基于本体和局部共现的查询扩展方法

王旭阳 尉醒醒
(兰州理工大学计算机与通信学院 兰州 730050)

摘 要 将语义扩展与统计扩展相结合,提出了一种基于本体和局部共现的查询扩展方法,该方法利用本体和局部共现分别得到语义候选扩展概念集和统计候选扩展概念集,对这两个扩展集进行二次筛选以得到最终的查询扩展概念;并给出了一种计算扩展词权重的方法。实验结果表明,扩展后的查询更能反映用户的查询请求,在设计的语义检索系统中,该方法能有效提高查全率和查准率。

关键词 本体,局部共现,查询扩展,语义检索
中图法分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.01.041

Query Expansion Method Based on Ontology and Local Co-occurrence

WANG Xu-yang WEI Xing-xing
(College of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract Combining the semantic expansion and the statistics expansion, a method of query expansion based on ontology and local co-occurrence was proposed. The set of semantics candidate expansion concepts and the set of statistics candidate expansion concepts are respectively obtained by ontology and local co-occurrence. Then the final query expansion concepts are got by secondary filter. The method was presented to calculate the weight of expansion words. The results show that the query after expanding can reflect the user's query better. In the design of semantic retrieval system, the method can effectively improve recall and precision.

Keywords Ontology, Local co-occurrence, Query expansion, Semantic retrieval

1 引言

传统的检索系统大多是基于关键字匹配进行检索的,不可避免地出现了用户的查询请求与文档之间词不匹配的问题,造成检索系统的查全率和查准率较低,检索结果不能满足用户的需求。针对该问题,早在1986年Van Rijsbergen就提出了所谓的"查询扩展"[1],即在原查询的基础上,通过某种方法加入与之相关的词,使检索结果更满足用户的需求。查询扩展技术是解决词不匹配、信息迷向等问题的有效手段。

2 研究背景

目前国内外专家和学者提出了许多查询扩展技术,按照扩展词来源的不同主要分为以下几种方法:

(1)基于全局分析法的查询扩展[2]。它需要对整个文档集中的词进行相关性分析,构造叙词表,当接收到用户的检索词时,根据叙词表将与之相关度高的词扩展进来。该方法可以最大限度地挖掘词之间的关系,但是计算开销很大。

(2)基于局部分析的查询扩展[3,4]。它解决了全局分析计算量大的问题,其检索性能依赖于初始检索结果,当初始检索结果与原查询相关度不高时,扩展后会严重降低查准率。

(3)基于关联规则的查询扩展[56]。该方法利用数据挖掘技术挖掘词之间的关联规则,将关联规则的后件结论作为扩展词的来源。该方法解决了局部分析的不稳定性,但是扩展词的质量依赖于使用的挖掘技术,且关联规则的形成也比较困难。

(4)基于用户日志的查询扩展[7,8]。该方法是在大量用户查询日志的基础上建立用户空间和文档空间,将用户的查询和所点击的文档以条件概率的方式连接起来,当新的查询到来时选择相应的条件概率最大的文档用词加入查询中。它需要大量的用户日志存在,查询和文档较复杂时并不能保证较高的查全率和查准率。

以上传统的基于关键词的查询扩展方法不仅存在自身不可避免的缺点,而且对查询请求也缺少语义上的理解。本体作为一种很好的概念建模工具,能够描述概念及概念之间的关系,所以基于本体的查询扩展[9-12]成为了近几年的研究热点。然而有学者质疑基于本体的扩展方法脱离了待检索的语料集以及存在语义边界确定困难的问题。

为此,本文提出了一种基于本体和局部共现性分析的查询扩展方法。该方法将基于语义的查询扩展和基于统计的查询扩展相结合,即在语义扩展的基础上加入了概念之间在局部文档中的统计信息作为约束,同时为避免局部分析的不稳定性,统计信息中也加入了概念之间的相似度作为筛选函数。

