CN 43-1258/TP
ISSN 1007-130X

计算机工程与科学
Computer Engineering & Science

第 37 卷第 11 期 2015 年 11 月
Vol. 37, No. 11, Nov. 2015

文章编号:1007-130X(2015)11-2188-07

甲骨文知识图谱构建中的实体关系发现研究*

熊　晶[1],钟　珞[2],王爱民[1,2]

(1.安阳师范学院计算机与信息工程学院,河南 安阳 455000;2.武汉理工大学计算机科学与技术学院,湖北 武汉 430070)

摘　要:为解决甲骨文信息处理研究中对专家的依赖性高而专家知识共享率却很低这一突出矛盾,提出构建知识图谱的方案,其核心问题是如何发现实体和实体间的关联关系。基于甲骨文的古籍特征,先通过知识图谱 MKD 的共词分析发现实体及其关系,再利用甲骨文本体和规则推理发现隐含的语义关系,最后将其与 MKD 共词关系进行融合。实验结果表明,该方法可以有效解决甲骨文知识图谱中实体关系发现问题。

关键词:甲骨文;知识图谱;本体;语义相似度;甲骨文知网

中图分类号:TP391.1　　**文献标志码**:A

doi:10.3969/j.issn.1007-130X.2015.11.029

Research on entity relation discovery for Oracle Bone Inscriptions knowledge mapping construction

XIONG Jing[1], ZHONG Luo[2], WANG Ai-min[1,2]

(1. School of Computer and Information Engineering, Anyang Normal University, Anyang 455000;
2. School of Computer Science and Technology, Wuhan University of Technology, Wuhan 430070, China)

Abstract: During the research process of Oracle Bone Inscriptions (OBI) information processing, there is a prominent contradiction between high dependence on experts and low sharing of their OBI knowledge. In order to solve this problem, we propose an OBI knowledge mapping scheme, the crucial issue of which is how to find the ontology relations. Firstly, based on the OBI features, the co-word analysis of Mapping Knowledge Domains (MKD) is done to find the ontology relations. Secondly, OBI ontology and reasoning are used to find the implicit semantic relations. Finally, the co-word relations and ontology relations are fused to describe the OBI knowledge mapping relations. Experimental results show that the method can effectively solve the problem of entity relation discovery in OBI knowledge mapping construction.

Key words: Oracle Bone Inscriptions; knowledge mapping; ontology; semantic similarity; OBIHownet

1　引言

甲骨文距今已有 3 500 多年的历史,记载了商代王室的占卜记录,其历史内容非常丰富,具有重要的史料价值[1]。自 1899 年甲骨文首次发现以来,经过 110 多年的发展,甲骨文的研究已形成一门具有严密规律、有丰富研究资料和多方面研究课题的学科—甲骨学。甲骨学和语言文字学、历史学、考古学、古代科学史、历史文献学、人类学等学科有着紧密的联系[2],因此研究甲骨文不能孤立地以释读甲骨片上的文字为对象,而应充分考虑其与相关学科的关系。然而,传统的甲骨文研究方式难度极大,培养一名甲骨文专家需要一二十年甚至更长的时间[2],这严重阻碍了甲骨文的研究进展。计算甲骨学利用信息技术改善传统的研究方式,综合

* **收稿日期**:2015-07-18;**修回日期**:2015-09-26
基金项目:国家自然科学基金资助项目(U1504612,61572037);河南省高校科技创新人才支持计划资助项目(15HASTIT023);河南省教育厅科学技术研究重点资助项目(14A520038)
通信地址:455000 河南省安阳市安阳师范学院计算机与信息工程学院
Address: School of Computer and Information Engineering, Anyang Normal University, Anyang 455000, Henan, P. R. China

