

事件本体模型及事件类排序

仲兆满^{1,†} 刘宗田² 李存华¹

1. 淮海工学院计算机工程学院, 连云港 222005; 2. 上海大学计算机科学与工程学院, 上海 200072;
† E-mail: zhongzhaoman@163.com

摘要 给出了事件、事件类的定义, 分析了事件类之间的关系, 在此基础上形成了面向事件的本体模型, 为基于语义的知识处理提供了新的方法与技术。在事件本体模型的基础上以事件类排序为实例, 研究了基于 HARank (Hubs-Authorities Rank) 算法的事件类排序方法, 给出了实验结果, 并对其进行了评价。与传统本体相比, 事件本体以更高粒度的事件作为知识表示单元, 更适合表示运动世界的知识。

关键词 事件; 事件类; 事件类关系; 事件本体; 事件类排序

中图分类号 TP311

Event Ontology Model and Event Class Ranking

ZHONG Zhaoman^{1,†}, LIU Zongtian², LI Cunhua¹

1. School of Computer Engineering, Huaihai Institute of Technology, Lianyungang 222005; 2. School of Computer Science and Engineering, Shanghai University, Shanghai 200072; † E-mail: zhongzhaoman@163.com

Abstract The authors define an event and an event class, and illustrate the relations between event classes. Then event-oriented ontology model is constructed, which will contribute new methods and technologies to semantic-based knowledge processing. Furthermore, taking event class ranking as an example based on event ontology, the method of HARank-based event class ranking is introduced for event ontology, and corresponding experimental and evaluating results are presented. In comparison with conventional methods, event ontology represents knowledge with a higher granularity, and it will be more suitable for the representation of the knowledge about the movement and state changes in the real world.

Key words event; event class; event class relation; event ontology; event class ranking

许多哲学家认为, 物质世界是由事物和事件构成的, “事件”是自然界一去不返的具体事实, “事物”则是事件永恒不变的特质, 事件之间存在着本质的内在的联系^[1]。一些认知科学家认为, 人类的命题记忆是以“事件”为存储单位的, 存储的是组成事件的概念及其之间的关系以及事件及其之间的关系^[2]。近年来, “事件”的概念逐渐被计算语言学、信息检索、信息抽取和自动文摘等知识处理领域所采用^[3-7], 主要围绕事件的表示、事件的推理和事件的提取展开了研究。

“本体”最初是一个哲学上的术语, 十多年前被

引入计算机领域。本体对于探索人的认知原理、发展自然语言理解技术和人机交互技术有重要的意义。传统本体模型仅仅反映了客观世界中事物存在规律, 对于事物的分类与非分类关系, 它存在以下明显缺陷。1) 人类主要是以“事件”为单位进行记忆和理解现实世界的, 事件关系到多方面的概念, 是比概念粒度更大的知识单元。传统本体所使用的概念模型难以反映事件这一更高层次和更复杂的语义信息, 模型缺少了更高层次结构。2) 传统本体定义为概念及其之间的关系, 事件只是作为概念的一类或关系的一类。概念之间的非分类关系往往是由动

国家自然科学基金(60975033)资助

收稿日期: 2012-03-01; 修回日期: 2012-07-03; 网络出版日期: 2013-01-08

网络出版地址: <http://www.cnki.net/kcms/detail/11.2442.N.20130108.1539.002.html>

作性的词来描绘的。由于动作的多样化导致概念之间关系的复杂性, 而且关系的重要性难以衡量。但事件之间的非分类关系相对固定, 种类明确, 而且重要程度有明显的区别, 用描述概念关系的方法描述事件之间的关系不仅没能揭示事件之间具有的特殊关系, 而且使事件之间的关系复杂化。

