语义 Web 搜索技术研究进展

叶育鑫 欧阳丹彤
(吉林大学计算机科学与技术学院 长春 130012)
(符号计算与知识工程教育部重点实验室 长春 130012)

摘 要 语义 Web 搜索技术是综合本体论、信息检索、自然语言处理等多学科理论和方法的新兴技术。介绍了语义 Web 和语义 Web 搜索的现状。在此基础上,给出了实现语义 Web 搜索技术的一般体系结构,并进一步分析了各组成模块的基本任务、现有技术和评价体系。最后给出了所做的相关工作和对语义 Web 搜索技术的展望。

关键词 语义 Web 搜索,本体,信息检索,智能搜索引擎
中图法分类号 TP18 **文献标识码** A

New Research Advances in Technologies of Semantic Web Search

YE Yu-xin OU YANG Dan-tong
(School of Computer Science and Technology, Jilin University, Changchun 130012, China)
(Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China)

Abstract Technologies of semantic Web search are very novel and young. There are many theories and approaches contributed to semantic Web search, such as ontologies, information retrieval, natural language processing, and so on. Situations of semantic Web and semantic Web search were introduced dividually. Then, the general architecture of semantic Web search engine was concluded through analyzing research at home and abroad in detail. This architecture is composed of four models, including information extraction models, intelligent interface models, information query models, and user query models. The functions, technologies, and evaluation systems of models which are different components of this general architecture were introduced in further. At last, the conclusion about our work was given. At the same time, the future challenge of semantic Web search technologies was discussed.

Keywords Semantic Web search, Ontology, Information retrieval, Intelligent search engine

1 引言

Tim Berners-Lee 提出万维网(World Wide Web)这一概念,给 20 世纪末期的互联网带来了革命。万维网的 Web 信息在以几何级数增长的同时,也给互联网上的信息检索带来了困扰。Yahoo!, Google, MSN 等搜索引擎在一定程度上缓解了 Web 信息检索与海量信息之间的矛盾,但这些基于关键字的技术无法从根本上解决搜索的准确率和召回率较低的问题[1]。其原因在于:文档的关键字未必一定和文档有关,而相关的文档也未必一定显式地包含此关键字。此外还存在同义词引起的召回率问题、歧义词引起的准确率问题、语境相关的上下位、近义、反义等一系列语义问题。为此,W3C (World Wide Web Consortium)提出将语义 Web 作为新的 Web 规范来表达互联网上的语义信息。目前已经完成的标准有 RDF (Resource Description Framework), RDF Schema, Ontology Web Language 等。语义 Web 通过给出 Web 信息的清晰语义和定义扩展万维网,实现用户与机器的协同工作[2,3],它的终极目标是通过制定技术标准使计算机能够更好地理解 Web 信息,支持语义搜索、数据整合、导航和自动化等。本体[4]是语义 Web 应用中实现知识表示、知识推理、知识共享和知识重用的重要技术。它一方面有效地组织与整合了语义 Web 上的资源,另一方面克服了传统搜索引擎无法逾越的鸿沟[5]。另外,自然语言处理在信息检索领域中的早期研究工作集中于标记、短语探测、语法成分分析等[6]。本体刻画了概念和概念间的关系,语义 Web 赋予了 Web 信息语义,它们恰恰为自然语言处理的进一步研究带来了契机。

语义 Web 为 Web 搜索搭建了新的平台,本体、自然语言处理等技术的渗透促成了语义 Web 搜索技术。语义 Web 搜索技术有赖于各个学科和各项技术的发展,反过来语义 Web 搜索技术也促进着各个学科和各项技术的进步。

2 语义 Web 对传统搜索技术的挑战

工作在 Web 环境下的传统搜索引擎,无法利用语义提升搜索精度,更无法识别语义进行语义搜索。语义 Web 为万维

到稿日期:2009-02-20 返修日期:2009-04-27 本文受国家自然科学基金重大项目基金(60496320, 60496321),国家自然科学基金(60773097, 60873148),新世纪优秀人才支持计划项目基金,吉林省科技发展计划项目基金(20060532, 20080107)和欧盟项目基金 TH/Asia Link/010 (111084)资助。

叶育鑫(1981-),男,博士生,主要研究方向为语义 Web 及其相关技术,E-mail:yuxin.ye@hotmail.com;欧阳丹彤(1968-),女,教授,博士生导师,主要研究方向为人工智能、自动推理等。

