



《语言学研究新视界文库·语言工程与技术系列》

主编：萧国政

A Study on the Formalized Description of
the Event Semantics in E-commerce Domain

电子商务领域 事件语义形式化描写研究

高精鍊／著



中国出版集团



世界图书出版公司



《语言学研究新视界文库·语言工程与技术系列》 主编：萧国政

国家社科基金重大项目课题（项目编号11&ZD189）阶段性成果

A Study on the Formalized Description of
the Event Semantics in E-commerce Domain

电子商务领域 事件语义形式化描写研究

高精鍊／著

中国出版集团
世界图书出版公司
广州·上海·西安·北京

图书在版编目 (CIP) 数据

电子商务领域事件语义形式化描写研究 / 高精鍊著. — 广州: 世界图书出版广东有限公司, 2015.11

ISBN 978-7-5192-0486-0

I. ①电… II. ①高… III. ①电子商务—研究 IV. ①F713.36

中国版本图书馆CIP数据核字 (2015) 第287284号

电子商务领域事件语义形式化描写研究

责任编辑 李 瑞

出版发行 世界图书出版广东有限公司

地 址 广州市新港西路大江冲25号

[http: // www.gdst.com.cn](http://www.gdst.com.cn)

印 刷 北京振兴源印务有限公司

规 格 710mm×1000mm 1/16

印 张 11.5

字 数 205千

版 次 2015年11月第1版 2015年11月第1次印刷

ISBN 978-7-5192-0486-0/H·0996

定 价 36.00 元

版权所有, 翻版必究

《语言学研究新视界文库》编委会

主 编 萧国政 胡 憐

副 主 编 陈海庆 李春玲

执行副主编 宋 焱

委 员 (以姓氏笔画为序)

马道山 王相锋 尤翠云 左步雷 占俊英

冯 曼 刘长江 刘宏刚 许宏晨 李爱国

张立新 陈宏俊 周 刚 金克中 孟凡君

俞建梁 聂中华 徐以中 唐承贤 章礼霞

詹 成 薛宏武 戴卫平

《语言学研究新视界文库》总序

语言是人类社会最为辉煌的产物，它见证和承载了人类心智的进化、人类社会的变迁和人类文明的发展。

语言学是最古老的学科之一，人类对语言的研究可以追溯到上古时期。古印度早在吠陀时期（约前 12 世纪）即开始对梵语进行系统的研究。古希腊在前 6—前 5 世纪的时候就开始有了对语言科学的记录。^[1]

语言学是一门复杂的科学，它跟许多其他学科密不可分。语言并不仅仅是语言学的研究对象，也是哲学、心理学、教育学、逻辑学、文学、人类学、历史学、医学、社会学、信息科学、人工智能学等学科的重要研究对象。这些学科不仅借鉴语言学的研究成果为其所用，同时它们的研究方法、研究发现和研究成果反过来也补充、修正和丰富了语言学的理论和应用研究，从而产生了一大批交叉学科，如心理语言学、认知语言学、社会语言学、计算语言学、病理语言学等，其中不少已成为语言学的新分支。另一方面，语言学在跟这些学科相结合的同时，还源源不断地从其他更多的学科，如数学、化学、物理学等吸取、借鉴新的理论和方法，派生出很多新的语言学理论和流派，如结构主义语言学、形式语法学（转换生成学派）、数理语言学、配价语法等。可以说，人类历史上迄今为止还没有哪一门学科像语言学一样，能够跟如此广泛的学科群产生如此深入的实质性的交叉和结缘。

语言学是一门充满朝气的学科，其他学科有新的理论和学说产生，常会给语言学带来新的血液。数千年来，这门古老的学科一直在与时俱进，不断焕发新的生命力。进入 21 世纪以来，科学技术成为社会发展的主要动力，创新、创造和发现成为新的时代特色。互联网及其快速发展使语言使用、语言信息化和智能化成为时代突

[1] R. H. Robins: *A Short History of Linguistics*, Addison Wesley Longman Limited, 1997.