Sánchez等^[8]使用事件三元组模型, 通过事件辅助传统本体的构造和学习。Lee等^[9]提出一种用于自动文摘基于本体的事件提取方法, 并提出一个4层(领域层、分类层、事件层和扩展概念层)面向对象的领域本体模型。Lin等^[10]提出一种事件本体检索技术, 该本体的顶层概念为事件的要素, 将事件的构成要素作为该本体中的主要分类, 进而构建出本体。实际上, Lin等的事件本体仍然是概念与概念关系的集合, 是现有的本体模型, 没有真正将事件的理论整合进去, 但具备了一定程度上按事件要素对主题词进行扩展的能力, 故在一定程度上提高了检索结果的质量。Hsu等^[11]建立了事件结构框架信息映射, 将其作为一种知识表示模式。对于查询概念A, 在事件结构框架信息映射中不仅会给出它的下位概念、同义概念和相关概念, 还将会给出A的相关行为。Vargas-Vera等^[12]将事件看做本体中的类, 在KMI本体中, 定义了41种不同的事件类, 指定了每一事件类所具有的槽, 还给出了事件类间的层次关系, 以此包含事件层次的本体辅助识别新闻报道中的事件。Mendez-Torreblanca等^[13]提出了一种事件本体构建的初步设想: 先从文本中获取事件, 再将事件间存在的层次关系找出来, 这样的具有层次关系的事件集称为本体。

1 事件本体模型

1.1 事件及事件类

定义1 事件(Event)^[14]: 指在某个特定的时间和环境下发生的, 由若干角色参与, 表现出若干动作特征的一件事情。形式上, 事件可表示为 e , 定义为一个六元组: $e = \langle A, O, T, V, P, L \rangle$ 。其中, 事件六元组中的元素称为事件要素, 分别表示动作、对象、时间、环境、断言和语言表现。

A (动作): 事件的变化过程及其特征, 是对程度、方式、方法、工具等的描述。

O (对象): 指事件的参与对象, 包括参与事件的所有角色。对象可分别是动作的施动者(主体)和受动者(客体)。

T (时间): 事件发生的时间段, 从事件发生的起点到事件结束的终点, 分为绝对时间段和相对时间段两类。

V (环境): 事件发生的场所及其特征等。

P (断言): 断言由事件发生的前置条件、中间断言以及后置条件构成。前置条件指为进行该事件, 各要素应当或可能满足的约束条件; 中间断言指事件发生过程的中间状态, 各要素满足的条件; 事件发生后, 事件各要素的变化和变迁后的结果, 成为事件的后置条件。

L (语言表现): 事件的语言表现规律, 包括核心词集合、核心词表现等。核心词为事件在句子中常用的标志性词汇, 核心词表现则为在句子中各要素表示与核心词之间的位置关系。

定义2 事件类 EC : 具有共同特征的事件的集合, $EC = \{E, C_A, C_O, C_T, C_V, C_P, C_L\}$ 。其中: E 是事件的集合, 称为事件类的外延; $C_i = \{C_{i1}, C_{i2}, \dots, C_{im}\}$ ($i \in \{A, O, T, V, P, L\}, m \geq 1$) 称为事件类的内涵, 是 E 中的每个事件在第 i 个要素上具有的共同特性的集合, C_{im} 是事件类中每个事件在第 i 个要素上具有的一个共同特性。

定义3 事件类影响因子 α : 事件类 EC_i 的事件的发生以一定的概率 α 触发事件类 EC_j 的事件的发生, 这个概率称为事件类影响因子, 记为 $EC_i \xrightarrow{\alpha} EC_j$ 。事件类影响因子用 $[0, 1]$ 间实数表示, 数的大小表示影响的强弱。

1.2 事件类间的关系

定义4 事件类分类关系。

$EC_1 = \{E, C_{1A}, C_{1O}, C_{1T}, C_{1V}, C_{1P}, C_{1L}\}$ 和 $EC_2 = \{E, C_{2A}, C_{2O}, C_{2T}, C_{2V}, C_{2P}, C_{2L}\}$ 是两个事件类, EC_1 和 EC_2 间存在着分类关系, 当且仅当 ($E_2 \subset E_1$ 或 $C_{1A} \subset C_{2A}$) 或 ($E_1 = E_2$ 且 $C_{1j} \subset C_{2j}$ ($j \in \{O, T, V, P, L\}$)), 其中, EC_2 称为 EC_1 的下位事件类, EC_1 称为 EC_2 的上位事件类, 表示为 $R_H(EC_1, EC_2)$ 。例如, “突发事件”是“地震”和“火灾”的上位事件类, “地震”和“火灾”是“突发事件”的下位事件类, 可以表示为 R_H (突发事件, 地震)和 R_H (突发事件, 火灾)。

