

工学博士学位论文

基于自然语言处理的
空间概念建模研究

**Research on Spatial Conceptual Model
Based on Natural Language Processing**

李晗静

哈尔滨工业大学

2007年 月 日

图书分类号：TP391.2

U. D. C. : 681.37

工学博士学位论文

基于自然语言处理的
空间概念建模研究

博士 研究生 : 李晗静
导 师 : 李 生 教授
申请学位级别 : 工学博士
学 科 、 专 业 : 计算机应用技术
所 在 单 位 : 计算机科学与技术学院
答 辩 日 期 : 2007 年 1 月 15 日
授 予 学 位 单 位 : 哈尔滨工业大学

Classified Index: TP391.2

U. D. C. : 681.37

Dissertation for the Doctoral Degree in Engineering

**Research on Spatial Conceptual Model
Based on Natural Language Processing**

Candidate:	Li Hanjing
Supervisor:	Prof. Li Sheng
Academic Degree Applied for:	Doctor of Engineering
Specialty:	Computer Application Technology
Affiliation	School of Computer Science and Technology
Date of Oral Examination:	Jan., 2007
University:	Harbin Institute of Technology

摘要

空间概念建模是文景转换研究中一个十分基本的问题。研究自动、有效的自然语言描述空间概念建模方法,对于自然语言理解及文景转换研究具有重要的理论意义和实用价值。本文以自然语言描述空间关系为研究对象,提出了一整套基于自然语言描述的空间概念建模方案。作为概念建模的一个重要环节,研究了空间本体库构建以及空间关系抽取技术。在此基础上,研究并且实现了基于两步法的物体空间摆放。基于 t 检验的目测评价方法对空间概念建模原型系统的场景进行了评价,验证了小规模空间本体库的有效性,以及本文提出的基于自然语言描述空间概念建模方案的可行性和重要价值。

具体地讲,本文从如下几个方面进行了研究:

1. 研究了空间本体库构建技术。分析了空间本体库构建需要解决的关键问题,提出了基于 SUMO 的空间本体概念、关系定义以及基于半自动化实例获取方法,有效的解决了本体实例化以及间接相关问题。在此基础上,综合利用汉语语义库、图形资源构建了小规模的空间本体库。

2. 研究了篇章级空间关系抽取技术。分析了目前信息抽取方法特点以及待解决问题,提出了基于线性分类器的空间关系抽取方法。提出汉语描述空间关系的形式化描述,以及抽象空间关系抽取为二值分类问题。在词性标注的基础上,结合词性、空间语义等特征,分层次地实现了篇章级空间关系抽取。该方法充分利用了汉语词性标注研究成果,避开了汉语句法分析以及语义分析的难题,有效地实现了空间认知概念的抽取,同时保证了正确率。

3. 在小规模空间本体库基础上研究了物体空间摆放方法。分析了目前概念建模方法特点以及待解决关键问题,提出基于区域确定实现物体空间摆放方法。区域确定是基于小规模空间本体库和聚类策略确定实体摆放区域方法,平滑地实现定性信息到定量信息的转换,离散自然语言系统到连续图形系统的过渡,同时有效地缩小了物体摆放搜索空间。以此为前提,提出基于最优保持简单遗传算法的物体精确摆放方法,该方法保证了搜索全局最优性。该方法有效地降低了问题难度,同时保证了物体空间摆放的合理性,为概念建模提出了一个崭新的思路。

4. 建立了基于自然语言描述的空间概念建模原型系统。探讨了空间概念建模系统的基本框架、理解可视化流程,以及提出了基于 t 检验目测评价方

案。基于 t 检验目测评价方法，通过对系统可视化场景的评价验证了基于自然语言描述的空间概念建模方案的有效性和可行性。

关键词：本体论；空间关系；信息抽取；概念建模；最优解问题；文景转换

中国知网
CNKI

Abstract

Spatial conceptual model is a fundamental problem in the text-to-scene conversion tasks. The research on effective automatic spatial conceptual model which is described in natural language is of great theoretical and practical significance. This paper presents a method for automatic spatial conceptual model based on natural language descriptions. As a crucial step, it first studies methods for the definition and creation of the spatial ontology and the extraction of spatial information. Then, based on these, the layout of 3D objects based on two steps is researched and realized. The prototype of spatial conceptual model is evaluated based on t-test we proposed. The evaluation suggested methods are useful in real model.

This thesis is arranged as follows:

1. The definition and acquisition of the spatial ontology is studied. On analysis the issues in spatial ontology creation, the methods of the definition of spatial concepts and relations based on SUMO and the acquisition of spatial ontology based on clustering are proposed. The methods solves the difficulty of the spatial ontology instantiation automatically and the problem of indirect association in cluster process. On the basis of these methods, the small spatial ontology is created on the resources including HowNet and graphics.

2. The extraction of spatial relationships in text level is studied. On analysis the issues in spatial information extraction, we propose the method extracting spatial relationships based on binary-classification. At first, we give definitions on spatial relationships in Chinese formally and modeled the extraction of the spatial relationship as a binary classification problem. And then, on the basis of part-of-speech tagging, the extraction of spatial relationships in text level is realized step by step on the features on part-of-speech and spatial semantics. The method takes advantage of the results on part-of-speech tagging in Chinese, keep away from the usage of syntactic parser and semantic parser, and realized the extraction on the cognitive level with the good performance.

3. Based on the small spatial ontology, the layout of 3D objects is studied.

We proposed the layout of 3D objects on the qualification of the layout area on the analysis the issues on objects layout. The qualification of the layout are is based on spatial ontology and clustering, crucial to turn the qualitative information into quantitative information, and the bridge between the symbolic system and the numeric system and reduces the search space for the exact layout. After that, exactly layouting of 3D objects based on OMSGGA is done. The OMSGGA can ensure to search the global optimization. The method reduces effectively the difficulty of objects layout and resolve objects layout deeply and soundly. It opens new way to conceptual model.

4. A prototype of Chinese spatial conceptual model system is implemented. The frame of the system and the modeling procedure are discussed. The evaluation based on t-test is proposed. Evaluation on the visualization results demonstrates the good performance of the schema of spatial conceptual model.

Keywords Ontology; spatial relationships; information extraction; conceptual model; the optimization problem; text-to-scene conversion

目录

摘要	I
Abstract	III
第 1 章 绪论	1
1.1 本文研究的背景和意义	1
1.1.1 研究背景	1
1.1.2 研究意义	2
1.2 相关研究综述	3
1.2.1 相关研究术语分析	3
1.2.2 虚拟角色系统发展概况	4
1.2.3 文景转换系统发展概况	7
1.3 本文的主要研究内容	13
第 2 章 空间本体库构建的研究	15
2.1 引言	15
2.2 空间常识研究概述	16
2.2.1 空间本体库现状	16
2.2.2 汉语语言学本体库现状	17
2.2.3 文景转换中空间知识管理分析	19
2.2.4 空间知识管理需要解决的关键问题	23
2.3 空间本体库分类体系的定义	23
2.3.1 顶层分类原则与定义	24
2.3.2 底层分类定义	26
2.3.3 空间规则定义	31
2.4 空间本体库自动实例化方法	32
2.4.1 基于知网实例化方法	33
2.4.2 基于认知实例化方法	37
2.4.3 基于规则实例化方法	39
2.4.4 基于经验手工实例化方法	40
2.5 小规模空间本体库实例	41
2.6 本章小结	41
第 3 章 基于线性分类器空间关系抽取	43
3.1 引言	43

3.2 空间信息研究概述	44
3.2.1 语言学界空间表达研究概述	44
3.2.2 空间信息抽取方法与分析	45
3.3 基于线性分类器的空间关系抽取模型	47
3.3.1 SNoW 概述	47
3.3.2 空间关系抽取问题的形式化	47
3.4 空间表达式抽取	49
3.4.1 空间表达式边界抽取的特征空间	49
3.4.2 界标识别	50
3.4.3 介词和方位词抽取	51
3.5 射体抽取	51
3.5.1 空间描述策略	52
3.5.2 射体抽取的定义	55
3.5.3 射体抽取的特征空间	55
3.6 语料库的建设	56
3.6.1 人称代词消解	57
3.6.2 空间表达式标注	58
3.6.3 界标标注	59
3.6.4 射体标注	60
3.6.5 空间关系标注	61
3.7 实验以及分析	61
3.7.1 参数训练	62
3.7.2 评价指标	62
3.7.3 空间表达式抽取实验	63
3.7.4 射体抽取实验	67
3.7.5 空间关系抽取实验	70
3.8 本章小结	71
第 4 章 基于区域确定方法的三维物体自动摆放	72
4.1 引言	72
4.2 摆放技术的概述	73
4.2.1 摆放技术的主要方法	73
4.2.2 物体摆放需要解决的关键问题	76
4.3 基于空间本体库的摆放区域确定	77

4.3.1 空间语义消歧	77
4.3.2 方位词计算模型	78
4.4 基于 OMSGGA 的优化摆放	80
4.4.1 优化摆放问题编码	81
4.4.2 适应度函数定义	82
4.4.3 操作算子定义	83
4.5 基于受限语义的图形检索	84
4.5.1 检索词语和文件名的结构	84
4.5.2 词语相似度	85
4.6 实验以及分析	86
4.6.1 区域确定实验	86
4.6.2 优化摆放实验	86
4.6.3 图形检索实验	90
4.7 本章小结	91
第 5 章 基于自然语言描述的空间概念建模原型	93
5.1 引言	93
5.2 基于自然语言描述空间概念建模原型框架	94
5.3 空间概念建模过程的实现	96
5.4 空间概念建模场景的评价	97
5.4.1 TTS 系统的评价研究概述	97
5.4.2 基于 t 检验的评价模型	99
5.4.3 评价及分析	101
5.5 本章小结	105
结论	107
附录	119
攻读博士学位期间发表的论文	120
哈尔滨工业大学博士学位论文原创性声明	121
哈尔滨工业大学博士学位论文使用授权书	121
哈尔滨工业大学博士学位涉密论文管理	121
致谢	122
个人简历	123

Contents

Chinese abstract	I
English abstract	III
Chapter 1 Introduction	1
1.1 The background and significance	1
1.1.1 Background	1
1.1.2 Significance	2
1.2 A summary of relevant researches	3
1.2.1 A analysis of relevant researches	3
1.2.2 Researches on virtual agents systems	4
1.2.3 Researches on text-to-scene conversion systems	7
1.3 The content of the dissertation	13
Chapter 2 Creation of the spatial ontology	15
2.1 Introduction	15
2.2 State-of-the-art of spatial common knowledge	16
2.2.1 The actualities of spatial ontologies	16
2.2.2 The actualities of linguistics ontologies	17
2.2.3 Main approaches for spatial management in TTS	19
2.2.2 Key issues in spatial management in conceptual model	23
2.3 The categories for spatial ontology	23
2.3.1 The rules and definitions for top category	24
2.3.2 The definitions for low category	26
2.3.3 The definitions for spatial relations	31
2.4 Clustering-based spatial entities acquisition	32
2.4.1 Method of HowNet-based instantiation	33
2.4.2 Method of cognition-based instantiation	37
2.4.3 Method of rule-based instantiation	39
2.4.4 Method of experience-based instantiation	40
2.5 Samples in the small spatial ontology	41
2.6 Summary	41
Chapter 3 Linear classifier-based extraction of spatial relations	43

3.1 Introduction	43
3.2 State-of-the-art of spatial information research	44
3.2.1 State-of-the-art of spatial descriptions in linguistics	44
3.2.2 Main approaches of spatial information extraxtion	45
3.3 Linear classifier-based spatial relations extraction model	47
3.3.1 State-of-art of SNoW.....	47
3.3.2 The abstraction of spatial relations extraction problem.....	47
3.4 Spatial expressions extraction	49
3.4.1 Features space for boundary extraction of spatial expressions	49
3.4.2 Extraction of lamdmarks	50
3.4.3 Extraction of prepositions and localizers.....	51
3.5 Trajectories extraction	51
3.5.1 Sptially descriptive strategy	52
3.5.2 Abstraction of trajectories extraction	55
3.5.3 Features space for trajectories extraction	55
3.6 Creation of corpus.....	56
3.6.1 Pronominal anaphora resolution.....	57
3.6.2 Annotation of spatial expressions.....	58
3.6.3 Annotation of landmarks	58
3.6.4 Annotation of trajectories.....	60
3.6.5 Annotation of spatial relations	60
3.7 Experiment and discussion	61
3.7.1 Parameters training.....	61
3.7.2 Evaluation metrics.....	62
3.7.3 Experiment of sprial expressions extraction	62
3.7.4 Experiment of trajectories extraction	67
3.7.5 Experiment of sprial relations extraction.....	69
3.8 Summary	70
Chapter 4 3D object layout based on the qualification of the layout area. 71	
4.1 Introduction	71
4.2 State-of-art of object layout.....	72
4.2.1 Main approaches of object layout.....	72
4.2.2 Key issues for object layout	75

4.3 Spatial ontology-based object layout boundaries acquisition	76
4.3.1 Spatial semantics disambiguation.....	76
4.3.2 Localizers computable models	77
4.4 OMSGGA-based object layout	79
4.4.1 Coding for object layout.....	80
4.4.2 Definitions for fitness functions	81
4.4.2 Definitions for operators.....	82
4.5 Limited knowledge-based graphics retrieval	83
4.5.1 Structures of key words and file names	83
4.5.2 Definition of word similarity.....	84
4.6 Experiment and discussion	85
4.6.1 Experiment of layout boundaries acquisition	85
4.6.2 Experiment of object layout	85
4.6.3 Experiment of graphics retrieval	89
4.7 Summary	90
Chapter 5 A spatial conceptual model prototype based on natural language	
descriptions.....	92
5.1 Introduction	92
5.2 Framework of spatial conceptual model prototype	93
5.3 Implementation of model procedure	95
5.4 Evaluation on spatial coceptual model scenes	96
5.4.1 State-of-art of evaluation in TTS.....	96
5.4.2 Evaluation model based on t-test.....	98
5.4.3 Evaluation on model scenes	100
5.5 Summary	104
Conclusion	105
References.....	107
Appendix.....	117
The published or accepted papers written by the candidate	118
Acknowledgement	120
Resume	121

第1章 绪论

1.1 本文研究的背景和意义

1.1.1 研究背景

随着多媒体时代的到来,人机交流方式可视化、便捷化的需求迅猛发展,如何消除自然语言与图形场景自动转换障碍成为了二十一世纪的研究热点。从上个世纪 50 年代以来,自然语言处理和计算机图形学分别在各自的领域曲折中前进,各自形成了众多理论、方法和技术,但是它们之间自动转换技术仍然不能令人满意,仍有诸多难题需要解决,其中,空间概念建模是困扰文本到场景自动转换发展的一个基础问题。

概念建模(Conceptual Modeling)就是从高级描述(High-level Description)生成二维或者三维的场景,并使场景展现的景象和高级描述表达的情景一致。例如:把“乌鸦站在大树上”的描述生成一个包含“乌鸦”和“大树”图形,且乌鸦和树的空间关系是“上”的场景,如图 1-1所示。从多样性角度看,高级描述具有概念性(Conceptual)^[1]。这里的多样性是指高级描述所对应的场景不是唯一的,而是多个实例。高级描述包含自然语言和类自然语言。我们把类高级语言说明、脚本语言、以及高级任务规范说明等统称为类自然语言,以和真实的、自由的、超模糊的自然语言相区别。而空间概念建模强调的是,把高级描述表达的实体间空间关系生成能够展示这样空间关系的二维或者三维的场景。



图 1-1 “乌鸦站在大树上”的概念建模结果

Figure 1-1 The scene of “The crow is on the tree” on conceptual modeling

计算机学家,尤其人工智能研究者,关注于高级描述和场景之间的关系由来已久,但是多数高级描述着眼于类自然语言。虽然这种类自然语言描述

能够满足场景生成的要求，但是类自然语言描述是由领域人员编写的，不具有通用性。这些手工编写描述的方式消耗大量的人力、物力和财力，而且没有经过训练的人员是不能胜任的。随着自然语言处理研究的发展，人们逐渐认识到这种基于类自然语言建模的方式已经成为制约全自动场景生成研究发展的瓶颈。如何利用有效的手段从真实自然语言出发建模引起了国际学术界的热心关注。

二十世纪九十年代以来，计算语言学(Computational Linguistics)迅速发展令人兴奋。人们可以对自然语言进行词性标注、句法分析及语义框架标注等处理，发展了信息检索、信息抽取等技术。人工智能专家充分注意到这些成果，并且积极用于概念建模中，实现基于自然语言的场景生成系统(Language-based 3D Scene Generation)，例如，九十年代初开始的，由陆汝钫院士率领开发的“天鹅”系统，实现中文故事到动画片，全自动计算机辅助动画自动生成^[2]；另一个雄心勃勃的系统，是AT&T实验室开发的WordsEye系统，其能够根据文本的简单描述生成静态的三维场景^[3]；瑞典Lund大学研制的 Carsim系统，是专门恢复交通事故报告描述场景的平台^[4]。与基于类自然语言概念建模不同，这些系统都能够一定程度的理解自然语言，具有句法分析、依存关系分析、实体关系抽取等功能，在一定程度上，直接处理解决了自然语言的歧义性、部分性。

然而，到目前为止，概念建模研究仍然不具有普遍性，只有 Carsim 系统所基于的描述是真实文本，但也局限于交通事故领域，而绝大多数的描述仍然只是属于自然语言子集合。同时，相对来说，基于汉语自然语言描述生成场景的研究还不是十分广泛。本文面向现代汉语，在真实自然语言描述文本基础上，研究实体之间空间关系三维建模技术，探索篇章级空间关系抽取、场景摆放的理论和方法。本研究的开展将有助于在空间建模这个文景转换研究中的基础问题上谋求突破，其研究成果将直接为基于自然语言的、全自动的、三维的、动态场景生成等研究提供坚实的基础。

1.1.2 研究意义

二维、三维场景对读者理解事物有不容置疑的帮助。正像我们经常注意到的那样，用图形、图像比用文字更加容易解释清楚物理现象、数学理论以及任何类型系统的结构，并且这一点已经在各类教科书中得到验证。然而，不容忽视，手工建立场景的工作是极其辛苦、枯燥，而且通常要求必须由有一定艺术基础的人员完成。概念建模是场景自动生成领域中最先锋的研究，

处理的对象是自然语言和类自然语言。空间概念是人类基本经验之一^[5, 6], 是自然语言概念体系的基础, 研究基于自然语言的空间概念建模技术, 对于自然语言理解和场景自动生成研究具有重大的理论意义和广泛的实用价值。

在理论方面, 本研究的开展将有助于在自动化空间概念建模这个基础问题上谋求突破。文景转换研究是一门与自然语言处理和计算机三维图形学相关的新兴交叉学科, 集中研究自然语言文本的实体关系抽取、故事情节理解、视点抽取, 以及三维场景的空间规划、路径规划、摄像机规划、声音规划等方法, 其目标是研究自然语言离散系统和场景连续系统之间的自动转换关系。面向文景转换的概念建模研究方法, 是理性主义方法和经验主义方法的有机结合, 对于认知语言学、自然语言理解、计算机动画和人工智能等许多领域研究都将产生重要影响和巨大的推动作用。

在应用方面, 随着多媒体计算机普及和网络化程度的日益提高, 人们对三维场景自动生成技术的需求越来越大。基于自然语言的空间概念建模的研究, 对于加快空间本体库的建设, 提高场景生成的自动化程度和智能性, 促进文景转换系统的实用化具有重要意义。这些研究同时对信息抽取、计算机辅助设计及增强现实等自然语言理解及计算机图形学应用提供重要的资源和技术支持。

1.2 相关研究综述

1.2.1 相关研究术语分析

目前, 从文献上看, 关于从文本生成场景的研究, 还没有相应的国际规范或者术语对其研究范畴、研究体系进行统一界定。与这个研究有关联的研究者在各自的文章中, 从各自领域角度出发, 使用各自认可的术语来描述其研究特点、目标以及范畴, 大致包含以下 8 个: 文景转换(Text-to-scene Conversion)^[3]、基于语言三维场景生成系统(Language-based 3D Scene Generation System)^[3]、全过程计算机辅助动画自动生成^[2]、自然语言故事到计算机动画翻译器^[2]、智能故事讲述系统(Intelligence Storytelling System)^[7, 8]、故事图片化(Story Picturing)^[9]、概念建模(Conceptual Modeling)^[1]、虚拟角色系统(Virtual Actor System)^[10, 11, 12]。

虽然这些术语从词语表面上看各有特色, 但是它们阐述的研究范畴是相同的。这些术语都是指从一种高级描述自动生成场景, 而且这个场景能够反

映高级描述的内容，其目的都是增强场景生成自动化和智能性。其中，高级描述包含：类高级语言描述、电影脚本、戏剧脚本、对话或者自然语言命令、故事、文本及报告等；场景分为：二维静态场景、二维动态场景、图片或者图像、三维静态场景、及三维动态场景。因此，上述术语阐述的研究范畴、研究体系是如出一辙的，都是属于人类控制虚拟世界领域的。

当然，这些术语阐述的研究所具有的明显特色是不容忽视的。从交互的角度看，可以分为两个类型，一个是虚拟角色系统，在已有的虚拟环境中，使用高级描述控制场景；另一个是文景转换系统，其场景是在高级描述驱动下，从无到有构建出来。从虚拟世界角度看，可以分为静态场景和动态场景两大类型，静态场景多是以生成空间关系为主要表达内容的场景系统；动态场景多是生成全方位表达内容的场景系统。内容包含空间规划、情节规划、动作、事件等等。另外，还可以从输入分析它们之间的差别，一种是以高级描述为输入；另一种是多媒介输入，包含文本、语音等。

另外，这些术语暗示了该类别研究是多个学科的交叉研究。其涉及的学科、方向有：人工智能、语言学、认知语言学、计算语言学、可视化研究等^[13]，其中最突出的是，自然语言处理研究和计算机图形学研究在虚拟世界人类控制领域的交叉。具体研究内容和系统的概述，更能佐证这个研究是多学科交叉的观点。

1.2.2 虚拟角色系统发展概况

虚拟角色系统是指，其通过高级描述，在已存在的虚拟环境中，控制智能人的行为，完成空间定位、特定任务的动画模拟等。

早期的虚拟角色系统主要集中对微观世界的控制。典型的系统有：二十世纪七十年代初的，麻省理工大学人工智能实验室的SHRDLU系统^[14, 15]，其可以通过智能人和用户的简单对话，控制积木块的空间摆放，图 1-2给出其使用的对话片段和控制的场景。这个期间的另一个典型代表CLOWNS系统^[16]，它能够处理关于木偶的简单的空间描述以及他们的动作，但只能生成句子中动词的简单动画，能够处理的文本是受限格式书写的英文，例如“A clown holding a pole balances on his head in a boat”。同期，Kehn在他的博士论文中也提到了，可以从故事生成动画的系统，实现微观世界物体的空间变化的动态场景^[17]。

二十世纪九十年代以来，研究者开始把目光由微观世界转移到宏观世界，控制更加贴近客观世界的场景。瑞典Lund大学主持开发的Ulysse系统^[18]，

^{19]}，它是一个使用户可以用自然语言，在虚拟现实环境中，漫游浏览的智能体，而且能和语音识别系统交互，主要包括语法分析器、语义模块、语言到虚拟实体的映射器、几何推理器等系统模块，使用规划规则来构造动画。微软的Persona项目^[20]，它致力于开发拟人格化的动画角色，这个角色能够和用户进行自然语音对话，同时有动作相配合，该项目基于该公司开发的连续语音识别系统和英语理解系统。



(a) 70 年代的场景界面

(b) 近期颜色渲染后场景界面

(a) Original screen display

(b) Later color rendering (Univ. of Utah)

Person: PICK UP A BIG RED BLOCK.

Computer: OK. (does it)

Person: GRASP THE PYRAMID.

Computer: I DON'T UNDERSTAND WHICH PYRAMID YOU MEAN.

Person: FIND A BLOCK WHICH IS TALLER THAN THE ONE YOU ARE HOLDING AND PUT IT INTO THE BOX.

Computer: BY "IT", I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I AM HOLDING.

Computer: OK. (does it)

图 1-2 SHRDLU 系统的对话和场景

Figure 1-2 The dialog and the scene of SHRDLU

美国海军研究室致力于研究的Nautilus系统(Navy AUTomated Intelligent Language Understanding System)^[21, 22]，其拥有面向浸入式虚拟现实环境的自然语言和语音理解界面，通过使用口头命令，充分利用语言学的定量、定性及指代特性，对对话进行逻辑分析，用户可以在模拟的三维环境中漫游、移动或隐藏虚拟物体、控制模拟回放等，这些功能的实现依赖于包含有关虚拟世界和物体信息的知识库。宾夕法尼亚大学人体建模中心研发许多基于指令的动画生成系统^[23, 24, 25]，其中AnimNL(Animation from Natural Language Instructions)系统的目标是，生成真实的动画，表现人执行任务的过程，这些任务是由自然语言指令描述，如图 1-3所示。它的功能流程包括自然语言语法分析、语义分析、规划推理、规划、模拟和人物动画，所生成的动画不仅

包括人的一般肢体动作，还有人的面部表情、手势、和语音合成，提供了各种知识库以支持各部分的功能。

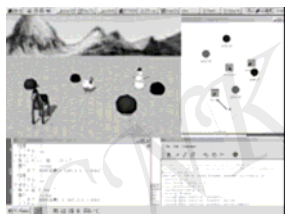


执行脚本：“If a military vehicle enters the checkpoint, open the gate,”

图 1-3 宾夕法尼亚大学研发的虚拟训练环境

Figure 1-3 The virtual environment for training by Upenn

东京工业大学研发的Kairai系统^[10, 12, 26]，研究者明确指出它是一个虚拟角色系统，能够理解日语自然语言口头命令，并且通过软件机器人，在三维虚拟环境中，展示其对任务的理解，完成向前移动、翻转、推动物体的动作，利用上下文信息，重点解决了物体参照问题，图 1-4是系统界面和执行指令。



- (1) “Niwatori wa sono kyu wo hidari kara osite.”
(小鸡，推你右侧的球.)
- (2) “Mousukosi osite.”
(再推远点.)
- (3) “Yukidaruma mo sore wo osite.”
(雪人，你也推它.)

