

# 面向事件的本体研究

刘宗田 黄美丽 周文 仲兆满 付剑锋 单建芳 智慧来

(上海大学计算机工程与科学学院 上海 200072)

**摘要** 世界是物质的,物质的世界是运动的,运动是绝对的,静止是相对的。事件作为人类知识的单元,反映了事物的运动,受到了学术界多领域专家的高度重视。在研究面向事件的知识表示的基础上,给出了事件定义及其六元组的形式化表示,并提出了事件本体模型;最后,以基于事件本体的篇章理解为例,简要说明了篇章与句子的事件网络表示、生成及其应用。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,更符合客观实际,更便于计算机模拟大脑工作。

**关键词** 事件,本体,事件本体,事件网络,知识表示

**中图分类号** TP311 **文献标识码** A

## Research on Event-oriented Ontology Model

LIU Zong-tian HUANG Mei-li ZHOU Wen ZHONG Zhao-man FU Jian-feng SHAN Jian-fang ZHI Hui-lai  
(School of Computer Engineering and Science, Shanghai University, Shanghai 200072, China)

**Abstract** Event-oriented knowledge representation method is reasonable according to the philosophical viewpoints that the world is of material, material is of movement, movement is absolute and static is relative. Nowadays, the academic community has paid more and more attention to event that is the unit of knowledge. By studying of the Event-oriented Knowledge representation, the definition of event was given and the representation of event in a 6-tuple formalized structure was presented. Then, an event ontology model was proposed. Taking text understanding as an application domain, the representation, building and application of the event network of texts and sentences were briefly introduced. In comparison with the conventional method, event ontology model represents knowledge with a higher granularity, meanwhile, it'll be much more suitable for computers to simulate human brain.

**Keywords** Event, Ontology, Event ontology, Event network, Knowledge representation

## 1 引言

世界是物质的,物质的世界是由实体组成的。实体是可以感知的、相对独立的、相对静止的存在。亚里士多德认为,实体有第一实体和第二实体之分。个别物体是第一实体,个别物体具有的属或种是第二实体,称之为静态概念。这里的静态概念不完全等同于哲学中的概念,是哲学中概念的一类。静态概念的包含关系构成了静态概念层次,理论上这是一个格结构。

世界是运动的,运动的世界是由事件组成的。事件是可以感知的、相对独立的、运动着的存在,它不同于静态概念。事件涉及多方面的实体,或称要素,包括时间要素、地点要素、参与者要素、过程状态要素。具有共同属性事件的集合的最大闭包称为事件类,或称为动态概念。事件类之间也具有包含关系,即由每个要素的属性的包含关系组合,这是一个多格结构。

世界的运动是绝对的,静止是相对的。任何实体都可以是事件要素的构成元素,不构成事件要素的实体是不存在的。事件是随着时间变化的具体事实。事件与事件之间具有本质的内在联系。

一些认知科学家从大脑的记忆原理和事件的结构方面去研究事件。斯坦福大学的 Zacks 认为事件是被旁观者所观察到的对现实世界所产生的行为,时间序列是事件因果关系的基本要素<sup>[1]</sup>。文献[2]介绍了 Lindsay 和 Norman 等提出的一个综合记忆模型(Elinor 模型),它也是一个网络模型,由结点和连线组成,结点代表概念、事件、情景等,连线表示两者之间的意义联系。Nelson 在脚本理论的基础上,形成了通用事件表示方法(GER)<sup>[3]</sup>,认为事件是包含了对象和关系的一个大的整体。

一些语言学家从语言学的角度给出了事件的定义和事件结构。Davidson 提出语义逻辑事件,认为事件不仅包括了表示动作行为的动词,还包括与之相关的名词和修饰成分。

到稿日期:2008-12-15 返修日期:2009-05-12 本文受国家自然科学基金项目(60575035)和上海市重点学科建设项目(J50103)资助。

刘宗田(1946--),男,教授,博士生导师,主要研究方向为人工智能、软件工程等,E-mail: ztliu@shu.edu.cn;黄美丽(1978--),女,博士生,讲师,主要研究方向为人工智能;周文(1979--),女,博士,讲师,主要研究方向为人工智能;仲兆满(1977--),男,博士生,主要研究方向为人工智能;付剑锋(1978--),男,博士生,主要研究方向为人工智能;单建芳(1979--),女,博士生,主要研究方向为人工智能;智慧来(1981--),男,博士生,主要研究方向为人工智能。