3 基于本体和局部共现的查询扩展方法

3.1 基本思想及扩展模型

用户的初始查询经预处理后得到查询概念集,根据本体

到稿日期:2015-11-24 返修日期:2016-03-20

王旭阳(1974-),女,硕士,副教授,主要研究方向为智能信息处理、数据挖掘、模式识别;尉醒醒(1989-),女,硕士生,主要研究方向为智能信息处理。

中概念之间语义相似度大小得到各查询概念的本体候选扩展概念集，由局部文档中的共现性分析得到统计候选扩展概念集；再利用共现性分析对本体候选概念集进行二次筛选，同时利用概念在本体中的相似度对统计候选概念进行二次筛选；最后将筛选后的概念作为最终的扩展概念。其扩展模型如图 1 所示。

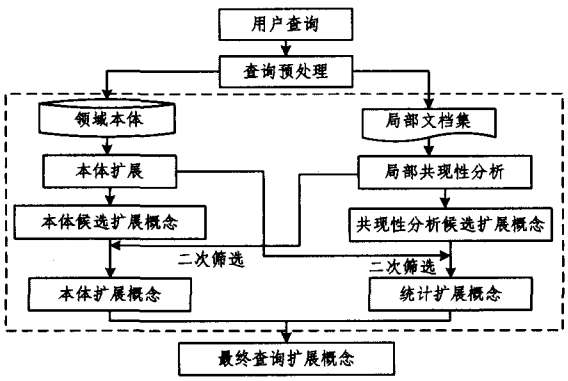

图 1　查询扩展模型

3.2　本体候选扩展概念集

基于本体的查询扩展主要根据概念在本体中的语义相似度进行扩展，将与查询概念相似度大的概念作为扩展概念。语义相似度指概念在语义上相符合的程度，其计算方法主要包括基于距离的方法、基于内容的方法和基于特征的方法。本文采用基于距离的方法来衡量概念之间的语义相似度。

假设给定概念 C_1 和 C_2，它们之间的相似度记为 $Sim(C_1,C_2)$：

$$Sim(C_1,C_2)=\begin{cases}0, & \text{两概念没有任何关系}\\ P, & P\in(0,1)\\ 1, & \text{两概念完全相同}\end{cases}$$

对于领域本体而言，概念之间都存在着某种关系，而对于两个完全相同的概念，在定义本体时其等价关系已标明，所以我们主要是计算相似度为 P 的这个值的大小。

通常把本体描述成一个层次结构，上层概念比下层概念更抽象，而下层概念都是对上层概念的具体化。当概念之间最短路径上的有向边个数相等时，下层概念之间的相似度要大于上层概念之间的相似度。例如概念对(RAM、ROM)和概念对(DRAM、SRAM)，其路径距离都是 2，显然后一对概念的相似度较大。

定义 1　本体层次结构中，根节点的深度定义为 1，即 $Deep(root)=1$，非根节点的深度定义为 $Deep(s)=Deep(f)+1$，其中 s 为任意子节点，节点 f 为节点 s 的父节点。

给定有向边 $f\rightarrow s$，它的深度为节点 f 的深度，则该有向边所表示的语义距离 $Dist(s,f)$ 为：

$$Dist(s,f)=\frac{1}{\frac{1}{2^{Deep(f)}}+\frac{1}{2^{Deep(f)-1}}+\cdots+\frac{1}{2}} \tag{1}$$

给定任意概念 C_1 和 C_2，其总的语义距离可表示为：

$$Dist(C_1,C_2)=Link[C_1,Nc(C_1,C_2)]+Link[C_2,Nc(C_1,C_2)] \tag{2}$$

$$Link[C_1,Nc(C_1,C_2)]=\sum_{n\in path(C_1,Nc(C_1,C_2))}Dist(n,parent(n)) \tag{3}$$

其中，$Nc(C_1,C_2)$ 表示节点 C_1 和 C_2 的最近公共祖先节点，$path(C_1,C_2)$ 表示在本体层次结构中 C_1 到 C_2 的最短路径上所有有向边的集合，$Link[\cdot,Nc(C_1,C_2)]$ 表示节点与最近公共祖先节点的语义距离。

