

基于本体和相似图的概念语义相似度计算

张晓宁, 王西锋

(宝鸡文理学院 计算机科学系, 陕西 宝鸡 721016)

摘要:概念语义相似度计算的研究是人工智能最基础和最重要的课题之一, 借鉴现有的概念建模思想和工具, 提出一种综合的计算形式概念分析中概念间语义相似度的算法。通过分析传统的计算方法, 对存在的问题进行改进, 结合领域本体和 FCA 的思想, 通过相似图和候选属性对集合等定义计算 FCA 中概念间的语义相似度。应用实例的结果表明计算结果与人类的主观判断基本一致。文中的方法对概念间语义相似度计算是可行的, 使用该方法可以获取在语义上和用户请求最接近的结果。

关键词:领域本体; 形式概念分析; 相似图; 语义相似度

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2011)08-0101-04

Concept Semantic Similarity Computation Based on Ontology and Similar Graph

ZHANG Xiao-luan, WANG Xi-feng

(Department of Computer Science, Baoji College of Arts & Science, Baoji 721016, China)

Abstract: Research on concept semantic similarity computation is one of the most basic and important subjects in artificial intelligence. Referencing the existed concept modeling ideas and tools, a competitive semantic similarity computation method in formal concept analysis between concepts was proposed. By analyzing the traditional computation methods and modifying the problems in them, integrates the ideas of domain ontology and FCA to compute the semantic similarity by the definitions of the similar graph and candidate attributes set. The results of applied case shows the computation results are identical with human subjective judgment. This method is effective for concept semantic similarity computation and get the closest semantics results from users' requests.

Key words: domain ontology; formal concept analysis; similar graph; semantic similarity

0 引言

概念就是反映事物类的本质属性及其分子的思维形式, 是人们认识世界的基础和一种形式化的规范说明, 可以将概念看作汉语语言文字中的词语^[1]。词语是汉语最基本的语法和语义的单位, 词语的语义相似度计算是计算句子相似度的基础。概念相似度是一个主观性较强的概念, 没有非常明确的客观标准可以衡量。目前, 概念语义相似度计算已广泛应用在数据挖掘、信息提取、自动问答系统、文本分类、检索等诸多领域, 成为当今人工智能领域研究的一个热点。

多年来, 对概念的语义相似性计算的研究一直属

于自然语言处理领域的基础课题, 是中文信息处理应用中迫切需要去解决的关键技术。传统的基于本体的概念语义相似度计算方法主要分为 2 种^[2-4]: 一种是基于信息论的方法, 该方法利用信息论来计算 2 个概念共享信息的程度, 具有较高的理论严谨性, 但是只能粗略地量化概念之间的语义相似度, 不能实现概念语义相似度的细致区分^[5]; 另一种方法是基于语义距离的方法, 该方法以概念之间路径的长短作为衡量语义距离的长短, 通过计算两个概念之间的语义距离来实现概念语义相似度的计算, 该方法简单、直观, 但忽略了影响语义距离的其他很多因素^[6]。

领域本体和形式概念分析虽然两者不同, 但它们建立都是对概念进行建模, 差异主要有三个方面:

(1) 两者建模的对象不同, 前者为现实建模, 后者为人工世界建模;

(2) 两者强调的内容不一样, 前者强调概念的内涵, 后者认为概念的外延和内涵都很重要, 将概念分别从内涵和外延两方面进行描述;

收稿日期: 2011-01-13; 修回日期: 2011-04-14

基金项目: 陕西省自然科学基金研究计划基金资助项目(2005F11); 陕西省教育厅专项科研计划基金资助项目(05JK137); 宝鸡文理学院院级重点科研项目(ZK10167)

作者简介: 张晓宁(1978-), 女, 陕西宝鸡人, 讲师, 硕士, 主要研究方向为人工智能; 王西锋, 讲师, 硕士, 主要研究方向为网络信息处理。

(3)目的不同,前者支持知识密集型应用,后者是支持用户在给定数据的基础上进行领域分析和建模。

FCA 作为一种对人工世界进行建模的工具无可替代的优势,它具有分析能力、描述能力和图形化能力。文中的研究目的在于探索 FCA 中概念语义相似度的计算方法,为进一步的研究奠定基础。

