data Model*. W3C Working Draft.
- [4] Jeff Heflin; Raphael Volz; Jonathan Dale. *Requirements for a Web Ontology Language*. W3C Working Draft.

- [5] DAML+OIL. <http://www.daml.org/2001/03/daml+oil-index>.
- [6] SHOE (Simple HTML Ontology Extension). <http://www.cs.umd.edu/projects/plus/SHOE>.
- [7] Dan Brickley; R.V. Guha. *RDF Vocabulary Description Language 1.0: RDF Schema*. W3C Working Draft.
- [8] Patrick Hayes. *RDF Model Theory*. W3C Working Draft.
- [9] OML (Ontology Markup Language). <http://www.ontologos.org/OML/OML.html>.
- [10] OIL (Ontology Inference Layer). <http://www.ontoknowledge.org/oil>.
- [11] 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*. <http://www.scientificamerican.com/2001/0501issue/0501berners-lee.html>, Maio 2001.
- [12] Protégé-2000. <http://protege.stanford.edu>.
- [13] SESAME. <http://sesame.aidministrator.nl>, Maio 2001.
- [14] Ora Lassila; Ralph R. Swick. *Resource Description Framework (RDF) Model and Syntax Specification W3C Recommendation*. W3C Working Draft, Fevereiro 1999.