

本期目录 | 下期目录 | 过刊浏览 | 高级检索

[打印本页] [关闭]

软件技术与数据库

电子目录本体合并中的实例消重机制

刘平峰, 章佩璐, 陈冬林

(武汉理工大学经济学院电子商务研究所, 武汉 430070)

摘要: 为解决电子目录集成中的本体实例冗余问题, 提出面向本体合并的电子目录本体实例消重机制, 综合考虑实例的名称、属性和关系设计实例语义相似度算法。结合字符串匹配和基于Wordnet计算名称语义相似度, 综合数据类型属性和对象类型属性计算属性语义相似度, 依据类的多重继承关系计算关系语义相似度。当2个实例的语义相似度大于事先设置的阈值时, 需删除其中一个以降低目标本体库的冗余度。实验结果验证了该机制的有效性。

关键词: 电子目录集成 本体合并 实例消重 语义相似度

Instance Duplicate Elimination Mechanism in Electronic Catalog Ontology Merging

LIU Ping-feng, ZHANG Pei-lu, CHEN Dong-lin

(Research Institute of E-business, College of Economics, Wuhan University of Technology, Wuhan 430070, China)

Abstract: To solve the problem of ontology instance redundancy in heterogeneous electronic catalog integration, this paper proposes a mechanism of instance duplicate elimination for electronic catalog ontology merging. This mechanism incorporates ontology concepts, attributes relationships into the calculation of semantic similarity between two ontology instances, while semantic similarity on concepts is calculated by string matching and Wordnet, that on attributes is calculated according to both data-type and object-type, and that on relationships is calculated based on multi-inheritance. If two instances have a semantic similarity above the preset threshold, one of them is deleted from ontology base to ensure the efficiency of operation on ontology instances. Experimental results show the validity of this mechanism.

Keywords: electronic catalog integration ontology merging instance duplicate elimination semantic similarity

收稿日期 2011-07-12 修回日期 网络版发布日期 2012-01-20

DOI: 10.3969/j.issn.1000-3428.2012.02.010

基金项目:

国家自然科学基金资助项目(70972094); 中央高校基本科研业务费专项基金资助项目(2010-IV-082, 2010-II-11)

通讯作者:

作者简介: 刘平峰(1972—), 男, 副教授、博士, 主研方向: 语义网, 知识工程; 章佩璐, 硕士研究生; 陈冬林, 副教授、博士

通讯作者E-mail: zpllovelxy@foxmail.com

参考文献:

- [3] 丁政建, 张 路. 一种改进的本体相似度计算方法[J]. 计算机工程. 2010, 36(24): 39-41 [浏览](#)
- [4] Zadeh L A. The Concept of Linguistic Variable and Its Application to Approximate Reasoning[J]. Information Science. 1975, 8(3): 199-249 [crossref](#)

扩展功能

本文信息

- ▶ Supporting info
- ▶ PDF(322KB)
- ▶ [HTML] 下载
- ▶ 参考文献[PDF]
- ▶ 参考文献

服务与反馈

- ▶ 把本文推荐给朋友
- ▶ 加入我的书架
- ▶ 加入引用管理器
- ▶ 引用本文
- ▶ Email Alert
- ▶ 文章反馈
- ▶ 浏览反馈信息

本文关键词相关文章

- ▶ 电子目录集成
- ▶ 本体合并
- ▶ 实例消重
- ▶ 语义相似度

本文作者相关文章


- ▶ 刘平峰
- ▶ 章佩璐
- ▶ 陈冬林


PubMed

- ▶ Article by Liu, B. F.
- ▶ Article by Zhang, P. L.
- ▶ Article by Chen, D. L.

[5] Budanitsky A, Graeme H. Evaluating WordNet-based Measures of Semantic Distance

[J].Computational Linguistics.2006, 32(1): 13-

[6] 聂规划, 左秀然, 陈冬林. 本体映射中一种改进的概念相似度计算方法[J]. 计算机应用. 2008, 28(6): 1563-1565 

[7] 梅 翔, 孟祥武, 陈俊亮. SSCM: 一种语义相似度计算方法[J]. 高技术通讯. 2007, 17(5): 458-463 

### 本刊中的类似文章

1. 王立霞, 淮晓永. 基于语义的中文文本关键词提取算法[J]. 计算机工程, 2012, 38(01): 1-4
2. 王艳娜, 周子力, 何艳. WordNet中基于IC的概念语义相似度算法[J]. 计算机工程, 2011, 37(22): 42-44
3. 蔡宇, 杨广超. 基于大规模问答对数据的问题检索模型[J]. 计算机工程, 2011, 37(21): 29-31
4. 李文杰, 赵岩. 基于本体结构的概念间语义相似度算法[J]. 计算机工程, 2010, 36(23): 4-6
5. 吕林涛, 董迎. 基于上下文的概念语义相似度计算模型[J]. 计算机工程, 2010, 36(21): 59-61
6. 魏韡, 向阳, 陈千. 基于基因本体的语义相似度研究[J]. 计算机工程, 2010, 36(20): 209-210
7. 吕刚, 郑诚. 改进的基于概念相似度的文本检索[J]. 计算机工程, 2010, 36(12): 55-57
8. 余 宏; 万常选. 基于XML的检索结果聚类方法[J]. 计算机工程, 2010, 36(1): 85-86, 9
9. 杨先娣; 何 宁; 吴黎兵. 基于范畴论的本体集成描述[J]. 计算机工程, 2009, 35(6): 76-78
10. 史 斌; 闫健卓; 王 普; 方丽英. 基于本体的概念语义相似度度量[J]. 计算机工程, 2009, 35(19): 83-85

### 文章评论

反馈人	<input type="text"/>	邮箱地址	<input type="text"/>
反馈标题	<input type="text"/>	验证码	<input type="text" value="5163"/>
<input type="text"/>			