鉴于上述原因,提出了一种概念语义相似度的计算方法,借鉴现有概念的建模思想和工具,应用本体的思想和形式概念分析(Formal Concept Analysis, FCA)的思想计算 FCA 中的概念语义相似度。首先分析了领域本体中的相似关系,然后结合相似图和候选属性对集合等定义提出概念语义相似度计算方法,最后通过一个应用实例验证算法的可行性。

1 预备知识

1.1 本体 (Ontology)

近年来,本体在知识共享、智能信息检索、数字图书馆、语义 Web 和信息集成等方面有着广泛的应用,对实现知识的重用和共享、提高系统间互相通讯、互操作、可靠性的能力有很大的意义。本体的哲学本原可以理解为对一个系统的说明或者解释,是对客观存在的现实本质的抽象。不同的人对本体有着不同的理解,现在广为认可的是由 Borst 提出的一种观点^[7]:“本体是共享概念模型的明确的、形式化、规范说明”。本体由概念、属性、关系、层次结构和实例组成,通过对领域知识的共同理解和确定该领域内共同认可的词汇来捕获相关领域的知识,并形式化地给出这些词汇和词汇间相互关系的明确定义。本体最鲜明的特点是对大量的概念属性以及实例进行层次化的管理。

1.2 形式概念分析(FCA)

FCA 是基于数学中的序和完备格理论提出的一种概念知识形式化的方法,它是一种进行数据分析和规则提取的有效工具,主要思想是从由二元关系构成的形式背景中提取概念层次结构及概念之间的泛化、特化关系构成一个概念格,从形式背景中生成概念格的过程实质上是一种概念聚类过程。概念格是 FCA 的基本数据结构,可通过哈斯图反映出概念层次结构,实现了对数据的可视化应用^[8],图形化表示简练、直观、形象,有助于用户对概念模型和关系的交互与理解。目前,概念格以其独特的优势在数据挖掘、信息检索、知识发现、机器学习和软件工程等方面获得了广泛应用^[9-12],正在赢得许多学者的关注和研究。

Definition 1:形式背景是一个三元组 $T:=(D,S,G)$, D 表示对象的集合, S 表示属性的集合, G 表示对象和属性之间的关系,如果 dGs 或 $(o,s) \in G$,表示对象 d 具有属性 s 。

Definition 2: D 和 M 表示对象的集合,其中 M 是 D 的子集 $M \subseteq D$,集合 M 中的所有对象具有的共同属性可以表示为 $M':=\{s \in S \mid \forall d \in M, dGs\}$; S 表示属性的集合,集合 N 是 S 的子集 $N \subseteq S$,具有 N 中所有属性的对象的集合可以表示为 $N':=\{d \in D \mid \forall s \in N, dGs\}$ 。

Definition 3:形式概念简称概念,是一个序偶 (E, I) ,其中 $E \subseteq D, I \subseteq S$,如果满足 $E' = I$ 且 $I' = E$,则 E 称为形式概念 (E, I) 的外延, I 称为形式概念的内涵。形式背景 (D, S, G) 所有形式概念的集合表示为 $\wp(D, S, G)$ 。

Definition 4 :对于 (M_1, N_1) 和 (M_2, N_2) 两个概念,若 $E_1 \subseteq E_2$,则 (M_1, N_1) 被称为 (M_2, N_2) 的子概念, (M_2, N_2) 被称为 (M_1, N_1) 的超概念,记作 $(M_1, N_1) \leq (M_2, N_2)$,偏序关系 \leq 称为形式概念之间的序。按此方式建立的 (D, S, G) 所有形式概念的集合表示为 $\wp(D, S, G)$,称做形式背景 (D, S, G) 的概念格。