语言学、逻辑学、哲学、计算机科学、人工智能、数学和统计学等学科来进行甲骨文信息处理的研究,是解决或缓解这一问题的有效途径。

目前,甲骨文信息处理研究取得了较多成果,包括甲骨文字库、计算机辅助甲骨缀合、甲骨文语料标注、甲骨文字编辑、甲骨文机器翻译等多方面[3],安阳师范学院在这些方面做了一些有意义的工作,积累了大量的多源异构的甲骨文数据集[1]。但是,这些研究均是针对甲骨文字本身的,而忽略了与甲骨文息息相关的其它学科背景知识,也忽略了甲骨文的知识演化历程,如甲骨文、金文、小篆、隶书、楷书、草书、行书等书体变化,汉字结构上的自然流变,以及与现代汉语的传承关系等。目前这一问题的解决依赖于甲骨学专家的研究经验及其学术积累,但这是一个长期的过程;而且,专家的知识无法有效地共享和重用,往往专家需要重复解决同一个问题。如何采取有效的手段充分共享已有的研究成果和专家知识?知识图谱具有"图"和"谱"的双重特性,既是可视化的知识图形,又是序列化的知识谱系,可以胜任这一工作。因此,本文的研究目标是通过发现实体及实体间的关联关系,构建甲骨文知识图谱 KM (Knowledge Mapping)。

2　知识图谱研究现状

目前,知识图谱的研究有两大主流:基于文献计量学的科学知识图谱 MKD (Mapping Knowledge Domains)和以 Google 为代表的知识图谱 KG (Knowledge Graph)。

2.1　科学知识图谱 MKD

MKD 是将应用数学、图形学、信息可视化技术、信息科学等学科的理论与方法与计量学引文分析、共现分析等方法结合,用可视化的图谱形象地展示学科的核心结构、发展历史、前沿领域以及整体知识架构的多学科融合的一种研究方法[4]。MKD 是目前计量学等学科关注的前沿学术领域之一,被广泛应用于社会科学与自然科学领域,并已从情报学迅速扩散到其它学科领域[5]。国外具有影响力的机构有德莱克斯大学、布鲁内尔大学、美国的圣蒂亚国家实验室、荷兰的伊拉兹马斯大学等[6],其中德莱克斯大学的陈超美[7]被认为是该领域的领军人物,其开发的知识图谱分析软件 CiteSpace[8]获得了广泛的应用。

国内,大连理工大学、武汉大学、中国科学院、浙江树人大学、南京大学、天津师范大学、河北大学等单位研究实力较强[9],其中大连理工大学[10]最早从科学计量学视角引进 MKD 方法,成为推动国内 MKD 研究的先行者[9]。研究表明,MKD 既可从时间轴纵向揭示特定领域的不同子领域的研究演化模式及其相互关系;也可以从横向比较中揭示研究主题接近所属领域热点问题的程度,进而预测领域知识的发展趋势。但是,一方面,目前相关研究均以文献计量为本,侧重分析学科结构及布局、研究领域进展、重点研究方向及热点、主流研究机构和学者及其合作网络等宏观知识群,而在分析领域内部微观知识方面研究较少;另一方面,MKD 的分析关系大多是直接或间接关联关系,无法表达甲骨文知识中深层次的语义关系,如商王世系关系、贞人与商王的关系、方国地理位置关系、祭祀对象关系等。因此,仅采用 MKD 不足以表达甲骨学领域中的微观知识。

2.2　Google 知识图谱 KG

自 2012 年 Google KG 融入 Google 搜索引擎之后,迅速成为研究热点,引发了大规模知识库的又一轮研究热潮。KG 是一个描述客观世界有关人物、地点和事物以及它们之间的关联关系的系统,其结点表示实体或概念,通过边连接实体/概念之间的语义关系。因此,构建 KG 的关键是确定实体及其关系。目前,大量研究是利用网络资源尤其是 Wiki 类资源和链接开放数据 LOD (Linking Open Data)[11],通过抽取实体及其关系来构建 KG。文献[12]基于 Wikipedia 实现了社交媒体的实体抽取、链接、分类及标注;文献[13]基于 Wikipedia 实现概念、实体及关系的抽取;文献[14]通过抽取网页中的文本、表格数据、网页结构和人工标注信息进行融合,构建了一个网络级的概率知识库 Knowledge Vault;Zhishi. me[15]通过整合百度百科、互动百科和中文维基百科资源构建了中文 LOD;文献[16]基于中英文维基百科以及百度百科和互动百科构建了跨语言知识图谱;文献[17]研究了维基百科中实体间缺失语义关系的发现算法;基于跨语言知识库;文献[18]通过先给定少量的种子链接,再利用概念标注方法发现新的链接,实现了知识的扩充。