网资源扩展了语义信息,语义的注入给传统搜索能力的提升带来了空间。语义 Web 搜索门户网站纷纷在实验室中诞生。

MindSwap, OntoWeb, Esperonto 和 Knowledge Web 都是语义 Web 搜索门户网站研究中的典范。其它颇具特色的有:CS AKTive Space[7]能够持续获取英国计算机领域的有关数据信息;MuseumFinland[8]是第一个用于聚合、关联博物馆领域异构信息资源的语义 Web 搜索门户网站;Flink[9]是 2004 年 Semantic Web 挑战评比的优胜者,它可以在线抽取、聚合社会网(social networks)数据资源;KMi[10]是专注于处理查询接口为关键字输入的语义 Web 搜索门户网站。

3 语义 Web 搜索相关技术及评价体系

在给出语义 Web 搜索引擎的一般体系结构模型基础上,进一步说明各组成部分的工作特征及设计原理,并归纳总结可采用的现有技术,给出各模块的评测体系。

3.1 语义 Web 搜索引擎的一般体系结构模型

语义 Web 搜索的一般体系结构模型及各模块之间的组织和联系如图 1 所示。

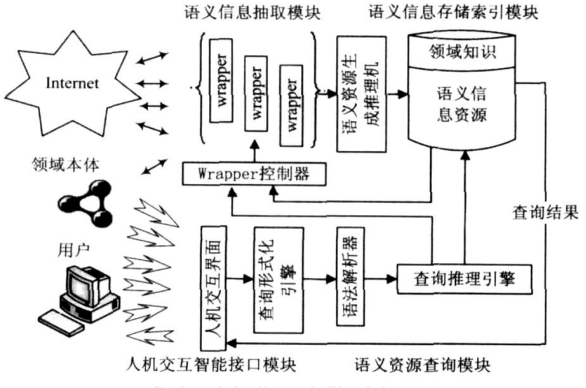

图 1 语义 Web 搜索引擎体系结构

图 1 中的语义信息抽取模块由 Wrapper 控制器和 Wrapper 工具工作组构成,负责网页爬行和信息抽取;图 1 中的人机交互智能接口模块负责处理用户查询请求的输入和输出;图 1 中的语义信息存储索引模块一方面用于接收来自信息抽取模块的数据,另一方面运行自身的语义资源生成推理机生成潜在的语义信息,一并存入知识库;图 1 中的语义资源查询模块负责接收人机交互智能接口模块的形式化语义查询语句、翻译查询请求或推理算法函数、执行查询推理操作,最后将运算结果返回给人机交互智能接口模块,由人机交互智能接口处理并呈现给用户。

3.2 语义信息抽取技术及评价

3.2.1 语义信息抽取

互联网向人们展示了海量的知识和信息,然而不幸的是人们无法直接(目前可用的搜索引擎都是间接达到目的)从当今的互联网上查询和使用他们所想要的知识和信息。这一问题需要开发用于访问、抽取、转换和整合数据的语言和工具来解决[11]。从语义 Web 的观点出发,未来 Web 上所有的可用信息都应该是结构化、标注化和高质量的。

信息抽取工具是用于从特定的信息源(结构化、半结构化和非结构化)中抽取相关内容,并以特定形式加以表示的程序。它由一系列的抽取规则以及应用这些规则的代码组成。在这里,语义 Web 搜索引擎中的 Web 信息抽取与传统搜索引擎中的网络爬行不同,它的功能已不仅仅是抓取网页。

3.2.2 语义信息抽取工具

目前有很多方法用于处理 Web 信息抽取问题,这些方法应用了自然语言处理、语言学与文法、机器学习、信息检索、数据库和本体等一系列领域。因而在处理 Web 信息抽取问题上,各种方法表现出完全不同的特征和能力。

首先可以通过设计各种特殊的程序语言辅助开发者组装 Wrapper。目前流行的 Perl 和 Java 程序语言都是 Wrapper 开发的有效语言。其它较著名的专门由 Wrapper 开发的语言有 Minerva, Tsimmis, Web-OQL, Florid 和 Jedi。利用 HTML 敏感性方法中比较有代表性的是 W4F[12], XWRAP[13], RoadRunner[14]等。还有一类方法应用自然语言处理技术学习抽取规则,它们通常采用诸如过滤、词性标注、词义标注来构建短语及句子间的关系,用于发现抽取规则。代表性的工具有 RAPIER[15], SRV[16]和 WHISK[15]。基于归纳方法用于发现描述数据潜层结构的形式化特征,该类工具的典型代表有 WIEN[17], SoftMealy[18]和 STALKER[19]。基于模型方法通过给定感兴趣对象的目标结构,尝试在网页中定位与目标结构隐式地匹配的部分。该类工具主要有 NoDoSE[20]和 DEByE[21]。另外,Brigham Young University 的数据抽取研究组给出的基于概念模型的数据抽取方法[22],是不同于以前基于依赖文档中数据表达特征的结构的全新理念。