2 概念语义相似度的计算

2.1 相似图的定义

Definition 5: 本体中的概念间存在多种语义关系,例如 ISA(是), part of(部分), whole(整体), relatedness(相关), similarity(相似)等,其中,相似关系定义为 $\text{similarity}(a, b, \text{val}(a, b))$,表示概念 a, b 具有相似关系,相似关系的度量由 $\text{val}(a, b)$ 给出,表示概念 a, b 之间的自相似度,一般为区间 $[0, 1]$ 之间的一个小数。

一般情况下,概念 c_i, c_j 之间的自相似度的大小由该领域内的专家给出。

Definition 6: 本体 T 和形式背景 (T, S, G) 的相似图定义为 $\Phi_{(T,S)}, \Phi_{(T,S)}$ 是所有 $\text{similarity}(a, b, \text{val}(a, b))$ 的集合且 $\Phi_{(T,S)}$ 满足以下的四个条件:

- (1) 对于 $c_i, c_j \in \xi$,若在本体 O 中具有关系 $\text{similarity}(a_i, b, \text{val}(a, b))$,那么 $\langle a, b, \text{val}(a, b) \rangle \in \Phi_{(T,S)}$;
- (2) 对于任意概念 $c \in \xi$,如果 $\text{val}(c, c) = 1.0$,那么 $\langle c, c, \text{val}(c, c) \rangle \in \Phi_{(T,S)}$;
- (3) 对于任意一对概念 $a, b \in \xi$,如果在本体 T 中没有定义自相似度,当 $\text{val}(a, b) = 0.0$ 时,那么 $\langle a, b, \text{val}(a, b) \rangle \in \Phi_{(T,S)}$;
- (4) 对于任意一对概念 $a, b \in \xi$,因为自相似度满足对称性, $\text{val}(a, b) = \text{val}(b, a)$,且 $\langle a, b, \text{val}(a, b) \rangle = \langle b, a, \text{val}(b, a) \rangle$ 。

2.2 算法描述

文中基于图论中的二分图最大权匹配的^[13]理论

提出了形式概念分析中的概念语义相似度计算方法,首先需要找到候选属性对集合 $\xi(I_1, I_2)$, 然后在此基础上计算相似度。

Definition 7: 如果 $(P_1, Q_1), (P_2, Q_2)$ 是形式背景的概念, 集合 Q_1, Q_2 的基数分别为 n 和 m , 即 $n = |Q_1|, m = |Q_2|$, 且 $n \leq m$ 。所有的 n 组属性对的集合定义为候选属性对集合, 定义为:

$$\xi(Q_1, Q_2) = \{ \{ \langle a_1, b_1 \rangle \dots \langle a_n, b_n \rangle \} \mid a_h \in q_1, b_h \in q_2, \forall h = 1, \dots, n, \text{ 并且 } a_k \neq a_l, b_k \neq b_l, \forall k, l \neq h \}$$

计算的步骤为: $(P_1, Q_1), (P_2, Q_2)$ 是本体 T 和两个形式背景 (D_1, S_1, G_1) 和 (D_2, S_2, G_2) 的概念, 令 $\Gamma_{(T, S, U, S)}$ 为本体 T 和给定的形式背景的相似图, 分别计算下列的值:

- (1) 计算集合 Q_1, Q_2 的笛卡尔积 $Q_1 \times Q_2$;
- (2) 在 Q_1, Q_2 的笛卡尔积中找到一个集合中的元素之间没有重复的候选的子集, 并将该集合标记为 $\xi(Q_1, Q_2)$;
- (3) 对 $\xi(Q_1, Q_2)$ 中的每一个候选集合, 根据 2.1 节中相似图的定义计算属性对的自相似度的和;
- (4) 属性对的自相似度是步骤 3 中的最大的值;
- (5) 将属性集合的交集的基数与上述步骤 4 的结果进行加权平均。

因此在以上定义的基础上, 具有相同(或不同的)形式背景的两个概念 $(P_1, Q_1), (P_2, Q_2)$ 的相似度定义为:

$$\text{Similarity}((P_1, Q_1), (P_2, Q_2)) = \frac{|(P_1 \cap P_2)|}{r} * w + \left[\frac{1}{m} \max_{p \in \xi(Q_1, Q_2)} \left(\sum_{\langle a, b \rangle \in p} \text{val}(a, b) \right) \right] * (1 - w) \quad (1)$$