计算得到概念 X 和 Y 之间的语义距离后，用下面的公式将语义距离转化成语义距离相似度：

$$Sim(C_1,C_2)=1-\frac{Dist(C_1,C_2)}{2(Maxlen-1)} \tag{4}$$

其中，$Maxlen$ 表示本体层次结构的最大深度。

假设用户的某次查询经预处理后得到查询概念集 $Q(q_1,q_2,\cdots,q_n)$，对其中的每一个概念 $q_i(i=1,2\cdots,n)$ 在本体中找到与其相关的概念集 $Set(C_1,C_2,\cdots,C_j)$，用式(4)计算 q_i 与 Set 集合中每一个概念 $C_j(j=1,2,\cdots,n)$ 的语义相似度 $Sim(q_i,C_j)$，将相似度大于 λ_1 的概念作为 q_i 的候选扩展概念，从而得到每个查询概念 q_i 的候选扩展概念集 $q_iOE(q_{i1},q_{i2},\cdots,q_{ik})$，其中 $k\leqslant j$。

3.3　统计候选扩展概念集

局部分析法通常用词项在局部文档中的词频信息来扩展初始查询。如果选出的局部文档与查询的相关性很大，那么该方法就能很好地扩展查询；反之，则会得到大量与原查询不相关的词语。所以文献[3]利用词语之间的共现性来选取扩展词。

两概念在一定的文本窗口范围内的共现性分析可以从某种程度上反映它们之间的相关度，共现度越大，则相关度越大。本文以一篇文档为有效窗口，则在有效窗口范围内概念 C_1 和 C_2 的共现频度为：

$$cf(C_1,C_2|d)=\log(tf(C_1|d)+1.0)\times\log(tf(C_2|d)+1.0) \tag{5}$$

其中，$tf(\cdot|d)$ 表示概念在文档 d 中出现的次数。

由于要在整个局部文档集 S 中分析概念 C_1 和 C_2 的关系，因此在 S 中的平均共现频度为：

$$acf(C_1,C_2|S)=\frac{\sum_{d\in S}cf(C_1,C_2|d)}{N} \tag{6}$$

其中，N 为共同出现概念 C_1 和 C_2 的文档数。

文献[2]分析得到：概念之间的相关度与它们在文档中的距离、共同出现的次数以及文档集中出现两者的文档数有关。因此用文献[13]给出的公式计算概念之间的共现度权值：

$$cw(C_1,C_2)=\frac{acf(C_1,C_2|S)\times\log(f_{C_1,C_2}+1.0)}{\log(avgs(C_1,C_2)+2.0)} \tag{7}$$

其中，f_{C_1,C_2} 为出现概念 C_1 和 C_2 的文档数占局部文档集总数的比率，$avgs(C_1,C_2)$ 为两者在局部文档集中的平均距离。

根据概念之间的共现性分析，同样为查询中的每一个概念 q_i 选取共现度权值大于 λ_2 的概念构成候选扩展概念 $q_iCE(q_{i1},q_{i2},\cdots,q_{im})$，$m=1,2\cdots$。

3.4　候选扩展概念的筛选

为了更准确地反应用户的查询请求以及确定扩展的边界，对两部分候选扩展概念 $q_iOE(q_{i1},q_{i2},\cdots,q_{ik})$ 和 $q_iCE(q_{i1},q_{i2},\cdots,q_{im})$ 进行筛选。为了使本体的扩展不仅仅依赖于所构建的本体结构，该筛选与局部文档相结合，即利用原查询概念与候选概念在局部文档中的共现性分析作为筛选函数，将满足如下条件的候选概念作为最终的本体扩展概念：