3.2.3 语义信息抽取工具质量评价

Web 信息抽取模块工作质量的好坏和工作效率的高低由构成该模块的信息抽取工具组决定,信息抽取工具质量的参考标准可以从表达能力、健壮性、运行时间等方面衡量。

工具应该具有对网页进行复杂、结构化模式定义的能力,并将相关数据转换成对应的结构化格式,即工具的表达能力。Web 信息特征之一是频繁的变更性,高质量的信息抽取工具应该能够自适应网页结构的简单改变,即工具的健壮性。需要提供高效的算法实现信息抽取原理,而且抽取工具应该有效地实现算法以更准确和更大范围地对 Web 资源进行信息抽取,即工作时间。高质量的信息抽取工具应该提供语义 Web 接口用于丰富本体资源。最理想的情况是使用逻辑语言作为信息抽取语言,而且最好能够与本体的逻辑体系相兼容,以便于推理,即工具的语义 Web 接口能力。

3.3 人机交互智能接口

3.3.1 人机交互智能接口

语义搜索引擎中的人机交互智能接口的主要目的是对终端用户隐藏复杂的语义搜索过程,并能够为初级用户提供简单、有效和友好的服务。目前存在的接口都融入了本体技术或自然语言处理技术来增强智能性。

人机交互智能接口首先要完成理解用户查询请求语义的任务。从语义 Web 的角度来看,用户请求的术语可以和概念、实例、关系匹配,最理想的情况是找到用户请求对应的所有语义。在完成终端用户查询请求语义辨析之后,人机交互智能接口需要将其翻译成恰当的形式化语义查询语句。

3.3.2 人机交互智能接口现有技术

关键字接口:该类接口的优点在于它为终端用户提供了直接指定问题的方式,终端用户一般都熟悉基于关键字的检索模式。TAP 搜索引擎[23]是最早的基于关键字的语义搜索系统之一。此外,为了选择合适的关键字完成搜索,Sem-

Search 系统[24]接口提供处理机制对查询请求重定位。

控件接口:控件是所有计算机用户熟悉的界面隐喻(interface metaphor),它们通过从合法术语列表中选择术语来避免处理语义映射问题,为用户提供领域解释[25]。Magnet[26]是 Haystack 搜索系统的一个基于控件接口,它除生成结构化信息外,还能够动态生成内容列表和搜索可选项。

视图接口:基于视图的系统在领域理解、查询结构化和查询重定位方面提供优质的性能。其典型代表是 GRQL[28]。该类中的 SEWASIE[29]接口存在用户操作最终定位查询请求繁琐的问题,随后的 Ontogator[30]提出了加速基于视图系统查询请求形式化的解决方案,一定程度上缓解了该问题。

自然语言接口:自然语言接口对终端用户查询使用的意义是显而易见的。使用自然语言方法实现的 question answering(QA)系统关注从原始文本检索答案和信息检索的查询扩展[31]。与传统的 QA 系统相比,这些新型 QA 系统能够利用语义信息提供更准确的答案。与其它的人机交互模式相比,自然语言接口的优点在于提供带参数的形式化查询。AquaLog[32]提供了一个简单的、基于本体的 QA 系统范例。

3.3.3 人机交互智能接口评价

从应对复杂并且规模巨大的语义搜索环境能力角度来看,接口应当能够在搜索环境下访问不同的本体而无需更改指定的领域配置;应该具有同时访问不同领域本体的能力。

从查询请求的术语和语义实体的匹配问题角度来看,需要智能接口有寻找映射和定位语义的能力;其次还需要能够根据用户满意度次序返回查询结果;再者,接口能否为不熟悉领域知识的用户提供一定的帮助并辅助完成正确的语义查询,也是匹配问题中的一个要素。