Chung<sup>[4]</sup>将事件定义为是由3个部分组成的术语：“谓词、事件框架(即谓词发生的时间段)、事件界(即谓词发生的情况或者条件)”。近年来,一些学者将语言学中的SVO(Subject-Verb-Object)结构和事件结构对应<sup>[5]</sup>:主语对应着事件的发起者,宾语对应着事件的承受者,谓语对应着行为。

在计算机信息学领域,事件研究越来越受到重视。知网(HowNet)中事件定义为“事情”,并将它分为静态和行动两大类。WordNet中给出了很宽泛的“事件”定义:“在特定地点和时间发生的事情”。在信息检索领域,“事件”被认为是“细化了的用于检索的主题”。在美国国防高级研究计划委员会(Defense Advanced Research Projects Agency, DARPA)主办的话题识别与跟踪(TDT)评测会议上,事件定义为“特定时间特定地点发生的事情”,认为事件是小于话题的概念,多个事件组成一个话题。话题识别与跟踪的5个主要子任务中包括事件识别这一重要的任务。信息提取是近年来自然语言处理领域的又一个热点,信息提取的3大主要任务中包括事件提取。推动事件提取领域发展的ACE(Automatic Content Extraction)评测会议,将事件定义为包含参与者的特殊的事情,事件通常可以描述为一种状态的改变。在自动文摘领域, Filatova等定义了称之为“原子事件(Atomic Events)”的概念<sup>[6]</sup>。它是动词(或者动名词)及其连接起来的的行为的主要组成部分(如参与者、地点、时间等)。周文和刘宗田等提出了基于事件和FCA的多文档自动文摘方法,并在DUC2005的语料上进行了评测,获得了较好的效果<sup>[7]</sup>。

由上可见,尽管各领域对事件的理解并不完全一致,但在认定事件是知识的重要组成这点上却是一致的。本文研究了传统本体缺乏对事件有合适表示的缺陷,提出以“事件”作为知识表示单元,构建面向事件的本体模型,并分析其在语义计算中的应用。

## 2 传统本体模型存在的缺陷

本体(Ontology)最初是一个哲学上的术语,10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着重要的意义。在国内,中科院的陆汝钤院士等许多学者也相继开展了本体方面的研究<sup>[8-20]</sup>,并取得了丰富的研究成果。

传统本体模型部分地反映了客观世界中事物的存在规律,特别是事物的分类与非分类关系,但也存在明显缺陷:

(1)认知科学家认为,人类主要是以“事件”为单位进行记忆和理解现实世界的;事件是随着时间变化的具体事实;事件与事件之间具有本质的内在的联系;事件关系到多方面的静态概念,是比静态概念粒度更大的知识单元。而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以及它们之间的关系,而概念定义为具有共同属性的对象的集合,对象由属性表示,这对于表示静态概念是合适的。但事件是动态的,是状态变化的。在传统本体中,事件只是作为概念的一类或关系的一类。用表示概念的方法表示事件类,忽视了事件类的动态特性,而用表示关系的方法表示事件类,不仅忽略了其动态特性,也忽略了事件的其它要素。

例如在文献[19]的本体中,概念project与developer之

间有developed by关系。但是,这个关系实际上应当是事件。概念project和project\_admin之间的manage也是如此。将develop作为关系,只表示了与主体和客体的关联,完全未能表示develop过程及作用等信息。

又例如在文献[20]中,将“理赔”、“交费”与“客户”、“保单”等同作为概念。但像“客户”这样的概念是静态的,用静态的属性可以表示它们。而“理赔”、“交费”这样的概念是动态的,它不仅包括静态属性,还包含动态属性,是一个运动过程,用描述静态概念的方法很难描述动态的概念。