$$cw(q_i,q_{ij})>\lambda_2(q_{ij}\in q_iOE) \tag{8}$$

同样，为了降低局部分析带来的不稳定性以及扩展缺少语义上的理解，对于局部文档根据共现性分析得到的统计候

选扩展概念 $q_iCE(q_{i1},q_{i2},\cdots,q_{im})$，利用本体中概念之间的相似度作为筛选函数对统计候选概念进行二次筛选，将满足下列条件的概念作为最终的统计扩展概念：

$$Sim(q_i,q_{ij})>\lambda_1(q_{ij}\in q_iCE) \tag{9}$$

两部分候选扩展概念经筛选后，将满足条件的概念合并作为最终的查询扩展概念。

3.5 扩展词的权重计算

对于查询和文档来说，不同的词语或者概念具有不同的重要程度。查询中词语或概念权重的分配是计算查询-文档相关度的基础。大多数查询扩展方法选择文献[3]中提到的传统的 Rocchio 方法计算扩展词的权重，然而该方法并没有把扩展词与原查询的相关度因素考虑进去，为此本文给出一种简单的扩展词权重计算方法：

$$weight(q_i)=\frac{Nq_iE}{NE} \tag{10}$$

其中，$weight(q_i)$表示原查询概念 q_i 的权重，Nq_iE 表示概念 q_i 的扩展概念个数，NE 表示扩展后的查询概念总个数，根据原查询概念的权重区分原查询概念中不同概念的重要程度。

q_i 的扩展概念 q_{ij} 的权重不仅考虑了原查询概念 q_i 的权重，也考虑了扩展概念与 q_i 在本体中的相关度和在局部文档中的相关度，其计算方法如下：

$$weight(q_{ij})=\alpha\cdot weight(q_i)+\beta\cdot Sim(q_i,q_{ij})+\gamma\cdot cw(q_i,q_{ij}) \tag{11}$$

4 实验及结果

4.1 实验设计

为了验证本文方法的有效性，设计了一个小型检索系统，其结构如图 2 所示，主要包括本体的构建及推理、查询和文档预处理、查询扩展和检索模块。

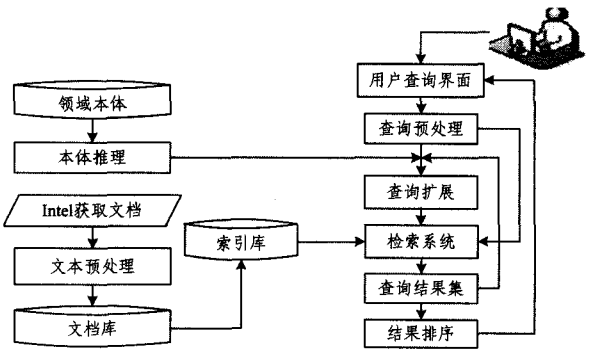

图 2 检索系统体系结构

实验中采用的领域本体是用 OWL 语言描述的计算机领域本体，该本体描述了计算机领域内的有关概念及概念之间的关系。本体的构建采用斯坦福大学医学院生物信息研究中心开发的 Protégé 工具，利用美国 HP 实验室提供的开源工具 Jena 对此领域本体进行相应的推理操作，部分本体结构图如图 3 所示。