综上所述,目前 KG 的构建方法,大多是综合 Wiki、网络知识库、搜索日志、LOD、社会网络、众包等资源实现实体抽取和实体链接,并利用本体进行知识映射或知识融合。但是,在线的甲骨学数据及知识描述资源极少,绝大多数数据均以不同的形式存储在各研究机构,因此针对这类线下数据需要

重新考虑知识实体的发现及关系挖掘方法。

3 实体关系发现

由于甲骨文的古籍特性,使得甲骨文研究必须充分依赖大量的文献资料,而 MKD 在文献计量方面极具优势。甲骨文研究离不开相关的辅助学科,如借助考古学,去解决甲骨出土问题;借助历史学和文献学,去解决甲骨文中的殷商历史问题;借助语言学,去解决甲骨文字的问题。同时,甲骨文研究又促进了相关学科的发展,并延伸到其它学科领域。这些学科的文献资料对甲骨文的研究起到积极作用。

但是,MKD 的分析关系大多是直接或间接的关联关系,无法表达甲骨文知识中深层次的语义关系,如商王世系关系、贞人与商王的关系、方国地理位置关系、祭祀对象关系等,而且甲骨文中还存在大量需借助推理获取的潜在关系,如甲骨字考释、甲骨文分期断代、残辞拟补等。因此,引入本体对 MKD 进行优化,提出图 1 所示的实体关系发现框架。

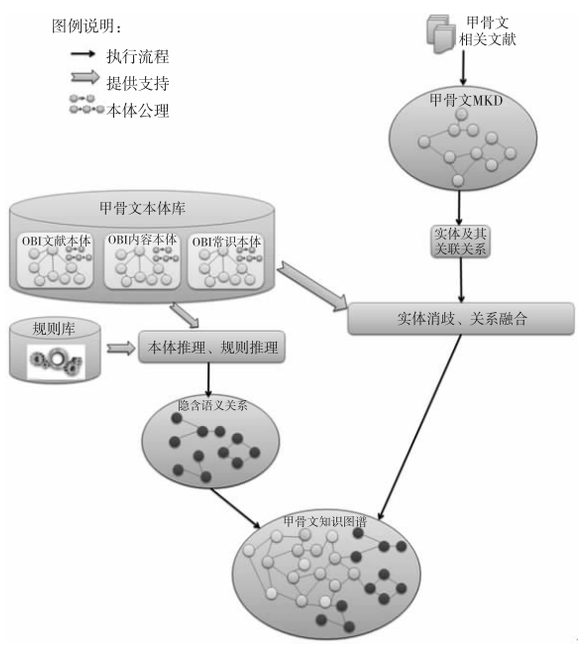


Figure 1 Entity relation discovery framework of OBI knowledge mapping

图 1 甲骨文知识图谱实体关系发现框架

3.1 利用 MKD 发现实体关系

现有的 MKD 绘制方法较多,如共引分析法、共词分析法、聚类分析法、社会网络分析法等。在众多的方法中,具有知识表达功能的元素只有引文、分类和词(短语),如基于引文的共引分析和基于主题词的共词分析等。引文指向的是一篇文献,代表的是一条法则、规律或一个问题,是若干知识单元的集合,所以基于引文的分析是不能反映微观层次的知识关联的。分类一般代表的是一个综合的知识领域,即使是比较详细的分类,对微观知识的反映也是有限的[19]。共词分析方法通过分析在同一个文本主体中的关键词对共同出现的形式,确定文本所代表的学科领域中主题间的关系,从而分析该领域的科学发展[20]。为了明确分析知识之间的关联关系,只有选择知识继承与发展的最小功能单元——词作为分析对象[19]。因此,选择共词分析法描述知识之间的联系。