定义5 事件类非分类关系。

1) 组成关系: 当事件类 EC_1 中的每一个事件都由事件类 EC_2 中的某个事件组成时, 则称两事件类具有组成关系, 称 EC_1 为大事件, EC_2 为小事件, 表示为 $R_l(EC_1, EC_2)$ 。例如, “盖房子”由“打地基”和“筑

墙”事件类组成,“打地基”和“筑墙”都是“盖房子”事件类的组成部分,可以表示为 R_1 (盖房子,打地基)和 R_1 (盖房子,筑墙)。

2) 因果关系: 事件类 EC_1 的事件的发生以一定的概率导致事件类 EC_2 的事件发生,发生的概率大于给定的阈值,则称两事件类之间具有因果关系,称 EC_1 是 EC_2 的因, EC_2 是 EC_1 的果,表示为 $R_{CE}(EC_1, EC_2)$ 。例如,事件类“地震”是事件类“死亡”的因,事件类“死亡”是事件类“地震”的果,可以表示为 R_{CE} (地震,死亡)。

3) 跟随关系: 在一定长度的时间段内,事件类 EC_1 的事件发生后, EC_2 中的事件可能跟随发生,若跟随发生的概率大于给定的阈值,则称两事件类之间具有跟随关系,称 EC_1 是 EC_2 的前件, EC_2 是 EC_1 的后件,表示为 $R_F(EC_1, EC_2)$ 。例如,事件类“起床”是事件类“刷牙”、“洗脸”的前件,事件类“刷牙”、“洗脸”是事件类“起床”的后件,可以表示为 R_F (起床,刷牙)和 R_F (起床,洗脸)。

4) 并发关系: 在一定长度的时间段内,事件类 EC_1 的事件和 EC_2 的事件同时或先后发生,同时或先后发生的概率大于给定的阈值,则称两事件类之间具有并发关系,表示为 $R_C(EC_1, EC_2)$ 。例如,事件类“下雨”和事件类“刮风”是并发关系,可以表示

为 R_C (下雨,刮风)。

5) 条件关系: 事件类 EC_2 的事件的发生,需要满足事件类 EC_1 的事件的发生,则称两事件类之间具有条件关系,称 EC_1 为 EC_2 的条件, EC_2 为 EC_1 的结论,表示为 $R_{Con}(EC_1, EC_2)$ 。例如,事件类“通过考试”、“发表论文”和“毕业”都是条件关系,可以表示为 R_{Con} (通过考试,毕业)和 R_{Con} (发表论文,毕业)。

6) 排斥关系: 事件类 EC_1 的事件和 EC_2 的事件不可能同时发生,即一个事件类的事件的发生隐含了另一个事件类的事件的不发生,表示为 $R_{Dis}(EC_1, EC_2)$ 。例如,事件类“生”和“死”是排斥关系,事件类“胜利”和“失败”也是排斥关系,可以分别表示为 R_{Dis} (生,死)和 R_{Dis} (胜利,失败)。

1.3 事件本体模型

定义 6 事件本体: 事件本体是共享的客观存在的事件类系统模型的明确的形式化规范说明,表示为 EO 。事件本体的逻辑结构可定义为一个五元组 $EO = \langle ECS, EIS, R, W, Rules \rangle$, 其中: 1) $ECS = \{EC_1, EC_2, \dots, EC_n\}$, 是事件类集合; 2) $EIS = \{EI_1, EI_2, \dots, EI_n\}$, 是事件实例集合; 3) $R = \{r | r \text{ 是 } \langle EC_i, EC_j \rangle \text{ 上的关系}, r \in \{R_H, R_1, R_{CE}, R_F, R_C, R_{Con}, R_{Dis}\}\}$; 4) $W = \{w_{ij} | w_{ij} \text{ 是有向边 } \langle EC_i, EC_j \rangle \text{ 上的影响因子,即事件类 } EC_i \text{ 对事件类 } EC_j \text{ 的影响因子的大小}\}$; 5) $Rules$ (规则)