其中 ξ 和 m 在前面已经定义过, $\text{val}(a, b)$ 是相似图 $\Gamma_{(T, U, S)}$ 中的 a, b 的自相似度。 r 是集合 P_1, P_2 的基数中的较大者, w 是闭区间 $[0, 1]$ 之间的一个权值, 一般由用户给出, 对于任何一对概念, 都满足以下关系: $\text{Similarity}((P_1, Q_1), (P_2, Q_2)) = \text{Similarity}((P_2, Q_2), (P_1, Q_1))$

3 应用实例分析

上文中的定义给出了用户在概念语义相似度计算的理论依据, 本小节给出一个简单的实例来说明上述算法的实际应用。我们设计了一个有关说明欧洲城市特点的形式背景验证文中提出的概念相似度算法的可行性。在这个欧洲城市的形式背景中, 共给出了 7 个欧洲城市: Athens (雅典)、Courmayeur (库马约尔)、Innsbruck (因斯布鲁克)、London (伦敦)、Paris (巴

黎)、Reykjavik (雷克雅末)、Rome (罗马) 具有的特点, Archeological_Site (考古遗址)、Beach (海滩)、Capital (首都)、Europe (欧洲)、River (江河)、Skiing_Area (滑雪区), 城市的特点在形式背景中表示为属性。形式背景的构造结果如表 1 所示。

表 1 欧洲城市的形式背景

City Name	Arc	Bea	Cap	Eur	Riv	Ski
Athens (A)	×	×	×	×		
Courmayeur (C)				×		×
Innsbruck (I)				×	×	×
London (L)			×		×	
Paris (P)			×	×	×	
Reykjavik (Re)			×			×
Rome (Ro)	×		×	×	×	

为了方便, 用城市英文单词的第一个大写英文字母表示表示欧洲的城市, 即对象名; 用特点英文单词的前三个英文字母表示对象具有该属性。

$T = \{ \text{Athens, Courmayeur, Innsbruck, London, Paris, Reykjavik, Rome} \}$

$S = \{ \text{Archeological Site, Beach, Capital, Europe, River, Skiing_Area} \}$

例如, 概念 $((\text{Athens, Paris, Rome}), (\text{Capital, Euro}))$ 记为 $((A, P, Ro), (\text{Cap, Eur}))$, 表示的含义是“雅典、伦敦和罗马都是欧洲国家的首都”。概念 $((C, I, Re), (\text{Ski}))$ 表示的含义是“在库马约尔、因斯布鲁克和雷克雅末这些城市, 可以滑雪。”从形式背景构造的概念格如图 1 所示。

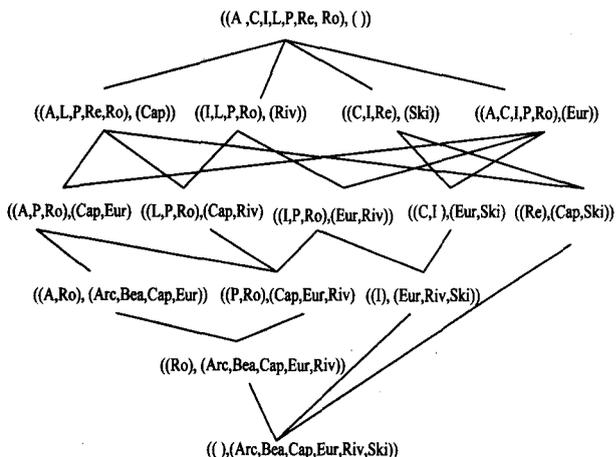


图 1 表 1 对应的概念格

对于上面给出的欧洲城市的形式背景, 本体 T 中满足以下的 similarity 关系:

$\text{similarity}(\text{City, Capital}, 0.8)$, $\text{similarity}(\text{River, Stream}, 0.9)$, $\text{similarity}(\text{Beach, Seaside}, 0.9)$ 。

本体 T 和此形式背景的相似图 $\Gamma_{(T, S)}$ 如下, 其中自相似度值为 1.0 和 0.0 的三元组采用了简单写法。

<City, Capital, 0.8>, <River, Stream, 0.9>, <Beach, Seaside, 0.9>, <City, City, 1.0>, <Archeological_Site, Archeological_Site, 1.0>, ……,<City, Archeological_Site, 0.0> <City, Beach, 0.0>, …….