本文选择余弦函数作为共词关系 RC(Relation of Co-word)的计算方法。设存在关键词集合 $D_i=(d_{1i},d_{2i},\cdots,d_{ni})^{\mathrm{T}}$ 和关键词集合 $D_j=(d_{1j},d_{2j},\cdots,d_{nj})^{\mathrm{T}}$,其中 d_{1i} 和 d_{1j} 均为文献中的关键词,则 D_i 与 D_j 之间由余弦函数计算的 RC 表示为:

$$RC=\frac{\sum_{k=1}^{n}d_{ki}d_{kj}}{\sqrt{\sum_{k=1}^{n}d_{ki}^2\sum_{k=1}^{n}d_{kj}^2}}$$

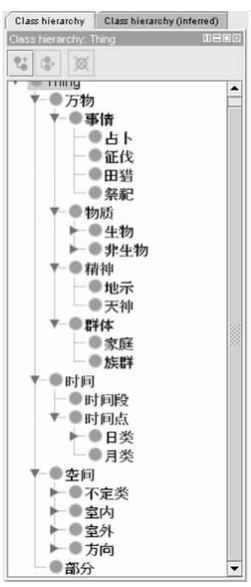
3.2 甲骨文本体库

用于建立 RC 的词往往是来自于文献的关键词列表,或者从标题、摘要或全文中提取的主题词,它们体现的是词的共现关系,在一定程度上可以描述同一或相近主题之间的知识关联。但是,RC 缺乏深层次的语义描述功能,尤其是在描述甲骨文领域知识中人、事、物、时空之间的语义关系方面无能为力。而本体作为共享概念模型的形式化规范说明[21],可以精确定义领域内的概念及它们之间的关系,可以很好地弥补 RC 在语义表达能力上的不足。

在甲骨文信息处理的研究成果基础上构建甲骨文本体库,可为 MKD 提供概念模型和逻辑基础。甲骨文本体库主要包括甲骨文文献本体、甲骨文内容本体和甲骨文常识本体三个本体[1]。其中,甲骨文文献本体描述的是关于甲骨文研究论文及专著的资源本体,如图 2a 所示;甲骨文内容本体是描述经甲骨文专家及历史学家考释出来的,反映商代社会国家机构、政治、社会、军事、文化、科技、经济等内容及其相互关系的知识库,如图 2b 所示;甲骨文常识本体描述的是甲骨文基础知识,包括甲骨文发现历史、考古记录、文字特征、语法知识等,如图 2c 所示。

a 甲骨文文献本体

b 甲骨文内容本体



c 甲骨文常识本体

Figure 2　OBI ontology library

图 2　甲骨文本体库

3.3　语义相似度和相关度计算

基于本体实现实体消歧和语义扩展,需要考虑语义相似度和语义相关度的计算。语义相似度是指两个实体(或概念)之间存在某些共同特性。相似度计算通常考虑 is-a 关系;语义相关度是指两者之间可能不存在相似性,但是可以通过某些其它关系关联起来。相关度计算通常考虑 is-a 之外的关系,如 part-of 关系等[22]。

甲骨文专家研究发现甲骨文同现代汉语有许多相同之处,如造字法、用字法、词的分类和句型等。商代人对事物的认识也与现代人大同小异:甲骨文所载卜辞大体可分为名物类(包括地理、天象、建筑、时间、空间、方位、物品、人、鬼神、动植物、组织、称谓等)、事件类(包括生产、生活、军事行动、占卜、祭祀等)、性质状态类、数量类等。HowNet[23]规定了现代汉语最基本的运算单元,即万物(包括物质和精神)、部件、属性、属性值、事件、时间和空间等。HowNet 在语义相似度和相关度计算方面有着广泛的应用,因此扩展 HowNet 体系的描述范围,构建"甲骨文知网 OBIHownet (Oracle Bone Inscriptions HowNet)",将甲骨文同现代汉语在语义上进行融合,实现 HowNet 在甲骨文方面的语义计算。OBIHownet 构建步骤参见文献[1]。