出的标志,知识增长的速度和规模盛况空前,各个学科互相结合、互为借鉴,取长补短,学科的大交叉、大融合成为当前学术发展的整体趋势。语言学在这一新的历史洪流中,跟其他周边学科的结合越来越紧密,新的研究领域不断得到拓展。比如,研究语言的生物学特性的生物语言学,研究语言符号组合倾向和概率的计量语言学,研究语言认知基础、模式与规律的认知语言学,研究语言与环境关系的生态语言学,研究语言经济功用与规律的语言经济学,研究语言个性与普遍共性规律的类型学,等等^[1],无一不展现出大学科融合背景下的新的研究视野,昭示着语言学新的研究趋势和方向,同时大批新的作品和学者不断涌现和诞生。

在这样的背景下,为了尽量及时、全面地向广大读者推介我国学者在语言学各个领域的最新优秀研究成果,经反复论证和广泛征求各方意见,我们决定编辑出版《语言学研究新视界文库》大型系列丛书。该文库由世界图书出版广东有限公司、中国语文学会语言信息化与智能化研究会和湖北省语言与智能信息处理研究基地共同策划,文库共分为如下六个子系列:

- (1) 语言本体与应用
- (2) 语言教学与传播
- (3) 语言认知与发展
- (4) 语言工程与技术
- (5) 翻译理论与实践
- (6) 外文博士论文库

这六个子系列既能相互独立,又相辅相成。

文库的出版为长期开放项目,拟连续滚动出版符合选题要求的优秀著作。我们诚邀海内外语言学界精英和有识之士携手合作,共襄新世纪我国语言学科发展之盛事!

文库的出版得到了中国图书出版集团、世界图书出版广东有限公司、中国语文学会语言信息化与智能化研究会以及湖北省语言与智能信息处理研究基地的各位领导和同仁的大力支持,在此谨致谢忱!

因编者水平所限,文库的缺点与不足在所难免,恳请广大读者批评指正。

萧国政 胡 憐
2015 年仲夏

[1] 当然,这些学科,有的已经有一定的研究历史和基础,有的才刚刚起步。

摘 要

电子商务是移动互联网时代商业活动的重要模式，电子商务的智能化是电子商务技术与应用发展的必然方向。实现电子商务的智能化涉及多种理论与技术，其中，智能语义计算是最重要也是最难的问题之一。

事件语义学作为一种新的语言学理论被引入到语言信息处理中，为篇章的语义分析带来了新的思路。对事件与事件结构要素的理解，是事件语义分析的基础；而事件分析技术在具体领域的应用，是事件语义学对自然语言处理的价值体现。目前，国内外在这两个方面的研究，都尚处于探索阶段。本书以事件语义学为基础，通过定义事件描写模板，建构电子商务领域核心事件形式化描写模板库，为实现电子商务的智能语义计算提供基础理论和资源。

全书共分6章，各章主要内容和观点如下。

第1章：绪论。概要阐述本书的研究背景、研究对象、研究内容与研究意义。通过分析电子商务的发展现状与技术趋势，阐明语义计算对电子商务智能化的主要意义及其实现途径。

第2章：事件语义学研究概述。全面梳理国内外事件语义学研究现状，探讨现有研究成果对语义计算应用的局限性，阐述自然语言处理领域事件语义学的研究方法及其应用价值。

第3章：事件要素与事件结构描写。从事件本身的语义特征出发，通过深入分析事件的内涵与外延来定义事件并对事件进行分类，阐明了每一类事件的语义特征。分别从状态变化及其所涉及的对象和具体语境中事件所凸显的语义信息重心出发，将事件分为两种结构类型（简单事件、复杂事件）、四种语义类型（时间变化型事件、空间变化型事件、时空并重型事件与时空无关型事件）、两种体特征类型（过程型



事件和结果型事件)。通过分析事件的语义结构,提取了事件的6个基本要素:事体、参与者、时间、空间、数量、事件环境,在这6个一级基本要素中进一步提取25个二级子要素,并以这些要素为基础构建了事件的通用描写模板。

第4章:电子商务领域事件本体建构。参照知识本体建构的通用方法,探讨电子商务领域事件本体的建构方法和步骤,重点探讨了电子商务领域核心事件概念体系的建构。通过复用专业词典和网络资源,结合电子商务交易真实文本语料,提取电子商务事件概念1057个。以商务交易流程与电子商务系统的技术特点为依据,对这些事件概念进行分层、分类,建构了电子商务基本事件概念分类体系,共分为三层:第一层包括信息流事件、物流事件、资金流事件、综合事件、通用事件,共5大类;第二层包括25个小类;第三层为词群,即概念节点,共452个。