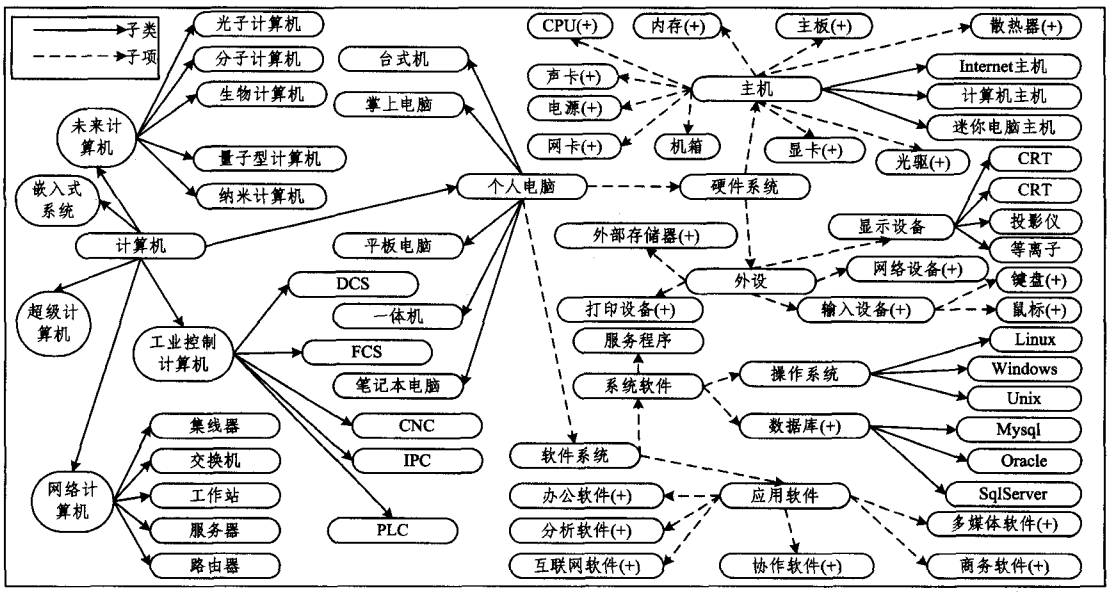

图 3 计算机领域本体部分片段

查询预处理主要是对初始查询进行概念提取，主要过程为：

(1)导入计算机领域本体相关概念和关系。使用 Jena 提供的接口和方法将本体内的所有概念和属性导出并生成文本文件，并将其作为领域本体概念词典 C。

(2)使用 ICTCLAS 自带的方法将概念词典 C 写入用户词典，并更新词库。

(3)利用 ICTCLAS 对用户的查询请求进行分词和去停用词处理。

(4)对分词结果与本体概念词汇集进行比较，提取出查询概念及该概念的同义概念，以供查询扩展使用。

文档预处理主要包括分词、概念提取和建立索引。分词和概念提取类似于查询预处理的分词和概念提取，从而完成文档到本体的映射，并对抽取的概念建立倒排索引，以便检索系统进行检索，其中文档特征向量中概念的权重采用 TF-IDF 方法得到。

由于标准的测试文本集 TREC 包含的领域范围较广，建立一个对应的领域本体比较困难，因此很难应用标准的测试文本集来验证该扩展方法的有效性。实验中将采用与所构建的领域本体库相对应的文档作为测试集，即利用网络爬虫工具从 Web 上获取 2800 篇有关计算机领域的文档作为测试集。

4.2 实验结果与分析

采用标准的查全率 *Recall*、查准率 *Precision* 和 *F*-measure

作为主要的评价指标,以 $Pr@n$ 为辅助评价指标,它表示检索出的前 n 篇文档中的查准率。

$$Recall=\frac{\text{检索出的相关文档数}}{\text{文档库中全部相关文档数}} \tag{12}$$

$$Precision=\frac{\text{检索出的相关文档数}}{\text{检索出的文档总数}} \tag{13}$$

$$F\text{-}measure=\frac{2\times Recall\times Precision}{Recall+Precision} \tag{14}$$

$$Pr@n=\frac{\text{检索结果的前 } n \text{ 篇中相关文档数}}{n} \tag{15}$$

实验中将基于共现性分析的查询扩展方法、基于本体的方法和本文查询扩展方法分别进行了 30 次查询,其中基于本体的查询扩展方法以 λ_1 为阈值,基于共现性分析的扩展以 λ_2 为阈值,两者又同时为本文方法的筛选阈值。查询概念主要包括"系统软件"、"Intel"、"编程语言"、"硬盘"、"SRAM"、"缓存"、"数据库语言"等,3 种方法的平均查询性能比较如表 1 所列。