图 1-4 Kairai 的系统界面和接受指令

Figure 1-4 The screen shot and input sentences of “Kairai”

对于虚拟角色系统来说，重要的研究问题就是，如何识别用户要求的动作以及关心的物体。因为，场景已经初始存在，所以系统不必关心初始信息获取，这些信息反而为场景的演化提供了充分的、确定的约束信息，而且系统又把任务执行对象聚焦在虚拟角色上，这样大大避开了自然语言不确定性、歧义性带来的许多难题。近些年来，使用高级描述直接生成场景的研究引起了国内外学者的广泛关注。

1.2.3 文景转换系统发展概况

“文景转换”这个术语最早在文献[3]中使用，其是指从自然语言自动生成场景，而且这个场景能够表达自然语言，其目的是使建模更加容易。在本文中，我们把具有从高级描述直接生成场景的研究都统一称为文景转换。从场景角度，这类系统包含以下两类：从高级描述到图片系统和从高级描述到三维场景系统。

1.2.3.1 从高级描述到图片系统 从高级描述到图片系统，主要是把高级描述表达的内容，生成静态的、二维图片组成的场景。相关研究多着眼于图片和其说明文字之间的语义相似度，用于图片检索、标注等方面^[27, 28]。

真正实现系统的研究并不是很多，文献[9] 阐述的“故事图片化”(Story Picturing)系统是目前较成型的系统。故事图片化就是用几幅有代表性的图片描述一篇文本传达的事件、活动和观点的过程。首先，对文本处理，识别确定的描述关键字和专有名词，基于WordNet计算关键字之间的词语相似度，用于图片初始筛选；然后，设定图片等级(Image Ranking)，把候选图片集合看成一个图，每幅图片是一个节点，图片间的相似度作为边的权重；然后，基于互加强策略选取图片；最后按照顺序组成故事图示。其中，图片库带有手工标志。该文用图片展现其系统效果，如图 1-5所示，对于故事Vermont，(1)、(2)和(3)是系统获取最好的图片，(4)、(5)和(6)是系统获取最差图片。

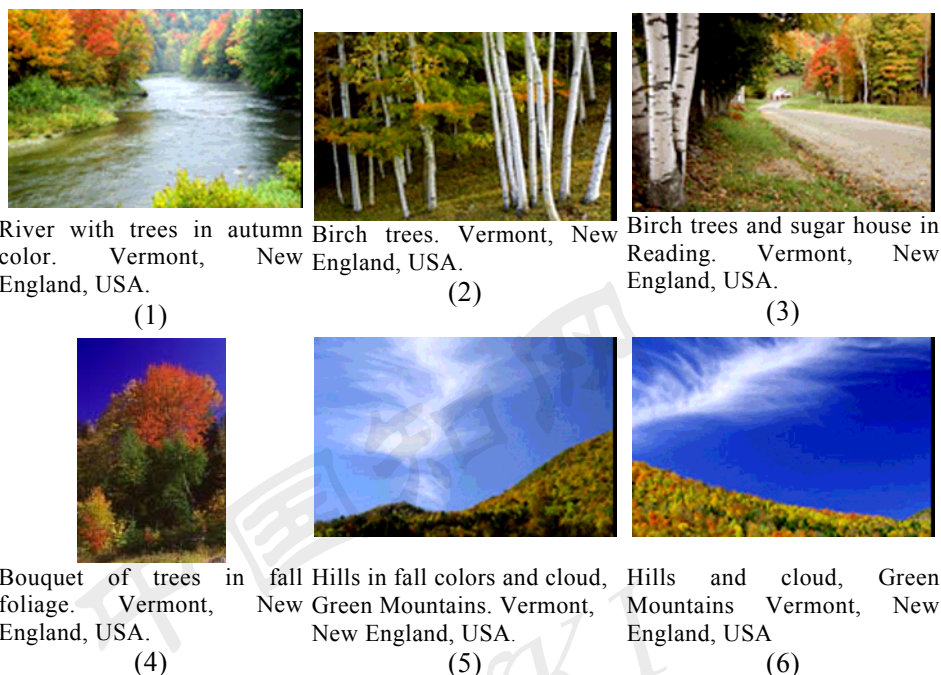
这个方面的研究成果典型的应用是为新闻、报道、游记以及说明等匹配图片插图。但是这样的系统目前强烈的依赖大规模的标注图片库，手工标注图片不仅成本很高，而且主观倾向性极强，这在很大程度上制约了从高级描述到图片系统的规模化、实用化。

1.2.3.2 从高级描述到三维场景系统 从高级描述到三维场景系统与从高级描述到图片系统的主要区别在于，前者生成的场景是三维的，而且在某些研究中是动态的。

到目前为止，对于这类系统的研究有两个方面的倾向，一个方面是把大部分研究重点放置于场景建模部分研究，极少或根本不关心高级描述的模糊性，文献[29]把这个方面的研究称为陈述式建模(Declarative Modeling)；另一个方面是同时兼顾高级描述理解和场景建模两个方面的研究，这里把这个方面的研究明确称为基于语言的场景生成系统(Language-based 3D Scene Generation System)。

(1) 陈述式建模 陈述式建模系统大体包含三个模块：描述单元。任务是

描述问题，使用约束描述场景，使用一种高级描述，例如类Java符号语言、结构化文档等^[30]，表达物体之间的关系来阐述约束；生成单元。任务是生成一个或者多个和描述匹配的场景；显示单元。任务就是把得到的场景可视化。



Vermont. Vermont is mostly a rural state. Many of the towns and villages, churches, covered bridges, weathered farms and barns date back to the 17th century, where the ancestors of many Americans lived before moving west. The countryside has the cozy feeling of a place which has been lived in for a long time, before the age of the machines. Each autumn, the landscape is transformed in a vibrant palette of colors that few places can rival, for the Green Mountains have a great variety of trees, many of which turn red or orange. (by:Q-T. Luong)

图 1-5 故事图片化系统得到的与 Vermont 故事匹配的图片

Figure 1-5 Pictures provided by the Story Picturing Engine to illustrate the Vermont

生成单元是这个方向的最核心问题，是系统性能表现的关键所在。其实质是在约束下寻找策略，属于NP-难问题。因此不能在多项式时间内被精确地解决^[31, 32, 33]。目前，用于生成的方法可以被分为两大类：构造方法(Constructive Approach)和迭代方法(Iterative Improvement)^[34]。构造方法是根椐一系列条目完成一个摆放。实质是枚举过程，具有确定性和完整性。对于处理低约束问题十分有效，耗时短，但是容易导致不令人满意的摆放；而迭代方法开始于一个初始摆放布局，然后在以降低代价为前提下，重复修正摆

放,直至代价不再改变为止。对复杂情况下物体摆放有效、灵活,缺点是耗时。

早期的构造方法通常都是基于原始规则,例如Fukunaga等人在各自的文献中提到的方法^[35,36,37,38,39,40,41]。典型过程是,选择一个种子物体摆放在区域内,然后其它物体按照和种子物体某种关系的关联度排序,而且按照这个顺序被选择出来,放在与已经摆放物体邻近的空白区域内。近年来,又出现了诸如分割摆放(Placement by Partitioning)以及定向力摆放(Force-directed Technique)、数字最优化(Numeric Optimization)^[1,42,43,44]等具体的构造方法,建模效果有所提高,同时耗时也增加了。分割摆放方法实质是把摆放区域重复不断的沿着不同方向分割,直至得到的子区域最小。定向力摆放是指计算每个物体应该处于的位置,且这个位置是最理想的。数字最优化方法是把物体摆放问题抽象成一个求方程组最小值和特征值计算的问题,一般使用线性或者非线性的方程来解决最小问题。其只有在极其严格的、精确的约束存在的前提下,才能保证得到结果。

迭代方法也被称为元启发方法(Meta-heuristics),属于随机方法类别(Stochastic Approach)。最简单的迭代改进方法就是交换随机选择的物体对的位置,如果新的位置降低了代价,那么这个位置就被接受了,直至代价不再改变,算法停止^[45,46]。目前最流行的迭代算法包含模拟退火(Simulated Annealing)^[47]、禁忌搜索(Tabu Search)以及进化论策略(Evolutionist Strategies)^[48,49]等。模拟退火算法使用 $(-\Delta C/T)$ 的比值作为摆放的接受概率, ΔC 是代价变化值, T 是称为温度的参数。起初被设置为很高的值,以保证大多数结果被接受;然后 T 呈梯度下降以保证使代价升高的结果被接受的概率减少。其保证算法找到全局最优,而不是像最简单迭代方法那样轻易陷入局部最优,此方法的缺点就是耗时很长,参数不好控制,也没有固定规则可循。

(2) 基于语言的场景生成系统 其实质目标是用图形、场景真正地取代语言。在没有文字说明情况下,完全可以胜任事件、故事、概念等交流的任务。

到目前为止,国内外已经有几个基于语言的场景生成原型系统。意大利Genoa大学研究的NALIG系统(NATURAL Language driven Image Generation)^[50]是我们关注到的早期系统。它能处理用意大利语描述的物体间的空间关系的短语。基于一流的定性空间推理理论,生成由积木组成的微观三维世界。短语的格式是: <主语> <介词> <宾语> [<参照>]。但是其应用局限于微观世界。Silicon公司和加利福尼亚大学联合研究的Put系统是一个直接使用自然语言生成物体空间摆放场景的系统^[51]。其研究者受到认知语言学中意象图示理论

(Image Scheme)的启发,把少数的空间关系,“in”、“on”、“at”,用有限的环境变量参数化,能够很好地区别‘on’(“上”)在“on the wall”(“墙上”)和“on the table”(“桌子上”)不同的空间布局。但是,它只关心存在物体的空间布局,能够接受的输入是属于英语的人工子集合,包含的表达式形如: $Put(X P Y)+$, 其中 X 和 Y 是物体名称, P 是空间介词。这两个系统都对物体空间摆放进行了研究,但是都没有脱离对输入格式严格要求的限制。

1993年日本京都大学YAMADA在博士论文中阐述的系统,以及印度IIT机器人研究中心的研究对输入格式限制有一定的突破,两者都可以接受较真实的自然语言文本。YAMADA阐述的系统^[44],可以把一个日语句子中描述的空间关系,通过空间信息抽取,势函数技术,显示在三维场景中。其中空间信息抽取技术是面向句子的,基于手工构建的辞典以及规则库,包含空间表达式单元、修改器和修改条件等。得到的空间信息作为约束信息,使用势函数仿真空间关系,多类势能函数的叠加实现一个具体的空间摆放,势能函数最小值,即势能为零的点就是实体的空间坐标,图 1-6显示了其势能函数。印度IIT机器人研究中心的研究^[52],也是采取连续区域,椭球域函数定量化空间区域。其研究不仅可以实现空间关系和动作建模,同时可以实现基于轴的形状建模(Axis-based Shape Modeling),也就是物体模型可以自动匹配文本描述。而且其研究者明确说明,类似于自然语言的描述,从可以有多个不确定场景对应的角度看,其具有概念性,可以统一称为概念描述,生成场景过程称为概念建模。这两个研究可以从自然语言生成三维场景,这是比早期系统进步的地方,但是,对模糊函数的严格要求,以及选取十分困难是制约它们发展的难题,而且不能精确的消歧空间语义。

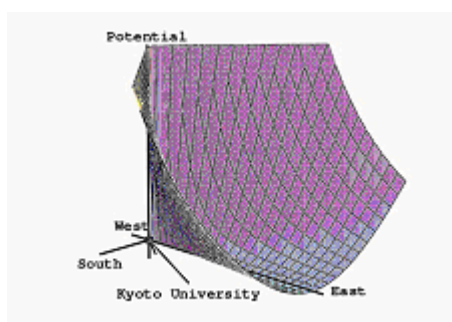


图 1-6 空间关系 East(Ginkakuji Temple, Kyoto University)的势模型

Figure 1-6 Potential energy function of East(Ginkakuji Temple, Kyoto University)

WordsEye、天鹅、CarSim和Seanchai/Confucius系统是目前这类系统中最

令人振奋的。AT&T实验室研发的WordsEye系统目前已经成为美国Semantic Light公司的主打产品。不像目前其他的文景转换系统，该系统的研究者对它未来应用的设想是史无前例的，它将被用于“母语和二外之间的指导师、电子贺卡、虚拟聊天、故事图是生成、游戏、烹饪指导或者产品安装等特殊领域”^[53]。它有一个含有几千个三维图形的模型库，能够根据文本的简单描述生成静态的三维场景。这个系统的流程是，首先，输入文本被词性标注、基于头驱动统计模型的句法分析，转换成依存结构；然后，依存结构被解释为语义表示(Semantic Representation)。其中，根据语言学家的空间介词语义研究，手工编写语义函数(Semantic Functions)，此函数在依存结构的帮助下，得到空间关系，并把此关系表示成为一个语义表示片段；其次，把语义表示转换为低级三维描述器，这些描述器表示了物体、姿势、空间关系以及其它属性；最后，根据描述器生成场景。图 1-7所示就是一个典型的转换。这个系统能够有这样的效果，大大得益于大规模的图形库，其中每个图形都附带各类标记，如使用长方体模型指明的空间标记，“上”、“下”等，如图 1-8所示，同时包含大量的纹理、姿势和动作。系统所处理的文本大部分是由一系列物体摆放的空间描述组成的，而且是研究者自己制造的，也就是说该系统不打算处理真实文本，而是打算做一个自然语言用户界面的三维建模工具。这一点可以在网站<http://www.wordseye.com>提供的在线软件得到证实。



“The mug is on the table. The picture of Rembrandt is on the left of the wall”

图 1-7 WordsEye 处理的文本和产生的场景

Figure 1-7 The input and its corresponding scene of WordsEye

天鹅系统是由我国中科院陆汝钤院士带领研究的另一个目标远大的系统^[54]。其技术路线是，将人工智能技术和基于知识的方法引进动画生成的全过程，是一种新的计算机动画技术——全过程计算机辅助动画制作。这实质上是一条动画片生产自动化的技术路线，它的目标是只要有了一个适当的故

事，以受限自然语言的形式把它输入计算机里，从此时开始，直到最终生成动画，每一步都是在计算机辅助下完成的。大致包含 5 个步骤：对故事文本作自然语言理解。故事文本是受限自然语言，句子的结构主要为“主语+谓语+宾语”的形式；使用自然语言单句理解系统对其分析，包含基于最大匹配法的分词、基于CATN的语法分析、格框架树表示语义分析；基于常识的指代分析和常识检查；故事情节理解。提出角色图的浅方法、基于高维和上下文有关文法或者角色动态特征的深方法；把故事改编为分场景剧本；根据分场景剧本作动画设计，包括角色、背景、动作等设计，其中体现了时间、空间规划；根据上述设计和规划，利用事先构造好的动画素材库和声音库，生成完整的动画。这个技术为计算机动画自动化的实现提供了保证，但是其中涉及的大量的常识库，包含情节常识库、摄影机常识库、是一般研究者望尘莫及的。



图 1-8 带有空间标记的模型

Figure 1-8 The models with spatial tags

隆德大学开发的CarSim系统是专门用于恢复交通事故报告描述的动态场景的平台^[4]。该系统能够处理的文本是交通领域的真实事故报告。包含三个部分：自然语言处理模块。首先，对文本进行浅层句法分析，然后基于决策树和启发规则识别实体、指代消解以及识别时间时序，最后，填写中间表示结构；时空规划和推理模块。基于符号表示得到充足的图形描述，符号表示包含四个概念类别：物体、事件、关系和定量值以及环境，其具体槽值由交通安全专业人员填写，其实质是一个领域本体库；最后是图形可视化模块。该技术研究者明确提出系统评价方案，用标准偏差定量衡量评价。该技术路线在交通安全领域，有很高的实用价值，但是其依赖领域本体库的特性导致移植性较差。

Seanchai/Confucius是一个智能的故事讲述系统^[8]。该系统的研究者似乎有更加雄伟的目标，要把他们的表示根治于语义和认知理论。然而，从相关的文献只能得到该系统能够处理的文本的几个例子，而且没有给出结果，同

时可以看出它能够解释清楚的文本似乎是由只包含一个事件的句子组成的。

综上所述,从人类控制虚拟世界研究角度看,自然语言处理和计算图形学两个似乎不太相干的研究领域找到了切实的交叉点。目前,专注于这类研究的系统,或者说技术方法,可以从交互的角度分为两类,虚拟角色系统和文景转换系统。不论哪类系统,不论生成动态还是静态场景,上面讨论的系统、研究技术路线中,都一致性的谈到物体空间关系建模研究,同时认知语言学家理论也清楚阐述空间概念是人类认知的基本概念。因此,空间关系是自然语言描述阐述的最基本关系,是理解其它各类关系的基础,研究和探索基于自然语言描述的空间概念建模方法,有望在空间关系场景生成这个基础问题上谋求突破。

1.3 本文的主要研究内容

本文研究的核心内容是:以小规模空间本体库为基础,以线性分类器和遗传算法混合为策略,实现自然语言文本中空间关系抽取、空间物体最佳摆放,最终生成三维静态场景。同时,鉴于目前没有发现针对本研究的公共测试集合,本文提出空间信息标注方案。要求语料库文本侧重描述真实空间关系,而非抽象或者隐喻空间关系,把中文版《伊索寓言》作为本文实验用语料库。最后,本文还将针对文景转换系统,初步探讨基于 t 检验的目测评价体系。本文研究力求达到,从真实文本中获取空间关系,生成与其匹配的三维场景布局,解决文景转换中空间建模基本问题,为其它概念建模提供有价值的研究基础。

本文内容包括三个主要部分:空间本体库的构建、空间概念建模和基于自然语言描述的空间概念建模原型系统。其中,空间本体库的构建为本文研究提供资源保障,是保证空间信息由定性到定量转换的基础。本文进行了基于 SUMO 的空间本体库体系确定,以及基于半自动方法实例化空间本体的研究(第 2 章);空间概念建模是本文研究的目的。第一、研究空间关系抽取(第 3 章),实现无结构空间信息到有结构空间信息的转换,标注语料库《伊索寓言》432 篇;第二、研究空间物体摆放(第 4 章)。搭建自然语言离散信息和图形系统连续信息的桥梁,确定物体精确摆放方位,生成可视化三维场景;最后,根据空间概念建模技术路线,实现了一个基于自然语言描述的空间概念建模原型系统(第 5 章),提出基于 t 检验的目测评价体系,验证了利用本文方

法空间概念建模的有效性。

具体的讲，各章研究内容如下：

第一章，首先介绍了本文研究背景，探讨了研究意义，然后对从高级描述到场景的研究和发展概况进行了简单地回顾，着重介绍了该研究术语使用现状，综述了虚拟角色系统和文景转换系统的技术路线和研究方法，最后概述了本文的主要内容。

第二章，简单回顾了现有本体库构建的主要方法，分析了空间本体库构建需要解决的关键问题。针对过多依赖手工建立本体库的问题，提出了常识和聚类方法相结合的半自动化本体实例化解决策略。在语言学知识库基础上，使用聚类算法有效地解决了汉语空间概念归类问题。

第三章，简单回顾了语言学界对空间概念的研究成果，讨论了目前信息抽取的特点，提出了基于线形分类器的空间关系抽取方法、空间关系标注方案。以语言学研究为基础，提出汉语空间关系形式化描述，把空间关系抽取抽象为二值分类问题，使用层次法实现空间关系抽取。该方法避开了汉语句法分析的难题，很好的解决了篇章级别的、认知层面的空间概念识别。基于一致通过原则，提出空间关系标注方案，并且实际在中文版《伊索寓言》中标注空间表达式、界标以及射体，这个方案适用于所有侧重描述物体空间关系，而非抽象或者隐喻空间关系的文本。

第四章，讨论了目前物体摆放方法的特点，分析了需要解决的关键问题，提出基于区域确定实现空间物体摆放的方法，并且在 ContonaSDK®辅助下实现 3 维可视化。区域确定是利用空间本体库的知识，把空间关系抽取得到的结构化离散空间信息定量化、连续化，得到定量约束信息，确定摆放区域的边界值，达到缩小下面步骤搜索空间的目的。然后，基于最优保持简单遗传算法，在有限约束下，实现物体的精确定位，完成物体优化摆放。这个方法为物体摆放提供了新的研究思路，达到了保证甚至加深任务难度的同时，很好地完成物体摆放。

第五章，在空间概念建模技术路线指导下，实现了一个基于自然语言描述的空间概念建模原型系统，给出了系统的基本框架，介绍了建模过程的详细流程及空间信息数据结构。最后在讨论目前文景转换系统评测方案的基础上，提出基于 t 检验的目测评价体系，并且使用这个策略验证了利用空间概念建模的有效性和可行性。

第2章 空间本体库构建的研究

2.1 引言

在空间概念建模系统中,有效、自动地组织以及管理世界常识,对于弥补自然语言表达本身固有的模糊性、部分性,提升场景的可接受性和缩短文景转换系统构建的周期具有重要意义。因此,将空间描述所涉及的自然现象层面的空间常识以及人类认知层面的空间常识进行组织管理是实现管理世界常识的一个基本层面。空间本体是指包含这两个知识层面的、从本体角度对空间本质的表述,包含具体空间物体建模、维数、形状和空间关系等等^[59]。空间本体库构建就是研究空间本体的组织管理。其中,本体是面向具体问题集合、拥有概念定义和它们相互关系的可计算资源^[55, 56]。

一直以来,由于真正从事和关注文景转换研究的研究者在国内外总计不过数人,因此相应研究也没有形成热点。但是,如果扩大应用领域,那么该研究得到很多领域研究者的一定程度重视。主要着眼点集中在空间常识内容、空间常识表示,对于空间常识自动获取技术只是初步涉猎。这几个方面比较而言,空间常识表示的研究最为深入,已经在各自的系统中发挥了较基础性的作用。本体表示是当前的研究热点。目前研究多是以认知主义(Cognitivist)、非现实(Nonrealist)主义^[57, 58]为基点考虑空间表示体系,方法也主要集中在基于集合理论(Set theory)和基于部分整体学(Mereology)的标准本体建模方法。近年来,空间常识内容的研究主要分别集中在两个方面,一是从本体论的角度来考虑空间、空间中的位置以及空间中的运动等内容,另一个是从语言学的角度来分析具体自然语言中的空间语言的语义等内容。空间常识内容的侧重直接涉及到该类本体的具体应用领域。第一类内容的本体主要应用的领域包含认知建模、人工智能以及地理信息系统等^[59],而第二类内容的本体则主要应用在自然语言生成、自然语言理解以及分析等领域。世界常识自动获取一直以来是知识工程领域的难题,目前,对于空间常识自动获取的研究也只是起步阶段。有的研究者们正在尝试从本体融合角度入手研究自动获取涉及的一些子方面问题。

本章在已有的空间本体和语言本体的基础上研究面向空间概念建模的

领域空间本体。针对自然语言空间描述和三维场景空间布局之间的背景知识缺乏的现状，综合利用已有空间常识本体和机器学习方法构建受限领域空间本体。重点讨论了空间本体实例化相关问题的解决策略。

本章第 2 节首先介绍了现有的空间本体和语言本体的体系以及文景转换系统中空间常识的组织，分析了领域空间本体需要解决的关键问题；第 3 节提出基于 SUMO(Suggested Upper Mergerd Ontology)的领域空间本体模型。领域空间本体提供了自然语言的语义知识、空间认知的定量知识以及实体的图形知识，是从自然语言空间描述到三维场景空间布局的桥梁。第 4 节提出半自动化领域空间本体实例化方法。第 5 节给出空间本体库中的一个具体实例。机器学习方法使构建成本降低，使对权威的依赖度降低，提高实例化的效率。

2.2 空间常识研究概述

所谓空间常识研究，是指从本体的角度，为问题解决、空间推理、空间数据检索、空间可视化、一致性检查、空间导航、普遍意义路径规划以及解释人类认知能力等问题，提供明确的、空间上适当的说明^[59]。到目前为止，人工智能研究者和语言学家从各自领域研究观点出发，构建了一些具体的，管理空间常识的本体库，可以分别称为空间本体库(Spatial Ontology)和语言学本体库(Linguistic Ontology)。这两类本体库都是从较普遍意义出发存在的，针对于文景转换这个具体研究方向，研究者结合具体问题，对空间常识进行了组织管理。

2.2.1 空间本体库现状

在目前大多数的大型本体库中，都包含对空间本体的组织管理。从它们应用角度看，主要从人工智能角度出发，以普遍意义为应用对象。

空间本体库从空间物体、空间属性以及空间关系三个方面具体组织管理空间本体。目前这个类别本体库的典型实例有：SUMO、OpenCyc 以及 DOLCE。

SUMO是当前被IEEE标准上层本体工作组考虑的本体库，是一个 3D本体库^[60]，包含物体实体、整体部分拓扑、空间关系和空间领域属性、函数的各种分类体系。它的空间关联实体的思想根源于Casati和Varzi的洞理论(Theory of Holes)^[61]，以及Smith和Guarino的区别整体部分拓扑学(Respective Meretopologies)^[62, 63]。它根植于多种理论的特性，使其可以作为空间领域本

体规划的很好参照。

OpenCyc是美国Cyc项目的开放资源版本^[64]，号称是世界上最大、最完成的常识推理系统。在1.0版本中就包含6000类别、关系和6000个实体。SpatialThing是它最基本的空间分类，表示在一些内嵌空间或者和其他空间事物相关的，具有空间扩展和位置关系属性的所有事物的集合，下面包含了表示空间区域、空间特征及空间关系的子分类。有些研究者认为它并没有考虑到专业知识和常识的界限，只是把收集来的大批技术报告整本整本地录进知识库中，这种做法并不能提高人们对常识性知识本质的认识，也难以使常识性知识库真正达到实用的地步。事实上，我们至今未见到CYC实际应用的报告^[65]。

DOLCE(The Descriptive Ontology for Linguistic and Cognitive Engineering)的目标是为语义网络建立一个基础的本体库^[66, 67]。相对而言，它把空间常识非常深入地整合在基本公理系统中，这些公理是物体实体和事件整体的描述。还给出基本的实体及位置关系，使用物理实体、特征和空间依赖表达具体的空间关系。“上”、“下”等传统空间关系属于空间区域的属性。同时，这些定义都是开放的。该本体库被一些研究者视为很好的具体空间本体库设计的基础。

2.2.2 汉语语言学本体库现状

目前，针对汉语，应用较广泛的，能够表达汉语词语语义的本体库典型实例有知网、同义词词林和同义词词林(扩展版)。

2.2.2.1 知网概述 知网(英文名称为HowNet)是一个以汉语和英语的词语所代表的概念为描述对象，以揭示概念与概念之间以及概念所具有的属性之间的关系为基本内容的常识知识库^[68]。知网作为一个知识系统，是一种网状结构，它所着力要反映的是概念的共性和个性。同时知网还着力要反映概念之间和概念的属性之间的各种关系。为了实现这样的目标，知网2005版本主要利用了以下基本概念：概念、义原、记录以及关系。

其中，“概念”是对词汇语义的一种描述，每一个词可以表达为几个概念。“概念”是用一种“知识表示语言”来描述的。这种“知识表示语言”所用的“词汇”叫做“义原”。在知网中，“义原”是用于描述一个“概念”的最小意义单位。

从表2-1中就可以看出来中文词语总数是81062，而概念总数是24089个，说明概念阐述的是词语的意义。而概念是抽象的，在知网中使用多级网

状结构来具体阐述概念。首先在知识词典中利用义原、动态角色和语用特征，借助记录这个载体来描述概念。然后再利用关系把义原联系起来，构成网状结构。

表 2-1 知网的一般统计

Table 2-1 The Statics on HowNet 2005

中文词语数	81062
英文词语数	76526
中文义项数	95690
英文义项数	95935
概念总数	24089
记录总数	156442

表 2-2给出了知网中定义的一个概念实例。对于“部分”一词，分别给出该词的汉语词性、汉语的例子、对应的英语词、英语的词性、英语的例子、关于该词语的语义标注等信息。其中NO.为顺序的概念编号，W_C为汉语的词语，G_C表示汉语的词性，E_C是部分汉语常用搭配的例子；W_E是英语的词语，G_E表示英语的词性，E_E是部分英语常用搭配的例子；DEF是知网对于该概念的利用义原而作的语义定义，这里“部分”这个词的义原是“部分”，动态角色是“整体(whole)”。

表 2-2 知网对概念的定义示例

Table 2-2 A definition of a Chinese word in Hownet

项目	示例
编号	NO.=012064
汉语词	W_C=部分
汉语词性	G_C=N [bu4 fen1]
汉语例子	E_C=身体各~, 这个~, 某~, 其中一~, 哪~更重要些
英语词	W_E=component
英语词性	G_E=N
英语例子	E_E=
语义定义	DEF={component 部分:whole={entity 实体}}

2005 版本中共有 2347 个义原，并且被分为 5 类义原，包含事件类义原(812 个)、实体类义原(151 个)、属性类义原(247 个)、属性值类义原(889 个)

以及专有名词义原(248个)。

关系反映了义原之间的联系,构成了知网的网状结构的脉络。2005版本中关系主要分为4大类:动态角色和特征(113个)、同义关系(484个)、反义关系(224个)以及第二特征(121个)。具体的关系说明请查阅知网2005版本的说明文档,此处不再赘述。