第5章:电子商务领域核心事件语义知识描写。选取电子商务基本事件概念体系中的核心事件,分析每类事件所涉及范围及其凸显的语义要素,运用六元组事件通用描写模板对电子商务领域核心概念进行详细的形式化描写,建构电子商务领域事件模板库。

第6章:总结与展望。对整个的研究工作进行总结,归纳取得的成绩与存在的不足,展望下一步的研究计划。

关键词: 电子商务; 语义计算; 事件本体; 事件要素; 事件描写模板

Abstract

E-commerce is a significant commercial pattern in the mobile Internet era and the intelligentization of E-commerce has been an inexorable direction of the development of its technology and application. The realization of its intelligentization involves various theories and techniques among which intellectual semantic computing ranks one of the most tricking and crucial questions.

Being a new theory introduced into the processing of natural language processing, event semantics has shed some light on discourse semantic analysis. The understanding of events and the elements in event structures lays the foundation for semantic analysis, while the concrete application of event semantic analysis is the embodiment of the value of natural language processing. Nowadays, the domestic and international studies of semantic analysis and its applications are still under exploration. Based on event semantics, this dissertation constructs a template bank for the formalized description of the core events in E-commerce by defining the template of event description, which aims at providing the theoretical foundation and resource for the realization of the intellectual semantic computation in E-commerce.

The book consists of six chapters and the main ideas for each chapter are as below.

Chapter 1: Introduction. This chapter briefly introduces the research background, objects, contents and its significance. By analyzing the overall development of E-commerce and its future trends, the significance of semantic computation for the intelligentization of E-commerce is illustrated.

Chapter 2: An Overview of Event Semantics Research. This chapter sorts out both



domestic and overseas research situation of event semantics and discusses the limitation of the research findings for semantic computation. Furthermore, it also expounds the research methods and application value of field event semantics in processing natural language.

Chapter 3: Event Elements and Event Structure Description. From the perspective of semantic attributes, this chapter thoroughly analyzes the connotation and denotation of events in order to make definition and classification of them as well as elaborates the feature of each event category. According to the changes of state and the involved participants, as well as the meaning in specific context, events are divided into two structure patterns (simple event and complex event), four semantic types (time-oriented event, space-oriented event, time-space dual-relevant event and time-space free event) and two aspect patterns (processing event and resultative event). By analyzing the semantic structures, six basic elements are extracted, namely event body, participant, time, space, number and situation. Based on the six primary elements, 25 sub-elements are further defined. These elements as a whole lay the basis of the general event description templates.

Chapter 4: The Construction of the Event Ontology in E-commerce Domain. Referring to the common approach of ontology construction, this chapter discusses the methods and procedures to construct event ontology in E-commerce with emphasis on the conception construct of the core events. 1066 event concepts are extracted from the specialized dictionaries, online resources, and the corpus of real business transactions. These concepts are further classified and stratified according to the transaction process and the technologies in E-commerce, and thus to construct the concept system of essential events in E-commerce. The system consists of three levels, the first level includes 5 categories of events, information flow events, logistics events, capital flow events, synthetic events and general events; the second level includes 25 subcategories and the third level covers 453 synsets, i.e. conceptual nodes.

Chapter 5: Semantic Description of the Core Events in E-commerce. In this chapter, some of the core events in E-commerce selected from the basic concepts list, the scope and semantic elements of each event category are analyzed. By using the universal six-tuple template, core concepts in E-commerce are made a detailed formalized description and thus

the event template bank in E-commerce is constructed.

Chapter 6: Conclusion and Outlook. The achievements and limitations of this research are concluded, and the outlook for further research is also remarked.