表 1　查询性能比较

查询扩展方法	Recall	Precision	F-measure	Pr@10	Pr@20
基于共现性分析的扩展	0.615	0.594	0.604	0.651	0.626
基于本体的查询扩展	0.618	0.602	0.610	0.712	0.643
本文查询扩展方法	0.706	0.690	0.698	0.736	0.706

从表 1 中可以看出本文查询扩展方法在查全率和查准率方面均有所提高,与基于共现性分析的查询扩展方法相比,增加了语义扩展信息,而与基于本体的查询扩展方法相比,又加入了统计扩展概念并以共现性分析作为约束条件,所以查询性能有所提高。F-measure 综合反映了检索性能,因此选取其中的 10 次查询进行 F-measure 的比较,结果如图 4 所示。

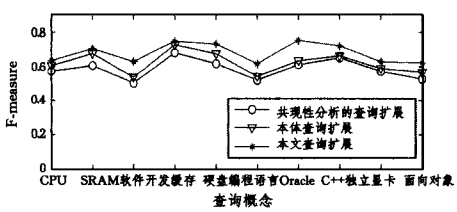

图 4　部分概念的平均 F-measure 比较

在查询扩展中扩展词的个数也是影响查询性能的重要因素,扩展词个数若太多则影响查准率,若太少则不能理解和满足用户的查询,所以对扩展词数量进行了实验对比。利用本文的查询扩展方法,通过参数调整来选取不同数量的扩展词,图 5 反映了扩展规模与查准率的关系。

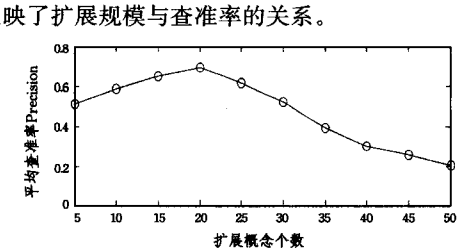

图 5　扩展规模对查准率的影响

从图 5 中可以看出,该查询扩展算法在扩展概念个数为 20 个左右时可以获得较高的查准率,当扩展概念个数超过 40 时查准率明显下降,甚至可能低于未扩展时的查准率。

上述的实验结果分析都是在固定最优实验阈值 λ_1 和 λ_2 的基础上得到的,这两个参数也是决定本文查询扩展方法性能的主要因素。为了对比本文扩展方法与上述两种方法的优缺点,对不同的参数值进行了实验比较,由于 λ_1 和 λ_2 是不同方法选定的不同大小的值,用百分比表示两者同时增加或降低的幅度。对应的查准率和查全率的实验结果如图 6 和图 7 所示。图中(λ_1,λ_2)分别为查询性能最优时对应的值,正负百分比分别表示相对于最优值增加或减少的量。

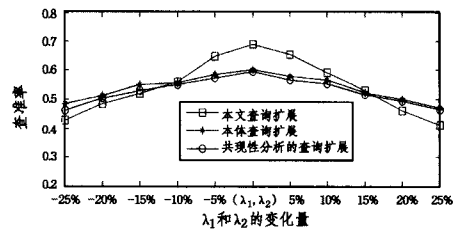

图 6　不同阈值的查准率对比

从图 6 的实验结果可知,当阈值相对较小或较大时,本文查询扩展方法的查准率低于其他两种方法,即该方法对 λ_1 和 λ_2 的大小比较敏感。当阈值 λ_1 和 λ_2 较小时,两者相互筛选的力度不够,扩展词就相当于是本体扩展和局部共现分析扩展的并集,导致扩展词个数较多,查准率下降;当阈值 λ_1 和 λ_2 较大时,过度筛选,导致扩展词个数较少,不能很好地表达用户的查询需求。