2.2.2.2 同义词词林概述 《同义词词林》是现代汉语比较常用的一部义类词典^[69]。它所收词语全部按语义分类编排,共分为12大类,94中类,1428小类。它描述了一个从上到下,由宽泛概念到具体词义的语义分类体系,并将所收的词分门别类组织在其中。在词条中用语义代码来表示这种语义分类体系。例如,“农民”的词义代码为Ae07,表示其处于A大类、e中类、07小类。

2.2.2.3 同义词词林扩展版概述 为了弥补、扩充《同义词词林》,哈尔滨工业大学信息检索实验室利用很多词语相关资源,并投入了大量的人力和物力,完成了一部具有汉语大词表的《同义词词林(扩展版)》^[70]。最终的词表包含77,343条词语。为了使用上的方便,对于词群、词义相关词---以下称第四级和第五级---的分类也编码。新增的第四级和第五级的编码与原有的三级编码和并构成一个完整的编码,唯一地代表词典中出现的词语。例如,此时的“农民”的词义代码是Ae07A01。

《同义词词林(扩展版)》的编码是,前三级编码同《同义词词林》相同。第四级用大写英文字母表示,第五级用二位十进制整数表示。由于第五级的分类结果需要特别说明,例如,有的行是同义词,有的行是相关词,有的行只有一个词,可以分出具体的三种情况,分别代表这三种情形。具体的标记参见表2-3。

表 2-3 《同义词词林(扩展版)》词语编码表

Table 2-3 The coding rules on Tongyici Cilin (extension edition)

编码位	1	2	3	4	5	6	7	8
符号举例	D	a	1	5	B	0	2	= \ # \ @
符号性质	大类	中类	小类	词群	原子词群			
级别	第1级	第2级	第3级	第4级	第5级			

2.2.3 文景转换中空间知识管理分析

基于语言的三维场景生成系统的关键问题就是处理语言的模糊性,例

如，针对句子“乌鸦站在树上”进行文景转换的时候，必须决定放置乌鸦的位置，但是在语言中没有关于“乌鸦”在树的哪个具体位置的明确信息，这是符号系统的显著特点。然而，每个人都有这样的常识，两个人之间用语言交流，想象或者画出来上面这个句子表示的场景，是一点障碍都不存在的，究其原因在于，人类具有管理分析空间常识的能力。因此，在目前的文景转换系统中，对这类空间知识都进行了管理，弥补符号系统和连续系统之间的模糊与精确的沟壑。现有的空间常识管理方法主要集中在知识库管理，根据知识库的具体形式可以分为：基于认知理论的空间知识库管理、基于常识的空间知识库管理和基于标记的空间知识库管理。

(1) 基于认知理论的空间知识库管理。Clay等人在Put系统中提出基于意象图示(Image Schemas)理论的空间知识管理的方法^[51]。这个方法是文景转换研究中一个很典型思路。意象图示是由Langacker提出的^[71]，是基于人类和物理世界交互的模型的一些结构，这些模式包含身体运动、与物体的交互以及空间朝向等。其中，最基本的意象图示是与空间关系和运动相连的。Clay等人在Put系统中，使用意象图示的很多研究成果，在相关的意象图示上执行空间摆放操作。

对于物体摆放最基本的意象图示有：坐标系统(Coordinate System)、中心(Center)、容器(Containment)以及接触面(Contact)。Clay首先定义了一些空间关系的具体意象图示结构，如表 2-4所示；然后，认为“in”、“at”及“on”三个空间概念时是大多数空间关系的来源，是任何摆放系统的基础；最后，在这个观点下，主要依照意象图示完成这三个空间关系摆放。例如：对于空间关系“in”，使用容器和接触面意象图示，射体被放置在界标内部底面中心点上，具体做法是射体的最低的、与世界坐标系相关的表面和界标的最低的、与世界坐标系相关的表面接触。另外，基于物体维数(Dimensionality of Objects)，根据缺省行为(Default Placement Behaviours)消歧空间语义。首先，物体维数由用户自定义，同时缺省行为由系统预先定义，例如：消歧“on”空间语义，缺省行为包含有对于 3D射体，物体被放置在界标垂直相关的表面等等类似的行为说明；然后，系统可以根据目前物体的维数选择合适的行为。

由于对认知理论的依赖，空间知识的管理与分析是一个十分专业的过程，需要用户对认知理论以及计算机图形学一些术语有所掌握，这样限制了其发展。

表 2-4 Put 系统的空间关系意象图示

Table 2-4 The image schemas of spatial relationships in Put

空间关系	坐标系系统	物体明确信息	附加的意象图示
in	世界, 物体	中心、底部、容器	接触面、向上(up)/向下(down)
on	世界, 观察者, 物体	焦点面	接触面、向上(up)/向下(down)
at	世界, 物体	位置(Location)	中心
above/below	世界	顶(top)或者底(bottom)	向上(up)/向下(down), 内(in), 上(on)
next-to	世界, 观察者, 物体	尺寸(size), 前(front), 侧(sides)	内(in), 上(on)
.....			

(2) 基于常识的空间知识库管理。从人工智能角度, 陆汝钤院士提出基于常识的计算机动画技术^[2]。所谓常识, 是相对于专业知识而言的, 是存在于地地道道的日常生活中的, 包含自然现象中的常识和社会生活中的常识^[65]。该方法就是把文景转换中需要的常识, 按照不同方面组织成各个常识知识库。由于这个方法提出的目的是, 服务于全过程计算机辅助动画自动生成研究, 所以空间常识与其它常识交织在一起存在, 主要由情节知识库表达。在情节知识库中, 针对空间知识的核心概念是基本场景, 它们是从那些出现在电影或计算机动画中的基本场景总结出来的, 核心内容包括主要角色的数量、这些角色的空间分布以及他们彼此之间的行为动作等。基本场景分为两层: 定量层和定性层, 定量层是指, 主要角色和次要角色数量, 按模糊及相对原则处理, 确定(一个或多个)注意中心; 定性层是指角色的空间分布和彼此之间的基本行为, 指出各注意中心的相对位置及其变动。

定量层有下列八种模式: **Solo**: 只有一个角色; **Double**: 两个角色并发操作; **Few**: 多于两个角色, 但不是太多(比如说少于 10 个); **Few-Times-Few**: 几小组人, 每组为 **Solo**, **Double** 或 **Few**; **Many**: 一大群人; **Few-Plus-Many**: 一小组人加一大组人; **Few-Times-Many**: (不多的)几个大组; **Many-Times-Few**: 许多小组。

对于每一种定量层, 都有几种定性层来描述人物的静态分布。例如, 与定量层 **double** 相对应的定性层则有 **facing**(两人面对面)、**siding**(两人并排)、

following(一先一后)、collecting(两人相聚)、separating(两人分开)等等。double-facing(张三,李四)是两人面对面基本场景实例。

在这样的基本场景知识基础上,首先基于启发式规则,确定人物的定性空间分布;然后,依据人物、物体和环境的几何特征、基于路径规划算法,把定性位置转化为定量位置,包含出场位置和运动终结位置。其中的规则以及特征都是常识知识库的组成部分。

Nugues 等人在 CarSim 系统中使用的方法也是这个空间常识管理方法的另一个实例。Nugues 在交通安全专家帮助下,使用面向对象框架知识表示,具有继承层次,组成一个本体库,其中槽关系和定量表示了物体空间信息,相关文献指出“Behind”和“FromLeft”是空间关系的例子,有预先定义的属性值。从目前的文献中,没有看到关于这个本体库的更详细的介绍。

可以看出,基于常识的空间知识管理方法是十分有效的,能够很好的弥补语言的模糊性,赋予系统丰富的背景知识。但是,目前依赖专家建设知识库的手段导致了其局限性。

(3) 基于标记的空间知识库管理。Sproat等人提出的使用图形标记表达空间概念的方法令人耳目一新^[3]。该方法的核心是,在每个3D模型上附加空间标记(Spatial Tags)。空间标记的目的是描述空间关系,具体是用一个简单的能够填充空间的3D物体,表示一个空间区域。空间标记包含:天篷区域(Canopy Area)、上表面(Top Surface)、底部(Base)、杯状内部(Cup)、墙(Wall)、盖扣(Cap)、包围区域(Enclosed-area)、背脊(Ridge)、顶点(Peak),例如:图 1-8 用一个长方体确切的表达了天篷区域和上表面空间标记;图 2-1中的长方体表达的分别是底部和杯状内部空间标记,其中杯状内部可以被用于描述“in”和“on”空间关系。

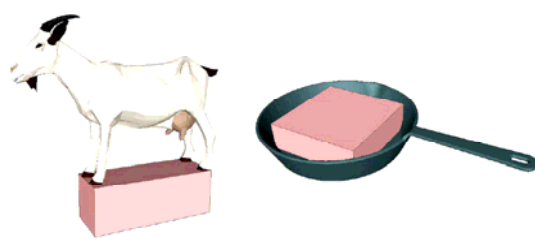


图 2-1 底部(Base)和杯状内部(Cup)空间标记

Figure 2-1 Spatial tags for “base” and “cup”

基于标记的空间知识管理的方法的优点是十分明确的,能够十分精确的解决语言的模糊性以及空间概念的歧异性,描述任意的物体空间关系,达到

目前最好的视觉效果。但是缺点也是明显的，空间标记是需要手工附加的，这意味着有多少个 3D 模型，就需要至少一倍的工作量添加标记，而这个工作要求专业软件和专业人员支持。

2.2.4 空间知识管理需要解决的关键问题

通过上述分析，本文认为在空间概念建模中，对于空间知识管理需要解决如下关键问题：1) 认知理论和常识的有机结合问题；2) 空间常识的具体定义问题；3) 空间常识的自动获取问题。基于以上认识，本章利用聚类、规则等方法和语言学知识相结合的方法构建空间本体库。聚类等方法可以脱离手工操作，能够自动学习空间知识，为空间知识定义提供依据。但是聚类等方法需要先验知识，而基于词典和语言学知识的方法正好为其提供知识资源。因此，综合利用多种方法有望得到更好的空间知识管理方法。

2.3 空间本体库分类体系的定义

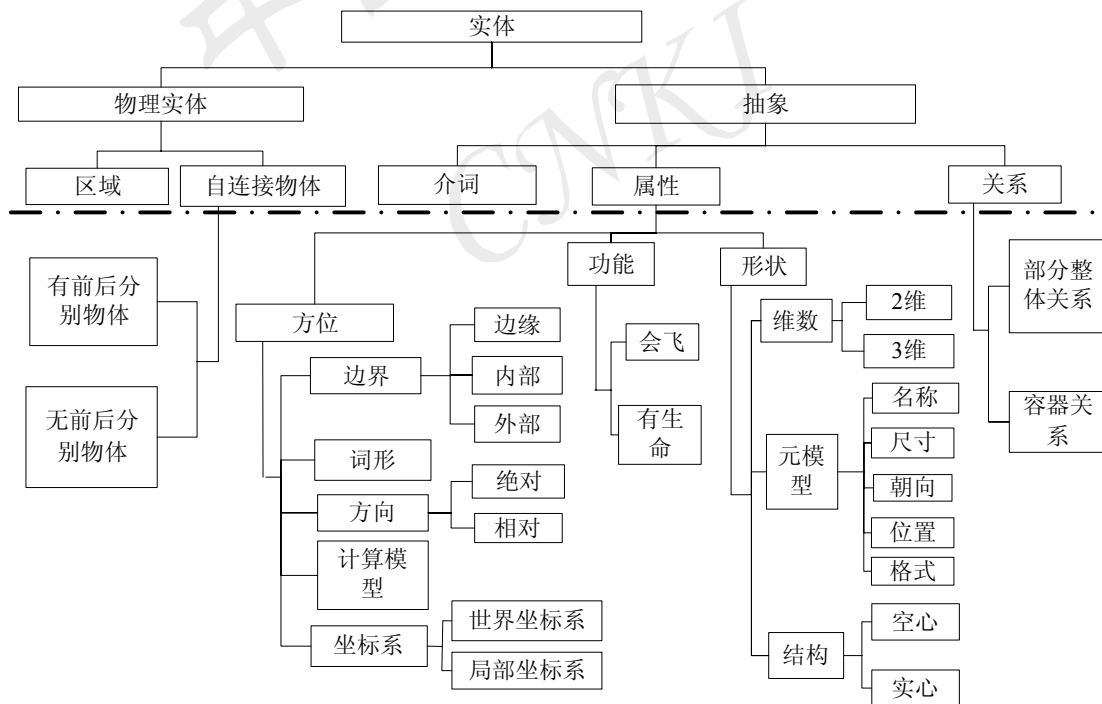


图 2-2 空间本体库的分类体系

Figure 2-2 The spatial ontology categories

本文构建了一个领域空间本体库，包含了空间关系可视化所需要的概念和关系。这个本体库是基于集合论思想，在SUMO本体库基础上建立的，首先，顶层分为两大类，一类是实体模型知识，另一类是实体空间常识，然后再分别细化。空间本体库的分类体系结构如图 2-2所示，虚线以上部分是顶层分类部分，以下部分是底层分类部分。

下面具体给出空间本体库分类中的每个概念、关系的具体含义。

2.3.1 顶层分类原则与定义

借鉴 SUMO 本体库的顶层分类体系，结合空间可视化的具体领域，形成了空间本体库的顶层分类体系，其中包含三层，共 8 个节点。

2.3.1.1 SUMO本体库的顶层体系结构 SUMO是一个 3D本体库，描述了许多类别，与这些类别相关联的方面包含：物体实体、整体部分拓扑学、广泛的空间关系、以及大量的和空间领域相关的属性和功能。SUMO本体库的顶层分类体系显示了最基本分类之间的关系，目的是很好的分解实体，这些实体组成世界或者具体领域，具体体系参见文献^[59]。

2.3.1.2 空间本体库的顶层概念的定义 本文空间本体库的顶层概念的含义以及分类沿用SUMO的顶层相关概念的具体含义和分类。把SUMO的四层顶层结构缩减为本文的三层空间本体库顶层结构。表 2-5列出本文空间本体库和SUMO顶层概念术语名称和分类的区别。

具体概念含义阐述如下：

SUMO 最顶层的概念是 Entity(实体)，一定要用最大可能中立的方式理解它的含义，它是在一个领域中某物位置的持有者。另外，从形式化上看，它仅代表一个层次的顶层节点。Entity(实体)被分解为两个基本子类，Phycal(物理实体)和 Abstract(抽象实体)。Phycal(物理实体)是那些在时间、空间上必要存在的实体，相反，Abstract(抽象实体)是那些在时间、空间上不存在的实体。进而，从概念的原子级别定义的角度，Phycal(物理实体)分为 Object(物体)和 Process(过程)两类。根据这些含义，本文中的空间本体库包含 Process(过程)以外的其它基本类别，并且采用和 SUMO 本体库一样的定义。

Abstract(抽象实体)的含义由其四个子类别描述。这四个子类别是：Attribute(属性)、Relation(关系)、Propositon(介词)和 Quantity(数量)。SUMO用 Attribute(属性)这个概念看待空间的两个方面，形状和位置；在广泛意义上，Relation(关系)是空间的，这个类包含部分整体关系和拓扑关系，空间本体库在原含义的基础上使用这两个概念。对于 Propositon(介词)，因为本文只

考虑空间信息，所以空间本体库中的介词是指，那些能够对表达空间信息起作用的介词。另外，在空间本体库中剔出了 Quantity(数量)概念。

SUMO本体库中的空间实体是物理世界部分，被Phycycal(物理实体)子类描述。很显然，在Phycycal(物理实体)包含的子节点中，Region(领域)和SelfConnectedObject(自连接物体)两个节点和空间概念紧密联系，而其它两个节点，Collection(聚集)和Agent(代理)，和本文的研究内容相距比较远，因此，本文中的空间本体库只包含前两个节点。在空间本体库中，Region(区域)和SelfConnectedObject(自连接物体)两个节点的含义与SUMO一致，Region(区域)是指地志学上的位置和空间，例如亚洲、天王星；SelfConnectedObject(自连接物体)的含义是指，其实例的组成部分在部分整体逻辑学意义下连接，例如，使整个物体分解的所有部分都是相连的^[61]。

表 2-5 SUMO 和本文空间本体库顶层概念区分表

Table 2-5 The difference between SUMO and spatial ontology in the top level

第一层概念	第二层概念	第三层概念	第四层概念	SUMO	本文
Entity(实体)	Physical(物理实体)	Process(过程)		√	
		Object(物体)	Region(区域)	√	√
			SelfConnectedObject(自连接物体)	√	√
			Collection(聚集)	√	
			Agent(代理)	√	
	Abstract(抽象)	Proposition(介词)		√	√
		Attribute(属性)		√	√
		Relation(关系)		√	√
		Quantity(数量)		√	

总之，顶层分类采用了SUMO本体库中相应概念，并且沿用了具体含义，但是，它们的子类别和SUMO不完全一致，详细情况在第2.3.2节介绍。

2.3.2 底层分类定义

空间本体库中，底层分类体系，也就是从图 2-2第四层开始的所有层次，不再沿用SUMO的体系，而是结合本研究定义。第三层节点中的自连接物体、关系和属性三个节点进行了再分类。而区域和介词这两个第三层节点不包含子节点，直接利用实例来说明其具体含义。

首先，以表格形式分别给出底层概念和分类的定义；然后，在各个小节具体阐述这样定义概念和分类的依据。

2.3.2.1 自连接物体节点 根据相对方向可视化的具体要求，自连接物体节点由两个子节点描述：有前后概念的物体和无前后概念的物体，具体定义见表 2-6。

例如：“一个男孩”就是有前后概念的物体，而对于“一块石头”，当没有上下文的时候，从普遍意义上讲，它就是一个没有前后概念的物体。对于这样的两个句子，“男孩的前面有棵树”和“石头的前面有只鸟”，要把它们描述的空间方位分别可视化，那么，对于第一个句子，很容易确定“前面”的具体方向，即男孩面向的方向，但是对于第二个句子，只是单纯的从这个句子得到“前面”所指的具体方向是困难的，原因在于，“石头”是没有前后概念的实体，需要借助于方位认知等常识。因此从这个角度把自连接物体分为这样的两个类别。

表 2-6 自连接物体节点定义和分类体系

Table 2-6 The definitions and classifier of connection-self node

概念名称	父概念	子概念	概念定义	正例	反例
有前后概念的物体	自连接物体		物体的本身具有前后之分	一个男孩	一块石头
无前后概念的物体	自连接物体		物体的本身不具有前后之分	一块石头	一个男孩

2.3.2.2 关系节点 关系概念从空间方位的角度考虑物理实体之间的联系，这个节点被分为部分整体节点、容器节点。具体定义见表 2-7。

部分整体关系对于确定物理实体的具体空间位置起到很重要的作用。例如“翅膀”是一个独立的物体实体，同时是“乌鸦”这个独立的物理实体的一个组成部分，因此，“翅膀”和“乌鸦”的关系就是部分整体关系。

容器节点关系可以用于消歧“里”、“外”等方位概念。例如：“飞机”这个独立的物理实体可以作为“男孩”、“行李”等独立物体的容器，也就是

可以把他们装起来。

表 2-7 关系节点定义和分类体

Table 2-7 The definitions and classifier of relation node

概念名称	父概念	子概念	概念定义	正例	反例
部分整体	关系		一个物理实体是否可以作为另一个物理实体的组成部分	“翅膀”和“乌鸦”	“桌子”和“椅子”
容器	关系		一个物理实体是否可以作为另一个物理实体的容器	“行李”和“飞机”	“鼠标”和“显示器”

2.3.2.3 属性节点 本研究从形状、方位以及功能三个方面描述实体的空间属性，这三个子属性又进一步细分为多个子属性，下面分别定义。

表 2-8 形状节点定义和分类体系

Table 2-8 The definitions and classifier of shape node

概念名称	父概念	子概念	概念定义	正例	反例
形状	属性	维数、元模型、结构	描绘物理实体的集合造型		
维数	形状	2 维、3 维	从相对的角度看，物理实体的每个实例可以看成点、线、面和体四类，即 0 维、1 维、2 维和 3 维		
元模型	形状	名称、尺寸、朝向、位置、格式	一个物理实体唯一对应的一个基本几何模型	渔夫	蹲着的渔夫
结构	形状	实心、空心	从每个实例是空心的还是实心的角度来定义其结构		

形状节点 形状属性描绘了物理实体的几何造型，具体由元模型、结构和维数三个子节点说明。形状节点和其儿子节点定义见表 2-8。

提出元模型概念的原因在于，客观世界中的每个物理实体都对应一个或者多个具体的几何造型，显然这个数量是无穷的，是不可计算的。例如：“渔夫”这个实体只对应于一个基本几何模型，而不是同时有“蹲着的渔夫”、“打鱼

的渔夫”等多种造型，显然，后面的造型是基本几何模型的变形，利用图形坐标变换技术就可以自动实现它们之间造型的转换，因此，这个概念的提出有利于用有穷模型表示无穷物理实体，节省建模时间以及存储空间。

元模型的具体含义由名称、数据格式、尺寸、朝向和位置四个属性阐述，具体定义见表 2-9。

表 2-9 元模型节点定义和分类体系

Table 2-9 The definitions and classifier of meta-model node

概念名称	父概念	子概念	概念定义	正例	反例
名称	元模型		随机的从实体可以接受的名称中选择一个词语作为其名称	椅子	“凳子”和“椅子”同时存在
格式	元模型		有效描述元模型几何形状的数据格式	VRML	
尺寸	元模型		元模型在三维空间中的实际大小	乌鸦长、宽、高分别是 0.1 米、0.05 米、0.05 米	乌鸦长、宽、高分别是 2 米、1 米、1 米
朝向	元模型		对于有前后概念的元模型，朝向是指其正面的方向，对于无前后概念的元模型，我们规定它的一个面的方向为其朝向；		
位置	元模型		每个元模型的初始位置是预先设定的，在坐标原点	(0, 0, 0)	

在客观世界中，实体的名称是不唯一的，即每个物理实体的名称可以用多个词语描述，如果一一列举每个名称，显然不符合本体库的含义。在具体空间关系可视化中，本文 4.5 节利用语义相似度查找实体的名称。

在计算机中，需要采用一种有效的数据格式描述元模型的几何形状，这样的数据格式可以由研究驱动定义，也可以使用标准语言描述。目前，本研究采

用 VRML 通用建模语言描述元模型的数据格式。

在具体的空间中，使用尺寸、朝向和位置三个属性定位元模型。具体空间下的位置由具体的算法确定，这些算法在第4章中阐述。

结构节点是形状的另一子节点，用于计算物理实体每个实例的中心点，有空心、实心这两个子节点具体描述结构节点的概念。具体定义见表 2-10。

表 2-10 结构节点定义和分类体系

Table 2-10 The definitions and classifier of structure node

概念名称	父概念	子概念	概念定义	正例	反例
实心	结构		物体内部和外部是一体	石头	水瓶
空心	结构		物体内部是空的，可填充	水瓶	石头

维数节点是形状的第三个子节点。具体定义见表 2-11。“田野”、“天空”等实例从相对的角度，可以看成是 2 维物体，相对比“房子”就是三维物体。

表 2-11 维数节点定义和分类体系

Table 2-11 The definitions and classifier of dimension node

概念名称	父概念	子概念	概念定义	正例	反例
2 维	维数		从相对的角度看，可以被看成面的物理实体的每个实例	田野	房子
3 维	维数		从相对的角度看，可以被看成体的物理实体的每个实例	房子	田野

方位节点 方位概念包含两个方面的含义，即位置和方向。在汉语语言学中，方位词是一组既能表示方向，又能表示位置的一组体词^[72]，因此，空间本体库中，位置概念主要是通过方位词概念来描述。

现代汉语方位词根据表示的方向和位置的性质，可以分为词形、边界、方向、坐标系和计算模型五类，也就是，在空间本体库中用这五类描述方位词概念。方位节点和其儿子节点的定义见表 2-12。

方位词在表示位置的时候，除了“东、南、西、北”这类之外，必须相对于语境中的某个物体或者几何图形的位置而言，这个位置称为位置参照点。而方位词在表示方向的时候，必须选择一个位置作为面对的方向，这个位置称为方向参照点。词形是指一个具体的汉语方位词，例如：“上”。

把方位词表述的方向和位置用三维图形形象的表现，称方位可视化，其关

键问题是如何把自然语言表达的离散信息转换成图形系统下的连续信息。计算模型概念的目就是把方位词的离散性连续化。

表 2-12 方位节点定义和分类体系

Table 2-12 The definitions and classifier of location node

概念名称	父概念	子概念	概念定义	正例	反例
方位	属性	边界、词形、方向、计算模型、坐标系	通过汉语空间方位词具体阐述物体方向、位置		
词形	方位		具体的汉语方位词	上	那儿
计算模型	方位		把方位词的离散性连续化,方位可视化的基础	见表 2-21、表 2-22	
边界	方位	边缘、内部、外界	揭示方位词表述的位置和位置参照点物理边界的关系		
方向	方位	相对方向、绝对方向	揭示方向参照点的特性		
坐标系	方位	世界坐标系、局部坐标系	是具体方位量化的参照		

边界概念进一步由边缘、内部、外界三个概念描述。具体定义见表 2-13。例如：在下面的短语中，“房子里”、“房子外”以及“房子周围”，“里”表述的位置是房子的内部，即位于房子边界之内，反之，“外”描述房子边界之外的位置，而“周围”描述的位置和房子边界的距离很近。

表 2-13 边界节点定义和分类体系

Table 2-13 The definitions and classifier of boundary node

概念名称	父概念	子概念	概念定义	正例	反例
边缘	边界		揭示方位词表述和位置参照点边界内或外临近的位置	周围	前
内部	边界		揭示方位词表述和位置参照点边界内临近的位置	里	外、周围
外部	边界		揭示方位词表述和位置参照点边界内临近的位置	外	里、周围

方向概念揭示了方向参照点的特性，即当方向参考点是语境之外的太阳的起落点，那么这样的方位词表述的方向是绝对的，例如，方位词“东”；而当方向参照点由语境之内某个物体的位置决定，那么这样的方位词表述的方向是相对的，例如，方位词“前”。具体定义见表 2-14。

坐标系节点在方位概念中，具有确定方位词表述的方位作用。根据方位的性质，坐标系分为世界坐标系和局部坐标系两个子概念。具体定义见表 2-15。例如：方位词“东”表述的方位应该由世界坐标系限定，而方位词“上”则由局部坐标系限定。

表 2-14 方向节点定义和分类体系

Table 2-14 The definitions and classifier of direction node

概念名称	父概念	子概念	概念定义	正例	反例
相对方向	方向		当方向参照点由语境之内某个物体的位置决定时候，方位词表述的方向是相对的	前	东
绝对方向	方向		当方向参照点是太阳的起落点的时候，方位词表述的方向是绝对的	东	上

表 2-15 坐标系节点定义和分类体系

Table 2-15 The definitions and classifier of coordinates node

概念名称	父概念	子概念	概念定义	正例	反例
世界坐标系	坐标系		唯一的、固定的、以参照点为原点	全局坐标系	
局部坐标系	坐标系		不唯一、可变的、以被显示物体为原点		

功能节点 最后，功能概念从空间概念的角度考虑物理实体的功能，包含飞概念和生命概念。具体定义见表 2-16。

2.3.3 空间规则定义

规则是空间本体库的另一个组成部分，其作用是消歧概念或者的关系的意义，检验常识一致性。本体的固有特性是导致歧义、矛盾的重要原因。空间本体库中的各个概念、关系的含义不是完全排斥的，这是符合本体的含义的，然而，在上下文的情况下，某些概念、关系含义必须是排斥的，例如，方位词“上”，