目前我们已添加 300 个甲骨词条到 HowNet,扩充后的 HowNet 在语义相似度方面的对比结果如图 3 和图 4 所示。

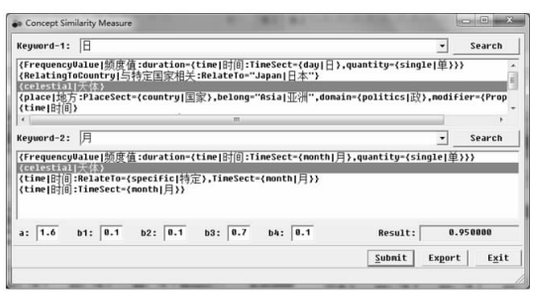


Figure 3　Semantic similarity computing based on HowNet

图 3　HowNet 语义相似度计算

图 3 显示,HowNet 中"日""月"的语义相似度达到 0.95,这是因为该例中"日"和"月"均是"天体"的实例,因此相似度很高。

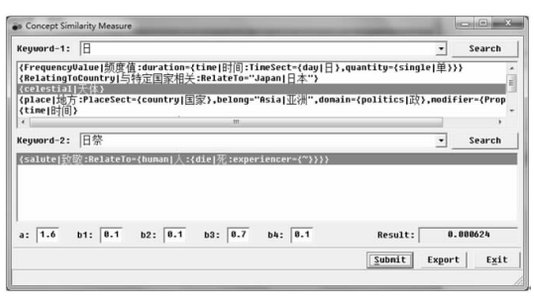


Figure 4　Semantic similarity computing based on OBIHownet

图 4　"甲骨文知网"的语义相似度计算

图 4 显示,OBIHownet 在 HowNet 中加入了甲骨文词条"日祭"后,与"日"的相似度只有 0.000 624。这个结果与甲骨文知识是吻合的,因为"日祭"是一种祭祀事件,是指对太阳神的祭拜,从而与作为"天体"实例的"日"相似度很低。

实验表明,通过构建 OBIHownet 进行语义相似度和相关度的计算是可行的。

3.4　隐含语义关系的发现

利用知识推理可以发现更多的隐含知识,并挖掘实体间更多的潜在语义关联。本文的知识推理

包含两个部分:一是基于本体关系的推理,即利用甲骨文本体已经定义好的关系(既有 kind-of、instance-of、property-of、part-of、equivalence 等通用关系,也有甲骨学领域涉及的商王世系、占卜事件、地理位置、时间空间等复杂的语义关系)和公理进行推理,还可以充分利用关系的传递性、自反性等进行推理;二是基于规则的推理,即需要书写显式的规则来表示本体无法直接完成的推理,如因果关系、甲骨字考释、甲骨文分期断代、残辞拟补等。

基于本体关系的推理参见作者的前期研究[24]。基于规则的推理需要建立有效的甲骨文知识规则库,利用本体推理机,获取更多的隐含语义信息。例如,假设主祭者 x 的祭祀对象是 y,而 y 是 z 的子孙后代,则 x 也会祭祀 z。例如,"祖甲"祭祀"武丁",而"武丁"又是"小乙""祖丁""祖辛"等商王的后代,则"祖甲"也会祭祀"小乙""祖丁""祖辛"等。规则表示如下:

RULE:(?x sacrifice ?y),(?y hasAncestor ?z)→(?x sacrifice ?z)

将规则写入 obi.rules 文件,再利用 Jena 实现推理,关键代码如下:

```
List rules = Rule.rulesFromURL("file:obi.rules");
  //obi.rules 文件中存储推理规则
GenericRuleReasoner reasoner = new GenericRuleReasoner(rules);
OntModel om = ModelFactory.createOntologyModel
  (OntModelSpec.OWL_MEM, model);
Resource configuration = om.createResource();
InfModel inf = ModelFactory.createInfModel(reasoner, om);
```