我们设计了两种计算方法做一个对比效果。第 1 种方法采用文中第 2 部分提出的概念相似度计算方法 ($w = 0.5$), 计算两组概念的相似度; 其中前三对概念都来自欧洲城市的形式背景, 后四对概念是来自欧洲的城市和世界的城市的形式背景。

例如, 计算图 1 中在同一层直接相邻两个兄弟概念((L, P, Ro), (Cap, Riv)) 和 ((A, P, Ro), (Cap, Eur)) 的相似度。根据前面的相似图定义, 因为 $r = 3, m = 2, w = 0.5$, 由公式(1)得:

$$\text{Sim} [((A, P, Ro), (Cap, Eur)), ((L, P, Ro), (Cap, Riv))] = \frac{2}{3} * \frac{1}{2} + \frac{1}{2} * (1.0 + 0.0) * (1 - \frac{1}{2}) = 0.58$$

计算在图 1 中不直接相邻的概念 ((P, Ro), (Cap, Eur, Riv)) 和 ((A, L, P, Re, Ro), (Cap)) 的相似度, 由公式 1 ($w = 0.5$) 得到下式:

$$\text{Sim} [(((P, Ro), (Cap, Eur, Riv)), ((A, L, P, Re, Ro), (Cap)))] = \frac{2}{5} * \frac{1}{2} + \frac{1}{3} * 1.0 * (1 - \frac{1}{2}) = 0.36$$

其他概念的计算方法和上面的计算过程相似, 在此就不再赘述了。具体结果见表 2。

第 2 种方法是分别让 30 个人对这两组概念相似度进行主观判断, 并求出结果的平均值。两种方法的结果如表 2 所示。

表 2 计算结果对比

No	Concept1	Concept2	方法 1	方法 2
1	((L, P, Ro), (Cap, Riv))	((A, P, Ro), (Cap, Eur))	0.58	0.51
2	((L, P, Ro), (Cap, Riv))	((P, Ro), (Cap, Eur, Riv))	0.66	0.56
3	((P, Ro), (Cap, Eur, Riv))	((A, L, P, Re, Ro), (Cap))	0.36	0.43
4	((Ro, Rio), (Sea, Cit, Str, Air))	((I, P, Ro), (Eur, Riv))	0.28	0.36
5	((Ro, Rio), (Sea, Cit, Str, Air))	((A, Ro), (Arc, Bea, Cap, Eur))	0.46	0.51
6	((Ro, Rio), (Sea, Cit, Str, Air))	((A, C, I, L, P, Re, Ro), ())	0.07	0.04
7	((Ro, Rio), (Sea, Cit, Str, Air))	((), (Arc, Bea, Cap, Eur, Riv, Riv))	0.22	0.34

考查两种方法的结果, 可见按照文中所提出的办法是有效、可行的, 因为方法 1 得到的结果表明两个概念间的语义相似度与方法 2 的人类的判断较为吻合。

4 结束语

概念的语义相似度计算的研究是许多中文信息处理及自然语言理解领域应用最终得以实现的理论基

础, 尽管已有很多学者花费了大量时间与精力进行了大量的研究工作, 但由于概念表达的主观性和复杂性, 具体应用领域的专业性等因素影响, 目前概念相似度计算仍然是人工智能长期需要深入研究的内容。它在信息整合、信息检索、基于实例的机器翻译、问答系统等许多领域有广泛的应用。