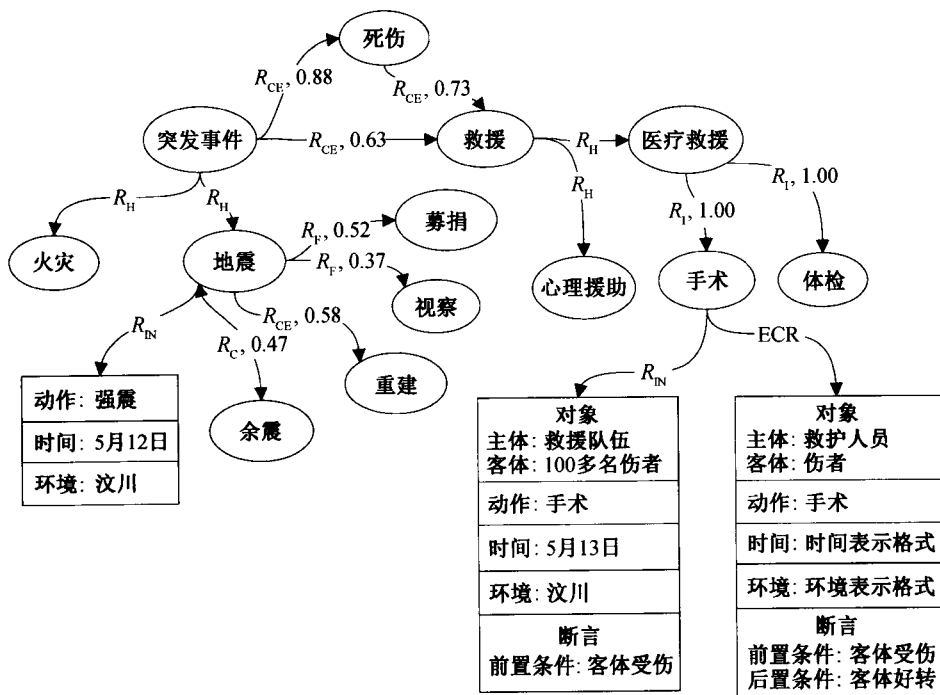


图 1 事件本体实例(片断)

Fig. 1 An example of EO (fragment)

由逻辑语言表示,可用于事件断言所不能覆盖的部分事件与事件间的转换与推理。

图1所示的是“地震”领域事件本体的片段。事件类用椭圆形标记表示,事件类之间若具有关系则有一条有向边相连,边上标记出关系的类型以及影响因子。事件类的实例为该事件类的具体事件,也用边的形式与事件类相连,用“实例(符号为 R_{IN})”表示。事件类的六元组表示形式与事件类之间也用边相连,用“事件类表示(符号为 ECR)”为边的标记。图1中共包含13个事件类(“突发事件”、“地震”、“火灾”、“死伤”、“救援”、“募捐”、“视察”、“重建”、“余震”、“医疗救援”、“心理援助”、“体检”和“手术”),16条边,边上标注有分类、因果、跟随、并发、组成等事件类关系,1个事件类“手术”,2个事件实例“地震”和“手术”。

事件类影响因子刻画的是一个事件类的事件实例的发生所蕴涵的其他事件类的实例发生的概率大小。对于图1所示的事件本体中的事件类影响因子,因果关系(R_{CE})、跟随关系(R_F)以及并发关系(R_C)是基于文本语料借助统计的方法获取的,具体

的统计计算方法参见文献[15]。组成关系(R_I)的影响因子统一标示为1.00,因为一个事件类实例的发生必定蕴涵了其组成事件类实例的发生,比如,图1中的“医疗救援”事件的发生必定包含“体检”和“手术”事件,故 $w(\text{医疗救援}, \text{体检})=1.00$, $w(\text{医疗救援}, \text{手术})=1.00$ 。对于父子关系(R_H)、实例关系(R_{IN})以及事件类的表示(ECR)不需要标记事件类影响因子。

定义7 事件本体邻接矩阵 $W=(w_{ij})_{n \times n}$, 称为事件本体的邻接矩阵, 其中,

$$n=|ECS|, 1 \leq i, j \leq n, w_{ij} = \begin{cases} w_{ij}, & \langle EC_i, EC_j \rangle \in R, \\ 0, & \langle EC_i, EC_j \rangle \notin R, \end{cases}$$

R 是事件本体关系的集合。

根据图1所示的事件本体(片段),用不同的数字编号代表不同的事件类。(1)“突发事件”,(2)“地震”,(3)“火灾”,(4)“救援”,(5)“死伤”,(6)“募捐”,(7)“视察”,(8)“重建”,(9)“余震”,(10)“医疗救援”,(11)“心理援助”,(12)“体检”,(13)“手术”。图1所示的事件本体的邻接矩阵为