Key words: E-commerce; Semantic Computation; Event Ontology; Event Element; Template for Event Description

目 录

第1章 绪 论	001
1.1 研究背景与研究意义	001
1.2 研究目标与研究内容	010
第2章 事件语义学研究概述	013
2.1 事件的定义	013
2.2 语言学中的事件研究	016
2.3 面向自然语言处理的事件研究	028
2.4 述评与讨论	030
第3章 事件要素与事件结构描写	038
3.1 事件语义资源库的内容与建设现状	038
3.2 事件的分类	041
3.3 事件要素提取	052
3.4 事件结构描写	060
3.5 事件模板与子事件关系描写	064



第4章 电子商务领域事件本体建构	075
4.1 本体建构方法概述	075
4.2 移动电子商务事件本体建构步骤	078
4.3 电子商务领域事件概念体系	079
第5章 电子商务领域核心事件语义知识描写	089
5.1 信息流核心事件模板	089
5.2 物流核心事件描写	108
5.3 资金流核心事件描写	121
5.4 综合事件描写	128
第6章 总结与展望	142
6.1 主要研究工作	142
6.2 主要创新点	143
6.3 尚待进一步研究的问题	144
附录：电子商务基本事件概念体系	145
参考文献	155
后 记	168

第 1 章 绪 论

1.1 研究背景与研究意义

1.1.1 我国电子商务的发展现状

近年来,电子商务的发展和应用不断普及和深化,从政府到企业再到消费者个人,都越来越重视电子商务。2012年3月,国家工业和信息化部发布了《电子商务“十二五”发展规划》,制定了发展总体目标:“到2015年,电子商务进一步普及深化,对国民经济和社会发展的贡献显著提高。电子商务在现代服务业中的比重明显上升。电子商务制度体系基本健全,初步形成安全可信、规范有序的网络商务环境。”^[1]

2014年9月,以电子商务为主要业务的中国互联网领军企业阿里巴巴公司在纽约证券交易所上市,成为全球头条新闻。这一全球互联网发展史上具有里程碑意义的事件表明,对于全球和中国市场而言,电子商务的地位日益凸显。近10年来,我国电子商务经历了高速发展期,目前正以一种崭新的商业模式加速与传统企业的深入融合。

根据2014年8月27日中国电子商务研究中心发布的监测数据,我们可以清晰地看到,近几年来我国电子商务的交易额与从业人数逐年大幅度稳步递增。^[2]

[1] 引自国家工信部文件《电子商务“十二五”发展规划》, <http://www.miit.gov.cn/n11293472/n11293832/n11293907/n11368223/14527814.html>.

[2] 数据来源:中国电子商务研究中心:《2014年(上)中国电子商务市场数据监测报告》, http://www.100ec.cn/zt/upload_data/down/2014shch.pdf, 其中2015年数据为预测数据。



图 1-1: 2010—2015 年中国电子商务市场交易规模

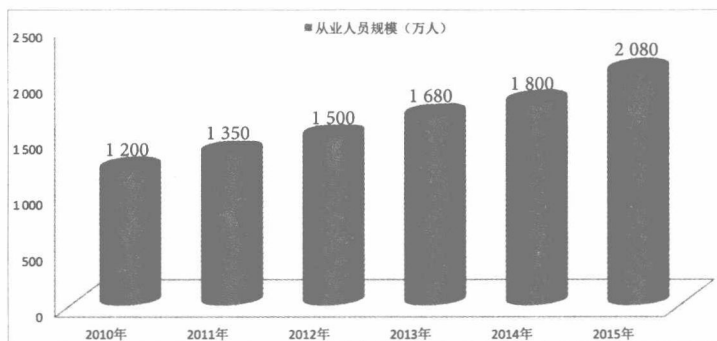


图 1-2: 2010—2015 年中国电子商务服务企业带动从业人员规模

目前, 电子商务已经成为互联网时代人们日常生活中必不可少的部分, 成为引领国民经济和社会发展的重要力量。

1.1.2 电子商务系统的技术发展趋势

作为互联网时代新兴的商业形式, 电子商务系统的运作高度依赖于多种信息技术, 如网络技术、移动通信技术、数据库与数据挖掘技术、信息检索技术、自然语言处理技术、在线支付技术、可信计算与信息安全技术、物联网技术等。信息技术的发展日新月异, 引领电子商务朝着越来越智能、越来越便利、越来越安全的方向发展。从信息技术的角度而言, 电子商务呈现出移动性、智能化的发展趋势。