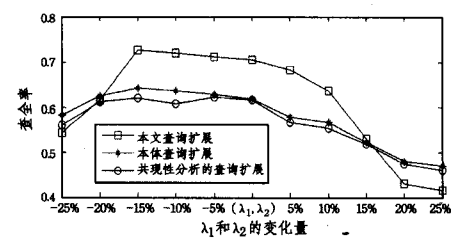

图 7　不同阈值的查全率对比

由于本文扩展方法是本体扩展和局部共现性分析的结合,因此图 7 的实验结果也表明,在很大范围内,本方法的查全率明显高于其他两种方法,但是当参数过大、两部分候选扩展词相互筛选时,大量相关的词语被过滤掉,导致相关的文档不能被检索出来,查全率明显下降。

结束语　本文提出了基于本体和局部共现的查询扩展方法,从语义和统计上均实现了查询扩展,并采用相互约束条件对候选扩展概念进行二次筛选,实验结果验证了该方法的有效性。考虑到概念之间的相关度是查询扩展和检索的基础,如何更加准确地衡量概念之间的相关度是下一步需要研究的内容。

参 考 文 献

[1] LIN Guo-jun, YE Fei-yue, et al. Concept Query Expansion based on Semantic[J]. Computer Engineering and Design, 2009, 30(6):1502-1509. (in Chinese)
林国俊,叶飞跃,等. 基于语义的概念查询扩展[J]. 计算机工程与设计,2009,30(6):1502-1509.

[2] TIAN Xuan, DU Xiao-yong, LI Hai-hua. Computing Term-concept Association in Semantic-Based Query Expansion[J]. Journal of Software, 2008, 19(8):2043-2053. (in Chinese)
田萱,杜小勇,李海华. 语义查询扩展中词语-概念相关度的计算

[J]. 软件学报,2008,19(8):2043-2053.

[3] DING Guo-dong,BAI Shuo,WANG Bin. Local Co-occurrence Based Query Expansion for Information Retrieval[J]. Journal of Chinese Information,2006,20(3):84-91. (in Chinese)
丁国栋,白硕,王斌. 一种基于局部共现的查询扩展方法[J]. 中文信息学报,2006,20(3):84-91.

[4] WU Qin,BAI Yu-zhao,LIANG Yong-zhen. A Local Query Expansion Method Based on Semantic Dictinary[J]. Journal of NanJing University(Nature Sciences),2014,50(4):526-533. (in Chinese)
吴秦,白玉昭,梁永祯. 一种基于语义词典的局部查询扩展方法[J]. 南京大学学报(自然科学),2014,50(4):526-533.

[5] LATIRI C,HADDAD H,HAMROUNI T. Towards an Effective Automatic Query Expansion Process Using an Association Rule Mining Approach[J]. Journal of Intelligent Information Systems,2012,39(1):209-247.

[6] LIU Cai-hong,QI Rui-hua,LIU Qiang. Efficient Query Expansion Based on Positive and Nagative Association[J]. China Science Paper,2013,8(1):51-56. (in Chinese)
刘彩虹,祁瑞华,刘强. 一种正负关联规则的快速查询扩展算法[J]. 中国科技论文,2013,8(1):51-56.

[7] TANNEBAUM W,RAUBER A. Using query logs of USPTO patent examiners for automatic query expansion in patent searching[J]. Information Retrieval,2014,17(5/6):452-470.

[8] DING Xiao-yuan,GU Chun-hua,WANG Ming-yong. Query Expansion of Local Co-occurrence Based on Query Log[J]. Computer Application and Software,2013,30(12):22-27. (in Chinese)
丁晓渊,顾春华,王明永. 基于查询日志的局部查询扩展[J]. 计算机应用与软件,2013,30(12):22-27.

[9] WANG Lu,WANG Guo-chun,GUI Jin-hua,et al. A Semantic Retrieval System Based on Ontology[J]. Journal of Changchuan University of Technology (Natural Science Edition),2013,34(6):726-730. (in Chinese)
王璐,王国春,桂金花,等. 本体语义检索系统[J]. 长春工业大学学报(自然科学版),2013,34(6):726-730.