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1)	0.00	0.00	0.00	0.63	0.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(2)	0.00	0.00	0.00	0.63	0.88	0.52	0.37	0.58	0.47	0.00	0.00	0.00	0.00
(3)	0.00	0.00	0.00	0.63	0.88	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(4)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(5)	0.00	0.00	0.00	0.73	0.00	0.00	0.00	0.00	0.00	0.73	0.73	0.00	0.00
(6)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(7)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(8)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(9)	0.00	0.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(10)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00
(11)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(12)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(13)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

在矩阵 W 中,由于“地震”和“火灾”都是“突发事件”的子事件类,所以它们都继承了“突发事件”的属性,即“地震”和“火灾”对“死伤”和“救援”都是因果关系,且影响因子同样继承。“死伤”对“救援”的影响因子为0.73,由于“医疗救援”和“心理援助”都是救援的子事件类,所以“死伤”对“医疗救援”和“心理援助”的影响因子也都是0.73。

2 基于 HARank 算法的事件类排序

2.1 事件类重要度的计算

事件本体中事件类关联的特殊性:1)事件类 EC_i 的发生经常引起其他事件的发生,这时事件类 EC_i 表现出 Hubs 值的特点,比如图1中的“地震”事件;2)事件类 EC_j 的发生经常是由其他事件引起的,这时事件类 EC_j 表现出 Authorities 值的特点,比如

图 1 中的“运输”事件。事件类既可能表现出 Hubs 值的特点,也可能表现出 Authorities 值的特点,所以可以使用综合考虑事件类的 Hubs 值和 Authorities 值的链接分析算法 HARank (Hubs-Authorities Rank)^[16]计算事件类的重要度。

在某领域事件本体邻接矩阵上, $\text{In}(\text{EC}_i)$ 表示指向 EC_i 的事件类的集合, $\text{Out}(\text{EC}_i)$ 表示 EC_i 指向的事件类的集合, $R(\text{EC}_i)$ 表示事件类 EC_i 的重要度, n 是事件本体中事件类的个数。则 $R(\text{EC}_i)$ 的迭代计算公式为

$$R(\text{EC}_i)_k = R(\text{EC}_i)_{k-1} + d \left(\alpha \sum_{j \in \text{In}(\text{EC}_i)} R(\text{EC}_j)_{k-1} w_{ji} + \beta \sum_{g \in \text{Out}(\text{EC}_i)} R(\text{EC}_g)_{k-1} w_{ig} \right) + \frac{1-d}{n}, \quad (1)$$

式(1)中的 d 相当于 PageRank 中的阻尼系数,取值范围为 $0 \sim 1$, 通常取 $d = 0.85$ 。 $R(\text{EC}_i)_k$ 为事件类 EC_i 在第 k 次迭代后的重要度。 w_{ji} 为事件本体邻接矩阵中事件类 EC_j 与 EC_i 的关联强度。 α 和 β 是调整 Hubs 值和 Authorities 值的两个参数 ($\alpha, \beta \geq 0$, 且 $\alpha + \beta = 1$)。如果 $\alpha = 0$, 就以 Authorities 值作为迭代计算的依据, 类似于 PageRank 算法; 如果 $\beta = 0$, 就以 Hubs 值作为迭代计算的依据, 类似于 Reverse

PageRank 算法。式(1)综合考虑了事件的 Hubs 值与 Authorities 值, 可以根据本体的实际情况调整这两个参数的大小, 在本体的 Hubs 值与 Authorities 值不存在明显差异的情况下, 可以取 $\alpha = \beta = 0.5$ 。

2.2 实验及结果分析

目前, 用于计算对象重要度的算法主要有 PopRank^[17], Swoogle^[18], Reverse PageRank^[19] 和 CARRank^[20]等, PopRank 和 Swoogle 侧重于对象的 Authorities 值, 此算法是 PageRank 的改进, Reverse PageRank 和 CARRank 侧重于对象的 Hubs 值。本文提出的计算事件重要度的算法简记为 HARank。每类选一个代表, 共用 3 种算法(HARank, PopRank 和 Reverse PageRank)进行了实验比较。