1.1.2.1 移动性是电子商务发展的必然方向

随着 3G、4G、WIFI 城域网等无线移动通信技术的高速发展, 以及智能手机、

平板电脑、笔记本电脑、车载电脑、可穿戴设备等智能移动终端的迅速普及，整个社会已经从互联网时代全面进入移动互联网时代。

中国互联网络信息中心于2015年1月最新发布的第35次《中国互联网络发展状况统计报告》的数据显示，我国移动上网的网民规模稳步增长，已经远远超过采用台式电脑等固定设备上网的数量（如图1-3、图1-4）。^[1]

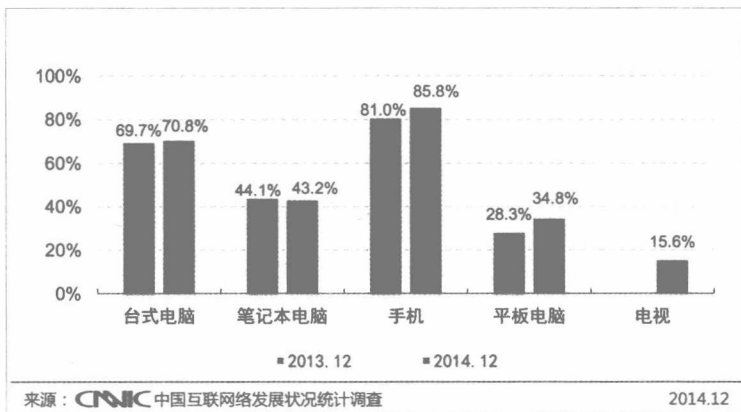


图 1-3：2013—2014 年中国互联网接入终端使用情况



图 1-4：2007—2014 年中国手机网民增长情况

上网方式和上网终端的移动化，导致互联网上的各种应用业务也迅速向移动化转向。在这种技术环境下，电子商务应用必然会逐渐从传统的互联网搬迁到移动互

[1] 数据来源：中国互联网络信息中心第35次《中国互联网络发展状况统计报告》，<http://cnnic.cn/gywm/xwzx/rdxw/2015/201502/W020150203456823090968.pdf>。



联网。移动通信技术、移动信息终端和移动网络业务相互促进,使得电子商务应用突破了以往的技术限制,为用户带来低成本、个性化、使用便捷、不受时空局限的应用模式。因此,移动性是电子商务系统和电子商务业务发展的必然方向。

1.1.2.2 智能化是电子商务发展立根之本

信息技术带给我们的直接利益之一,就是生活的智能化。智能化生活体现在很多方面,比如自动化的智能家务管理、交通的智能化解决方案、运动与健康的智能监测、智能化的娱乐休闲活动、智能化的经济与消费活动等。这一切,无不涉及电子商务,而且主要是通过移动电子商务系统解决方案来实现的。

移动电子商务的智能化,涉及多个领域的技术应用。比如,通过云存储技术,实现大规模数据采集与交换;通过数据挖掘技术,实现用户行为建模与个性化信息推送服务;通过图像识别与二维码技术,实现商品快速搜索与定位;通过语义识别,实现智能信息检索;通过无线射频识别(Radio Frequency Identification,简称RFID)技术,实现智能物流管理;通过基于位置服务(Location Based Service,简称LBS)技术,实现用户跟踪与点对点互动营销。

在信息技术革命带来的智能化应用已经逐渐渗透社会生活各个环节的今天,移动电子商务已经成为联系这些应用、服务与用户之间的纽带。因此,智能化是电子商务发展的立根之本,若不能实现智能化,电子商务将失去其生存的土壤。

1.1.3 电子商务系统中的自然语言处理技术

自然语言处理技术是人工智能的核心技术之一,是目前最重要的智能信息处理技术。它是联结各类信息媒体技术和知识工程的纽带,为未来基于数字网络信息的知识服务奠定了重要的技术基础。自然语言处理技术也是实现电子商务智能化不可替代的核心技术。电子商务智能化的多个环节,都离不开自然语言处理技术。

1.1.3.1 知识本体

虽然从严格的学科分类角度而言,本体和本体工程(Ontologies and Ontological Engineering)是知识工程(Knowledge Engineering)的研究领域,并不直接隶属于自然语言处理。然而,由于自然语言处理的很多理论和应用,尤其是基于语义网的研究,