形式概念是现实世界中各种概念的抽象, 通过概念的内涵与外延之间的关系形式化地刻画抽象概念。文中着眼于对概念的建模, 在 FCA 中运用领域本体的思想并基于现有研究现状基础上, 提出了一种计算 FCA 中的概念语义相似度的方法。应用实例的结果表明, 文中提出的相似度算法得到的结果与人类的主观上的判断基本是一致的, 使用该算法用户可以获取在语义层面上与请求最接近的答案或结果。但文中的算法也不是十分完美, 在实际应用中也有它局限性, 下一步的研究方向是一方面是对文中提出的算法进行修改和优化, 因为概念往往是领域相关的, 可以通过提供足够的上下文语义信息, 使自相似度的计算结果尽量可以信赖, 尽量不依赖于用户的主观判断; 另一方面, 因为现实世界的的数据量很大, 为了减少数据处理的规模, 可以考虑使用概念聚类技术减小概念格规模, 缩短建概念格的时间, 提高计算效率。最后, 将本算法应用到具体的系统中, 进一步提高系统的效率。

参考文献:

[1] 林智超, 朱国进. 一种基于 FCA 的概念相似度算法[J]. 计算机技术与发展, 2008, 18(9): 112-114.

[2] Steichen O, Daniel-Le Bozec C. Computation of Semantic Similarity Within an Ontology of Breast Pathology to Assist Inter-observer Consensus[J]. Computers in Biology and Medicine, 2006, 36: 76-88.

[3] Gan K W, Wong P W. Annotation Information Structures in Chinese Texts Using How Net[C]//Second Chinese Language Processing Workshop. Hong Kong: [s. n.], 2000: 85-92.

[4] Budan Itsky A, Hirst G. Evaluating Word Net-based Measures of Lexical Semantic Relatedness [J]. Computational Linguistics, 2004, 1(1): 1-49.

[5] Dela Escalera A, Moreno L E, Sallchsm A. Road Traffic Sign Detection and Classification[J]. IEEE Transactions on Industrial Electronics, 1997, 44(6): 848-859.

[6] Cross V. Fuzzy Semantic Distance Measures Between Ontological Concepts[C]//Processing NAFIPS'04: IEEE Annual Meeting of the Fuzzy Information. Washington DC: IEEE Press, 2004: 635-640.

服务的 SOAP 请求消息封装起来并传送给 Web 服务提供端, Web 服务提供端从管道消息中提取出 SOAP 请求消息, 利用 RPC 客户端模块与 Axis 服务器交互, 获得 SOAP 响应消息, 该响应消息又被封装成管道消息返回给服务请求端, 后者从中提取出 SOAP 响应消息, 实现 Web 服务的调用。

3.3 JXTA Bridge 的不足与改进

通过分析 JXTA Bridge 的工作流程, 可以发现 JXTA Bridge 主要存在两方面缺陷: (1) Web 服务的发布机制是基于 JXTA 对等组服务形式, 由于 JXTA 对等组服务是与其关联的对等组紧密相关的, 在对等组通告发布之后, 就不能再在对等组通告中进行添加、修改或删除对等组服务的操作, 因此, Web 服务一旦发布之后就不能进行动态更新; (2) Web 服务的调用是基于底层 SOAP RPC 机制, 针对的是特定的 SOAP 服务器和客户端(如 Axis), 没有提供一种通用的机制来实现 Web 服务的调用, 这将导致 SOAP 编码的移植性较差, 因此 Web 服务的可用性较低。

通过分析 Web 服务和 JXTA P2P 的一些关键技术, 将采取如下策略解决其不足: 针对缺陷(1), 利用基于 JXTA 对等体服务机制来改进 JXTA Bridge 的 Web 服务发布, 由于对等体服务具有可动态发布特性, 因此改进后的 Web 服务发布可以动态地进行, 这样可以及时反映 Web 服务所做的添加、修改或删除等操作; 针对缺陷(2), 通过集成一种更为通用的且独立于协议的 Web 服务调用框架来实现非组内成员的 Web 服务调用, 提高 Web 服务的可用性。今后进一步的研究将根据上述策略进行。