从本体意义考虑，其即符合内部概念的意义，又符合外部概念的意义，但是，在具体的上下文中，如“火车上”，其只符合内部概念的含义，也可以说，此时其内部和外部这两个概念是矛盾的，如果同时满足，那么是不符合常识的，因此，此时需要消歧方位词“上”的含义。

目前，空间本体库中包含规则如下：

- (1) 如果一实例同时属于边缘概念和内部概念，那么在具体上下文中消歧其概念；
- (2) 如果一实例同时属于边缘概念和外部概念，那么在具体上下文中消歧其概念；
- (3) 如果一实例同时属于外部概念和内部概念，那么在具体上下文中消歧其概念；
- (4) 如果一实例同时属于二维概念和三维概念，那么在具体上下文中消歧其概念；
- (5) 如果一实例同时属于世界坐标系概念和局部坐标系概念，那么在具体上下文中消歧其概念。

其中的具体消歧策略将在4.3.1节阐述。

表 2-16 功能节点定义和分类体系

Table 2-16 The definitions and classifiers of function node

概念名称	父概念	子概念	概念定义	正例	反例
功能	属性	会飞、有生命	从空间概念角度考虑物理实体功能		
会飞	功能		物理实体能飞翔	乌鸦	兔子
有生命	功能		物理实体有生命	乌鸦	石头

2.4 空间本体库自动实例化方法

上面给出的空间本体库的分类体系结构相当于骨架，空间本体实例化技术就是要填充这个骨架，把领域中具体实体填充进去。本文中，根据概念、关系的固有特性，基于归并思想，首先，把空间本体库中的概念、关系分为四类；然后，提出相应的实例化方法：基于知网实例化、基于认知实例化、基于规则

实例化和基于经验手工实例化；最后，根据分类体系把实例化结果合并。

2.4.1 基于知网实例化方法

在空间本体库实例化中，针对词性、方向、边界和维数概念语义特性一致，把它们分为一类，根据知网的义原概念，计算实例间的语义距离，以最大限度实现实例的自动聚类。本文使用的聚类算法步骤描述如下：首先，设待聚类混合样本集合 $X = \{x_1, x_2, \dots, x_n\}$ ，以 n 表示 X 所含的样本数目；以 X_i 表示第 i 个类别；以 m_i 表示第 i 个类别对应的样本均值；以 n_i 表示第 i 个类别所含的样本数目；然后，具体聚类过程如算法 2-1所示。

算法 2-1:

- (1) 选择初始聚类中心 m_1, m_2, \dots, m_k 、样本集的初始划分 X_1, X_2, \dots, X_k ;
- (2) 若 $n \geq 1$ ，继续；反之，转(7)；
- (3) 取样本集 X 的第一个样本 x ，令 $n = n - 1$ ；
- (4) 计算 x 与划分 X_i 的语义距离 S_i ；
- (5) 根据 S_i 把 x 归类；
- (6) 转(2)；
- (7) 算法停止。

其中，语义距离 S_i 的计算过程如下：首先，使用知网 2005 提供概念相似度函数(2-1)，得到相似度 Sim_H ；然后，由公式(2-2)计算 x 与划分 X_i 的语义相似度 Sim_X ；最后，由公式 (2-3)得到 x 与划分 X_i 的语义距离 S_i 。

$$Sim_H(x, y) = HowNet_Get_Concept_Similarity(DWORD AdwUnitIDA, DWORD AdwUnitIDB, float AfA = 1.6, float AfB1 = 0.1, float AfB2 = 0.1, float AfB3 = 0.7, float AfB4 = 0.1)$$

$$(AdwUnitIDA: \text{记录 } x \text{ 的记录号}; AdwUnitIDB: \text{记录 } y \text{ 的记录号}) \quad (2-1)$$

$$Sim_X(x, X_i) = \frac{\sum_{j=1}^{n_i} Sim_H(x, x_{ij})}{n_i} \quad (x_{ij} \in X_i) \quad (2-2)$$

$$S_i = 1 / Sim_X(x, X_i) \quad (2-3)$$

由于初始聚类中心、概念相似度 Sim_H 的计算以及根据 S_i 把 x 归类的这 3 个步骤的具体计算过程是和具体聚类实例的语义密切相关的，因此，下面将结合具体实例聚类给出这 3 个步骤的具体实现过程。

2.4.1.1 词形体系实例化 本文以汉语语言学家研究为基础，辅以小样本统计，获取方位词词形概念的具体实例。

从现代汉语语言学的角度，文献[72]、[73]和[74]分别给出各自认可的方位词范围，我们把这些范围合并，并且去掉冗余，形成初始方位词集合。在语言学中，方位词表示的方位既包含空间上的，又包含时间上的，而在本文中只考虑能够表示空间上的方位的方位词，所以把只能表示时间上方位的方位词，包含“以上”、“以下”，从初始方位词集合中删除。但是，保留那些既能表示空间上的，又能表示时间上的方位词，例如“之前”等。至此，得到了语言学家认可的方位词词形概念的具体实例集合。

接下来，我们又考察了计算语言学家对方位词范围的界定。使用本实验室的词性标注工具，标注中文版《伊索寓言》，统计其中的所有被标注为方位词的词语，至此，得到了计算语言学家认可的方位词词形概念的具体实例集合。把这个集合和语言学家认可的集合进行比较发现，其包含形如“脚下”这样的词语，可以看出，这类词语既包含位置参考点，又包含方位词。由于本文中需要对位置参考点单独处理，因此，把这类词语从集合中剔除。同时发现，在语言学家认可的集合中没有出现“尾”，所以加入这个方位词。最后，得到方位词词形实例集合 S_{Loc} ，包含 194 个元素，其详细内容见附录 1。

2.4.1.2 方向概念体系实例化 在方位词词形实例集合 S_{Loc} 内，根据常识以及 2.3 节的概念定义，得到方向概念体系下的具体实例。对应于绝对方向概念的具体实例有 61 个，如表 2-17 所示，对应于相对概念的具体实例有 133 个，如表 2-18 所示。

2.4.1.3 边界体系实例化 根据统计，我们发现， S_{Loc} 集合中 45 个方位词没有被包含于知网 2005 版本中，如表 2-19 所示，其余 149 个既包含相对方向类别的实例(106 个)，又包含绝对方向类别的实例(43 个)。因此，基于层次法和 k-均值聚类相结合的模型，把附录 1 中的方位词聚类得到边界体系下的实例。

实例化模型步骤如下：

(1) 手工聚类绝对方位词。基于空间认知，把符合绝对方向概念的方位词，手工分类，分别归类到内部概念和外部概念实例集合中；

(2) 使用 k-均值聚类算法，聚类相对方向方位词。共有三个类别：外部类、内部类和外部类，分别以“里”、“外”和“附近”为种子，即初始聚类中心。使用算法 2-1 把相对方位词分成“里”、“外”、“附近”和“其它”四个子类别。属于其它类的所有实例，根据算法的定义，可知它们同时属于前三个类别，因此，把它们分别全部加入前三个类别。表 2-20 所示的聚类结果显示，此

方法较令人满意；

(3) 手工归类知网 2005 版本中不包含的 27 个相对方向方位词；

(4) 合并。把上面三个步骤得到的结果合并，那么就是位置体系下的具体实例集合。(由于篇幅，此处不再一一罗列实例)

表 2-17 绝对方向概念的具体实例

Table 2-17 The instants of the absolute direction

绝对方向概念的具体实例
东，西，南，北，东面，西面，南面，北面，东头，西头，南头，北头，之东，之西，之南，之北，东边，西边，南边，北边，以东，以西，以南，以北，南北，东部，西部，南部，北部，东端，西端，南端，北端，东侧，西侧，南侧，北侧，东南边，东北边，西南边，西北边，东南面，东北面，西南面，西北面，东南部，东北部，西南部，西北部，东南方，东北方，西南方，西北方，东南角，东北角，西南角，西北角，东南，东北，西南，西北，

表 2-18 相对方向概念的具体实例

Table 2-18 The instants of the relative direction

相对方向概念的具体实例
上，下，前，后，上头，下头，前头，后头，里头，外头，左，右，里，外，中，内，旁，之上，之下，之前，之后，之外，之内，之中，之间，上边，下边，前边，后边，左边，右边，里边，外边，上面，下面，前面，后面，左面，右面，里面，外面，以内，以外，上下，前后，左右，里外，内外，底下，顶上，头里，当中，当间，中间，内中，内里，面前，跟前，头前，背后，旁边，边上，面上，对面，这儿，这里，那儿，那里，以远，以近，附近，周围，上部，下部，前部，后部，中部，内部，左端，右端，上端，下端，前端，后端，左侧，右侧，外侧，内侧，左上头，左下头，右上头，右下头，左上部，左下部，右上部，右下部，左上方，左下方，右上方，右下方，前上方，后上方，左上角，左下角，右上角，右下角，一边，一面，一头，两边，两头，两旁，两旁边，四面，四边，四外，四周，前后左右，边，头，尾，面，顶，外边，侧，际，四围，四周围，中心，

表 2-19 没有包含于知网 2005 的方位词

Table 2-19 The localizers excluded from HowNet 2005

绝对方向类别的实例 (18 个)	相对方向类别的实例 (27 个)
东南边，东北边，西南边，西北边，东南面，东北面，西南面，西北面，东北部，西北部，东南方，东北方，西南方，西北方，之东，之西，之南，之北，	顶上，内中，头前，面上，上部，右端，里端，外端，上端，下端，前端，后端，左上头，左下头，右上头，右下头，左上部，左下部，右上部，右下部，前上方，后上方，两旁边，四边，两面，之上，之内

层次法第 2 个步骤中是使用算法 2-1 实现的聚类，其中涉及的概念相似度

Sim_H 的计算以及根据 S_i 把 x 归类的具体实现如下。

表 2-20 层次聚类结果

Table 2-20 The result of the layer classifier

边缘概念类实例	内部概念类实例	外部概念类实例	其它类实例
跟前, 以近, 附近, 旁, 旁边, 边上, 这儿, 这儿, 那儿, 那里, 以远, 周围, 一边, 一面, 两边, 两头, 两旁, 四面, 四外, 四周, 边, 侧, 四围, 四周围	里头, 里, 中, 内, 里边, 里面, 当中, 中间, 内里, 内部, 内侧, 际, 中心, 之中, 之间, 当间, 中部	外头, 外, 之外, 外边, 外面, 以外, 外侧, 外边	上, 下, 前, 后, 上头, 下头, 前头, 后头, 左, 右, 之下, 之前, 之后, 之中, 之间, 上边, 下边, 前边, 后边, 左边, 右边, 上面, 下面, 前面, 后面, 左面, 右面, 上下, 前后, 左右, 里外, 内外, 底下, 头里, 面前, 背后, 旁边, 边上, 对面, 这儿, 这儿, 那儿, 那儿, 以远, 下部, 前部, 后部, 中部, 左侧, 右侧, 左上方, 左下方, 右上方, 右下方, 左上角, 左下角, 右上角, 右下角, 前后左右, 头, 尾, 面, 顶

为了计算概念相似度, 我们提出根据义原概念消歧方位词词语语义。这里的语义消歧目的是获取表达空间概念的语义, 也就是在知网 2005 字典中得到描述空间概念的记录号。方位词词语有多种语义, 包含空间语义, 也包含其它语义, 例如: “上”有程度副词“高等”的修饰语义、有动词“做”的动作语义, 当然也有“位置”的空间语义, 在本文中需要得到其空间语义。语义消歧具体过程如算法 2-2所示。

算法 2-2:

- (1) 初始化空间方位义原集合 $SE = \{\text{空间、方位特性值、方向特性值、部件、界限、地方、距离值}\}$, 词语记录号集合 $R = \{r_1, r_2, \dots, r_n\}$ 空间语义记录号集合 $RS = \text{NULL}$;
- (2) 若 $n \geq 1$, 继续; 反之, 转(7);
- (3) 取记录号集 R 的第一个元素 r , 令 $n = n - 1$;
- (4) 利用知网 2005API^[75]计算 r 的上位义原集 H ;
- (5) 若 $H \cap SE \neq \text{NULL}$, 则将 r 移入 RS ;
- (6) 转(2);
- (7) 算法结束。

根据 S_i 把 x 归类的实现步骤是:

- (1) 若 $S_i = \min\{S_j\}, j = 1, 2, \dots, k$, 则将 x 移入 X_i 类中, 置其标志符 $\text{exact} = \text{TRUE}$, 并重新计算 m_i ;

(2) 若 $S_i = \forall S_j, j=1, 2, \dots, k$, 则将 x 移入 X_{k+1} 类中, 置其标志符 $\text{exact} = \text{FALSE}$;

2.4.1.4 维数概念实例化 由于维数这个概念具有可视化属性, 使用观测法可以直接得到实例, 但是这个方法时间复杂度太高, 耗时、耗力, 因此, 我们提出基于知网语义相似度, 使用聚类算法得到维数概念实例。聚类算法如算法 2-1所示, 其中, 初始聚类中心、根据 S_i 把 x 归类的这 2 个步骤的具体计算过程如下。

此处使用两步法选择种子, 即初始聚类中心。

第一步、基于知网语义相似度, 聚类实体名称。同样使用算法 2-1把《伊索寓言》中 430 个实体名词, 按照词语相似度, 得到多个划分。

第二步、基于观察, 选择种子。从物体维数角度观察, 划分可以被分为两类。一类是划分中的元素都有相同的维数, 例如“房子”、“舍”属于同一个划分元素, 而且它们都是三维物体; 另一类是划分中的元素在维数上没有共同的特点。因此, 在第一类划分中, 把其中都是三维物体的元素作为三维物体的种子, 组成三维种子集合 $xm_1 = \{ \text{“房子”、“蜂舍”、“渔夫”} \}$, 同时把都是二维物体的元素作为二维物体的种子, 组成二维种子集合 $xm_2 = \{ \text{“田野”、“海滩”} \}$ 。

根据 S_i 把 x 归类的具体方法是: 若 $S_i < \varepsilon$, 则将 x 移入 X_i 类中; 反之不归类, 进行下一个样本的聚类。 ε 表示语义距离阈值, 根据经验选取 3.3。

2.4.2 基于认知实例化方法

根据坐标系体系概念和计算模型体系概念的语义特性一致性, 即都和认知有密切的关系, 把它们归为一类, 且基于认知实现实例化。

2.4.2.1 坐标系体系实例化 根据坐标系概念下世界坐标系和局部坐标系两个子概念的含义, 从讲汉语的人群的认知角度看, 绝对方向概念下所有实例, 例如“东”, 都是参照太阳的起落点制定坐标系原点, 也就是说, 是以世界坐标系为参照系的; 而相对方向概念下的所有实例, 例如“前”, 则是把物体, 即位置参照物, 的中心为坐标系原点, 也就是说, 是以局部坐标系为参照系的。因此, 世界坐标系实例包含所有的绝对方向实例, 局部坐标系实例包含所有的相对方向的实例。

2.4.2.2 计算模型体系实例化 位置、方向等概念定性地描述了方位词的方向和位置。

但是, 空间关系可视化需要严格的定量分析, 而且, 定量分析每个方位词是不现实的。因此, 我们根据认知语言学家的理论, 即某些空间概念是其它空间概念的基础^[76], 以人类认知常识为前提, 提出基本计算模型和叠加策略结合的方法, 实现方位词的定量分析, 具体方法见4.3节。这里首先给出基本计算模

型。根据方位词概念体系，计算模型体系实例分为方向模型类、位置模型类两个类别。方向模型类包含“东”、“上”、“右”、“前”四个基本计算模型；位置模型类包含“里”一个基本计算模型。为了给出具体的计算模型公式，不妨规定：以界标的中心为坐标系原点；视点沿x轴负方向；界标和射体大小都用其包围盒近似表示，用 (lm, bm, hm) 、 (lt, bt, ht) 分别表示它们的长、宽、高；以 d 表示与界标外表面的距离； k 是系数，根据经验取值。方向模型类别的计算模型见表 2-21，以阴影区域表示空间。位置模型类别的计算模型见表 2-22。

表 2-21 方向基本计算模型

Table 2-21 The basic models on directional localizers

名称	模型	图示
“东”， <i>basic_east</i>	$basic_east = (y \leq 0)$ $d = k * ht / 2$	
“上”， <i>basic_on</i>	$basic_on = \begin{cases} x \leq \infty \\ y \leq \infty \\ z \geq 0 \end{cases}$ $d = k * ht / 2$	
“右”， <i>basic_right</i>	$basic_right = \begin{cases} x \leq \infty \\ y \leq 0 \\ z \leq \infty \end{cases}$ $d = k * ht / 2$	
“前”， <i>basic_front</i>	$basic_front = \begin{cases} x \leq 0 \\ y \leq \infty \\ z \leq \infty \end{cases}$ $d = k * ht / 2$	

表 2-22 位置基本计算模型

Table 2-22 The basic models on locationnl localizers

名称	模型	图示
“里” , <i>basic_in</i>	$basic_in = \begin{cases} x \leq \frac{lm}{2} \\ y \leq \frac{bm}{2} \\ z \leq \frac{hm}{2} \end{cases}$	

2.4.3 基于规则实例化方法

根据功能概念体系、自连接物体概念体系和关系概念体系三者概念特征规律可循的特性，把它们归为一类，且使用基于规则的实例化方法

2.4.3.1 功能概念体系实例化 功能概念体系中包含飞和生命两个子概念，我们提出，基于知网语义分类，利用一元逻辑获取这两个概念的实例。

逻辑公式如下：

- (1) $\exists y((y \in fs(x)) \wedge (y \leftrightarrow a)) \rightarrow LIFE(x)$
- (2) $LIFE(x) \wedge \exists y((y \in fs(x) \wedge (y \leftrightarrow b))) \wedge \neg z((z \in fs(x)) \wedge z \leftrightarrow d) \rightarrow FLY(x)$
- (3) $LIFE(x) \wedge \exists y((y \in fs(x) \wedge (y \leftrightarrow e))) \rightarrow FLY(x)$
- (4) $\exists y((y \in fs(x) \wedge (y \leftrightarrow c))) \rightarrow FLY(x)$

其中，*a*、*b*、*c*、*d*、*e* 是常量，在这里分别表示义原“animate|生物”、“bird|禽”、“aircraft|飞行器”、“domesticated|家养”和“fly|飞”；*x*、*y* 是变量，*x* 的域是空间关系可视化领域的所有的实体，*y* 的域是知网中的所有义原；*fs(x)* 是取得自变量 *x* 的所有义原以及它们的上位义原的函数；*LIFE* 和 *FLY* 两个谓词，一个用于判断是否有生命，另一个用于判断是否可以飞翔。

2.4.3.2 自连接物体概念实例化 自连接物体概念包含有前后概念物体和无前后概念物体两个子概念。

它们的含义是排斥的，因此，只需要确定前一个概念的实例。获取方法分为两个步骤，首先使用公式 $\exists y((y \in fs(x)) \wedge (y \leftrightarrow a)) \rightarrow BB(x)$ 获取有生命物体；然后对于无生命物体进行人工获取，例如：石头、显示器。其中，*a* 是常量，表示义原“animate|生物”；*x*、*y* 是变量，*x* 的域是空间关系可视化领域的所有的实体，*y* 的域是知网中的所有义原；*fs(x)* 是取得自变量 *x* 的所有义原以及它们的上位义原的函数；*BB* 用于判断是否有前后概念的物体的谓词。

2.4.3.3 关系概念体系实例化 关系概念包含两个子概念，部分整体、容器。基于知网的知识，结合人工获取，得到这个体系下的实例。

使用谓词逻辑公式 $\exists y((y \in fs(x)) \wedge (y \leftrightarrow a \vee y \leftrightarrow c)) \rightarrow PART(x)$ 和 $\neg PART(x) \rightarrow WHOLE(x)$ 得到部分整体实例。其中， a 、 c 是常量，分别表示义原“part|部件”、“component|部分”； x 、 y 是变量， x 的域是空间关系可视化领域的所有的实体， y 的域是知网中的所有义原； $fs(x)$ 是取自变量 x 的所有义原以及它们的上位义原的函数； $PART$ 和 $WHOLE$ 两个谓词，一个用于判断物体可以作为其它物体的部分，例如“翅膀”是“乌鸦”的一个组成部分，另一个用于判断物体是否是其它某些物体的整体，例如“乌鸦”可以作为“翅膀”和“爪”等的整体。

使用谓词逻辑 $\exists y((y \in fs(x)) \wedge (y \leftrightarrow a)) \wedge \exists z((z \in gd(x)) \wedge (z \leftrightarrow b)) \rightarrow CT(x)$ 得到容器关系实例。其中， a 、 b 是常量，分别表示义原“implement|器具”、动态角色“终处所 LocationFin”； x 、 y 是变量， x 的域是空间关系可视化领域的所有的实体， y 的域是知网中的所有义原； $fs(x)$ 是取自变量 x 的所有义原以及它们的上位义原的函数， $gd(x)$ 是取自变量 x 的所有动态角色； CT 用于判断物体是否可以作为其它物体的容器的谓词。

这个公式可以识别《伊索寓言》中的 70% 左右实体，其余不能被这个方法处理的实体，一部分是没有包含在知网中，例如：“跑马场”；一部分需要人工获取，例如：“家”。

2.4.4 基于经验手工实例化方法

由于结构、元模型以及介词三个概念具有或者对应实例有限、或者视觉判断更加简单、直接的特性，因此把三者归为一类，且基于经验手工实例化。

2.4.4.1 结构概念实例化 本文人工获取实体的结构特性，即空心还是实心。

2.4.4.2 元模型概念实例化 元模型概念包含名称、数据域、尺寸、朝向、位置 5 个子概念。

VRML 图形的命名以图形的内容为准，例如：若图形做的是一个老虎，则图形的名字为“老虎.wrl”。图形以汉字命名，不包含英文。

本文采用 VRML 虚拟建模语言描述元模型。目前，一部分元模型来源网络，一部分是自行构建，数据结构符合 VRML2.0 标准。

在本文中，实体的大小是用其最小包围盒(Bounding Box)来表示的。根据 VRML 场景的坐标系结构，实体在 (x,y,z) 三维坐标轴上的最大长度分别代表了实体的长、宽、高，这也就是实体最小包围盒的大小。模型大小与现实世界中的

相应的实际物体大小一致。


元模型的初始位置是模型的中心在 VRML 坐标系原点处，其朝向 x 轴正方向。

2.4.4.3 介词概念实例化 在汉语中能够出现在空间表达式中的方位介词词形以及数量都是确定的。文献^[1]和^[2]都从语言学的角度分别给出了这类词语的具体界定，集合 $S_{Pre} = \{\text{在, 到, 从, 于, 向, 朝, 往, 自, 至, 由, 沿着}\}$ ，在本文的空间本体库中介词概念实例都自来这个集合。

2.5 小规模空间本体库实例

使用以上阐述的方法，以中文版《伊索寓言》为基础，构建了一个小规模空间本体库。在2.4.1节和2.4.4.3节已经详细说明了这个空间本体库中的方位词和介词体系下的实例，因此，这里不再重复。其它体系下的实例，我们从具体物体的角度，以面向对象单层框架作为知识表示，把430个实体的本体分别描述，图2-3给出一个具体实例。这些实例显示本文空间本体库构建方法的有效性，其具体的应用将在下面两章具体阐述。

Name=乌鸦	//元模型名称
Dim=3D	//维数 3D
Fly_Attribute=fly	//功能会飞
Life_Attribute=life	//功能有生命
Solid_Attribute=solid	//结构实心
Center.x =0 , Center.y =0 , Center.z =0	//元模型位置
Front_Attribute =true	//有前后分别物体
Entire_or_Part=entire	//整体物体
Container_Attribute=false	//不能做容器
Related_OrienWord =(前,)	//关联方位词
towards.x=0 ,towards.y=0 ,towards.z=1	//元模型朝向
size.x=0.1 ,size.y=0.05 ,size.z=0.05	//元模型尺寸



//元模型 VRML 格式

图 2-3 空间本体库中实体实例

Figure 2-3 The sample on the object in the spatial ontology

2.6 本章小结

本章在回顾了典型的空间本体和语言本体的体系，以及文景转换系统中的空间知识管理方法的基础上，分析了空间概念建模中空间知识管理的难点。针

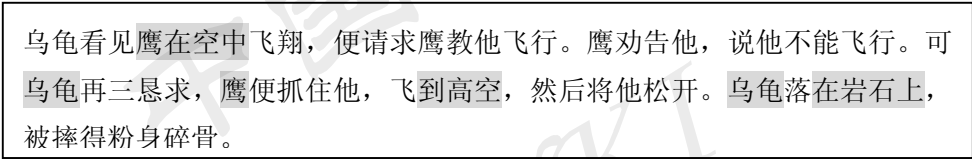
对空间概念建模具体应用，基于 SUMO 本体库体系，提出了空间本体库的体系结构，给出概念、关系和规则的具体定义，把认知理论和常识很好的结合在一起。同时，基于归类思想，依据聚类、规则等方法，提出空间本体半自动实例化，促进了自动本体实例化的发展。最后，以中文版《伊索寓言》为基础，构建了一个小规模空间本体库，在文景转换领域显示了比依赖专家手工构建空间常识库高的效率，好的可行性。本章为后面的两章工作提供了认知语言学空间知识，例如方位词的空间概念体系，以及实体和图形知识，例如元模型概念体系。

中国知网
CNKI

第3章 基于线性分类器空间关系抽取

3.1 引言

文本中空间信息抽取就是从真实文本中，把描述物体的空间方向、位置的信息抽取出来。这些信息从认知语言学角度看，包含界标(Landmark)、方位词(Localizer)、射体(Trajectory)以及这三个元素构成的空间关系，完成对物体空间信息的描述。从行文角度看，它们各自分布在整个篇章中，而不是仅局限一个句子中。本文主要研究的就是从真实文本中，在篇章级别把空间信息抽取出来。例如，从图 3-1所示的文本中抽取的空间信息包含：“鹰在空中”、“鹰到高空”、“乌龟到高空”以及“乌龟在岩石上”，图 3-1中阴影部分汉字表示要被抽出的，能够组成空间信息的关键字。



乌龟看见鹰在空中飞翔，便请求鹰教他飞行。鹰劝告他，说他不能飞行。可乌龟再三恳求，鹰便抓住他，飞到高空，然后将他松开。乌龟落在岩石上，被摔得粉身碎骨。

图 3-1 《鹰与乌龟》

Figure 3-1 “The eagle and the tortoise”

可以看到，与传统的信息抽取相比，组成空间信息的词语从句法上看，并不是紧邻的，而是处于篇章级别的；从概念上看，界标、射体是认知语言学家的研究对象，而非语法、语义语言学家的研究对象，可以说是认知层面的。因此，本文针对现代汉语空间描述的特点，首先，提出空间信息由两部分组成的观点，即空间表达式和射体两个部分；然后，把空间信息抽取任务形式化为二值分类问题；最后，在此基础上，只依赖于词性标注，提出基于线性分类器的空间信息抽取方法。初步尝试了在自然语言处理领域理解认知层面概念的技术，同时实现了篇章级别关系的抽取，而且最大限度的降低了对自然语言处理技术的依赖。

篇章级别的空间信息抽取可以为文景转换、地理信息系统以及其它自然语言应用提供更大程度的支持。此外，由于空间信息抽取把文本自由描述结构化，包含了大部分空间关系可视化需要的约束信息，因此对于实现空间物体自动摆放具有重要的意义。

本章第 2 节对语言学界空间概念的研究简介, 以及对现有的空间信息抽取方法进行了分析; 第 3 节给出空间信息抽取的形式化模型; 第 4 节提出基于线性分类器的空间表达式抽取, 这是空间关系抽取的第一步; 第 5 节提出基于线性分类器的射体的抽取, 这是空间关系抽取的第二步; 最后, 结合我们自己构建的语料库, 对实验结果进行了分析和总结。