用表示静态概念的方法表示事件,这比用关系前进了一步。但是事件所包含的知识是动态的,是用静态的方法所不能完全表示的。事件与静态概念的最大区别在于事件的参与对象要发生改变。例如“给”事件不是只在本体中给出注明“给”的节点以及“给予者”、“接受者”和“给予物”的联系就认为表示清楚了,实际上这远远不够,起码“给”事件应当反映“给予者”、“接受者”和“给予物”的状态改变,在“给”之前,“给予物”在“给予者”的掌控中;而“给”结束后,“给予物”在接受者的掌控中。除此之外,还可能还有其他知识需要表示。例如当“给予物”是“接受者”期望的时,“给”事件完成之后“接受者”会很高兴。显然,像这样的动态性知识,是静态概念表示方法很难表示的。

## 3 事件与事件类的定义

**定义1(事件)** 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:

$$e=(A,O,T,V,P,L)$$

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、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,例如快慢、使用什么、根据什么等等。

**O(对象)**:指事件的参与对象,包括参与事件的所有角色,这些角色的类型数目称为对象序列长度。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。例如:在小池塘里游泳,场所:小池塘,场所特征:水中。

**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,它们可以是事件发生的触发条件;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,事件各要素将引起变化或者各要素状态的变迁,这些变化和变迁后的结果,将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、核心词搭配等。核心词是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的表示与核心词之间的位置关系。核心词搭配是指核心词与其他词汇的固有的搭配。可以为事件附上不同语言种类的表现,例如中文、英文、法文等等。

**定义2(事件类, Event Class)** 指具有共同特征的事件的集合,用EC表示:

$$EC = (E, C_1, C_2, \dots, C_6)$$

其中,  $E$  是事件的集合, 称为事件类的外延;  $C_i = \{c_{i1}, c_{i2}, \dots, c_{im}, \dots\} (1 \leq i \leq 6, m \geq 0)$  称为事件类的内涵, 是  $E$  中每个事件在第  $i$  个要素上具有的共同属性的集合,  $c_m$  是事件类中每个事件在第  $i$  个要素上具有的一个共同属性。图 1 给出了一个事件类的例子。

/*	
事件名:	发现——主动类事件
动作特征:	程度: 很快、慢慢、突然、惊奇、最终、无意 方式: 通过分析、通过实验、 共性特征: 主体看、寻找
对象:	主体: 动物 客体: 物
断言:	前置条件: 客体不被主体感知 中间断言: 客体不被主体感知 后置条件: 客体被主体感知
语言表现:	“发现”、“发现”
*/	

图 1 事件类示例

事件各要素的属性值决定了这个事件的类别。6 个要素是冗余的, 但却是一致的, 这是因为有了语言表现要素。其他要素都是事件的本质属性, 惟有语言表现要素是人为属性。

## 4 事件的要素

### 4.1 事件的动作要素

事件是如何进行的, 我们的大脑能感知到。这很简单, 因为记住了这个事件的影视演变过程。例如我们看见了一个汽车交通事故事件, 记住了汽车冲向行人的画面、行人飞出跌倒在地的画面、汽车刹车的画面等等, 心理学称之为表象<sup>[21]</sup>。我们即便没有亲眼感受, 也可以利用已有的知识在头脑中制作出这样的影视。许多这样的事件的信息在头脑里整合的结果, 便有了事件类的动画, 即体现了许多同类事件的共同特点

的动画。但是对于计算机, 目前还无法模拟这种记忆过程。虽然其可以存储视频信号, 但是还不能像人一样能迅速判断出两个视频片段是否类似, 更无法将多个类似的视频片段归纳成一个具有共性的视频片段。

因此, 只好用属性值表示事件的动作。这是不得已的表示, 是符号主义的。将来的事件表示, 一定不仅仅用属性值表示, 而是主要用视频, 用可按照内容检索的视频表示。

动作属性又分为程度、方式、方法、工具等。

程度可以描述事件运行的激烈程度。例如交通事故的严重程度、跑事件的速度等。这是一个相对的值, 例如快的跑和快的走是不一样的, 因此还应当有事件类属性的平均值。事件实例的程度是相对于这个平均程度而言的。