从接口的迭代和探索能力角度来看,语义搜索引擎中需要辅助用户使用正确的领域知识。另外,需要查询请求形式化、查询请求重定位、查询请求重用和提供建议性查询请求。

3.4 语义信息存储技术及评价

3.4.1 语义信息存储技术

目前大多数语义存储模块的实现都借助于关系数据库或面向对象的关系数据库,尤其是对 RDF 数据的存储。合理设计数据库模式以及优化提高数据检索的操作,对 RDF 存储来说至关重要。

语义存储需要 RDF 解析、triples 输入和 triples 推理 3 个功能组件。RDF 解析器根据 RDF 语法规范[33]解析 RDF/XML 文件的陈述并将解析结果传给 RDF 引入器(RDF impoter)。目前有很多代码开源的 RDF 解析器,如 another RDF paser[34],Validating RDF parser[35]等。在推理组件中,基于形式化模型理论[36]的 RDF 语义被具体表达成一组推理规则。目前关于 triples 产生的推理引擎的设计主要有两类:一类是以 sesame[37]为代表的插入触发式推理引擎,另一类是以 Jena[34]为代表的查询触发式推理引擎。部分研究者[38]还将两种设计相结合,设计出混合型 triples 产生推理引擎,以改进推理机制,进一步提高推理效率。

3.4.2 代表性语义存储工具

目前语义信息存储受到国内外语义 Web 领域的高度重视,很多机构都在开展语义存储技术的研究。这里列举几个有代表性的工作,并给出它们的功能特征。

Sesame[37]是荷兰 Vrije University 开发的一款 RDF 和 RDFs 信息存储、查询工具,它是欧洲 IST 项目 On-To-Knowledge[39]成果的一部分。Sesame 的设计和实现是独立于任何存储设备的,它可以布置在任何类型存储设备之上,并且无需改变任何查询引擎或其它函数模式。

RDFStore[40]由一组 Perl 语言实现的模块构成,它通过一种简单直接的方式来管理 RDF 模型数据库。它的存储子系统允许对数据库中 RDF 的结点、边和标签进行透明的存储和检索操作。RDFStore 也同样支持各种类型数据库,包括内存存储结构、本地硬盘存储结构和大规模移动存储结构。

Jena[34]是一个技术领先的工具集产品,是用 Java 语言开发的一个开源项目。关于它的源码和文档可以从 SourceForge 上免费下载。它通过 JDBC 链接使用一个 SQL 数据库实现 RDF graphs 的持久存储。Jena1 支持很多 SQL 数据库引擎,如 Postresql,MySQL,Oracle,并且其体系结构可以灵活地布置在新的 SQL 数据库引擎和自由配置。

3.4.3 语义存储评价

从数据库模式设计角度看,不同的设计模式会产生不同的存储效果。平衡存储空间和查询效率,找到它们二者之间的最优点,是数据库模式设计的主要任务。

从 Table triples 的索引角度看,在实际应用中 Table triples 数据量巨大,合理的索引机制能够提升语义查询的效率。各种索引模式对不同具体应用有不同的效果,没有通用的检索设计适合所有的语义信息存储。语义存储索引设计的好坏是评价系统的一个重要标准。

从动态性和网络分布性来看,语义存储模块要能够及时删除 Web 上的无效资源的语义信息,更新被修改资源的语义信息,添加新产生的资源信息,同时要结合网络带宽、系统使用情况等管理存储系统对 Web 信息资源和动态信息快速反应和管理。语义信息存储的动态性和分布性是评价存储系统优劣的又一指标。

3.5 语义资源查询技术及评价

3.5.1 语义资源查询技术

语义资源查询的实现主要包含 3 部分:语义查询语言设计、查询语句语法解析器和查询算法。它们分别用于设计开发查询语言表述查询语义;解析查询语句,并生成内部数据结构;在知识库中查找符合要求的数据。

任何一门计算机语言都有严谨的形式化语法定义[41],语义查询语言的设计也不例外。查询语句解析器可以手工设计,也可以使用 JavaCC,Lex 和 Yacc 等解析器生成程序快速实现。通过在语法定义中嵌入与语义相关的程序,对每个语法单位均生成与之相对应的对象,并执行相关的语义动作。对象之间以引用保存上下级语法单位之间的关系,从而实现了语法制导翻译。解析完查询语句后,生成的一系列互相联系的对象可由查询算法处理。