[10] MENG Hong-wei,ZHANG Zhi-ping,ZHANG Xiao-dan. Research on Intelligent Information Retrieval Model Based on Domain Ontology[J]. Journal of Intelligence,2013,32(9):180-184. (in Chinese)
孟红伟,张志平,张晓丹. 基于领域本体的文献智能检索模型研究[J]. 情报杂志,2013,32(9):180-184.

[11] CHAUHAN R,GOUDAR R,SHARMA R,et al. Domain Ontology based Semantic Search for Efficient Information Retrieval through Automatic Query Expansion[C]// International Conference on Intelligent Systems and Signal Processing. 2013:397-402.

[12] WANG Hong,FAN Hong-jie,LI Jian,et al. Research on the Method of Semantic Query Expansion in Civil Aviation Emergency Domain Ontology[J]. International Journal of Digital Content Technology and its Applications,2014,8(5):128-135.

[13] WANG Xu-yang,XIAO Bo. Query Expansion Method Based on Ontology and Local Context Analysis[J]. Computer Engineering,2012,38(7):57-60. (in Chinese)
王旭阳,萧波. 基于本体和局部上下文的查询扩展研究[J]. 计算机工程,2012,38(7):57-60.

(上接第 202 页)

[4] POLL E. Subtyping and inheritance for categorical data types [C]// Proc. of Theories of Types and Proofs, Kyoto: Japan, RIMS Lecture Notes 1023. 1998:112-125.

[5] NOGUEIRA P,MORENO-NAVARRO J. Bialgebra views:a way for polytypic programming to cohabit with data abstract[C]// Proceedings of the ACM SIGPLAN Workshop on Generic Programming,New York,2008. NY:ACM,2008:61-73.

[6] SU J D,YU S S. Coinductive data types and their applications in programming languages[J]. Computer Science,2011,38(11):114-118. (in Chinese)
苏锦钿,余珊珊. 程序语言中的共归纳数据类型及其应用[J]. 计算机科学,2011,38(11):114-118.

[7] SU J D,YU S S. Comonadiccorecursions on strong coinductive data types[J]. Journal of South China University of Technology (Natural Science Edition),2014,42(1):128-134. (in Chinese)
苏锦钿,余珊珊. 强共归纳数据类型上的 Comonadic 共递归[J]. 华南理工大学学报(自然科学版),2014,42(1):128-134.

[8] DOREL L,VLAD R. Program equivalence by circular reasoning [J]. Formal Aspects of Computing,2015,27(4):701-726.

[9] GHANI N,REVELL T,ATKEY R,et al. Fibrational units of measure [EB/OL]. [2015-03-21]. http://personal. cis. strath. ac. uk/neil. ghani/pub. htm.

[10] KENNEDY A J. Relational parametricity and units of measure [C]// Proceedings of the 24th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages(POPL '97). New York:ACM,1997:442- 455.

[11] DENIS B,ALEXANDER D,ANNIE F. Categorical grammars with iterated types form a strict hierarchy of k-valued languages [J]. Theoretical Computer Science450,2012,450(13):22-30.

[12] USA:Department of Computer Science,University of Illinois at Urbana-Champaign[R]. Formal semantics and analysis of behavioral AADL models in real-time Maude,2010.

[13] BARR M,WELLS C. Category theory for computing science [M]. NewYork:Prentice-Hall,1990.

[14] QU Y W. Formal semantics foundation and formal description (2nd Version)[M]. Beijing:Science Press,2010. (in Chinese)
屈延文. 形式语义学基础与形式说明(第二版)[M]. 北京:科学出版社,2010.

[15] MIAO D C,XI J Q,JIA L Y,et al. Formal language algebraic model [J]. Journal of South China University of Technology (Natural Science Edition),2011,39(10):74-78. (in Chinese)
苗德成,奚建清,贾连印,等. 一种形式语言代数模型[J]. 华南理工大学学报(自然科学版),2011,39(10):74-78.

[16] LI W. Mathematical Logic[M]. Beijing:Science Press,2008. (in Chinese)
李未. 数理逻辑[M]. 北京:科学出版社,2008.