4 结束语

文中分析了目前 Web 服务和 P2P 技术存在的问题, 从中间件平台、服务发布和发现、服务组合以及服务社区管理四个方面讨论了 P2P 和 Web 服务技术融合的研究现状, 着重分析了 P2P 与 Web 服务融合的关键技术——JXTA Bridge 的体系结构、模块功能、存在的主要问题及其解决思路。P2P 与 Web 服务的技术融

合可以促进 Web 应用向更大规模的分布式发展, 对提高 Web 应用系统的健壮性、可靠性、安全性以及高性能具有重要研究价值。

参考文献:

- [1] 岳 昆, 王晓玲, 周傲英. Web 服务核心支撑技术: 研究综述[J]. 软件学报, 2004, 15(3): 428-442.
- [2] 朱红康, 余雪丽. 基于 P2P 的分布式 Web 服务挖掘技术[J]. 计算机工程, 2010, 36(9): 23-25.
- [3] 刘志忠, 王怀民, 周 斌. 一种双层 P2P 结构的语义服务发现模型[J]. 软件学报, 2007, 18(8): 1922-1932.
- [4] 龚世忠, 唐文忠. 一种基于 P2P 的两阶段 Web 服务发现研究[J]. 计算机技术与发展, 2010, 20(6): 121-125.
- [5] 徐志伟, 黄映辉. 基于 P2P 和服务质量的语义 Web 服务发现[J]. 计算机技术与发展, 2009, 19(12): 25-28.
- [6] 王 磊, 胡孝波. 基于 P2P 的 Web 服务聚类及组合[J]. 计算机工程, 2009, 35(17): 7-10.
- [7] Skoutas D, Sacharidis D, Kantere V, et al. Efficient Semantic Web Service Discovery in Centralized and P2P Environments[J]. Lecture Notes in Computer Science, 2008, 5318: 583-598.
- [8] Li Ruixuan, Zhang Zhi, Wang Zhigang. WebPeer: A P2P-based System for Publishing and Discovering Web Services[C]//The 2005 IEEE International Conference on Services Computing (SCC 2005). Orlando, Florida, USA: IEEE Computer Society, 2005: 149-158.
- [9] 战学刚, 郭 建, 迟呈英, 等. 应用移动 Agent 加强 JXTA 网络的搜索功能[J]. 计算机工程与设计, 2009, 30(24): 5622-5625.
- [10] Abiteboul S, Benjelloun O, Manolescu I, et al. Active XML: Peer-to-Peer data and Web services integration[C]//In: Proc. of the 28th Int'l Conf. on Very Large Data Bases. Hong Kong: Morgan Kaufmann Publisher, 2002: 1087-1090.
- [11] 胡江晖, 袁道华, 段 恺. 一种基于 JXTA 的 SOAP 消息通信实现模型[J]. 计算机技术与发展, 2009, 19(8): 45-48.
- [12] Benatallah B, Sheng Q Z, Dumas M. The Self-Serv Environment for Web Services Composition[J]. IEEE Internet Computing, 2003, 5(3): 40-48.

(上接第 104 页)

- [7] Borst W N. Construction of Engineering Ontologies for Knowledge Sharing and Reuse[D]. Enschede: University of Twente, 1997.
- [8] Ganter B, Wille R. Formal Concept Analysis: mathematical foundations[M]. Heidelberg: Springer, 1999: 17-57.
- [9] Dtintsch I, Gediga G. Algebraic aspects of attribute dependencies in information systems[J]. Fundamental in Formaticae, 1997, 29: 119-133.
- [10] Pagliani P. From concept lattices to approximation spaces: Algebraic structures of some spaces of partial objects[J]. Fundamental in Formaticae, 1993, 18(1): 1-25.
- [11] Yao Y Y. Concept lattices in rough set theory[C]//Proceedings of 23rd International Meeting of the North American Fuzzy Information Processing Society. [s. l.]: [s. n.], 2004.
- [12] Pressman R S. 软件工程: 实践者的研究方法[M]. 梅宏, 译. 北京: 机械工业出版社, 2004.
- [13] Galil Z. Efficient algorithms for finding maximum matching in graphs[J]. ACM Computing Surveys, 1986(18): 23-38.