3.5 共词关系与本体关系的融合

基于本体关系和规则的推理,可以得到层次关系 RH(Relation of Hierarchy)、语义关联关系 RR(Relation of Relevance),将其与共词关系 RC 进行融合。融合时 RH 的权重系数 W_{rh}、RR 的权重系数 W_{rr} 和 RC 的权重系数 W_{rc} 可以通过下列公式[19]计算:

$$W_{\mathrm{rh}}=\frac{rh\_max^2}{rh\_max^2+rr\_max^2+rc\_max^2}$$

$$W_{\mathrm{rr}}=\frac{rr\_max^2}{rh\_max^2+rr\_max^2+rc\_max^2}$$

$$W_{\mathrm{rc}}=\frac{rc\_max^2}{rh\_max^2+rr\_max^2+rc\_max^2}$$

其中,rh_max、rr_max 和 rc_max 分别为关键词集合的层次关系矩阵、关联关系矩阵和共词关系矩阵中非对角线的最大值,且满足 $rh\_max+rr\_max+rc\_max=1$。

关系融合即是将三类关系矩阵进行合并,合并操作时,先将 RC 分别与 RH、RR 进行合并,再将三者进行融合。

4 实验及分析

以 CNKI 收录的文献为数据来源,用"甲骨文"作为主题词检索 1986~2015 年的文献,从检索结果中随机选取 350 篇进行实验。由于文献的标题、关键词和摘要已经能反映出甲骨文知识的大部分内容,因此实验中只取文献的标题、关键词和摘要进行共词分析,而不需要文献全文。利用 CiteSpace 工具,采用余弦函数进行的共词分析片段如图 5 所示。图 5 所示的结果存在的最大问题是未考虑语义关系对关键词的优化,如"比较""字符""字频"等对表示甲骨文知识的意义不大,应该剔除;相反,有助于表示甲骨文知识的同义词、上位词、下位词等关系则没有体现出来。因此,利用本体进行优化并与 RC 进行融合,得到的结果如图 6 所示。

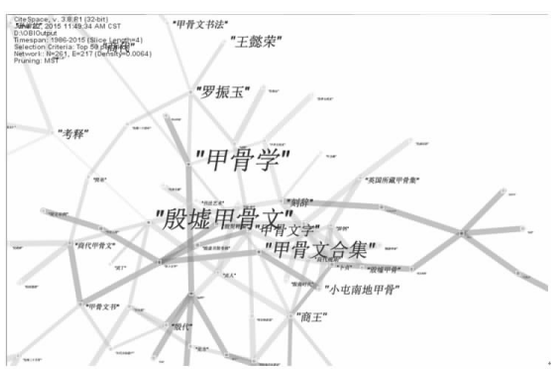


Figure 5 Co-word analysis of OBI knowledge mapping

图 5 甲骨文知识图谱的共词分析片段

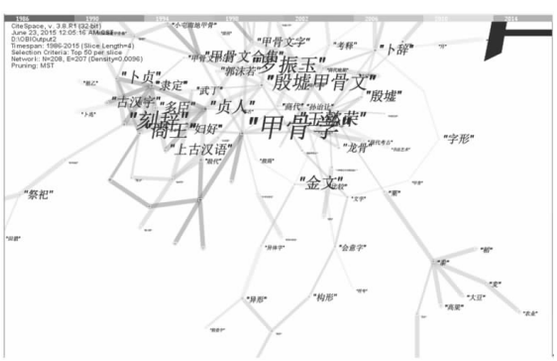


Figure 6 OBI knowledge mapping of co-word relations and ontology relations

图 6 共词关系与本体关系融合的知识图谱片段

图6结果较之图5有更为丰富的语义关系。如通过图5中的“武丁”可以获取图6中的“妇好”,因为本体中定义了“武丁”的妻子是“妇好”这一关系;利用本体关系,通过“刻辞”能获取“卜辞”;通过“粟”,可在本体中得到其属于“农业”,进而得到“麦”“大豆”“黍”等更多的农作物。

实验中还发现,利用本体可以剔除与甲骨文领域无关的信息。如“甲骨文软件中国公司的差异化竞争战略研究”“甲骨文公司的应用产品发展策略研究”等文献的信息不会出现在甲骨文KM中,原因是甲骨文本体限定了描述领域为古文字中的“甲骨文”而不是软件公司“甲骨文”。