方式指的是事件演化的方式。例如谈判事件, 可以是友好方式、激辩方式、粗野方式等。

方法是又一类属性。例如对于“搜索”事件, 有广度方法、深度方法、盲目方法、启发式方法等。

工具也是一类属性。例如“用铲子”挖地、“用毛笔”写字。

方式、方法、工具的取值是与事件类密切关联的, 不可能对于所有事件有同一值域。

方式、工具一般是实体, 但有时也可以是事件。例如“用战争的手段达到目的”, 这里战争是事件, 是作为另一个事件的方式。

### 4.2 事件的对象要素

对象是事件的主要参与者, 是事件中的主导作用者或被改变者。对象可分为主体和客体。主体是主导者, 是事件中的主角, 有时是事件的制造者或期望事件发生者。客体是事件中的被动者。按照分工, 事件中的对象分为不同的角色, 例如“选举”事件, 其中有主持组织者、监票者、投票者、候选人、当选者、落选者。

事件的对象一般是第一实体, 但也可能又是事件, 也可能是关系。还有的事件的对象是有关联的事件集合。

对于特定的事件类, 对象是有约束的, 例如选举事件的组织者可以是人或组织。利用这些约束, 可以判断句子中的主语或宾语。其他要素也有约束, 也可以利用。

### 4.3 事件的时间要素

每个事件都必须发生在一个时间段内。事件有开始、发展、结束。时间段由绝对时间表示, 也可以用相对时间表示, 也就是用另一个事件的时间的相对偏差表示。

### 4.4 事件的环境要素

任何事件必然发生在一定的环境中。环境由多个实体按照特定的关系组成。例如“上课”事件的环境有教室, 教室内有黑板、讲台、课桌、椅子等, 这些东西是有序存在的。

环境要素中的实体与对象要素中的实体都是事件涉及的, 还有动作要素中的工具。为什么有的属于对象要素、有的属于环境要素、有的属于动作要素? 关键在于它们在事件中的地位和作用。环境要素中的实体是相对静止的, 在事件中几乎是不变的, 或者虽然变化, 但并不重要。而动作要素的工具中的实体是运动的, 是对事件的实施有重要作用的。对象要素中的实体是最重要的, 或者是事件的主角, 或者是事件的驱动者, 或者是事件的主要被改变者。

### 4.5 事件的断言要素

断言描述的是事件的作用。前置条件表示的是事件发生之前的状态特点, 也是事件发生的条件。中间断言表示事件进行过程中的状态特点, 也就是在开始后和结束前各状态共同的特点。后置条件是事件结束状态的特点。

### 4.6 事件的语言表现要素

将事件与语言分离有很多优点。事件是客观的, 本来就不依赖于语言。语言是人类为了交流各自知道的事件发生而创造出来的符号表示, 因而事件就有了语言特性。交流和思考离不开语言, 然而很多研究工作往往将它们混为一谈。混为一谈的研究方法容易把问题搞复杂。

事件的语言表现是随着语言的变化而变化的。同一个事件, 对于不同的语言有不同的表现, 即使对于同一个语言, 也会有不同的表现。

语言表现的主要内容是核心词。核心词是对于事件的称谓、对于事件中各角色的称谓、对时间和环境的描述格式、对各要素结合的描述格式。

语言表现可以用文法表示, 也可以用规则表示, 或用模板表示。

### 4.7 要素的属性名和属性值

既然事件的特征只能用符号表示, 就必须给出属性名和属性值的符号约定。由于属性名和属性值数量庞大, 最直接的方法是借用自然语言中的词汇或句子表示。这就引起一个问题: 表示属性的词汇和句子又怎样理解呢? 如果计算机也像人一样具有感受, 这个问题就很好解决了。例如词汇“疼

痛”可以引起人对疼痛的感受的回忆,这个词的意思当然较容易理解了。文献[22]曾经提出过一个感受及活动、知觉、概念、语言的4层智能系统模型。但是很遗憾,到目前为止,我们还不知道如何让计算机有疼痛的感受。