3.5.2 语义查询语言

RQL[42]是第一个类型声明性的 RDF 查询语言,是在欧盟的项目 C-Web 和 MesMuses 中产生的。RQL 由一组基本的查询和能够通过功能合成而建立新查询的迭代器(iterators)所定义,支持一般的路径表达。RQL 依赖于一个形式化的图模型,并允许借助于一个或多个模式来解释双重的资源描述。RQL 的独特之处是完美地结合了 RQL 模式和数据路径表达。

Triple 是一个层次化、模型化的规则语言,目标是支持需要 RDF 查询、推理和转换的应用程序。Triple 语言基于 Horn 逻辑并且借用了 F-Logic 的许多特征,提供了一个类 Prolog 语法和基于 RDF 的语法。

SPARQL[1] 构建在以前的 RDF 查询语言(如 RDQL, SeRQL)基础上,拥有一些有价值的新特性。SPARQL 协议和 RDF 查询语言(SPARQL)目前已成为 W3C 推荐标准。SPARQL 支持各种平台和语言。编写更复杂的查询,可选匹配替换匹配、值约束条件、处理多个图、在特定图中查找匹配、查找包含某个模式的图、组合后台数据和命名图。

目前各大研究机构已发布了 30 余种语义查询语言,这些查询语言在数据模型、表达能力、模式信息支持和查询类型上各有特点,在此不再逐一描述。

3.5.3 语义查询语言评价

从关系操作角度来看,所有的查询语言都支持 3 种基本的代数操作:选择、投影、笛卡尔积,并用于路径表达查询。但对并和差的操作的支持能力不尽相同。对关系操作的支持程度是评价语义查询语言表达能力的指标之一。

从聚合、分组及递归功能角度来说,聚合功能是从一个多值集合中计算一个数量值,分组则可以实现按组计算。递归在语义查询中指子类的递归性。一个语义完备的语义查询语句应该具有聚合、分组和递归的功能。

从推理能力的强弱,评价查询语言性能。目前的许多 RDF 查询语言都缺乏完善的推理功能对显性知识进行推理而获得隐性知识。另外,缺乏一种像关系数据库的关系代数那样的综合、全面的 RDF 查询代数来提供一组语义明确的操作,作为 RDF 查询的形式化基础。

4 我们的工作

在参考国内外同行的工作成果和经验的基础上,结合自身技术优势和遇到的实际问题,取得了一系列研究进展:从语义 Web 技术兴起到现在短短几年时间内,我们率先完成第一个基于语义 Web 的中文搜索系统原型[43] 总体设计和实现;运用自然语言处理、机器学习、本体指导等方法构建功能不同的 Web 语义信息抽取工具,联合组成自动语义信息抽取和标注引擎,满足对 Web 异构资源的抽取和整合,实现 Web 到语义 Web 的延拓[44,45];提供自然语言查询接口,在自然语言处理技术的辅助下,采用一系列机制尽可能对用户隐藏复杂的语义查询语言和领域本体背景,为用户提供友好智能的人机交互接口[46];运用本体技术构建领域本体并检测领域本体一致性[47],并在一致性检测算法研究和效率改进方面取得了进展[48]。研究语义信息存储和查询机制,设计高效的存储模式和查询算法,实现语义查询功能[49,50]。

5 语义 Web 搜索技术展望

就当前研究的重点和热点来看,可以预见语义 Web 搜索技术将会在以下几个方面有所突破:首先,未来的语义搜索引擎应该不仅能够对没有语义的 Web 资源进行语义化,还应该对已有语义的 Web 资源进行识别、同化和重用。其次,未来的语义搜索引擎应该采用多本体架构,能够同时处理多个领域本体,并且应当具有同领域本体间的本体匹配和本体映射的良好机制。再次,目前的语义搜索引擎系统多从传统的智能信息系统演化而来,如何找到 Web 资源特性和语义搜索引擎系统的推理机制结合点,将是未来研究的重点。再一个值得关注的是 Web2.0 的兴起,在这一框架下需要为终端用户提供添加和标注数据机制,并且深度整合分布式资源。

结束语 可以看到无论是语义 Web 资源自动处理还是多本体支持机制,无论是 Web 资源特性和推理的结合还是 Web2.0 与语义 Web 的融合,都对未来的语义搜索技术提出了巨大的挑战,同时给语义搜索技术的发展创造了巨大空间。

参 考 文 献

[1] Fensel D. The semantic web and its languages[J]. IEEE Intelligence Systems, Nov/Dec 2000