5　结束语

本文针对甲骨文专家知识的共享及重用程度低的问题,提出了构建甲骨文知识图谱的解决方案,其关键技术是发现知识图谱中实体的关联关系。甲骨文文献是甲骨文研究的第一手资料,基于MKD的共词分析方法可以发现甲骨文知识实体及实体之间的关系,利用本体的语义优势,融合MKD共词关系,从而发现更多的隐藏语义关系。在下一步研究计划中,将扩充本体库及OBIHownet,并考虑利用规则挖掘技术来减轻手工书写规则的负担。

参考文献:

[1] Xiong Jing, Gao Feng, Wu Qin-xia. Research on semantic mining for large-scale oracle bone inscriptions foundation data [J]. New Technology of Library and Information Service, 2015, 31(2):7-14. (in Chinese)

[2] Jiang Ming-hu. Natural language processing[M]. Beijing: Higher Education Press, 2006. (in Chinese)

[3] Xiong Jing, Zhong Luo, Wang Ai-min. Example and ontology based machine translation for oracle bone inscriptions [J]. Journal of Huazhong University of Science and Technology (Nature Science Edition), 2013, 41(S2):222-226. (in Chinese)

[4] Qin Chang-jiang, Hou Han-qing. Mapping knowledge domain-a new field of information management and knowledge management [J]. Journal of Academic Libraries, 2009, 27(1):30-37. (in Chinese)

[5] Hu Ze-wen, Sun Jian-jun, Wu Yi-shan. Research review on application of knowledge mapping in China [J]. Library and Information Service, 2013, 57(3):131-137. (in Chinese)

[6] Liu Ze-yuan, Chen Yue, Hou Hai-yan, et al. Mapping knowledge domains methods and application [M]. Beijing: People's Publishing House, 2008. (in Chinese)

[7] Chen C, Ibekwe Sanjuan F, Hou J. The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis[J]. Journal of the American Society for Information Science and Technology, 2010, 61(7):1386-1409.

[8] Chen C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature [J]. Journal of the American Society for information Science and Technology, 2006, 57(3):359-377.

[9] Tang Jian-min, Yu Feng-min. Review and evaluation of knowledge mapping research in China:2004-2010 [J]. Information and Documentation Services, 2012(1): 16-21. (in Chinese)

[10] Chen Yue, Liu Ze-yuan. The rise of mapping knowledge domain [J]. Studies in Science of Science, 2005, 23(2):149-154. (in Chinese)

[11] Bizer C, Heath T, Berners-Lee T. Linked data-the story so far [J]. International Journal on Sematic Web and Information Systems, 2009, 5(3):1-22.

[12] Gattani A. Entity extraction, linking, classification, and tagging for social media: A wikipedia-based approach[J]. Proceeding of the VLDB Endowment, 2013, 6(11):1126-1137.

[13] Deshpande O, Lamba D S, Tourn M, et al. Building, maintaining, and using knowledge bases: A report from the trenches[C]// Proc of the 2013 ACM SIGMOD International Conference on Management of Data, 2013:1209-1220.

[14] Dong X, Gabrilovich E, Heitz G, et al. Knowledge vault: A web-scale approach to probabilistic knowledge fusion[C]// Proc of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014:601-610.

[15] Wu C, Wang H, Qu J, et al. ZhishiLink: entity linking on zhishi. me [C]// Proc of the 7th China Sematic Web Symposium, 2013:161-174.

[16] Wang Z, Li J, Wang Z, et al. XLore: A large-scale English-Chinese bilingual knowledge graph[C]// Proc of International Semantic Web Conference (Posters & Demos), 2013:121-124.

[17] Xu M, Wang Z, Bie R, et al. Discovering missing semantic relations between entities in Wikipedia[C]// Proc of the Semantic Web-ISWC 2013, 2013:673-686.