在计算机没有感受的前提下,这种用自然语言的小集合描述属性和属性值的方法可行吗?如果可行,这就是说,任何句子、篇章都可以映射为用这个小集合描述的一些事件以及它们之间的关系,而对这个小集合的理解假定是已知的。如果不可行,则意味着,在无感受能力的机器上实现类似人的本体是不可能的。

## 5 事件本体

**定义3(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  $EO$ 。

**定义4(事件本体的结构)** 事件本体的逻辑结构可定义为一个三元结构  $EO: = \{ECS, R, Rules\}$ , 这里,  $ECS$ (事件类集合)是所有事件类的集合。 $R$ (事件类之间的关系集合)包括事件类之间的分类关系和非分类关系。由分类关系可构成事件类层次。非分类关系上标明关系种类名和链接强度。链接强度用区间 $[0, 1]$ 之间的值来表示,可通过学习或遗忘改变。 $Rules$ (规则)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。

事件本体作为一种面向事件的知识表示方法,更符合现实世界的存在规律和人类对现实世界的认知规律。过去的世界发生了许多的事件,变成了今天的世界;今天的世界又将发生许多事件,变成明天的世界。描述变化的历史,就是描述这些发生的事件以及它们之间的关系。人类用话语文本表述,描述的只是事件及其关系的语言表现。要理解这些话语文本,就必须知道这些事件类丰富的内容,这些内容的绝大部分是不可能话语文本中叙述的,而是作为共同知识预先存在于每个交流者的头脑中。事件本体正是为计算机建造这样的共同知识。

## 6 基于事件本体的篇章理解

一个句子或篇章可以利用事件本体的先验知识映射生成一些事件以及它们之间的关系,我们称之为事件网络,看作篇章语义的一种特殊的表示形式。映射生成事件网络的步骤如下:

- Step1 扫描文本,找出事件的核心词;
- Step2 根据核心词,在事件本体中找到对应事件类,使它们连接;
- Step3 根据事件类各要素的约束及语言表现,匹配篇章的句子成分;
- Step4 找出事件和各要素的最合适的实际值;
- Step5 重复 Step1—Step4,直至所有可能获得的事件要素均填充完毕;
- Step6 通过句子或篇章中对各事件描述之间的连接词、各要素的关联、在事件本体中对应事件类之间的关系,确定事件之间的语义关系。

图2是从范文澜著的《中国通史》中的句子“[[居住 A 在北方、西方 V 的人 O]统被称为 A“狄族”、“戎族 O]”。[其中“犬戎族”O 自称 A[祖先 O 为二白犬],当是[以犬 O 为 A 图

腾 O]。”(进行了事件标注)提取出的事件网络(未表示与事件本体的联系)。

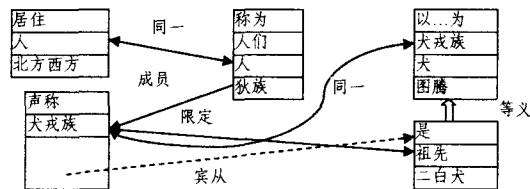


图2 基于事件本体的句子之事件网络表示

## 7 事件网络运算和应用

从句子或篇章中提取的事件网络,与事件本体有着千丝万缕的联系,事件与在事件本体中的所属事件类相连。事件网络是一种节点和边都含有许多信息的有向图。运用有向图的一些性质,结合这些信息,可以对事件网络做许多有趣的运算,例如事件网络内部聚类、事件网络约简、事件网络联想扩展、事件网络相似度的计算、多事件网络数据挖掘等。

在事件网络运算的基础上,可以对篇章进行逻辑层次划分、文本分类、自动摘要、对句子或篇章的自动阅读理解回答。例如根据节点之间的连接强度,可以将事件网络做内部聚类运算;选用不同的阈值,可以将事件网络分成层次结构,对篇章进行逻辑层次划分;根据事件在事件类的层次、关系的主次、与主题的贴切度等计算事件的重要度;可以根据重要度对事件网络进行各种约简;通过事件网络约简,可以对篇章进行自动摘要或简写;对于两个事件网络,可以根据它们包含的事件的相似程度以及顺序、关系等,给出两个事件网络的相似程度,通过对事件网络相似度的计算,可以对文本分类;因为事件网络与事件本体联系,事件网络之外的一些知识可以通过在事件本体中联想或推理得到,可以用于篇章的自动阅读理解的问题回答;从多个事件网络中,可以通过数据挖掘得到频繁且同时出现的事件要素,从而得到有关知识。