3.2 空间信息研究概述

3.2.1 语言学界空间表达研究概述

中外语言学家, 对自然语言如何描述空间概念都有深入的研究, 由于本文的处理对象是现代汉语, 所以着重介绍现代汉语空间概念的研究。对现代汉语空间表达的研究主要集中在空间关系的表达、空间方位参照的研究。

刘宁生在文献[77]中, 提出空间关系是人类语言的一项基本功能。汉语表达物体之间的空间关系的手段是多样化的, 主要由“目的物”、“参照物”和“方位词”三个重要的成分共同承担。目的物是人们目光集中的地方; 参照物的作用是确定目的物的位置和方向, 方位词表达了图景中这个物体之间的空间关系。与这三个成分对应的术语不尽相同, Talmy在文献[78]中, 把“目的物”和“参照物”称为Figure(图形)和Ground(地面); 而Langacker在他的认知语法框架中, 使用Landmark、Trjecotary这对术语, 蓝纯在文献[79]中把他们翻译为“路标”和“动体”, 张辉在文献[76]中把它们翻译成“界标”、“射体”, 界标就是具有或者被概念化内部、边界和外部的结构特性的物体, 射体就是被一个具体的、又界限的界标包围的物体^[80]。廖秋忠在文献[81]中认为现代汉语在给定一人、物或者处所定位时, 经常采用的格式是: (在+)名(+的)+空间方位词(+距离), 括号内的成份在一定的条件下可以省略或不提; “在”是个表示场所的介词, 同一类的还有“从、到、向”等; “名”是个名词或名词短语, 指某一已知的人、物或处所, 是它后面方位词的参考点。同时分析认为, 从篇章中某一句来看, 空间参考点一旦出现于上文或者可从上文会意出来, 那么在本句中如还用同一参考点来定位时, 就有省略不提的可能; 从篇章的层次来看, 某些句子中因为参考点支配范围的持续, 而省略。

方经民在空间方位参照研究中作了许多工作。他把空间方位参照定义为: 叙述者选择观察点, 利用方位词跟相关的方向参照点、位置参照点的关系确定空间的方位辖域^[82, 83]。而且认为空间方位参照是一种认知结构, 空间方位参照反映了语言社会对认知空间中方位关系的认知过程和认知方式, 它不同于一般

所认识的线性的、有形的、表层的句法结构，是一种立体的、抽象的、深层的认知结构^[84]。其结构要素包括方位词、叙述者、观察点、方向参照点和位置参照点。方位词的基本作用是表示方向；叙述者是说话者或写作者。说话或写作可以采取第一人称的叙述方式，也可以采取第二人称的叙述方式；位置参照点是运用方位参照定位时选择的参照点，方向参照点是运用方位参照定向时选择的参照点。是廖秋忠在文献[72]中最早提到方向参照点和位置参照点的不同，但在具体论述时没有加以区分。

语言学者对现代汉语空间表达的结构、成分及参照的研究是文本中空间信息定义、抽取方法确定的重要的理论基础，是本文空间信息抽取的重要思想来源。

3.2.2 空间信息抽取方法与分析

3.2.2.1 空间信息抽取主要方法

目前，基于自然语言的3D场景生成系统中，空间信息抽取方法主要为基于句法分析法。该方法就是依赖于自然语言处理技术，依次对文本进行词性标注、句法分析和语义分析，最后结合一定的知识得到空间信息。

基于句法分析空间信息抽取方法是现有空间信息抽取中普遍采用的方法。美国的Sproat等人在Wordseye系统中直接使用基于句法分析方法抽取空间信息^[3]。首先，文本初始标注、分析，使用Church的基于随机的词性标注工具^[85]和Collins的头驱动模型句法分析工具^[86]；然后，句法分析树转换为依存结构(Dependency Structure)，显示实体之间空间依存关系，格式如图3-2所示，其中箭头是从头(Head)指向依存者(Dependents)；最后，把依存结构转换为语义表示结构(Semantic Representation)，即语义表示片断列表，图3-3是图3-2依存结构对应的语义表示，其中“STATIVE-RELATION'on”是空间关系'on'的语义表示片断，它是在语义函数(Semantic Functions)控制下完成填充的。我国陆汝铃院士等人则在二十世纪九十年代就是基于句法分析，配合知识库进行定性空间信息规划^[2]。首先，他们对用受限中文描述的儿童童话故事进行自然语言单句理解，包含三个子步骤：1)分词，其在一个较大的汉语词汇底表基础上，采用最大匹配法分词；2)语法分析，其实质是基于中文扩展转移网络(Chinese Augmented Transition Network)，只针对结构为“主语+谓语+宾语”的句子，进行句法分析；3)用格框架树的形式表示语义分析结果，这个格框架树的基础是1968年Fillmore提出的以动词为中心分析句子语义的格框架模型。至此，抽取空间信息中的实体以及其动作；然后，结合知识库，就是在情节知识库中，采取系统建议、用户调整的方法，得到完整的定性空间信息，也就是基本场景。

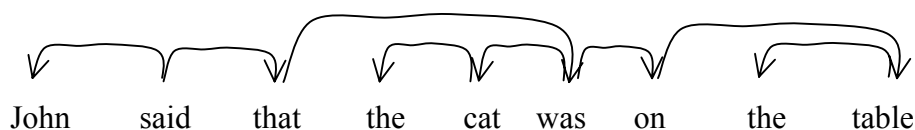


图 3-2 句子 “John said that the cat was on the table” 的依存结构

Figure 3-2 Dependency structure for “John said that the cat was on the table”

```
(("node2" (:ENTITY :3D-OBJECTS ("mr_happy")
:LEXICAL-SOURCE "John" :SOURCE SELF))
("node1" (:ACTION "say" :SUBJECT "node2"
:DIRECT-OBJECT ("node5" "node4" "node7")...))
("node5" (:ENTITY :3D-OBJECTS ("cat-vp2842")))
("node4" (:STATIVE-RELATION "on" :FIGURE "node5"
:GROUND "node7"))
("node7" (:ENTITY :3D-OBJECTS
("table-vp14364" "nightstand-vp21374"
"table-vp4098" "pool table-vp8359" ...))))
```

图 3-3 句子 “John said that the cat was on the table” 的语义表示结构

Figure 3-3 Semantic representation for “John said that the cat was on the table”

3.2.2.2 现有方法存在的不足 从具体的文景转换系统角度来看，基于句法分析空间信息抽取方法有其优点，但是该方法仍然有以下困难：

(1) 难以获得足够鲁棒的、精确的句法分析结果。自然语言语言句法分析技术尚有不足，难以找到一个可靠的句法分析工具，例如 WordEye 系统使用基于统计的句法分析工具，显然其精确不可能是 100%，而且和词性标注存在错误累积问题。但是在该方法中句法分析是整个过程的基础，直接决定着空间信息抽取的结果。

(2) 难以获得篇章级别空间信息结果。即使得到精确的句法分析结果，也只能局限在单句空间信息获取。依存结构和中文扩展转移网络本身固有特性决定了只能得到句子级别的空间关系。

(3) 手工获取知识具有一定的主观性，且耗时、耗力。在该方法中，都结合了一定的知识，以获取具体的空间关系。无论是空间介词的语义函数还是情节知识库的构造都需要一定水平的专业知识，而且还不能穷尽所有常识，其主观性和获取困难可见一斑。

基于线性分类器的空间信息抽取方法提出了一个崭新的思路，它把空间信息抽取抽象为二值分类问题，只基于汉语词性标注一种分析工具，最大限度的降低了错误累积和对知识的依赖，而且在保证精确性的前提下，实现真实文本

中篇章级空间信息抽取。

3.3 基于线性分类器的空间关系抽取模型

本文认为自然语言描述的空间关系具有三个部分、两个层次。从认知事物角度看,空间关系由界标、方位词和射体三部分组成:方位词描述了实体之间的空间关系;射体是空间关系的主体;界标则为射体的方位确定提供了参考。从认知层次看,空间关系由空间表达式、射体两个层次构成:空间表达式描述射体的方位,主要由方位词和界标组成;射体是空间关系的主体。

例如,“农夫坐在墙下”表述了一个具体的空间关系,包含界标“墙”、方位词“下”和射体“农夫”三个部分,组成空间表达式“墙下”和射体“农夫”两个认知层次。

由3.1节对空间信息抽取的表述来看,抽取空间信息就是要把其中的空间表达式、射体分别抽取出来,并且按照对应关系形成具体的空间关系,因此,本文把空间信息的抽取看成二值分类问题,在此基础上,提出基于线性分类器的空间关系抽取模型。

3.3.1 SNoW 概述

本文把SNoW作为分类器工具。文献[87]提出SNoW(Sparse Network Of linear separators)体系结构。这个网络是由两层节点组成的,第一层节点以数据驱动的方式接受特征空间;第二层节点是目标节点,其个数对应问题的分类数目。两层节点由权重联系起来。

SNoW采用在线学习方式。首先,根据胜者为王的策略,学习机给每一个带标记的学习样例分配最可能的类别;然后,视学习类别和标准答案的差别调整网络权重,SNoW使用Winnow和Perceptron等方法调整目标概念的子网络的权重;最后,在一定量的学习下,得到带有权重网络,可以用于日后的测试等实际分类任务。

SNoW的特点是:接受高特征空间学习;比基本的线性分类器完善,可实现多值分类;自主实现特征剪枝;可预测信任机制;数据驱动机制;决策支持机制以及相关特征信息抽取整合机制。其已经被成功的用于自然语言处理和视觉处理等领域,例如:词义消歧、浅层语义分析以及人脸识别等。

3.3.2 空间关系抽取问题的形式化

首先,给出空间信息抽取问题的形式化定义,各个组成部分的形式化定义,以及抽取问题的形式化定义。

3.3.2.1 空间关系的定义 语言学家已经对射体、界标、空间描述进行了定义，但是，从形式上看略有不同。因此，我们在这一小节，根据语言学家的定义，结合计算语言学的目的，给出上述术语较规范的描述。

在自然语言空间描述中，有这样一类名词或者名词短语，它指代某一已知的人、物、或处所，而且是后面方位词在空间上的参照物，其被称为空间方位参考点，也称为界标。

射体是被界标界定的物体，是空间描述的焦点物，标记为 TR 。

空间表达式限制着相对应的射体的位置。被抽象为一个三元组($[Pre]$, Loc , LM)，其中 Pre 表示方位介词，也就是介词集合 S_{Pre} 中的元素，在汉语中这个部分有时省略； Loc 表示方位词，也就是方位词集合 S_{Loc} 中的元素； LM 表示界标，例如：“在空中”被形式化为(在，中，空)。如果， Pre 部分不省略，那么这样的三元组称为标准空间表达式；否则，即 Pre 部分省略，得二元组(Loc , LM)，其对应的空间表达式被称为方位参考点省略的空间表达式，简称省略空间表达式。

如果篇章中存在省略空间表达式，那么这种现象被称为方位参考点省略。

一个空间描述阐述了射体和与它匹配的界标之间的空间结构。本文中，把空间描述称为空间关系，一个空间关系由两个部分组成：空间表达式和射体，抽象为一个二元组($[Pre]$, Loc , LM), TR)，例如：空间关系“鹰在空中”被抽象为((在，中，空)，鹰)。

3.3.2.2 基于层次法空间关系抽取步骤 由上一节的定义推出，空间关系的识别就是要分别识别空间表达式和其对应射体，因此，提出二步法识别空间关系。第一步，识别空间表达式，首先识别空间表达式边界，然后识别其中的三个组成部分：介词、方位词和界标；第二步，识别射体，即识别第一步得到的空间表达式对应的射体。

3.3.2.3 二值分类问题 如果把文本定义为集合 $S = \{s | s \text{ 是字符串}\}$ ，那么空间关系 sr 一定属于集合 S ，而且，识别空间关系等价于判别 s 是否是 sr 的组成部分，因此其实质是二值分类问题。

例如：对于空间表达式识别，不妨设谓词 P 表示字符串是否是一个空间表达式； $C = \{-1, 1\}$ 是 P 取的逻辑结果； $P(s) = 1$ 表示字符串是空间表达式，反之不是，因此，空间表达式边界识别就是找超平面函数 $h: S \rightarrow C$ 。本文使用 SNoW 找到这个超平面函数。

同理，界标识别和射体识别也可以被抽象成二值分类问题。

实验显示，在同一特征空间下，分别使用 Winnow 算法和 Perceptron 算法训练 SNoW 网络权重，得到的结果十分接近。由此可见，相比而言，影响实验

性能的主要因素是特征空间的选取，而不是同类别算法的更换。同时，实验显示，基于线形分类器的结果比较令人满意。

3.4 空间表达式抽取

由定义可以看出来，空间表达式由介词、界标和方位词三个部分组成，同时又是一个相对于射体独立的整体，因此，本文首先把空间表达式边界识别出来，然后再分别识别其组成部分：介词、界标和方位词。

3.4.1 空间表达式边界抽取的特征空间

下面阐述特征空间及其数值化。设字符串 $s=\{w_1/p_1, w_2/p_2, \dots, w_n/p_n\}$ ， w_i 表示词， p_i 是词对应的词性；特征空间 $SPACE=\{WORDPRE, LOCI, POSTRING, SYMSE\}$ ，其中 $WORDPRE$ 、 LOC 、 $POSTRING$ 、 $SYMSE$ 分别表示词形、位置、词性串以及语义四个类别的特征

(1)词形 在汉语中，能够出现在空间表达式中的方位介词、方位词的词性和数量都是确定的。根据第2章 阐述使用两个集合：方位介词集合 S_{Pre} 和方位词集合 S_{Loc} 。在这两个集合中每个元素的词形是一个特征。

下面通过介词“在”说明词性特征数值化。设 S_{Pre} 集合的势是 $|S_{Pre}|$ ，因为 $w_i = \text{“在”}$ ，并且 p_i 是介词词性符号，那么介词词形特征 $WORDPRE_i$ 的值是1，其它 $|S_{Pre}| - 1$ 个，即10个，介词词性特征的值是0。其它介词、方位词词形按照这个方式数值化。

(2)位置 方位介词和方位词在字符串中的位置是确定空间表达式边界的重要特征。方位介词在字符串中出现的位置有三种情况：头、中以及尾部。方位词的情况同理。集合 S_{Pre} 、 S_{Loc} 中的每个元素的每种位置情况是一个特征。例如，如果介词“在”出现在字符串的头部，那么头位置特征的值是1，其他位置特征的值都是0。其它词形同理。

(3)词性串 把每个字符串中的词性抽取出来，按照原序列形成词性串，那么训练集合中的每个正例的词性串构成一个有序集合，记 S_{Pos} ，势是 $|S_{Pos}|$ ，其中每个元素就是一个特征。例如，字符串 s' 的词性串是‘*pngf*’，并且是 S_{Pos} 中的第 i 个元素，那么词性串特征 $POSTRING_i$ 的值是1，其它的 $|S_{Pos}| - 1$ 个词性串特征的特征值为0。

(4)语义 欲正确识别空间表达式，那么就要把其与隐喻空间表达式、时间表达式区别开来，例如把“在桌子上”和“在脑海中”、“在公元前”区别开。显然，空间表达式中参考点，即界标，有明显的实体类语义，而其它两种没有，

因此利用参考点语义消歧。我们使用知网，结合极大似然比，得到上位义原集合 $Sym=\{\text{物质, 群体, 部分}\}$ ，其每个元素就是一个语义特征。例如，“桌子”的上位义原是‘物质’，那么对应的语义特征值是1，“群体”和“部分”对应的语义特征值都是0。

3.4.2 界标识别

界标识别，就是要在已经识别出来的空间表达式中，找到空间参考点，例如：空间表达式“在空中”已经识别出来，那么把空间参考点“空”识别出来就是界标识别的工作。分析汉语表达，发现在篇章中存在省略界标的情况，所以，界标识别包含两层含义，一、把没有省略的界标抽取出来；二、把省略的界标从篇章中恢复。因此，本文提出基于层次法识别界标，首先，判断界标是否省略，然后根据情况分别作界标抽取或者恢复。

3.4.2.1 界标省略判断 利用规则判断空间表达式中是否省略界标。把空间表达式符号化，即 $SExp=\{p_i|p_i\text{是词性, } i, j \in [0, n]\}$ ，那么使用的规则如下：

$$\neg \exists p((p \in PN) \wedge (p \in SExp) \wedge (f(p) \leftrightarrow a)) \rightarrow O(SExp)$$

其中， p 是变量，值域是汉语词性集合； $SExp$ 是变量，值域是空间表达式组成的集合； a 是常量，表示方位词词形's'； PN 是名词类词性组成的集合； $f(x)$ 函数求自变量 x 最右邻词语的词性，值域是汉语词性集合；谓词 O 表示是界标省略的空间表达式。

3.4.2.2 界标抽取 对汉语篇章进行小样本统计，97%的空间表达式没有省略界标，因此，界标抽取是识别任务中重要部分。

界标抽取的本质是二值分类问题。如果把无界标省略的空间表达式定义为集合 $Sexp=\{w|w\text{是字符串}\}$ ，而且界标 l 一定属于集合 S ，那么界标抽取等价于判别 w 是否是 l 。因此，设谓词 P 表示字符串是否是一个界标； $C=\{-1, 1\}$ 是 P 取的逻辑结果； $P(s)=1$ 表示字符串是界标，反之不是，界标识别就是找超平面函数 $h: S \rightarrow C$ 。当前词、其前一个词和其后一个词的词性构成该问题的特征空间。

3.4.2.3 界标恢复 本文提出一种基于有限知识的界标恢复方法，基于汉语句法分析，以知网为常识库，根据界标的词性、语义以及位置有限知识，恢复界标。优点在于：以语言学家研究成果为基础，利用有限知识解决语义层面问题。

根据经验选择优先指示器，包含词性、语义和位置。

界标词性特征集合为 $SPS\{(pos, Ppos) | pos\text{是词性, } Ppos\text{是该词性对应的频率}\}$ 。基于最大似然比得到词性的频率。

通过对训练语料中界标上位义原的统计得到语义类特征集合，表示为

$SMS\{(sym, Psym) \mid sym \in Sym, Psym \text{ 是该义原作为界标上位义原的频率}\}$,
 $Sym = \{\text{物质、群体、部分}\}$ 是实体语义类集合。

根据篇章的连贯性, 候选界标的位置表示为四元组 (s, i, j, k) , $s \in \{1, 0\}$,
 $i, j, k \in [1, n]$ 。候选界标是某个空间表达式的组成成分则 s 取 1; 反之, 取 0。
 i 表示候选界标所在的或者被统摄的空间表达式的编号。 j 表示篇章中句子的编
号。 k 表示篇章中自然段的编号。

同理, 从短语、句子和段落三个层次编号待恢复的空间表达式的位置, 表
示为三元组 (m, n, l) , 含义同 i, j, k 。

首先, 用公式(3-1)(3-2)计算候选界标 Obj_u 的权值 $weight_u$ 。其中, α_s 的值由 s
值决定, 当 s 的值是 1 的时候, $\alpha_s = 2$, 否则 $\alpha_s = 1$ 。

$$weight_u = Ppos + Psym + \alpha_s * Sim(Obj_u, Sexp_u) \quad (3-1)$$

$$Sim(Obj_u, Sexp_u) = 1/(m - i + 1) * 1/(n - j + 1) * 1/(l - k + 1) \quad (3-2)$$

然后通过公式(3-3)得到最大权重 $weight_w$, 得到候选界标在候选集合中的编
号 w 。最后, 根据公式(3-4)得到这个编号所对应的 $ObjWord$, 它就是先行界标。

$$weight_w = Argmax_i(weight_i) \quad (3-3)$$

$$ObjWord = word_w \quad (3-4)$$

3.4.3 介词和方位词抽取

介词和方位词是空间表达式的组成部分, 在空间本体库中, 介词和方位词
的范围是有穷的, 因此我们使用规则识别这两个成分:

$$\exists w((w \in SPre) \wedge (l(w) \leftrightarrow a)) \rightarrow P(x)$$

$$\exists w((w \in SLoc) \wedge (l(w) \leftrightarrow b)) \rightarrow LO(x)$$

其中, w 是变量, 表示候选词语; a 、 b 是常量, 分别表示“开始”和“结
尾”; 函数 $l(w)$ 求得候选词语在空间表达式中的位置, 值域是“开始”、“中
间”和“结尾”; 谓词 P 表示介词, 谓词 LO 表示是方位词。

3.5 射体抽取

射体抽取就是在篇章中把与空间表达式对应的射体全部找到, 并且每个射
体分别和该空间表达式构成一个具体的空间关系。例如: 在图 3-4中显示的寓
言中, 射体识别就是把与空间表达式“到高空”对应的射体“鹰”和“乌龟”
都准确的抽取出来, 并且构成两个空间关系: “鹰到高空”和“乌龟到高空”。
本文着眼于真实文本的处理, 在对汉语语言空间描述的分析基础上, 把射体抽
取建模成二值分类问题, 使用句法特征, 结合线性分类器识别射体。这个方法

有三个优点：一、把认知问题简化成句法级问题，即使用句法级别特征有效地识别认知层面的概念——射体；二、面向真实文本，即有效地在真实复杂的空间语义环境中识别射体；三、噪音减低到最小，只使用含有词性标注工具作为预处理。

乌龟看见鹰在空中飞翔，便请求鹰教他飞行。鹰劝告他，说他不能飞行。可乌龟再三恳求，鹰便抓住他，飞到高空，然后将他松开。乌龟落在岩石上，被摔得粉身碎骨。

图 3-4 《鹰与乌龟》

Figure 3-4 “The eagle and the tortoise”

3.5.1 空间描述策略

从空间表达式和对应的射体之间关系的角度，分析现代汉语空间描述的策略，以便给出射体识别的方法。我们把一个射体和其对应的空间表达式之间的捆绑方式称为一个描述策略，其分为三种类型：数量映射、空间焦点转移和句法位置。

为了分析描述策略，我们标注了中文版《伊索寓言》432 篇作为语料库，详细地标注方案见3.6 节。这个语料库分为两个部分，一、训练语料，包含语料库中随机选取的 325 篇；二、测试语料，包含语料库中其余的 109 篇，这两个部分同时作为本章实验的语料库。下面将针对这个语料库，对每个描述策略进行详细分析，并且分别给出两类百分比，一类是，在训练语料中，符合具体描述策略的空间关系占总空间关系的比例；另一类是，在测试语料中，符合这个描述策略的空间关系占总关系的比例，以此来定量说明描述策略。

3.5.1.1 数量映射 数量映射是指空间表达式和其匹配的射体之间的个数比值，例如：在图 3-4 《鹰与乌龟》中，空间表达式“到高空”和其射体，“鹰”和“乌龟”，之间的数量映射关系就是 1:2。

《伊索寓言》语料库中，有四种数量映射关系：1:0, 1:1, 1:2 和 1:3，其中，1:0 的含义是在上下文中，找不到和这个空间表达式匹配的射体，但是可以从语境中意会，我们称为射体语义省略。四种映射的统计数据见表 3-1，显然，1:1 映射占最大比重。不能肯定的说，在汉语描述中，只有这四种数量映射关系，但是，可以肯定随着映射关系中射体个数的增加，对应的空间关系的比例是极具下降的，这也符合人类对空间关系的认知。

表 3-1 数量映射的统计数据

Table 3-1 The statistic data on the mapping between *TR* and *Sexp*

语料库	1:0 的数量(%)	1:1 的数量(%)	1:2 的数量(%)	1:3 的数量(%)
训练	5 (8%)	539 (86.94%)	58 (9.35%)	18 (2.90%)
测试	2 (8.85%)	203 (89.82%)	18 (7.96%)	3 (1.46%)

3.5.1.2 空间焦点转移 空间焦点就是读者在空间上关注的当前的一个或者几个实体，以及它或者它们的空间位置^[88]，也就是本文中射体。在汉语中，如果在射体和它相对应的空间表达式中，存在其它完整的空间关系，那么就发生了空间焦点转移，例如，在图 3-5所示的寓言中，在射体“鹰”和与其匹配的空间表达式“朝下”之间，有一个完整的空间关系“农夫坐在墙下”，这里空间焦点从“鹰”转移到“农夫”，然后又回到“鹰”。我们认为空间焦点转移在语义上破坏了空间关系的连续性。

从空间焦点转移的角度看，射体和其匹配的空间表达式的关系可以分成两个大类，其中第一个大类，又可以分成两个小类。

第一大类：在射体和空间表达式之间，没有空间焦点转移。从句法上看，又把它分为两个小类，子类一，标记 *TypeSI*，即在射体和其匹配的空间表达式之间，不存在任何其它的空间关系或者空间表达式；子类二，标记 *TypeSII*，即在射体和其匹配的空间表达式之间，存在其它的空间表达式，而且它和这个射体匹配而构成空间关系，例如：在图 3-6所示寓言中，在射体“鸽子”和其匹配的空间表达式“地上”之间有一个空间表达式“画板上”，而且“鸽子”和“画板上”构成一个空间关系，在类型 *TypeSII* 中，空间焦点没有转移，但是空间关系在语义上也被破坏了连续性，这主要是由于主语或者谓语省略而造成射体省略，正如这段的例子就是省略了主语“鸽子”，我们称之为射体语法省略。

第二大类：在射体和空间表达式之间，存在空间焦点转移，也就是在射体和其匹配的空间表达式之间有一个或者多个空间关系，图 3-5中的“鹰”和“地下”就是这样的情况，这种类型标记为 *TypeSIII*。

.....。有/vg 一/m 天/q， /wo 鹰/ng[TR3] 看见/vg 农夫/nc[TR1][TR2] 坐/vg 在/d #l 将要/vz 倒塌/vg 的/usde #l 墙/ng[LM1][LM2] 下/f #r[SE1]#r[SE2]， /wo 就/d 立刻/d #l 朝/p 下/f #r[SE3]飞/vg 去/vg， /wo 用/p 脚爪/ng 抓起/vg #l 他/r[AR 农夫/nc] #l 头[LM4][LM5]上/s #r[SE4]#r[SE5]的/usde 头巾/ng[TR4][TR5]。 /wj.....

图 3-5 《农夫与鹰》片段

Figure 3-5 The sentence from the fable named “The farmer and the eagle”

.....他/r[AR 鸽子/ng][TR2][TR3][TR4][TR5] 立刻/d 呼呼/o 地/usdi 猛/d 飞/vg 过去/vq , /wo 不料/d 一头/d 碰撞/vg #l 在/p #l 画板/ng[LM2][LM3] 上/f #r[SE2]#r[SE3], /wo 折断/vg 了/ut 翅膀/ng , /wo 摔/vg #l 在/p #l 地[LM4][LM5] 上/s #r[SE4]#r[SE5], /wo.....