[18] Wang Z, Li J, Tang J. Boosting cross-lingual knowledge linking via concept annotation[C]// Proc of the 23rd International Joint Conference on Artificial Intelligence, 2013: 2733-2739.

[19] Wang Jian-fang. Analysis method of mapping knowledge domains based on ontology [EB/OL]. [2014-01-15]. http://ir.las.ac.cn/handle/12502/3837. (in Chinese)

[20] Xiao Zhi-xiong, Gu Jing. Co-word analysis-based study on research hotspots of synergetics in China [J]. Information Research, 2015(5):6-9. (in Chinese)

[21] Borst W N. Construction of engineering ontologies for knowledge sharing and reuse [D]. Ensched: Universiteit Twente, 1997.

[22] Liu Hong-zhe, Xu De. Ontology based semantic similarity and relatedness measures review [J]. Computer Science, 2012, 39(2):8-13. (in Chinese)

[23] Dong Z, Dong Q, Hao C. Hownet and its computation of meaning[C]// Proc of the 23rd International Conference on Computational Linguistics:Demonstrations, 2010:53-56.

[24] Xiong Jing, Wang Ai-min, Xu Jian-liang. Information retrieval optimization strategy based on domain ontology[J]. Computer Engineering and Design, 2011, 32(8):2695-2699. (in Chinese)

附中文参考文献:

[1] 熊晶, 高峰, 吴琴霞. 甲骨文大规模基础数据的语义挖掘研究[J]. 现代图书情报技术, 2015, 31(2):7-14.

[2] 江铭虎. 自然语言处理[M]. 北京:高等教育出版社, 2006.

[3] 熊晶, 钟珞, 王爱民. 基于实例和本体的甲骨文机器翻译方法研究[J]. 华中科技大学学报(自然科学版), 2013, 41(S2):222-226.

[4] 秦长江, 侯汉清. 知识图谱——信息管理与知识管理的新领域[J]. 大学图书馆学报, 2009, 27(1):30-37.

[5] 胡泽文, 孙建军, 武夷山. 国内知识图谱应用研究综述[J]. 图书情报工作, 2013, 57(3):131-137.

[6] 刘则渊,陈悦,侯海燕,等. 科学知识图谱:方法与应用[M]. 北京:人民出版社, 2008.

[9] 汤建民, 余丰民. 国内知识图谱研究综述与评估:2004-2010年[J]. 情报资料工作, 2012(1):16-21.

[10] 陈悦, 刘则渊. 悄然兴起的科学知识图谱[J]. 科学学研究, 2005, 23(2):149-154.

[19] 王建芳. 基于本体的科学知识图谱分析方法研究[EB/OL]. [2014-01-15]. http://ir.las.ac.cn/handle/12502/3837.

[20] 肖志雄, 谷静. 基于共词分析法的国内协同学研究热点分析[J]. 情报探索, 2015(5):6-9.

[22] 刘宏哲,须德. 基于本体的语义相似度和相关度计算研究综述[J]. 计算机科学, 2012, 39(2):8-13.

[24] 熊晶, 王爱民, 徐建良. 基于领域本体的信息检索优化策略[J]. 计算机工程与设计, 2011, 32(8):2695-2699.

作者简介:

熊晶(1979-),男,湖北团风人,博士,副教授,CCF 会员(E200018711M),研究方向为语义网和人工智能。**E-mail**:xiong-hb125@sohu.com

XIONG Jing, born in 1979, PhD, associate professor, CCF member(E200018711M), his research interests include semantic web, and artificial intelligence.

钟珞(1957-),男,湖南长沙人,博士,教授,研究方向为软件工程、人工智能和可视化研究。**E-mail**:zl@whut.edu.cn

ZHONG Luo, born in 1957, PhD, professor, his research interests include software engineering, artificial intelligence, and visualization research.

王爱民(1957-),男,河南长垣人,博士,教授,CCF 会员(E200012611S),研究方向为数据挖掘、智能技术和决策支持系统。**E-mail**:wam508@126.com

WANG Ai-min, born in 1957, PhD, professor, CCF member(E200012611S), his research interests include data mining, intelligent technology, and decision support system.