根据事件网络和事件本体连接,可以将它改写成另外的文本,可以翻译成其他语言的文本,可以推理判断出最初的句子或篇章中未直接描述但却隐含的意思。

**结束语** 传统本体作为一种有效地表现静态概念层次结构和语义关系的模型,在一定程度上能有效地解决目前计算机应用领域中存在的语义问题。但是,世界是物质的,物质的世界是运动的。因此,在对现实世界建模时,除了需要表示静态的物质概念及相互关系之外,还必须表示由事物参与的动态事件概念及相互关系。为此,本文在对面向事件的知识表示研究的基础上,以事件作为知识表示单元,提出了事件本体模型及利用事件本体将句子或篇章映射生成事件网络的思想,分析了事件网络上的运算和应用。

本文研究只是这项研究的开始,许多内容正在展开和深入。有关内容将在后续的文章中陆续展开讨论。

## 参考文献

- [1] Zacks J M, Tversky B. Event structure in perception and conception[J]. Psychological Bulletin, 2001, 127(1): 3-21
- [2] 潘云鹤,耿卫东. 面向智能计算的记忆结构理论综述[J]. 计算机研究与发展, 1994, 31(12): 37-42
- [3] Nelson K, Gruendel J. Event knowledge: structure and function in development[M]. Hillsdale, NJ: Erlbaum, 1986

(下转第199页)

环境	87.78%	90.13%	81.71%	83.67%
平均率	87.07%	89.32%	82.30%	84.69%
	CHI		DF	
类别	准确率	召回率	准确率	召回率
教育	84.52%	84.53%	83.34%	76.98%
经济	83.46%	86.58%	80.45%	80.54%
军事	84.39%	85.41%	82.37%	77.95%
体育	85.19%	82.71%	81.59%	80.27%
医药	81.57%	81.34%	79.48%	79.37%
交通	85.91%	85.47%	80.33%	80.09%
计算机	86.37%	84.96%	84.25%	81.31%
环境	82.77%	80.73%	82.43%	78.56%
平均率	84.27%	83.97%	81.78%	79.38%

从表 1 可以看出,在准确率和召回率方面,本文提出的特征选择方法优于 *MI*, *CHI*, *DF*,这说明本文方法选择的特征子集较优。

**结束语** 首先讨论了几种经典特征词选择方法,总结了它们的不足,然后给出了一个基于最小词频的文档频方法,紧接着把粗糙集引入进来并提出了一个基于 Beam 搜索的属性约简算法,最后把该属性约简算法同基于最小词频的文档频方法结合起来,提出了一个综合的特征选择算法。该方法在选择特征时不但考虑了特征词的权重而且还考察了特征之间的潜在关系,在某种意义上改变了传统的仅以词权重为标准来选择特征词的做法,不但降低了向量空间的维数,而且还使选出的特征子集具有较好的代表性。实验证明,本文特征选择方法同 3 种经典特征选择方法“互信息”、“ $\chi^2$  统计量”以及文档频相比有较高的准确率和召回率,为后续的知识发现算法减少了时间与空间复杂性,从而使得本文方法在文本分类

中有一定的使用价值。

## 参考文献

(上接第 192 页)

[4] Chung S, Timberlake A. Tense, aspect, and mood[J]. *Language Typology and Syntactic Description*, 1985, 3: 202-258

[5] Chang J. Event Structure and Argument Linking in Chinese[J]. *Language and Linguistics*, 2003, 4(2): 317-351

[6] Filatova E, Hatzivassiloglou V. Domain-independent detection, extraction, and labeling of atomic events[C]// *Proceedings of RANLP. Borovetz, Bulgaria*, 2003: 145-152