图 3-6 《鸽子和画板》中的片段

Figure 3-6 The parts from the fable named “The pigeon and the painting”

在语料库一对一数量映射中,对上面的三种情况进行了统计,结果如表 3-2 所示。

表 3-2 空间焦点转移的统计数据

Table 3-2 The Statistic Data on the Spatial Focus Shift

只含有 1:1 映射语料库	TypeSI 的数量(%)	TypeSII 的数量(%)	TypeSIII 的数量(%)
训练	523 (97.03%)	14 (2.6%)	2 (0.37%)
测试	190 (93.60%)	11 (5.42%)	2 (0.99%)

3.5.1.3 句法位置 最后一个描述策略是句法位置,也就是射体和其匹配的空间表达式是否是同一个句子。

这里,我们把由句号、问号和感叹号结束的句子称为句子,而由逗号和省略号结束的句子称为子句。

这个描述策略含有三种类型,一、子句位置,即射体和其匹配的空间表达式在同一个子句内部,标记为 *TypeLI*;二、句子位置,即射体和其匹配的空间表达式不在同一个子句内部,而是在同一个句子内,标记为 *TypeLII*;三、篇章位置,即射体和其匹配的空间表达式既不在同一个子句内部,又不在同一个句子内,而是处于同一段落,或者不同段落,标记为 *TypeLIII*。

在语料库中,对句法位置描述策略统计,结果见表 3-3,其中在同一子句的情况最多。

表 3-3 句法位置的统计结果

Table 3-3 The Statistic Data on the Syntactic Location Relations

语料库	TypeLI 数量(%)	TypeLII 数量(%)	TypeLIII 数量(%)
训练	439 (71.38%)	158 (25.69%)	18 (2.93%)
测试	159 (70.98%)	63 (28.13%)	2 (0.89%)

3.5.2 射体抽取的定义

通过3.5.1节的阐述，发现汉语中对射体的描述是复杂的，但是射体识别任务完全可以形式化为二值分类问题。

首先，对待识别射体的文本进行词性标注和空间表达式标注，得到初始文本，把这个文本抽象为含有两个层次的字符串，第一层：待匹配射体的空间表达式，记为 s_{exp_tr} ；第二层：有词性标注的词语，组成集合 S_{text_w} ，这个集合不包含 s_{exp_tr} 中的词语，那么，集合 S_{text_w} 中的每一个元素就是一个候选射体，记为 ctr 。设谓词 P 表示字符串是否是一个射体； $C = \{-1, 1\}$ 是 P 取的逻辑结果； $P(s) = 1$ 表示字符串 s 是射体，反之不是，因此，射体抽取就是找超平面函数 $h: S \rightarrow C$ 。

3.5.3 射体抽取的特征空间

根据节提到的描述策略的特点，从语法角度选择特征，并且可以分为四类，即标点符号类，标记为 pun ；词性类，标记为 pos_X ；以及距离类，标记为 dis_X 。它们组成的特征空间被抽象为单元向量，标记为 $f\{pun, pos_str_1, \dots, pos_str_m, pos_verb_1, \dots, pos_verb_n, dis_verb_1, \dots, dis_verb_o, dis_SEL_1, \dots, dis_SEL_p, dis_SE_1, \dots, dis_SE_q, dis_SEN_1, \dots, dis_SEN_l, dis_by_1, \dots, dis_by_u, dis_with_1, \dots, dis_with_l\}$ 。下面讨论特征及其取值。

3.5.3.1 标点符号类特征 当候选射体是标点符号的时候，特征 pun 的值是 1，其它所有的特征值都是 0；反之， pun 的值是 0，而其它特征的值由以下情况而定。

3.5.3.2 词性类特征 本文使用的词性标注工具含有 52 个汉语词性，为了降低特征空间维数，我们把其中的表示“其他符号”的词性“ sym ”、表示“标点符号”的所有词性都归为一类，视做一个词性。特征空间中有考虑两类词的词性，一类是候选射体的词性；另一类是动词的词性，这里的动词是指那么和候选射体处于同一个句子的动词。这两类词语的每个词性对应于特征空间中的一个具体的特征，例如：词性“ ng ”就是特征空间中的一个词性特征，因此，当某个候选射体的词性是“ ng ”的时候，这个词性特征值的就是 1，其它词性特征值都是 0，其它词性特征取值同理。

3.5.3.3 距离类特征 因为，初始文本中只有词形和词性两类简单句法特征，为了能够充分识别射体这个认知层面的概念，所以，在距离类特征中考虑三种情况：

一、候选射体和功能词之间的距离，功能词包含“把”和“被”，标记为 fw ，而且与候选射体在同一个子句；二、候选射体和动词之间的距离，这里的

动词是那些同候选射体出现在同一个句子的动词，标记为 ph ；三、候选射体和空间表达式之间的距离，这里的空间表达式包含两种情况，一个是紧邻待匹配的空间表达式之前，另一个是紧邻待匹配的空间表达式之后，空间表达式也被标记为 ph 。

根据公式(3-5)计算第一种距离，其中 $sign$ 的值由公式(3-7)确定； n 表示候选射体和功能词之间的词语的个数，当这样的功能词不存在的时候， n 的值就是零； n_0 的值由公式(3-8)确定； n_1 表示候选射体和动词或者空间表达式之间的子句的个数； n_2 表示候选射体和动词或者空间表达式之间的句子的个数； n_3 表示候选射体和动词或者空间表达式之间的段落的个数；每一个距离值对应一个特征，例如：候选射体和“被”的距离是 5，那么特征空间中就有一个对应的特征，同时其值是 1。其它的距离特征取值同理，并且距离特征取值不排斥。

$$dis(ctr, fw) = sign * n \quad (3-5)$$

$$dis(ctr, ph) = sign * \sum_{i=0}^3 (n_i * 10^i) \quad (3-6)$$

$$sign = \begin{cases} 1 & \text{当候选射体位于功能词、动词或者空间表达式之前;} \\ 0 & \text{当候选射体位于功能词、动词或者空间表达式之后;} \end{cases} \quad (3-7)$$

$$n_0 = \begin{cases} 1 & \text{当候选射体位于动词或者空间表达式同一个子句;} \\ 0 & \text{反之;} \end{cases} \quad (3-8)$$

3.6 语料库的建设

目前，没有见到国际上有针对性这个方面的公共测试数据集合，因此我们对文献[89]中《伊索寓言》共十卷，432 篇进行了标注，构建了本课题的实验语料库，然后，随机选择其中的 325 篇作为训练语料，其余的作为测试语料。

挑选语料的原则如下：第一，任务背景，既以满足文景转换研究需要作为选择语料的大背景。第二，真实性，即语料必须是没有被本课题组人员做任何修改的真实文本，该寓言从网上直接下载，我们没有进行任何人工改动。第三，广泛性，即最大限度地选取全面反映汉语空间概念的用法的语料。

为了胜任空间关系可视化任务的实验要求，伊索寓言原始文本经过了以下步骤的标注：一、自动标注词性，这个步骤使用本实验室研发的句子切分和词性标注工具对其处理，句子切分是基于错误驱动方法实现的，准确率超过 96%^[90]。基于二阶隐马尔科夫模型词性标注模块的准确率达到 95%左右，速度达到实时性能^[91]；二、人工消解人称代词；三、人工标注空间关系，其中依次

包含空间表达式标注、界标标注以及射体标注。

在第二和第三人工标注阶段中，标注的原则是：参与者分别独立标注，然后，依据全数通过原则综合结果，参与者包含一名博士生和四名硕士生，他们皆承担着与空间概念建模课题相关的不同子任务。

下面详细阐述第二步骤和第三步骤的标注方法。

3.6.1 人称代词消解

人称代词消解就是明确人称代词及其先行语的关系，从而确定人称代词的所指对象。人称代词的消解有助于空间关系的理解，经过自动词性标注步骤之后，文本的格式如图 3-7所示，其中含有词语和相匹配的词性——词性标注集合详情请见附录，在这样格式的文本中，我们发现存在一定量的指代现象，这种语法现象的存在能够导致人称代词独立的作为界标或者射体被识别出来，例如图 3-7中的“那里”就会被作为界标抽取出来，同时“你”也作为射体被识别出来，此时，脱离了语境，就很难理解这个界标或者射体的具体指代，造成空间关系理解的困难，因此，本课题中人工消解人称代词的指代。为了减少噪音，减少冗余工作量，在保证空间关系可理解的前提下，只对人称代词进行人工消解。

.....便/c 严厉/a 地/usdi 对/p 它/r 说/vg : /wj “/wyl 为什么/r 你/r 还/d 站/vg 在/p 那里/r 打呵欠/vg ? /wj 一切/r 准备/vg 妥当/a , /wo 只/d 等/vg 你/r 了/y , /wo 赶快/d 跟/p 我/r 走/vg 吧/y ! /wj ” /wyr.....

图 3-7 标注词性的《主人和他的狗》中的片段

Figure 3-7 The parts with POS tagging from the fable named “The host and his dog”

消解原则如下：在人称代词的前文找到其先行词语，如果前文没有，则在其后文找先行语；把找到的先行语以及对应的词性放在这个人称代词的右侧，且放在中括号中，用‘AR’标示。例如图 3-7的人称代词“他”的先行语是“一只屎壳郎”，那么就把标记“[AR一/m 只/q 屎壳郎/ng]”放置在“他”的右侧。经过人称代词消解的文本如图 3-8所示。

.....便/c 严厉/a 地/usdi 对/p 它/r[AR 狗/ng] 说/vg : /wj “/wyl 为什么/r 你/r[AR 狗/ng] 还/d 站/vg 在/p 那里/r[AR 门[LM3]口/s]打呵欠/vg ? /wj 一切/r 准备/vg 妥当/a , /wo 只/d 等/vg 你/r[AR 狗/ng] 了/y , /wo 赶快/d 跟/p 我/r[AR 主人/nc] 走/vg 吧/y ! /wj ” /wyr.....

图 3-8 人称代词消解后的《主人和他的狗》中的片段

Figure 3-8 The parts with pronouns resolution from the fable named “The host and his dog”

经过人称代词消解的伊索寓言文本被称为初始语料库。

3.6.2 空间表达式标注

空间表达式是空间关系的基本组成部分，更重要的是，从认知层面上讲，人类对空间表达式的识别要早于界标和射体，也就是人类确定一个具体的空间方位的时候，首先确定方位，然后才是射体的具体位置，因此，我们首先在初始语料库中标注空间表达式。

空间表达式标注原则：首先，确定空间表达式，即根据3.3.2.1节的定义找到篇章中的空间表达式；然后，标注空间表达式符号，即给这个空间表达式标注边界和编号。

在确定空间表达式的步骤中，有嵌套空间表达式情况发生，例如：在图 3-9 中，短语“在家中”根据定义含有两个空间表达式：“在家中”和“家中”，也就是空间表达式嵌套，此时，由于语料库是用于训练，而且这两个短语都符合空间表达式概念，为了保证学习的完备性，所以嵌套的空间表达式分别标记。

.....是/vx 他/r[AR 主人/nc] 教/vg 我/r[AR 守门/vg 狗/ng] 不/d 去/vq 打猎/vg ,/wo 坐/vg #l 在/p #l 家/ng 中/f #r[SE1]#r[SE2]享受/vg 别人/r 的/usde 劳动/ng 果实/ng 。/wj.....

图 3-9 标注空间表达式的《两只狗》中的片段

Figure 3-9 The parts with spatial-expression tagging from the fable named “The two dogs”

空间表达式标注符号制定如下：左边界符号是“#l”，放置在空间表达式最左汉字的前面；右边界符号是“#r”，放置在空间表达式最右词性标注的后面，如果有人称指代标记，则放在这个标记的后面；编号符号以“SE”开头，以空间表达式在篇章中出现的顺序编号，且把整个编号放置在中括号内，整体放在右边界标记的后面。例如：在图 3-10所示的篇章中，“在那里”是一个空间表达式，而且有人称指代标注，因此，左边界符号“#l”放在了“在”的前面，右边界符号“#r”放在人称指代标记“[AR门口/s]”的后面，编号“[SE3]”在“#r”的后面，且表示这是本篇章的第三个空间表达式；再例如：图 3-9中，“#l家/ng中/f #r[SE1]#r[SE2]”是典型的嵌套空间表达式的标记。

经过空间表达式标注的语料库就如图 3-9、图 3-10所示。

3.6.3 界标标注

界标是空间表达式的一部分，所以在带有空间表达式标注的语料库中标注界标，也就是在空间表达式内部标注界标。

.....便/c 严厉/a 地/usdi 对/p 它/r[AR 狗/ng] 说/vg : /wj “/wyl 为什么/r 你/r[AR 狗/ng] 还/d 站/vg #l 在/p #l 那里/r[AR 门口/s] #r[SE3]#r[SE4]打呵欠/vg ? /wj 一切/r 准备/vg 妥当/a , /wo 只/d 等/vg 你/r[AR 狗/ng] 了/y , /wo 赶快/d 跟/p 我/r[AR 主人/nc] 走/vg 吧/v ! /wi ” /wvr.....

图 3-10 标注空间表达式的《主人和他的狗》中的片段

Figure 3-10 The parts with spatial-expression tagging from the fable named “The host and his dog”

界标标注原则如下：首先，界标的确定，即根据3.3.2.1节定义在空间表达式中找到表示界标的词语；然后，标注界标符号，包含标示符号“LM”，以及编号，其和所在的空间表达式编号一致，并且界标标注符号都放在中括号中，例如：“家/ng[LM1]”表示“家”是本文本中第一个空间表达式的界标。

因为界标标注是在词性标注、人称代词消解以及空间表达式标注三个步骤的后面，所以在界标确定的步骤中，需要考虑四种情况：

一、当代表界标的词语同词性工具切词得到的词语一致的时候，那么这个词语就作为界标，其界标符号放置在这个词语对应的词性的后面，例如：在图 3-11所示的文本中，“家”被标注人员确定为界标，同时它也被词性标注工具标注为一个独立的名词，此时，界标标记符号[LM1]放在“家/ng”的后面；

.....是/vx 他/r[AR 主人/nc] 教/vg 我/r[AR 守门/vg 狗/ng] 不/d 去/vq 打猎/vg , /wo 坐/vg #l 在/p #l 家/ng[LM1][LM2] 中/f #r[SE1]#r[SE2]享受/vg 别人/r 的/usde 劳动/ng 果实/ng 。 /wj.....

图 3-11 标注界标的《两只狗》中的片段

Figure 3-11 The parts with landmark tagging from the fable named “The two dogs”

二、当代表界标的词语同词性工具切词得到的词语不一致的时候，那么把界标标记放在表示界标的词语的后面，而不是放在切词得到的整个词语后面，例如：在图 3-12所示的文本中，“门”被标注人员确定为界标，但是它没有被

切成一个独立的词，而“门口”被切在一起了，此时，界标标记符号[LM1]放在“门”的后面，而不是“门口/s”的后面；

三、当做界标的词语出现在人称指代消解中括号内部，而且属于第一种情况，那么标记原则同第一种情况，如果属于第二种情况，那么原则同第二种情况，例如：图 3-12所示的文本中，“门”是代表界标的词语，而且出现在人称代词消解符号“[AR门口/s]”内部，但是其符合第二种情况，所以界标标记符号[LM3]放在“门”的后面；

四、当同一个词语同时是多个空间表达式的界标的时候，界标标记分别给出，例如：图 3-12所示的“门”同时做空间表达式“门口”和“在门口”的界标，那么，“门”的后面有两个标记“[LM1]”和“[LM2]”。

.....他/r[AR 主人/nc] 看见/vg 他/r[AR 主人/nc] 的/usde 狗/ng 仍/d 站/vg #l 在/p #l 门[LM1][LM2]口/s #r[SE1]#r[SE2]打呵欠/vg , /wo 便/c 严厉/a 地/usdi 对/p 它/r[AR 狗/ng] 说/vg : /wj “/wyl 为什么/r 你/r[AR 狗/ng] 还/d 站/vg #l 在/p #l 那里/r[AR 门 [LM3][LM4]口/s] #r[SE3]#r[SE4]打呵欠/vg ? /wj.....

图 3-12 标注界标的《主人和他的狗》中的片段

Figure 3-12 The parts with landmark tagging from the fable named “The host and his dog”

界标标注后的语料库形如图 3-11、图 3-12所示。

3.6.4 射体标注

射体标注就是在上下文中找到空间表达式界定的表示焦点物的词语，因此，需要在全文范围内查找代表射体的词语。

射体标注原则如下：首先，射体的确定，即根据3.3.2.1 节定义在全文中找到与目前的空间表达式匹配的、能够代表射体的所有词语；然后，标注射体符号，包含标示符号“TR”，以及编号，其和匹配的空间表达式编号一致，并且射体标注符号都放在中括号中，例如：“狗/ng[TR1]”表示“狗”是与本文本中第一个空间表达式匹配的射体。

在射体确定这个步骤中，射体定义不是确定射体的充分条件，例如：在图 3-13所示的文本中，与空间表达式“在家中”匹配的射体根据定义应该是“守门狗”，但是代表这个射体的词语全文中有 6 处(这 6 处已经用符号⊙在图 3-13 标出)。由此看来，只借助定义是不能够完成射体确定的，因此，规定在确定射体的时候，本着最左近邻原则，即选择与空间表达式左侧最近邻的可以代表射体的词语作为标注对象，例如，在上例中，选择第六个词语“狗”作为标注对

象。

另外，与界标规定相同，当同一个词语同时是多个空间表达式的射体的时候，射体标记分别给出，例如：符号“狗/ng [TR1][TR2]”的含义是，其同时作为第一个空间表达式“在家中”和第二个空间表达式“家中”的射体。

3.6.5 空间关系标注

空间关系标注就是按照空间认知，把空间表达式、界标以及射体之间争取的搭配关系标注出来。通过上面分别对空间表达式、界标和射体的标注，空间关系已经由它们的一致标号显示的标注，例如：在图 3-13所示的文本中，空间表达式“在家中”、界标“家”和射体“狗”借助同一个标号“1”而构成了一个具体的空间关系。

有/vg 个人/r 养/vg 了/ut 两/m 只/q 狗/ng , /wo 他/r 驯养/vg 一/m 只/q 狗/ng 狩猎/vg , /wo 另/r 一/m 只/q 看家/vg 守门/vg 。/wj 每/r 次/q 猎人/nc 带/vg 着 /ut 猎狗/ng 出去/vq 打猎/vg , /wo 获得/vg 什么/r 猎物/ng , /wo 总是/d 分/vg 给/p 守门/vg 狗/ng⊙ 一些/m 。/wj 猎狗/ng 对/p 此/r 很/d 不/d 高兴/a , /wo 便 /c 指责/vg 守门/vg 狗/ng⊙ , /wo 说/vg 自己/r[AR 猎狗/ng] 每/r 次/q 出去/vg 打猎/vg 都/d 是/vx 四处/d 奔跑/vg , /wo 十分/d 辛苦/a , /wo 而/c 他/r[AR 守门/vg 狗/ng] ⊙ 什么/r 都/d 没有/d 做/vg , /wo 却/d 坐享其成/i 。/wj 守门/vg 狗/ng⊙ 对/p 猎狗/ng 说/vg : /wj “/wyl 你/r[AR 猎狗/ng] 别/d 责怪/vg 我/r[AR 守门/vg 狗/ng] ⊙ , /wo 应该/vz 去/vg 责怪/vg 主人/nc , /wo 是/vx 他/r[AR 主人/nc] 教/vg 我/r[AR 守门/vg 狗/ng⊙ [TR1][TR2]] 不/d 去/vq 打猎/vg , /wo 坐 /vg #1 在/p #1 家/ng[LM1][LM2] 中/f #r[SE1]#r[SE2]享受/vg 别人/r 的/usde 劳动/ng 果实/ng 。/wj ” /wyr

图 3-13 标注射体的《两只狗》¹

Figure 3-13 The fable named “The two dogs” with trajectory tagging

3.7 实验以及分析

在对文献 10 中《伊索寓言》10 卷共 432 篇进行标注基础上，随机选择其中的 325 篇作为训练语料，其余的作为测试语料，组织各类实验，分别给出实验结果，以便验证空间关系识别方法的有效性。

¹ 这个图中的符号⊙不属于语料库的标注集合，而表示可以被标注为射体的候选词语。

3.7.1 参数训练

在 SNoW 工具中，参数选取是实验的第一个步骤，因此，采用枚举方法得到参数。

对于 Winnow 算法，需要的参数有提升因子 α 、 β 和阈值 θ ，根据经验得到每个参数的取值范围以及变化步长，如表 3-4 所示。

对于 Perceptron 算法，需要的参数有学习率、阈值 θ 和缺省值，根据经验缺省值取 0.2，其它每个参数的取值范围以及变化步长，如表 3-5 所示。

在训练语料中，每个算法都分别按照上面的取值范围和步长设置参数，每一组参数对应一个训练结果，在所有结果中，取 F 测度最大对应的参数为测试参数值。

下面的每一个具体实验都是按照这个方法选择参数。

表 3-4 Winnow 算法参数的取值范围和步长

Table 3-4 The Numeric Areas and Steps on the Parameters for Winnow

参数类型	最小值	最大值	步长
α	1.0	2.0	0.1
β	0.0	0.9	0.1
θ	1	6	1

表 3-5 Perceptron 算法参数的取值范围和步长

Table 3-5 The Numeric Areas and Steps on the Parameters for Perceptron

参数类型	最小值	最大值	步长
学习率	0.1	0.9	0.1
θ	2	6	1

3.7.2 评价指标

空间关系抽取结果用正确率和召回率进行评价，正确率和召回率定义如下：

$$\text{正确率} = \frac{\text{正确识别结果}}{\text{得到的识别结果总数}} \times 100\% \quad (3-9)$$

$$\text{召回率} = \frac{\text{正确识别结果}}{\text{语料库标注结果总数}} \times 100\% \quad (3-10)$$

此外，使用 F-score 作为正确率和召回率的综合评价：

$$\text{F-score} = \frac{2 \times \text{正确率} \times \text{召回率}}{\text{正确率} + \text{召回率}} \quad (3-11)$$

3.7.3 空间表达式抽取实验

3.7.3.1 空间表达式边界识别实验 为了检验3.4.1节中各个特征的贡献度，空间表达式边界识别分别在三个特征空间下完成，定义如表3-6所示。

在上面定义的特征空间下，在训练集合上，分别得到的性能指标如图3-14所示，其中参数见表3-7所示。图3-14中显示在非词性特征空间下，整体性能最低，尤其召回率下降最快，这说明了词性特征对于空间表达式边界识别的重要性。然而，*Fea2*下和*Fea1*下的性能没有很大的区别，似乎表明语义特征的贡献度不是很大，因此，我们又分别在这两个特征空间作了开放测试，结果如图3-15所示，在开发测试情况下，*Fea2*的性能要比*Fea1*的坏，尤其召回率下降幅度较大。因此，可是说明语义特征能够较好的把空间表达式和时间表达式区别开。

表 3-6 空间表达式边界识别的特征空间

Table 3-6 The feature spaces on the extraction of spatial expressions

特征空间名称	特征空间内容
全空间，标记 <i>Fea1</i>	包含3.4.1节阐述的所有特征
非语义空间，标记 <i>Fea2</i>	在 <i>Fea1</i> 中去掉语义特征
非词性空间，标记 <i>Fea3</i>	在 <i>Fea1</i> 中去掉词性串特征

表 3-7 不同特征空间对应的参数值

Table 3-7 The parameters' values corresponding different feature spaces

特征空间	α	β	θ
<i>Fea1</i>	2.8	0.7	1
<i>Fea2</i>	1.6	0.8	1
<i>Fea3</i>	1.2	0.9	6

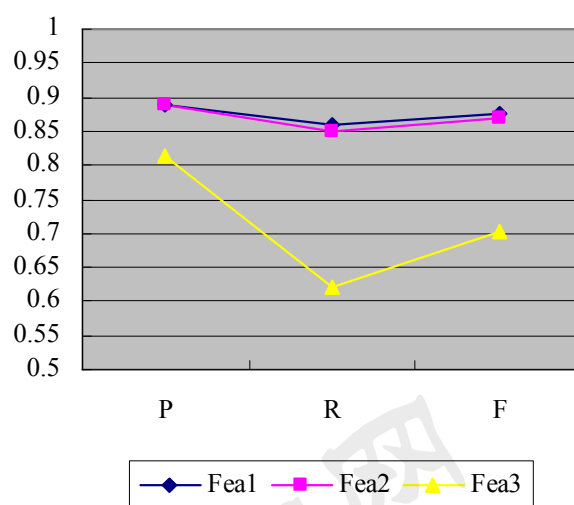


图 3-14 三个特征空间下的封闭测试结果
Figure 3-14 The open results on three feature spaces

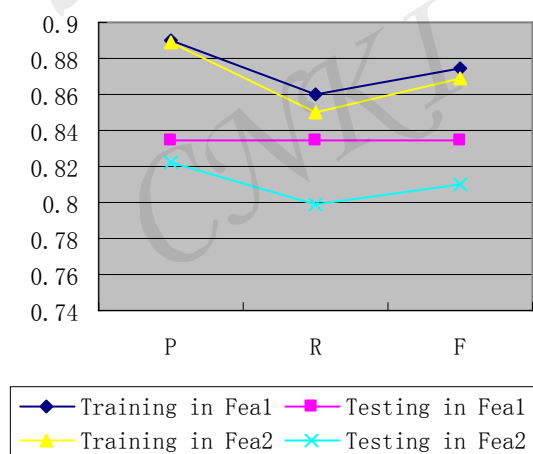


图 3-15 前两个特征空间下的封闭和开发结果
Figure 3-15 The close and open results on the first two feature spaces

下面，为了说明特征在识别任务中的重要性，做了如下封闭和开发测试：在语料上，使用特征空间 $Fea1$ ，参数如表 3-8所示，分别使用Winnow算法和Perceptron算法，结果如图 3-16所示，从图中可以看出，无论是封闭测试还是开放测试，不同算法下的结果相差在 1 个百分点左右，这说明对于空间表达式这个具体的识别任务来讲，特征空间的选取对性能的影响是绝对的，而具体算法

的选择不能产生质的变化，也进一步说明线性分类器对认知层面概念识别的有效性。

表 3-8 不同算法对应的参数值

Table 3-8 The parameters' values corresponding different classifiers

算法	α /学习率	β /缺省值	θ
Winnow	2.8	0.7	1
Perceptron	0.1	0.2	4

3.7.3.2 界标的抽取与恢复实验 首先，利用3.4.2.1节的规则识别界标省略问题。语料库中一共有空间表达式 786 个，省略界标的空间表达式有 30 个，规则可以 100%识别界标省略情况。

然后，做界标抽取实验。为了考察上下特征的有效性，在下面三个特征空间下，分别对界标的抽取作封闭和开放测试。特征空间的定义见表 3-9，参数见表 3-10，图 3-17显示了结果，显然，在上下文特征的特征空间 $FeaL1$ 和 $FeaL2$ 下的性能要远比无上下文特征的 $FeaL3$ 好，说明上下文特征的重要性，同时，在 $FeaL1$ 下的封闭测试的F值虽然比 $FeaL2$ 下的高，但是开放测试的时候反而下降，经过分析数据发现 60%左右的界标没有上一个词语上下文，因此，可以说是数据稀疏造成了 $FeaL1$ 特征空间性能的不稳定。

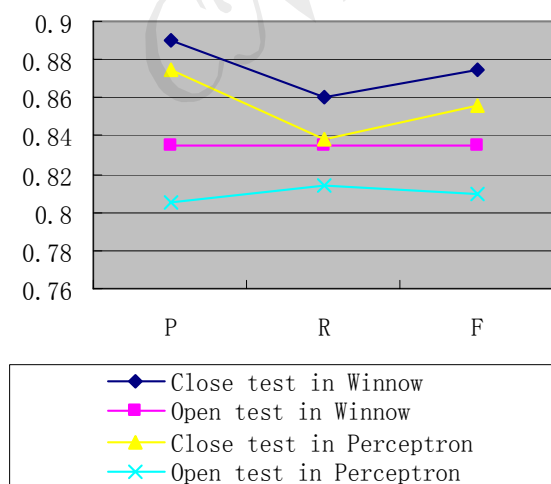


图 3-16 分别在 Winnow 和 Perceptron 算法下封闭和开发测试结果
Figure 3-16 The close and open tests results on Winnow or Perceptron

表 3-9 界标抽取的特征空间

Table 3-9 The feature spaces on the extraction of landmarks

特征空间名称	特征空间内容
上下文空间, 标记 <i>FeaL1</i>	包含3.4.2 节阐述的所有特征
下文空间, 标记 <i>FeaL2</i>	只含有当前词和下一个词词性
单空间, 标记 <i>FeaL3</i>	只含有当前词词性