基于完整的“地震”事件本体的邻接矩阵, 使用式(1)计算事件类重要度以及在指定的精度下迭代收敛的次数(取 $d = 0.85$, $\alpha = \beta = 0.5$), 当前后两次迭代误差 $\varepsilon \leq 1 \times 10^{-9}$ 时, 认为迭代过程收敛到不动点, 事件类重要度在输出时保留小数点后 6 位。取重要度最大的 10 个事件类与专家给出的结果做比较。实验结果如表 1 所示。

从准确率和偏差度两个方面对表 1 中的数据进行分析, 准确率 $P(\text{precision})$ 和偏差度 $D(\text{deviation})$ 的定义如下:

表 1 事件类重要度排序比较
Table 1 Comparison of event class importance sorting

比较项	HARank		PopRank		Reverse PageRank		专家评价
	事件类	重要度	事件类	重要度	事件类	重要度	
事件类及重要度	地震	0.069660	运载	0.035564	地震	0.237765	地震
	救援	0.039687	消杀	0.035564	救援	0.002230	救援
	医疗救援	0.024478	心理援助	0.029224	医疗救援	0.001306	死伤
	死伤	0.020935	医疗救援	0.029224	死伤	0.001120	疫情
	疫情	0.016484	募捐	0.019695	奔赴	0.000923	医疗救援
	心理援助	0.013749	疾病防疫	0.017774	疫情	0.000872	疾病防疫
	重建	0.010541	死伤	0.013029	运载	0.000785	心理援助
	募捐	0.010541	救援	0.013029	募捐	0.000714	重建
	运载	0.010378	疫情	0.013029	疾病防疫	0.000714	募捐
	疾病防疫	0.010079	地震	0.010887	重建	0.000714	派遣
P	90%		80%		80%		
D	1.5		3.79		2.21		

$$P = \frac{|CS \cap ES|}{|ES|}, \quad (2)$$

$$D = \frac{\sum_{|ES|} |Num_{CS} - Num_{ES}|}{|ES|}, \quad (3)$$

式(2)和(3)中, CS 是使用某种算法得到的事件类集合, ES 是专家给定的事件类集合, Num_{CS} 是事件类集合 CS 中某一事件类的序号, Num_{ES} 是专家给定的此事件类的序号。|ES| 是 ES 中事件类的个数。如果 CS 中的某事件类不在专家给定的事件类集合 ES 中, 则其偏差度为专家给定的事件类个数, 即 |ES|。

从表 1 可见, 本文提出的 HARank 算法对事件类重要性的排序与人工评测最为吻合, 效果最差的是 PopRank 算法。主要原因是 PopRank 算法侧重于事件类的 Authorities 值, Reverse PageRank 侧重于事件类的 Hubs 值。但在实际应用的本体中, 有的事件类的 Hubs 值和 Authorities 值都不大, 但其 Hubs 与 Authorities 值之和却可能很大。另外, 同一个事件本体中, 有的事件类的 Hubs 值偏大, 有的事件类的 Authorities 值偏大, 单纯的侧重于 Hubs 值或 Authorities 值的事件类排序算法都是不合理的。比如, 在“地震”领域事件本体中, “地震”发生后通常会引起其他事件发生, 则“地震”表现的 Hubs 值就非常大; 但如果是“救援”事件本体, 则引起“救援”的一些事件也非常多, 那么“救援”表现的 Authorities 值就非常大。

在不同的精度下(10^{-9} ~ 10^{-3}), 3 种算法迭代收敛的次数见表 2。

3 种算法执行效率和本体的结构密切相关, HARank 算法的执行效率居中。在图 1 所示的本体中, 几个关键事件类的 Hubs 值大于其 Authorities 值, 所以 PopRank 算法收敛速度最快。同样, 如果一个本体中关键事件类的 Authorities 值大于其 Hubs 值, 则 Reverse PageRank 算法收敛速度最快。当一个本体中事件类的 Hubs 值和 Authorities 值大致相

等时, 则 HARank 算法收敛速度是最快的。

3 结束语

现有本体模型对事件或者用概念间的关系, 或者以描述静态概念的方式表示, 这虽然在一定程度上可以表达事件, 但是存在诸多问题。为此, 本文在面向事件的本体研究的基础上, 以“事件”作为知识表示单元, 提出事件本体模型。事件本体模型是一种包含了许多有用信息的有向图, 可以应用于智能信息处理的诸多方面: 1) 借助事件本体可以辅助发现文本中事件之间存在的关系, 再添加事件实例之间的关系, 就可以用来描述文本, 是文本的一种新的表示模型; 2) 领域事件本体可以反映一个领域中事件类之间的关系及其关联强度, 借助事件本体可以对领域事件类进行排序; 3) 基于事件本体可以进行事件类与事件类之间的查询扩展, 是解决初始查询信息覆盖主题不全的一条新的途径; 4) 在事件本体上可以基于事件类之间的影响因子以及事件类诸要素之间的联系进行多层次的事件推理, 等等。