[2] Berners-Lee T. Semantic Web. W3C Recommendation [EB/OL]. http://www.w3.org/2000/Talks/1206-xml2k-tbl/, 2000

[3] Berners-Lee T, Hendler J, Lassila O. The Semantic Web[J]. Scientific American, May 2001

[4] Gruber T. A Translation Approach to Portable Ontologies[J]. Journal on Knowledge Acquisition, 1993, 5(2): 199-220

[5] Schreiber A T, et al. Ontology-based photo annotation[J]. IEEE Intelligent Systems, 2001, 16: 66-74

[6] Lewis D D, Jones K S. Natural Language Processing for Information Retrieval[J]. Communications of the ACM, 1996, 39(1): 92-101

[7] Schraefel M C, Shadbolt N R, et al. CS AKTive Space: Representing Computer Science in the Semantic Web[C]// Proceedings of the 13th International World Wide Web Conference (WWW 2004). 2004: 384-392

[8] Hyvonen E, et al. MuseumFinland-Finnish Museums on the Semantic Web[J]. Journal of Web Semantics, 2005, 3(2): 224-241

[9] Mika P. Flink: Semantic Web Technology for the Extraction and Analysis of Social Networks[J]. Journal of Web Semantics, 2005, 3(2): 211-223

[10] Lei Y, Lopez V, Motta E. An Infrastructure for Building Semantic Web Portals[C]// WISM '06. Grand Duchy, Luxembourg, 2006: 1019-1035

[11] Bollegala D, Matsuo Y, Ishizuka M. Measuring semantic similarity between words using Web search engine[C]// WWW 2007. Banff Alberta, Canada, 2007: 757-766

[12] Sahuguet A, Azavant F. Building intelligent Web applications using lightweight wrappers[J]. Data & Knowledge Engineering, 2001, 36(3): 283-316

[13] Liu L, Pu C, Han W. XWRAP: An XML-enabled Wrapper Construction System for Web Information Sources[C]// ICDE 2000. San Diego, California, 2000: 611-621

[14] Crescenzi V, Mecca G, Merialdo P. RoadRunner: Towards Automatic Data Extraction from Large Web Sites[C]// VLDS 2001. Rome, Italy, 2001: 109-118

[15] Soderland S. Learning Information Extraction Rules for Semi-

1) http://www.w3.org/TR/rdf-sparql-query/

structured and Free Text[J]. Machine Learning,1999,34(1-3): 233-272

[16] Embley D W,Jiang Y S,Ng Y K. Record-boundary Discovery in Web Documents[C]//ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery. Philadelphia, PA,1999:467-478

[17] Kushmerick N. Wrapper induction: Efficiency and expressiveness[J]. Artificial Intelligence,2000,118(1/2):15-68

[18] Hsu C N, Dung M T. Generating Finite-state Transducers for Semi-structured Data Extraction from the Web[J]. Information Systems,1998,23(8):521-538

[19] Muslea I,Minton S,Knoblock C A. Hierarchical Wrapper Induction for Semistructured Information Sources[J]. Autonomous Agents and Multi-Agent Systems,2001,4(1/2):93-114

[20] Adelberg B. NoDoSE-A Tool for Semi-automatically Extracting Semi-structured Data from Text Documents[C]//Proceedings ACM SIGMOD International Conference on Management of Data. Seattle,Washington,1998:283-294

[21] Li Z,Ng W K. WDEE: Web Data Extraction by Example[C]// DASFAA 2005. 2005:347-358

[22] Embley D W,Campbell D M,et al. Conceptual-model-based Data Extraction from Multiple-record Web Pages[J]. Data & Knowledge Engineering,1999,31(3):227-251

[23] Guha R V,McCool R,Miller E. Semantic search[C]//WWW 2003. Budapest,Hungary,2003:700-709

[24] Lei Y,Uren V S,Motta E. SemSearch: A Search Engine for the Semantic Web[J]//EKAW 2006. 2006:238-245

[25] Corby O,Dieng-Kuntz R,et al. Searching the semantic web: approximate query processing based on ontologies[J]. IEEE Intelligent Systems,2006,21(1):20-27

[26] Sinha V, Karger D R. Magnet: supporting navigation in semi-structured data environments[C]//ACM SIGMOD Conference 2005. Baltimore,Maryland,2005:97-106