[7] Zhou W, Liu Z. Event-based Knowledge Acquisition for Ontology Learning[C]// *Proceedings of The 6th IEEE International Conference on Cognitive Information (ICCI'07)*. Lake Tahoe, California, USA, 2007: 498-501

[8] 陆汝钤. 世纪之交的知识工程与知识科学[M]. 北京:清华大学出版社, 2001

[9] 金芝. 基于本体的需求自动获取[J]. *计算机学报*, 2000, 23(5): 486-492

[10] 李善平, 尹奇, 胡玉杰, 等. 本体论研究综述[J]. *计算机研究与发展*, 2004, 41(7): 1041-1052

[11] 史忠植, 董明楷, 蒋运承, 等. 语义 Web 的逻辑基础[J]. *中国科学: E 辑*, 2004, 34(10): 1123-1138

[12] 高济. 基于表示本体论的智能系统开发[J]. *计算机研究与发展*, 1996, 33(11): 801-807

[1] Delgado M, Martin-Bautista M J, Sanchez D, et al. Mining text data: special features and patterns [A]// *Proceedings of ESF Exploratory Workshop*[C]. London: U. K, Sept. 2002: 32-38

[2] 朱颢东, 蔡乐才, 刘忠英. 一种改进的文本特征选择算法[J]. *现代电子技术*, 2008(8): 97-99, 102

[3] Friedman N, Geiger D, Goldszmidt M. Bayesian Network Classifiers[J]. *Machine Learning*, 1997, 29(2): 131-163

[4] 张海龙, 王莲芝. 自动文本分类特征选择方法研究[J]. *计算机工程与设计*, 2006, 27(20): 3838-3841

[5] Pawlak Z. Rough sets [J]. *International Journal of Information and Computer Sciences*, 1982, 11(5): 341-383

[6] Jiye L, Chin K S, Chuangyin D, et al. A new method for measuring uncertainty and fuzziness in rough set theory [J]. *International Journal of General Systems*, 2002, 31 (4): 331-342

[7] 曾黄麟. 智能计算[M]. 重庆: 重庆大学出版社, 2004

[8] 刘振华, 刘三阳, 王珏. 基于信息量的一种属性约简算法[J]. *西安电子科技大学学报: 自然科学版*, 2003, 30(6): 835-838

[9] 杨胜, 施鹏飞. 前向 Beam 搜索粗糙集属性约简算法[J]. *计算机科学*, 2006, 33(5): 200-204

[10] 邹涛, 王继成, 朱华宇. www 上的信息挖掘技术及实现[J]. *计算机研究与发展*, 1999, 36(8): 1019-1024

[13] 王海涛, 曹存根, 高颖. 基于领域本体的半结构化文本知识自动获取方法的设计和实现[J]. *计算机学报*, 2005, 28(12): 2010-2018

[14] 眭跃飞, 高颖, 曹存根. NKI 中的本体、框架和逻辑理论[J]. *软件学报*, 2005, 16(12): 2045-2053

[15] 李曼, 王大治, 杜小勇, 等. 基于领域本体的 Web 服务动态组合[J]. *计算机学报*, 2005, 28(4): 644-650

[16] 吴健, 吴朝晖, 李莹. 基于本体论和词汇语义相似度的 Web 服务发现[J]. *计算机学报*, 2005, 28(4): 595-602

[17] 康达周, 徐宝文, 陆建江, 等. 支持语义 Web 模糊本体的描述逻辑[J]. *东南大学学报: 英文版*, 2006, 22(3): 343-347

[18] 强宇, 刘宗田, 李旭, 等. 一种基于模糊聚类的模糊本体生成方法[J]. *计算机科学*, 2006, 33(4): 148-150

[19] 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[J]. *计算机学报*, 2007, 30(9): 1490-1499

[20] 杨立, 左春, 王裕国. 面向服务的知识发现体系结构研究与实现[J]. *计算机学报*, 2005, 28(4): 445-457

[21] 刘宗田. 一种表象思维模型的研究[J]. *计算机科学*, 1996, 23(5): 57-59, 66

[22] 刘宗田. 人工神经系统的心理学方法[J]. *计算机科学*, 1994, 21(1): 44-46, 35