参考文献

- [1] Chen X. Why did John Herschel fail to understand polarization? The differences between object and event concepts. *Studies in History and Philosophy of Science*, 2003, 34(3): 491-513
- [2] 潘云鹤, 耿卫东. 面向智能计算的记忆结构理论综述. *计算机研究与发展*, 1994, 31(12): 37-42
- [3] Zhong Z Z, Liu Z T, Li C H, et al. Identifying key people from a single document using people event map. *Journal of Computational Information Systems*, 2010, 6(1): 17-23
- [4] 贾自艳, 何清, 张海俊, 等. 一种基于动态进化模型的事件探测和追踪算法. *计算机研究与发展*, 2004, 41(7): 1273-1280
- [5] 王伟, 赵东岩, 赵伟. 中文新闻关键事件的主题句识别. *北京大学学报: 自然科学版*, 2011, 47(5): 789-796
- [6] Makkonen J, Ahonen-Myka H, Salmenkivi M. Simple semantics in topic detection and tracking. *Information Retrieval*, 2004, 7: 347-368
- [7] 刘小明, 樊孝忠, 刘里. 融合事件信息的复杂问句分析方法. *计算机科学*, 2011, 39(7): 140-145

表 2 不同精度迭代收敛的次数

Table 2 Times of iteration convergence for different precision

算法	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
HARank	41	54	62	70	79	88	95
PopRank	33	45	53	58	64	69	75
Reverse PageRank	52	63	70	77	85	96	106

- [8] Sánchez D, Moreno A. A methodology for knowledge acquisition from the web. *International Journal of Knowledge-Based and Intelligent Engineering Systems*, 2006, 10(6): 453–475
- [9] Lee C S, Chen Y J, Jian Z W. Ontology-based fuzzy event extraction agent for Chinese e-news summarization. *Expert Systems with Applications*, 2003, 25(3): 431–447
- [10] Lin H F, Liang J M. Event-based Ontology design for retrieving digital archives on human religious self-help consulting // *Proceeding of the 2005 IEEE International Conference on e-Technology, e-Commerce and e-Service*. Hong Kong, 2005: 522–527
- [11] Hsu W L, Wu S H, Chen Y S. Event identification based on the information map-INFOMAP // *Proceedings of the 2001 IEEE International Conference on Systems, Man, and Cybernetics*. Tucson, 2001: 1661–1666
- [12] Vargas-Vera M, Celjuska D. Event recognition on news stories and semi-automatic population of an ontology // *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*. Beijing, 2004: 615–618
- [13] Mendez-Torreblanca A, Lopez-Lopez A. From text to ontology: the modeling of economics events. *Lecture Notes in Computer Science*, 2004, 3257: 502–503
- [14] 刘宗田, 黄美丽, 周文, 等. 面向事件的本体模型. *计算机科学*, 2009, 36(11): 191–195
- [15] 仲兆满, 刘宗田, 周文, 等. 事件关系表示模型. *中文信息学报*, 2009, 23(6): 56–60
- [16] 仲兆满, 刘宗田. 利用事件影响关系识别文本集合中重要事件的方法. *模式识别与人工智能*, 2010, 23(3): 307–313
- [17] Nie Z, Zhang Y, Wen J R, et al. Object-level ranking: bringing order to Web objects // *Proceedings of the WWW*. Chiba, 2005: 567–574
- [18] Ding L, Pan R, Fin T, et al. Finding and ranking knowledge on the semantic Web // *Proceedings of the ISWC*. Galway, 2005: 156–170
- [19] Fogaras D. Where to start browsing the Web? // *Proceedings of the IICS*. Leipzig, 2003: 65–79
- [20] 吴刚, 张阔, 李涓子, 等. 利用相互增强关系迭代计算本体中概念与关系的重要性. *计算机学报*, 2007, 30(9): 1490–1499