[27] Wu H,Cheng G,Qu Y. 2006 Falcon-S: An Ontology-based Approach to Searching Objects and Images in the Soccer Domain [C]//Supplemental Proceedings o4f ISWC. Nov. 2006

[28] Athanasis N,Christophides V, Kotzinos D. 2004 Generating on the fly queries for the semantic web: the ICS-FORTH graphical RQL interface (GRQL)[C]//ISWC'04. Hiroshima,Japan, 2004:486-501

[29] Catarci T,Di Mascio T,et al. An ontology based visual tool for query formulation support[C]//ECAI-04. Valencia, Spain, 2004:308-312

[30] Hyvonen E, Saarela S, Viljanen K. 2003 Ontogator: combining view-and ontology-based search with semantic browsing[C]// XML Finland 2003. Kuopio,Finland,2003:82-85

[31] Broekstra J, Kampman A. SeRQL: A Second Generation RDF Query Language[C]//Proceedings of SWAD-Europe Workshop on Semantic Web Storage and Retrieval. Amsterdam,2003:13-24

[32] Lopez V,Pasin M,Motta E. 2005 AquaLog: an ontology-portable question answering system for the semantic Web[C]//ESWC 2005. Heraklion,Crete, Greece,2005:546-562

[33] Klyne G,Carroll J. Resource Description Framework (RDF): Concepts and Abstract Syntax, W3C Recommendation [EB/OL]. http://www.w3.org/TR/rdf-concepts/,2004

[34] Wilkinson K,Sayers C,Kuno H,et al. Efficient RDF Storage and Retrieval in Jena2[C]//Proceedings of the 1st International Workshop on Semantic Web and Databases. Berlin, Germany, 2003:131-151

[35] Beckett D. The design and implementation of the Redland application framework[C]//WWW 2001. Hong Kong,2001:449-456

[36] Hayes P. Resource Description Framework (RDF): Semantics [EB/OL]. http://www.w3.org/TR/2004/REC-rdf-mt-20040210/#rdf-entail,2004

[37] Broekstra J,Kampman A,Harmelen F V. Sesame: An Architecture for Storing and Querying RDF Data and Schema Information[C]//Spinning the Semantic Web. MIT Press,2003:197-222

[38] Shen W, Qu Y. An RDF Storage and Query Framework with Flexible Inference Strategy[C]//APWeb 2006. Harbin,China, 2006:166-175

[39] Fensel D. Ontology-Based Knowledge Management[J]. IEEE Computer,2002,35(11):56-59

[40] Beckett D,Grant J. Mapping Semantic Web Data with RDBMSes [OL]. http://www.w3.org/2001/sw/Europe/reports/scalable-rdbms-mapping-report/,2003

[41] Miller L, Seaborne A, Reggiori A. Three Implementations of SquishQL: a Simple RDF Query Language[C]//International Semantic Web Conference 2002. Sardinia,Italy,2002:423-435

[42] Karvounarakis G,Alexaki S,Christophides V,et al. RQL: A Declarative Query Language for RDF[C]//WWW 2002. Budapest, Hungary,2002:592-603

[43] Che H Y,Sun J G,Jing T,et al. A Prototype of Semantic-based Intelligent Search Engine for Chinese Documents[C]//FSKD 2007. Haikou,China,2007:663-667

[44] 车海燕,孙吉贵,等. 一个基于本体主题的中文知识获取方法[J]. 计算机科学与探索,2007,1(2):206-215

[45] Sun J G, Bai X, et al. Towards a Wrapper-driven Ontology-based Framework for Knowledge Extraction[C]//KSEM 2007. Melbourne,Australia,2007:230-242

[46] Shi L, Sun J G, Che H Y. Populating CRAB Ontology Using Context-profile-based Approaches[C]//KSEM 2007. Melbourne,Australia,2007:210-220

[47] Shi L,Sun J G,Lu S,et al. Flexible Planning Using Fuzzy Description Logics[C]//ICAI 2007. USA,2007:306-312

[48] 叶育鑫,欧阳丹彤,等. 规则与本体整合的推理方法研究与设计[J]. 吉林大学学报:工学版(已录用)

[49] Bai X,Sun J G,Li Z H,et al. Domain ontology learning and consistency checking based on TSC approach and Racer[C]// WRRS 2007. Innsbruck,Austria,2007:148-162

[50] 叶育鑫,欧阳丹彤,等. 基于 SHOIQ(D) 的本体一致性检测[J]. 计算机工程与科学(已录用)