表 3-10 不同特征空间对应的参数值

Table 3-10 The parameters' values corresponding different feature spaces

特征空间	α	β	θ
<i>FeaL1</i>	1.3	0.9	1
<i>FeaL2</i>	1.6	0.1	3
<i>FeaL3</i>	1.8	0.1	5

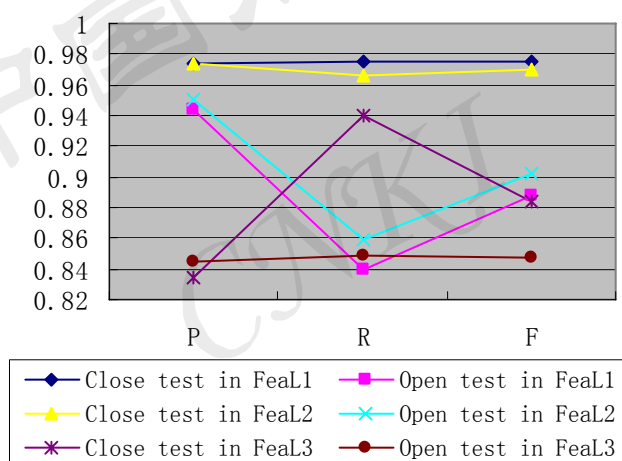


图 3-17 不同特征空间下的界标抽取测试结果

Figure 3-17 The test results on different features spaces for landmark recognition

最后, 做界标省略恢复实验。语料中有 3%, 30 个空间表达式省略了界标, 利用3.4.2.3 节方法可以恢复 30 个界标, 其中正确恢复 27 个, 正确率 90%, 召回率达 90%, F值 90%。

3.7.3.3 介词和方位词的抽取 使用3.4.3节中的规则，在测试语料中分别识别介词和方位词。把测试语料库中标注的最长的空间表达式作为测试对象，即不考虑重复，其中没有省略介词的空间表达式有112个，没有省略方位词的空间表达式有198个，识别结果如表3-11、表3-12所示，数据说明了方法的有效性。

表 3-11 介词识别结果

Table 3-11 The result on the recognition of spatial prepositions

空间介词个数	算法正确识别个数	识别的正确率
112	112	100%

表 3-12 方位词识别结果

Table 3-12 The result on the recognition of localizers

方位词个数	算法正确识别个数	识别的正确率
198	190	95.96%

3.7.4 射体抽取实验

3.7.4.1 射体抽取的特征空间 为了检验3.5.3节中各个特征的贡献度，射体抽取分别在三个特征空间下完成，定义如表3-13所示。

表 3-13 射体抽取的特征空间

Table 3-13 The definitions of feature spaces for the trajectory extraction

特征空间名称	特征空间内容
候选词性空间，标记 <i>Fea00</i>	包含3.5.3.2节阐述与候选射体相关的词性特征
空间表达式空间，标记 <i>Fea01</i>	包含候选射体词性、候选射体和空间表达式距离特征
全空间，标记 <i>Fea10</i>	包含3.5.3节阐述的所有特征

3.7.4.2 确定候选射体 我们使用了以下方法确定候选射体，首先，在与待匹配空间表达式不同的句子中，找到其紧邻的前一个空间表达式和后一个空间表达式，然后，候选射体就是这两个空间表达式之间的词语——这些词语包含出现在这两个空间词语，但是不包含出现在待匹配空间表达式的词语。这个方法可以在覆盖尽可能多的射体前提下，保证射体抽取的正确率。

3.7.4.3 不同特征空间下比较实验 首先，在特征空间 $Fea00$ 和 $Fea01$ 下，分别作封闭和开发测试，参数如表 3-14所示。分析图 3-18显示的实验结果，发现只有候选射体词性的特征空间下的射体抽取性能只有 20%左右，十分不好，而加入了和空间表达式相关的特征后，其性能明显上升，可以达到 60%，这说明射体概念不是句法层面的，而是认知层面的，是和空间表达式概念密切相关的。

表 3-14 射体抽取实验对应的参数

Values of the parameters on diferrent feature spaces

特征空间	α	β	θ
$Fea00$	1.2	0.2	2
$Fea01$	1.6	0.3	2
$Fea10$	1.7	0.5	2

接下来，由于图 3-18显示，在特征空间 $Fea01$ 下，其开放测试的召回率严重下降，比封闭测试降低了有 10%左右，因此，加入关于动词和功能词的特征，组成特征空间 $Fea10$ ，以此作了封闭和开放测试。从图 3-19所示结果可以看出，开放测试中，在特征 $Fea10$ 下的召回率要好于特征空间 $Fea01$ ，大约高出 20%多，而且 F 值性能也较好，由此可以说，动词和功能词的词性以及距离特征对于射体识别有一定的贡献度。

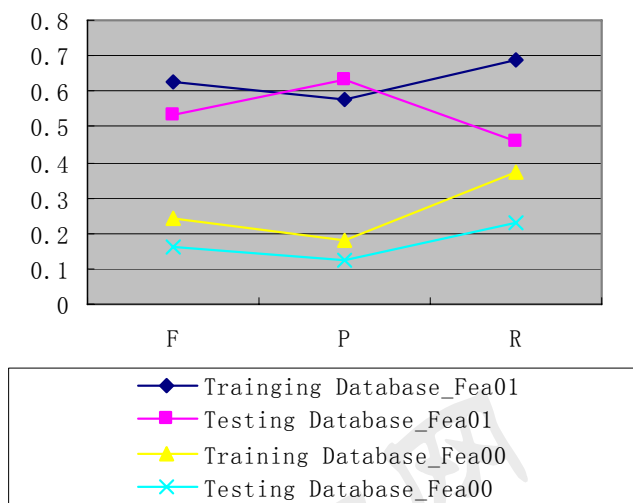


图 3-18 特征空间 Fea00 和 Fea01 下的实验结果

Figure 3-18 The results on the feature spaces Fea01 and Fea00

3.7.4.4 不同数量映射的实验结果 为了说明本方法对于不同数量映射情况的效果，在特征空间 *Fea10* 下，我们统计了各类数量映射的封闭测试和开放测试的召回率，从可以看出，对于 1:1 映射，本方法识别射体的召回率可以达到 70%，同时图 3-20 中显示 1 对多映射下，开放测试的召回率反而高于封闭测试，分析发现，这个原因非常可能是由于数据稀疏造成的，测试语料库中只有 12%，76 个空间关系属于 1 对多映射。

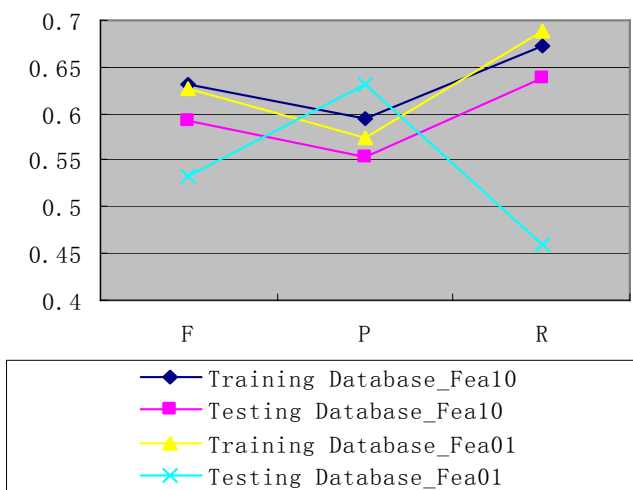


图 3-19 特征空间 Fea10 和 Fea01 下的实验结果

Figure 3-19 The results on the feature spaces Fea10 and Fea01

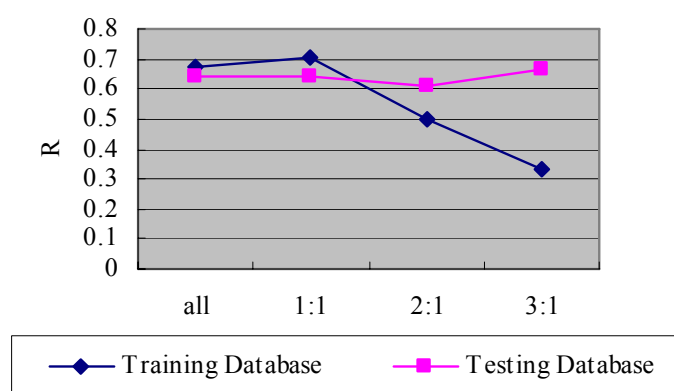


图 3-20 不同数量映射下的结果

Figure 3-20 The results on the different mappings

3.7.4.5 不同句法位置关系下的实验结果 为了说明本方法对于不同句法位置情况的效果，在特征空间 $Fea10$ 下，我们统计了各类句法位置的封闭测试和开放测试的召回率，从图 3-21可以看出，对于同一个子句的位置关系，本方法识别射体的召回率可以达到 85%；对于同一个句子的位置关系下，召回率不到 30%，一方面是由于 158 个空间关系导致的训练不够，另一方面也反映了目前特征对于这种情况的效果较低；不同句子的位置关系下，召回率几乎为 0，分析发现，这个原因非常可能是由于数据稀疏造成的，测试语料库中只有 3%，18 个空间关系是不同句子的关系。

3.7.5 空间关系抽取实验

根据3.3.2.2节的阐述，本文是采用分步法实现空间关系抽取的，因此，完全可以根据每个步骤的实验结果得到空间关系抽取的结果。

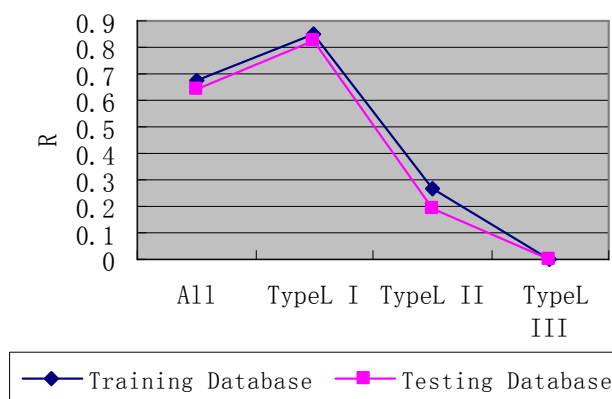


图 3-21不同句法位置下的结果

Figure 3-21 The results on the different syntactic location relations

前几个小节中，已经得到空间表达式、界标以及射体等步骤的实验性能，但是，每个步骤都是以无错误累计为前提的，例如：空间表达式边界抽取是界标抽取的前端步骤，然而在3.7.3.2节的实验中，是以空间表达式边界抽取100%正确为前提，得到在封闭测试下界标抽取最好的F值是97.49%，因此，根据乘法原理和加法原理，利用公式(3-12)和公式(3-13)计算空间关系抽取的性能。

$$P = \min_i(P_i) \quad (3-12)$$

$$P_i = \prod_{j=1}^n P_j' * P_i' \quad (3-13)$$

其中，符号' P '表示真实性能指标；' P' '表示无错误率下的性能指标；性能指标包含F值，准确率和召回率，例如：在封闭测试下，空间关系抽取的F值计算方法是：首先，得到各个步骤的无错误率下的F'值：空间表达式边界抽取的F'值87.47%，界标抽取的F'值97.49%，界标恢复的F'值90%，介词抽取的F'值100%，方位词抽取的F'值95.96%，射体抽取的F'值63.16%；然后，根据公式(3-13)得到每个步骤的真实F值：空间表达式边界抽取的F值87.47%，界标抽取的F值是85.25%(即87.47%*97.49%)，界标恢复的F值78.72%(即87.47%*90%)，介词抽取的F值87.47%(即87.47%*100%)，方位词抽取的F值83.94%(即87.47%*95.96%)，射体抽取的F值55.25%(即87.47%*63.16%)；最后，根据公式得到空间关系抽取的真实F值55.25%，即所有步骤的真实F值的最小值。同理可得，空间关系准确率和召回率两个指标，分别是52.93%、57.9%。

另外，上面的所有实验中都是以词性标注的真实实验性能为前提的，空间关系性能计算也是以其真实性能为前提的。

3.8 本章小结

在对简单句子级别空间关系抽取方法进行分析的基础上，本章提出了基于线性分类器的从真实文本中，在篇章层次上抽取空间关系方法。与传统的句子级别抽取方法相比，该方法只用词性标注的结果就可以达到篇章层次认知概念抽取的目的，特别适用于句法分析尚不深入的语言的关系抽取。与基于句子级别的抽取模型相比，该方法大大提升了关系抽取处理语料的真实性，同时在一定程度上保证了关系抽取的认知合理性。本文得到的空间关系抽取结果，可以用于实现多物体在三维空间中，自动优化摆放，为进一步的多物体空间优化摆放研究打下了很好的基础。

第4章 基于区域确定方法的三维物体自动摆放

4.1 引言

三维物体摆放是计算机辅助设计工具研究中的一个非常基础又极难解决的问题。对于面向文本的描述性建模系统来说，这一问题尤为突出。描述性建模系统的实现需要多方面的约束、涉及到拓扑关系生成、几何关系生成、以及物理关系等各个方面的处理。传统的描述性建模系统中，建模工具通常提供基于菜单的人工方式构建场景。随着场景规模的扩大，这种人工构建方式逐渐暴露出其效率低、成本高以及人员艺术素质要求高等弱点，已经成为制约场景自动生成系统发展的瓶颈问题。

随着自然语言处理技术的日益成熟，场景自动生成的研究者逐渐意识到，如果能够有效地、自动地摆放空间中的物体，将会对解决场景自动生成系统的场景建模瓶颈问题、提高场景自动生成的效率具有重要意义。在前面的两章中，本文研究了基于机器学习的空间本体库的构建以及基于线性分类器的空间关系抽取技术，其目的正是为了实现约束的自动获取，以能够真正实现三维物体自动摆放。本章将在此基础上研究多物体在三维空间中自动摆放策略。通过本文的研究，希望建立起一个面向真实文本描述性建模系统。这个系统输入是真实文本，通过对其进行篇章级别的空间关系抽取，结合空间本体库，进而实现文本描述的角色在三维空间中自动摆放，保证平台的输出是文本描述空间关系对应的静态三维场景。这样的静态的三维场景可以被虚拟现实以及增强现实系统所利用，最终达到提高场景建模效率、降低成本、缩短虚拟现实以及增强现实系统构建周期的目的。

在空间本体库和文本中空间关系抽取基础上，本章研究多三维物体自动摆放。针对约束类型的差别，提出了基于区域确定方法的自动摆放策略，解决了传统摆放策略中人工理解约束的难题，同时保证了多个物体在同一区域中自动优化摆放的可行性。本章第2节对现有的物体摆放方法进行了分析；第3节提出了基于区域确定的三维物体自动摆放策略。这个策略是指，首先基于空间本体库给出物体摆放区域的边界值，称摆放区域确定；然后基于最优保持遗传算法精确计算物体摆放方位，称优化摆放；最后对实验结果进行了分析和总结。

4.2 摆放技术的概述

物体摆放(Object Layout),就是指根据一定格式的约束生成一个或者几个物体在空间中具体方位的技术。约束表达了物体之间必须维持的某种描述性关系,包含拓扑约束(Topologic Constraints)、几何约束(Geometric Constraints)以及物理约束(Physical Constraints),例如:“书在桌子上”是一个拓扑约束;“狐狸在乌鸦前面”是属于几何约束中的朝向约束;而重力约束是典型的物理约束。

4.2.1 摆放技术的主要方法

目前,摆放方法可以被分为两大类:构造方法(constructive approach)和迭代方法(iterative inprovement)^[34]。构造方法就是根据一系列条目完成一个摆放,其实质是枚举过程,具有确定性和完整性,这种方法对于处理低约束问题十分有效,耗时短,但是容易导致不令人满意的摆放;而迭代方法开始于一个初始摆放布局,然后在以降低代价为前提下,重复修正摆放,直至代价不再改变为止,对复杂情况下物体摆放有效、灵活,缺点是耗时。

下面介绍四种典型的物体摆放方法:

(1) 基于遗传算法三维物体摆放方法。近年来,有些研究者都先后基于遗传算法,提出建筑上的、空间的进化系统^[49,92,,93,94,95,96,97],其中基于遗传算法实现的DEM²ONS系统和MultiCAD-GA系统较为典型。其主要经过两个步骤实现了摆放问题:第一步是,初始摆放布局的生成;第二步是,基于典型的遗传算法得到较满意的摆放布局。

生成初始布局的方法,有两种:基于约束满足问题和基于规则。在文献[49]中提到,系统MultiCAD-GA使用约束编程技术(Constraints Programming Techniques)得到遗传算法中的初始摆放布局,这样克服单纯依赖约束编程技术得到最终摆放布局的弱点:穷尽所有情况而带来的超常运行时间。文献[92]介绍系统DEM²ONS是使用域过滤技术(Domains Prefiltering)作为遗传算法的初始步骤,也就是物体是由层次包围盒组成,例如:房间由一个大的包围盒(Bounding Box)表示,而其中桌子等是这个大包围盒下面的小包围盒表示,因此,当摆放只涉及房间这一层,则域过滤技术就把桌子这一层次的包围盒过滤掉了,即不参与约束计算。这样可以使后期计算量减少,同时保证了对层次约束的处理。

基因编码(Genome Coding)是遗传算法中最核心的步骤,能够映射出所能处理的问题的本质。系统MultiCAD-GA中,对六个参数编码,即包围盒左下前定点的坐标(x, y, z)以及表示包围盒长、宽、高的 $\Delta x, \Delta y$ 和 Δz ,并且用二进制编码。图 4-1显示了在假设参数取值范围在 3 的情况下的基因编码。

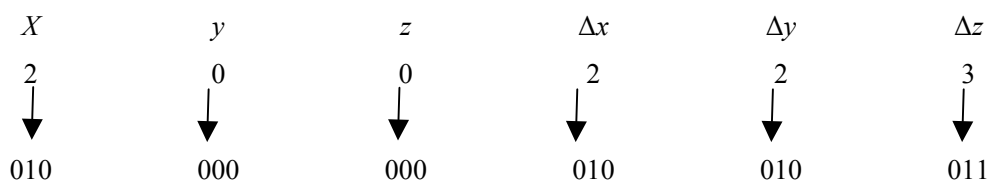


图 4-1 MultiCAD-GA 系统中一个场景对应的基因表示

Figure 4-1 Genome representation of a scene in MultiCAD-GA system

系统DEM²ONS对几何约束一区域(Zone)、朝向、距离和尺寸进行编码，而且这些约束组成一个二进制树，即约束树，其中叶节点是约束，其它节点是二值逻辑操作节点，下面的约束序列对应的约束树如图 4-2所示。

- C1 = Plate ZONE(on) table
- C2 = Plate ORIENTATION(pi) chair#1
- C3 = Plate ZONE(in_front) chair#1
- C4 = Plate ORIENTATION(pi) chair#2
- C5 = Plate ZONE(in_front) chair#2
- C6 = Plate NOT ZONE(ON) vase

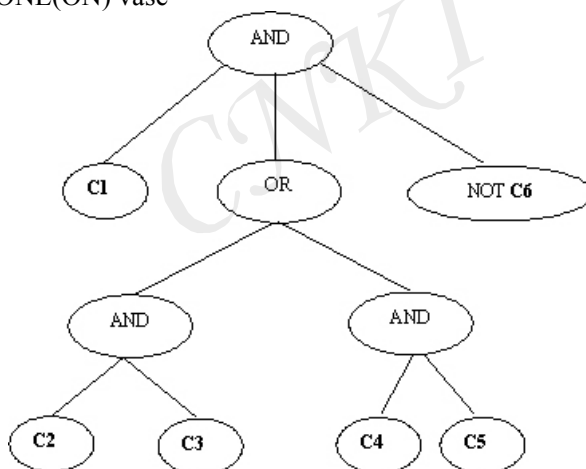


图 4-2 约束树示例

Figure 4-2 The sample on constraint tree

适应度函数(Fitness Function)选择是解决过程最关键的问题，是遗传算法能够收敛到有价值解的唯一保障。适应度函数一般包含三个层次：基因适应度函数、染色体适应度函数以及群体适应度函数。系统MultiCAD-GA使用交换方式，让用户给出本次摆放布局的满意度，即群体适应度值。而系统DEM²ONS中，一

个约束一个具体的适应度函数，例如图 4-2对应的基因适应度函数是公式(4-1)；染色体适应度函数是公式(4-2)；群体适应度函数是公式(4-3)。其中， C_i 是基因对于约束 i 的适应度值； p_i 是约束 C_i 的优先因子。

$$f_g = \text{AND}(p_{C1} * C1, \text{OR}(\text{AND}(p_{C2} * C2, p_{C3} * C3), \text{AND}(p_{C4} * C4, p_{C5} * C5)), p_{C6} * \text{NOT}(C6)) \quad (4-1)$$

$$f_{ch} = \frac{\sum f_g}{nb_{Genes}} \quad (4-2)$$

$$F = \frac{\sum p_{ch} * f_{ch}}{\sum p_{ch}} \quad (4-3)$$

如同所有的启发技术一样，这里的简单遗传算法的参数、群体大小、变异率等是问题相关的，同时不能保证全局收敛。

(2) 基于描述器摆放方法。描述器(Depictors)方法是B.Coyne和Sproat在WordsEye系统中提出的，其实质属于构造方法。文献[3]把由低级图形详细说明组成的集合称为描述器，它用于最终定义所有场景，当然也控制着三维物体的位置和朝向，也就是摆放问题。描述器中的空间标记(Spatial Tags)用于描述空间关系，标记实质是简单的、可填充的三维物体，其伴随数据库中的每个物体，表示“内部”、“上表面”等代表的空间，例如：对于“椅子上”这样的拓扑约束，在图形库中椅子这个三维模型的坐面上有一个长方体描述器，其直观精确的表示了“上”这个拓扑约束，如图 1-8所示。

在物体摆放的时候，首先，在图形库中找到对应的三维物体；然后，找到和这个物体关联的空间标记，如果存在与方位词对应的标记，那么就用射体把标记换下来；以此类推，最后得到一个摆放布局。

这种方法可以保证约束的极端精确性，能够最大可能的消歧方位词的含义，例如“墙上”，以提高场景的真实性。然而，通过手工来标记图形库中的所有物体，是一件十分耗时、令人不愉快的工作。

(3) 基于连续域模型摆放方法。一些研究者把用于摆放的区域形式化为一个连续域，而其中的每个点都附有和具体方位词相关的权重，例如：Mukerjee提出“前面”对应的区域是连续的，而且区域中的点属于“前面”模糊集合，有不同的隶属度，由此就可以得到摆放的全局最佳点。因此，此方法属于数字最优化方法。

构造连续域计算模型的典型方法有两种：Socher等人在文献[98]中是使用几个线形函数构造一个连续域；而Yamada、Mukerjee等人在文献[44]、[1]、[99]

和[100]中分别只使用一个非线性函数定义连续域。无论哪种方法，核心问题是：模型的具体定义以及其全局最小值的获取。Yamada基于胡克定律以及人类空间认知，使用弹簧弹性势能函数构造连续域，取弹簧势能最小点为全局最优点。Mukerjee提出使用椭球域模型，表示物体的位置以及朝向，并且通过心理实验来决定全局最小值，即物体摆放的点。

对于空间约束的处理，这两种具体模型都是经过两个步骤，首先，对每种空间关系，也就是日语或者英语中的空间介词，分别构建一个独立的、以界标为参照的连续域；然后，对于组合约束，Yamada使用势能函数叠加方法解决，Mukerjee的方法是在计算全局最小值的时候，考虑到相关域的全局最小值，而不是整个椭球域。

连续域模型方法把摆放问题形式化为求函数的极值问题，可以形式化表示模糊，如“附近”，和精确的空间关系。然而，连续域模型的具体定义是这个方法的关键，模型必须有能力把物体约束到空间具体部分、必须足够的平滑、必须减少陷入局部最优的概率等等。显然，找到这样一个非线性函数是相当困难的事情。目前，研究者们只是提出模型，然后用可视化场景说明其可行性，我们还没有发现关于这个问题较形式化的证明。

(4) 基于知识的摆放方法。陆汝钤院士提出基于知识的两步法完成场景布局^[2]，第一步是基于知识的启发式原理，设计定性的场景规划，包含：一、确定人物分布的基本场景，这些场景是从已有的电影和动画片总结出来的；二、根据基本场景确定人物的定性空间分布，也就是从基本场景库中得到空间分布；第二步是基于规则的定量位置规划，规则的依据是人物、物体和环境的几何特征，这些特征被保存在常识库中。

这个摆放问题只是陆汝钤院士带领开发的天鹅系统中的很小的一个功能。这个方法能够保证动画的真实性，避免失真。但是，建立这样的高标准常识库是一般研究者可望不可及的。

4.2.2 物体摆放需要解决的关键问题

基于描述的物体摆放要比基于菜单的，更加复杂、更加困难。主要由描述模糊性造成的。基于描述物体摆放的关键特征是从高级描述得到场景，例如脚本、类prolog语言等高级描述。模糊性是这类描述的本质属性，也就是给出空间关系信息都是不确定的。例如在句子“狐狸站在树下”中没有任何关于“树”、“狐狸”的具体位置、朝向的定量信息。然而，定量信息是摆放问题的必要条件。因此，从4.2.1节的概述可以看出，很多研究者从离散或者连续的角度出发，

提出不同方法把描述模糊性转换为精确性。

在任何摆放方法中都面临以下的难点：

(1) 背景知识的获取。基于多么高级的场景描述都不会造成人类之间交流的困难，但是却是人类和计算机之间交流的大障碍。这主要是由于人们之间在交流的时候拥有共同的情景、角色等背景知识。因此，如何准确的定义、规范以及获取这类背景知识是基于描述物体摆放问题的瓶颈问题。

(2) 约束信息的规范。物体摆放问题的实质就是基于约束的编程问题，那么，约束信息的数量、范围以及相互关系决定着最终摆放的合理性，也就是实例摆放布局是否能够尽最大可能反映描述的本意。约束信息的充分性和获取难度之间的矛盾是不言而喻的。因此，在不增加约束信息获取难度的前提下，如何在尽可能少的约束下，实现更加合理的摆放就成为了研究的难点。目前，还没有学者具体提出这样的矛盾。

从上面的分析可以看出，如果可以在得到摆放区域严格边界基础上，再结合优化问题求解方法，那么可以降低对背景知识涵盖度的依赖性、减少约束信息的数量，实现多物体的三维空间中自动摆放。在第2章已经阐述了空间本体库的构建，因此基于这个本体库，提出基于区域确定的物体摆放技术。首先，利用空间本体库，以空间关系抽取结果为前提，得到每个物体在空间中摆放区域的边界，即摆放区域确定；然后，以摆放区域为空间关系基本约束信息，结合其它的约束信息，基于遗传算法，确定每个物体空间中的准确方向和位置，即三维物体优化摆放。

4.3 基于空间本体库的摆放区域确定

摆放区域确定是在空间关系抽取提供的描述性脚本基础上，结合空间本体库，得到区域空间，缩小物体摆放区域约束。在本文中，描述性脚本就是空间抽取的结果，其格式就是3.3.2.1小节中空间关系的二元组($[[Pre], Loc, LM), TR$)，例如： $((在, 中, 空), 鹰)$ 就是典型的描述性脚本，确定“鹰”所在空间区域的实质也就是量化方位词“中”。因此，为了符合人类空间认知，引入空间语义消歧规则和基于空间本体库的方位词量化模型。

4.3.1 空间语义消歧

在上下文中，同一个空间语义可以用不同的方位词表达，反之，同一个方位词可以表达不同的空间语义，例如：在这两个句子中，“车里有五名乘客”和“车上有五名乘客”，“里”和“上”表达的一个空间语义，即车厢内部；

而在这两个句子中，“车上有五名乘客”和“山上有颗树”，“上”表示不同空间语义，前者表达车厢内部，但后者表达山上部。这里歧义是由于同一个方位词同时具有多重空间属性造成的，因此，在具体的上下文中，我们采取基于规则的方法，对方位词属性消歧。

规则如下：

- (1) If 界标是 2D then 方位词取外部属性；
 - (2) If 界标和射体是部分整体关系 or 界标可以做射体的容器 then 方位词取内部属性 else 方位词取外部属性
 - (3) If 界标是 2D and 法线方向= z 负向 then 方位词“上”替换为“下”；
- 空间语义消歧后，方位词表达的空间语义就是唯一的了，为下一步具体计算模型的选择提供了前提。

4.3.2 方位词计算模型

在汉语中，空间方向和位置主要是由方位词表达的。同时，根据人类空间认知，方位词各自表达的方位不是独立的而是关联的。因此，首先提出方位词分类体系，然后，以空间本体库(2.4.2.2 节)为基础，结合叠加策略，给出方位词计算模型之间的关系。

4.3.2.1 方位词分类 根据方位词在空间概念系统中地位[101,102]，本文把方位词分为单纯方位词和组合方位词。单纯方位词表述的方位是空间概念体系的基础，包含“东”、“南”、“西”、“北”、“上”、“下”、“前”、“后”、“左”、“右”、“里”、“外”和“附近”13个方位词，其中前10个表达的方位是对称的，最后3个表达相对范围。组合方位词一般是双音节，同时表达了两个基本空间概念，例如：“右上”是由“右”和“上”两个基本概念组成的。

为了得到这个分类结果，基于层次法和k-均值聚类相结合的模型，把附录1中的方位词进行分类。层次步骤如下：

- (1) 手工聚类绝对方位词。基于空间认知，把符合绝对方向概念的方位词，手工分类，得到“东”、“南”、“西”、“北”四个子类；
- (2) 使用k-均值聚类算法，聚类相对范围方位词。使用算法 2-1，以“里”、“外”和“附近”为种子，把步骤(1)中外的方位词分成“里”、“外”、“附近”和“其它”四个子类别；
- (3) 使用k-均值聚类算法，聚类对称性方位词。使用算法 2-1，以“上”、“下”、“前”、“后”、“左”、“右”为种子，把步骤(2)“其它”子类别中方位词聚类，分为“上”、“下”、“前”、“后”、“左”和“右”六个

类别;

(4) 手工归类知网 2005 版本中不包含的方位词。

最终得到 13 个子类, 如表 4-1 所示。其中只属于一个类别的方位词是单纯方位词, 而同时属于多个类别的方位词是组合方位词。

表 4-1 方位词空间概念体系下的聚类结果

Table 4-1 The results of classifiers of localizer on spatial concept

方位词实例 类别名称	
东类	东 东面 东头 之东 东边 以东 东部 东端 东侧 东南边 东北边 东南面 东北面 东南部 东北部 东南方 东北方 东南角 东北角 东南 东北 东西
西类	西 西面 西头 之西 西边 以西 西部 西端 西侧 西南边 西北边 西南面 西北面 西南部 西北部 西南方 西北方 西南角 西北角 西南 西北 东西
南类	南 南面 南头 之南 南边 以南 南北 南部 南端 南侧 东南边 西南边 东南面 西南面 东南部 西南部 东南方 西南方 东南角 西南角 东南 西南
北类	北 北面 北头 之北 北边 以北 南北 北部 北端 北侧 东北边 西北边 东北面 西北面 东北部 西北部 东北方 西北方 东北角 西北角 东北 西北
上类	上 上面 跟前 上端 左上方 右上方 左上角 右上角 上头 上边 上下 上部 面 顶 之上 顶上 面上 左上头 右上头 左上部 右上部 前上方 后上方
下类	下 下头 之下 下面 底下 跟前 下端 左下方 右下方 左下角 右下角 下边 上下 下部 左下头 右下头 左下部 右下部
左类	左 左边 左面 左右 跟前 左侧 左上方 左下方 左上角 左下角 前后左右 左端 左上头 左下头 左上部 左下部
右类	右 右边 右面 跟前 右侧 右上方 右下方 右上角 右下角 前后左右 右端 右上头 右下头 右上部 右下部
前类	前 前头 之前 前边 前面 前后 头里 面前 前端 一头 对面 前部 头 前后左右 头前 前上方
后类	后 后头 之后 后边 后面 背后 后部 尾 前后左右 后端 后上方
里类	里头 里 中 内 里边 里面 当中 中间 内里 内部 内侧 际 中心 之中 之间 里外 内外 当间 中部 之内 内中 里端
外类	外头 外 之外 外边 外面 以内 以外 外侧 外边 里外 内外 外端
附近类	跟前 以近 附近 旁 旁边 边上 这儿 这里 那儿 那里 以远 周围 一边 一面 两边 两头 两旁 四面 四外 四周 边 侧 四围 四周围 两旁边 四边

4.3.2.2 单纯方位词计算模型 依据的基本计算模型，通过集合计算得到方位词的计算模型。

根据对称性计算“南”、“西”、“北”三个单纯方位词的计算模型。它们与“东”模型(见表 2-21)之间的关系满足依次绕z轴顺时针旋转 $\pi/2$ 弧度，公式(4-4)完成模型之间的转换， (x, y, z) 是当前模型的点，例如：由*basic_east*定义的“东”模型中的点， (x_p, y_p, z_p) 是下一个模型的点，例如：即“南”模型中的点。“下”的计算模型与“上”模型之间满足平面对称，使用公式(4-5)，可以从空间本体库中，即*basic_on*定义模型推导出“下”计算模型。同理，“前”和“后”、“左”和“右”也满足平面对称，推导公式和公式(4-5)类似。

$$\begin{aligned} x_p &= x * \cos(\pi/2) + y * \sin(\pi/2) \\ y_p &= y * \cos(\pi/2) - x * \sin(\pi/2) \\ z_p &= z \end{aligned} \tag{4-4}$$

$$basic_under = \begin{cases} x_p = x \\ y_p = y \\ z_p = z \end{cases} \tag{4-5}$$

根据相对性，“外”表达的空间是“里”(见表 2-21)的补空间，且与界标外表面的距离满足 d ，由公式(4-6)得到，而“附近”表达的空间是“里”和“外”的合并空间由公式(4-7)得到。这里根据集合理论，引入交、并、补运算，采用的符号分别为 \cap 、 \cup 、 $-$ 。全集是整个三维空间；模型之间交运算计算结果是两个模型共同拥有的空间；并运算的结果是模型分别拥有的空间；补运算结果是不属于模型表达的空间。

$$basic_out = -basic_in \text{ 并且 } d = 0.6 * ht / 2 \tag{4-6}$$

$$Basic_near = basic_in \cap basic_out \tag{4-7}$$

4.3.2.3 组合方位词计算模型 根据组合方位词的定义，叠加其包含的基本空间概念对应的模型。叠加计算即交运算，计算规则如同4.3.2.2节，例如：“右上”的区域是 $basic_right \cap basic_on$ 。

4.4 基于 OMSGA 的优化摆放

物体优化摆放是物体摆放的第二个步骤，是以确定摆放区域步骤得到的区域，作为位置约束范围，使用最优保持简单遗传算法(Optimum Maintaining Simple Genetic Algorithm, 简称 OMSGA)，得到场景中每个物体的具体位置和朝向区域空间，为摆放可视化提供定量数据。

最优保持简单遗传算法就是在简单遗传算法中，加入精英模型。精英是当

代中具有最优适应度值的个体，精英模型就是把精英选择到下一代，把优秀的个体遗传的模式。

恽为民等人在文献[103]中已经证明OMSGA是全局收敛的，而简单遗传算法不是全局收敛的。

使用OMSGA求解物体优化摆放，需要解决的关键问题包含：物体优化摆放问题编码、适应度函数的选则以及遗传操作的具体定义。

4.4.1 优化摆放问题编码

物体优化摆放实质是在摆放解空间中寻找最优解。固定空间区域内，需要摆放多个物体，每个物体具体的位置和朝向组合在一起，构成了物体摆放的一个解，这样解构成了解空间。例如：对于“树下有乌鸦和狐狸”这样的摆放要求，乌鸦位于(0,0,0)，与 z 轴的夹角是12弧度，以及狐狸位于(0,5,0)，与 z 轴的夹角是-12弧度就构成了一个具体的摆放，也就是，这样两对位置、朝向的组合，以及无重叠、符合重力约束就是摆放解空间中的一个解。最优解就是满足一定约束的解。

在遗传算法中，有特定术语描述解、解空间等。解空间就是遗传算法要搜索的空间；解空间可以被分解为多个相交或者不相交的子解空间，作为每一次的搜索范围，这样的子解空间就是群体(Population)；解就是个体(Individual)，是一个符合一定码制的符号串，代表着最优问题的一个解决方案，组成群体；每个待摆放的物体就由一个染色体(Chromosome)描述；染色体由基因(Gene)组成，每个基因表达一个约束。

4.4.1.1 约束分类 受Sanchez等人在文献[104]中提出约束的启发，我们主要选择低级约束，利用这些约束的组合来对应高级描述。低级约束是指，其可以独立的描述一个局部摆放问题，可以被很容易的解决，分为两类：物理约束和几何约束。

物理约束能够保证生成场景的连续性和真实性，包含无重叠约束和重力约束。无重叠约束避免物体之间的碰撞，也就是物体没有互穿透性；重力约束是指，对于非飞行的物体，一定要有一个支撑面，而对于飞行的物体，则不能超出一定范围，即每个物体都受地球引力的限制。

几何约束的目的是描述每一个简单的、基本的空间问题，包含位置约束、朝向约束和尺寸约束，位置约束就是，在摆放区域确定步骤给出的取值范围内，物体中心点的三个三维坐标值；朝向约束对于都前后概念的物体，是指其正面和水平线正向的夹角，而对于无前后概念的物体，我们任意选择的面和水平线正向的夹角；尺寸约束就是指物体本身的大小。

4.4.1.2 个体编码 在摆放问题中，我们定义个体就是空间中，所有待摆放物体的一次摆放布局。

个体编码就是在十进制上，使用基因表达摆放。为了降低搜索空间，提高搜索速度，不是把每个约束都有相应的基因表达，而是针对部分约束编码。

优化摆放问题需要处理 5 种约束：无重叠约束、重力约束、位置区域约束、朝向约束和尺寸约束，由每种约束特性决定是否由一个或者多个基因表达。

尺寸不用基因表达。本文中，使用包围盒表达物体的尺寸，并且保存在图形库中，所以物体具体尺寸可以直接从图形库得到。

位置使用 2 位基因表达。首先，从摆放区域确定的空间范围中，可以直接得到高所在轴的坐标值；然后，利用十进制，对水平面的两个坐标编码，因此，位置用 2 位基因表达，取值范围由摆放区域确定步骤给出。

朝向使用 1 位基因表达。此基因采取十进制编码，取值范围在区间 $[0, 2\pi]$ 。

无重叠不使用基因表达。无重叠约束的本质不是针对独立个体的，而是描述物体之间的约束，所以不用独立基因表示，而是在适应度函数中体现。

重力约束也不使用基因表达。其实我们已经在摆放区域确定中考虑了重力约束，也就是摆放区域一定会提供一个摆放平面，即符合重力的支撑面。因此，个体编码步骤中不考虑重力约束。

总之，每个物体的一次摆放由 3 位十进制基因表达，如果一次摆放的物体个数是 N ，那么个体包含 $3N$ 位十进制基因，例如，本节的例子对应的个体是 $(x_1, y_1, \theta_1, x_2, y_2, \theta_2)$ 。可见，这种编码方式大大减少了搜索空间。

4.4.2 适应度函数定义

适应度函数，也就是评价，能够衡量摆放的合理性、可接受性以及最优性。在本文中，适应度函数由基因适应度函数和个体适应度函数组成。基因适应度函数，是评价每个物体的摆放优劣性的，定义为，物体在水平面上的投影面积与该物体和水平面相切表面的面积的比值，标记是 $F(O_i)$ ，计算公式是公式(4-8)，图 4-3 给出抽象图示， i 表示第 i 个物体。个体适应度函数，是评价一次摆放，即全体物体的摆放，优劣性，定义为，在符合无重叠约束情况下，个体中全部基因适应度的平均值，标记是 F_k ，计算公式是(4-9)， k 表示第 k 个个体， L 是无重叠约束因子，由公式(4-10)计算得到， v_i 是第 i 个物体投影平面内的点集合。

$$F(O_i) = S(OH) / S(OG) \quad (4-8)$$

$$F_k = L * \left(\sum_{i=1}^N F(O_i) \right) / N \quad (4-9)$$

$$L = \prod_{i=1}^N v_i = \begin{cases} 1; & \text{交集为空的时候} \\ 0; & \text{交集不为空的时候} \end{cases} \quad (4-10)$$

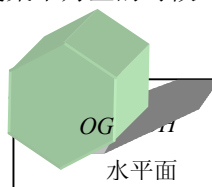


图 4-3 基因适应度公式的抽象图示

Figure 4-3 The graphic of the gene fitness function

4.4.3 操作算子定义

遗传算法在理论上包含许多操作算子，在摆放问题中，我们选用的是最优保持简单遗传算法，包含单点交叉、定概率变异和轮转法选择三个操作算子，具体计算方式如下：

(1) 单点交叉。首先，根据定概率 P_c ，选择当代中一对个体，即双亲；然后随意产生一个交互位置；最后，将双亲基因链上这个位置的基因交换，繁殖一个新的个体，即后代(Offspring)，如图 4-4所示。

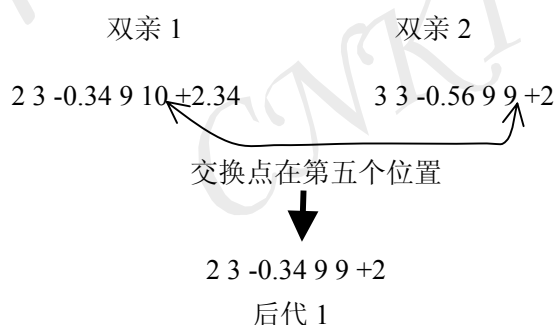


图 4-4 单点交叉运算图示

Figure 4-4 The traditional method of crossover

(2) 定概率变异。在交叉产生的后代中，首先，根据定概率 P_m 随机产生一个基因位置；然后，在取值范围内任意产生一个实数，替换这个位置的基因。变异算子在一定程度上，保证了新的摆放策略的生成，弥补了初始布局的局限性，减少了陷入局部最优的概率。

(3) 轮转法选择。选择算子，就是在由当代和后代组成的一代中，选择一定数量的个体组成下一代的操作。轮转法选择，就是个体按照与其相对适应度成正比的概率被选入下一代。个体相对适应度函数是公式(4-11)， F_i 表示第*i*个个

体适应度， M 表示群体的大小。这种选择算子，保证了最优个体的遗传，一定程度的保障了全局最优解的获得，是最优保持简单遗传算法的经典之处。

$$RF_i = F_i / \sum_{j=1}^M F_j \quad (4-11)$$

4.5 基于受限语义的图形检索

使用前面几节阐述方法可以完成物体空间概念建模，即得到待摆放物体在三维空间方位，然而，为了实现可视化，还必须得到可视表示物体的模型，即三维模型的检索。本文实现了基于受限语义的图形检索方法。

图形检索是个比较成熟的技术，在需要领域都有所应用，主要包含两种方法，一、基于内容检索。这种检索方法是，基于图形、图像的颜色、形状以及纹理等特征进行检索。显然，使用这种方法的前提是，可以提供受检索图形、图像的草图；二、基于语义检索。这种检索方法是，基于语义相似度进行检索，语义相似度是指，检索关键字和配在图形、图像上的文字说明之间的的语义相似度，例如：Lycos、Google系统都使用这种检索方式^[105,106]。针对本文中空间关系可视化系统是处理文本对象的，不能提供图形草图，而且，在图形数据库中，没有充分的配图文字说明，我们提出基于受限语义的图形检索方法。这个方法的核心是相似度，即计算检索词和图形文件名称之间的词语相似度，包含语义相似度和模糊相似度。这个方法大大降低了标注配图文字的工作量，提高了检索速度，保证了检索质量。

4.5.1 检索词语和文件名的结构

检索词语就是，用户用于描述要得到图形的特点的词语。在本文中，就是表达射体或者界标的词语，例如：“狐狸”、“小狐狸”等。文件名就是，每个三维图形文件的文件名称。因为，本文是以汉语为研究对象，所以文件名是以汉语命名的，而且其反映了图形的特点，例如：一个乌鸦三维物体，其文件名是“黑色乌鸦”。可见，检索词语和文件名实质是一个短语，为了区分修饰部分和核心部分，我们定义其结构。

首先，对检索词语和文件名分别分词、词性标注，然后进行短语分析。这样检索词语的结构是 $kw(ap, cw)$ ， ap 表示检索词语的附加部分， cw 表示检索词语的中心词；文件名的结构是 $fn(fap, fcw)$ ， fap 表示文件名的附加部分， fcw 表示文件名的中心词。

4.5.2 词语相似度

词语相似度是指检索词语和文件名之间的相似度，有中心词相似度和附加词相似度两部分组成。 $sim_w(c1, c2)$ 表示词语相似度、 $sim_w(cw, fcw)$ 表示中心词相似度、 $sim_w(ap, fap)$ 表示附加词相似度，它们之间的关系如公式(4-12)所示， α 表示权重，衡量各部分对相似度的贡献度，下面这两部分都统称为词相似度。由于语义库的容量是有限的，不能包含所有的词语，因此，我们提出分情况计算词相似度，当词语(中心词或者附加词)包含在语义库中，基于语义库的语义相似度得到词相似度；当词语(中心词或者附加词)不被包含在语义库中，基于词形的模糊相似度得到词相似度。

$$Sim_w(c1, c2) = sim_w(kw, fn) = sim_w(cw, fcw) + \alpha * sim_w(ap, fap) \quad (4-12)$$

4.5.2.1 基于语义库的语义相似度 首先，可以根据具体的语义库，计算词语之间的语义距离。如果同义词词林是选用的语义库，那么利用文献[107]方法计算语义距离。

《同义词词林》是现代汉语比较常用的一部义类词典。它所收词语全部按语义分类编排，例如，“农民”的词义代码为Ae07，表示其处于A大类、e中类、07小类。可以把整个语义分类体系想象成一棵语义树，根节点的儿子是所有大类，某个大类的儿子是他下属的中类，叶子节点为各个小类。两词义 S_1 与 S_2 之间的语义距离 $SenseDist(S_1, S_2)$ 可以定义为语义树中结点 S_1 到结点 S_2 的最短路径的长度，通过比较两个词的语义代码即可计算出它们的语义距离。如：

$$SenseDist(Aa02, Aa01) = 2;$$

$$SenseDist(Ab01, Aa01) = 4;$$

$$SenseDist(Ba01, Aa02) = 6;$$

显然， $SenseDist(S_1, S_2)$ 越小， S_1 与 S_2 就更可能同处于小的类中，因此他们在语义上越相似。定义词义 S_1 与 S_2 的语义相似度为：

$$SenseSim(S_1, S_2) = \begin{cases} 1 / SenseDist(S_1, S_2) & S_1 \neq S_2 \\ 1 & S_1 = S_2 \end{cases} \quad (4-13)$$

在此基础上，定义两个汉语词 c_1 和 c_2 的语义相似度为：

$$sim_s(c_1, c_2) = \max_{\substack{S_m \in \text{Classof}(c_1) \\ S_n \in \text{Classof}(c_2)}} SenseSim(S_m, S_n) + \beta * \left(\frac{\text{count}_{\substack{S_m \in \text{Classof}(c_1) \\ S_n \in \text{Classof}(c_2)}} (SenseSim(S_m, S_n) - h) - 1}{\text{count}_{\substack{S_m \in \text{Classof}(c_1) \\ S_n \in \text{Classof}(c_2)}} (SenseSim(S_m, S_n) - h) - 1} \right) \quad (4-14)$$

其中， $\text{Classof}(c)$ 函数返回词语 c 的语义代码集合， $\text{count}(c)$ 函数返回词 c_1 和 c_2 的语义相似度大于 h 的语义个数。

4.5.2.2 基于词形的模糊相似度 很多时候,所用的语义库中不含有检索词语或者文件名,例如:“鸟巢”就不被知网 2004 版收录,那么,我们提出根据词形的模糊相似度,定义两个汉语词 c_1 和 c_2 的模糊相似度为:

$$sim_m(c_1, c_2) = (2 * |c_1 \cap c_2|) / (|c_1| + |c_2|) \quad (4-15)$$

其中, $|c_1|$ 表示汉语词 c_1 中包含的汉字的个数, $c_1 \cap c_2$ 返回汉语词 c_1 和 c_2 共同含有的汉字的集合。例如:汉语词“鸟巢”和“鸟”的模糊相似度的计算是,首先,它们两个共同的含有的字是“鸟”,即“鸟巢” \cap “鸟”的结果是“鸟”;然后, $|鸟巢|$ 的结果是 2, $|鸟|$ 的结果是 1;最后, $sim_m(鸟巢, 鸟)$ 的结果,模糊相似度是 0.67, 即 $2 * 1/3$ 。

4.6 实验以及分析

物体摆放实验在 Inter® Pentium® 4, CPU 2.93GHz, 1GRAM 机器上运行。所用图形库见 2.4.4.2 节描述, 语料库见 3.6 节描述。

对于物体摆放实验, 结合语料库, 设计了两组空间关系: 第一组、((在, 桌子, 上), 2 个盆花)、((在, 桌子, 上), 5 个球)、((在, 桌子, 上), 3 个瓶子); 第二组、((在, 狐狸, 右), 老鼠)、((在, 狐狸, 右), 乌鸦)、((在, 狐狸, 右), 白鹭)、((在, 狐狸, 右), 牛)、((在, 狐狸, 右), 驴)。

4.6.1 区域确定实验

对于第一组和第二组空间关系, 分别使用 4.3 节的方法确定桌子上面的所有射体的摆放区域, 即图 4-5(a) 中球上面的黄色区域, 和狐狸右侧所有物体的摆放区域, 即图 4-5(b) 中球右侧的绿色区域。

4.6.2 优化摆放实验

为了突出最优保持简单遗传算法的优点, 我们分别基于最优保持简单遗传算法、简单遗传算法对同一组空间关系做实验, 而且比较分析了最优解保持简单遗传算法的参数性能。

4.6.2.1 基于最优解保持简单遗传算法实验 对本节中系统输入的两组空间关系, 基于最优解保持简单遗传算法, 分别运行系统, 实验结果如下。参数取值分别是: 群体大小 $M = 50$, 交叉概率 $P_c = 0.8$, 变异概率 $P_m = 0.15$, 个体适应度阈值 $F_m = 0.98$ 。

(1) 第一组空间关系。因为在“桌子上”区域内摆放的实体数目一共有 10 个 ($2+5+3=10$), 而每个实体的方位由遗传算法的 3 个基因表示, 所以, 每个个体的基因数为 30 个。在实验中, 当算法迭代到第 256 子代时, 当代最佳个体

的适应度值达到 0.9816，大于适应度阈值 F_m ，算法结束，运行时间为 23 秒。得到摆放方位计算结果如表 4-2所示，可视化结果如图 4-6所示。

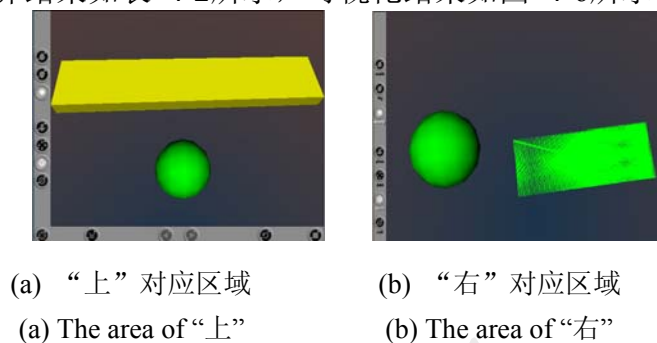


图 4-5 区域确定的实例

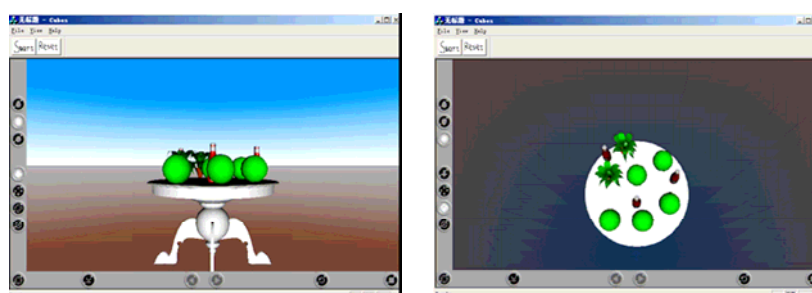
Figure 4-5 The samples of the area acquisition

表 4-2 第一组空间关系对应的摆放数据

Table 4-2 The data for the display of the first input

物体名称	水平面轴 1 坐标	水平面轴 1 坐标	朝向弧度值
盆花 1	-0.19	0.18	-0.22
盆花 2	0.37	0.28	-1.49
球 1	-0.08	-0.62	-1.45
球 2	-0.76	0.56	-0.57
球 3	0.65	0.79	-0.26
球 4	-0.50	-0.59	0.07
球 5	0.55	-0.53	1.35
瓶子 1	-0.76	-0.19	-1.40
瓶子 2	0.27	-0.18	0.087
瓶子 3	-0.11	0.62	-0.14

(2) 第二组空间关系。在“狐狸右”区域内摆放的实体数目一共有 5 个，而每个实体的方位由遗传算法的 3 个基因表示，因此，每个个体的基因数为 15 个。在实验中，当算法迭代到第 73 子代时，当代最佳个体的适应度值达到 0.9901，大于适应度阈值 F_m ，算法结束，运行时间为 9 秒。得到摆放方位计算结果如表 4-3所示，可视化结果如图 4-7所示。



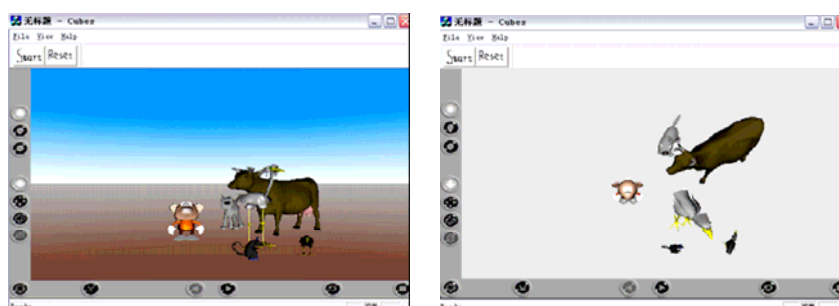
(a) 平视图 (b) 俯视图
(a) The horizontal viewpoint (b) The bird viewpoint

图 4-6 第一组空间关系的摆放场景
Figure 4-6 The scene of the first input

表 4-3 第二组空间关系对应的摆放数据

Table 4-3 The data for the display of the second input

物体名称	水平面轴 1 坐标	水平面轴 2 坐标	朝向弧度值
乌鸦	1.81	0.16	1.01
驴	0.83	0.435	-1.3
老鼠	0.20	0.15	1.41
白鹭	1.17	0.68	0.54
牛	1.65	0.73	-0.82



(a) 平视图 (b) 俯视图
(a) The horizontal viewpoint (b) The bird viewpoint

图 4-7 第二组空间关系的摆放场景
Figure 4-7 The scene of the second input

4.6.2.2 基于简单遗传算法的比较实验 对本节中系统输入的两组空间关系，基于简单遗传算法，分别运行系统，实验结果场景分别如图 4-8、图 4-9所示。其中约束包含物体中心点位于摆放区域和重力两条，

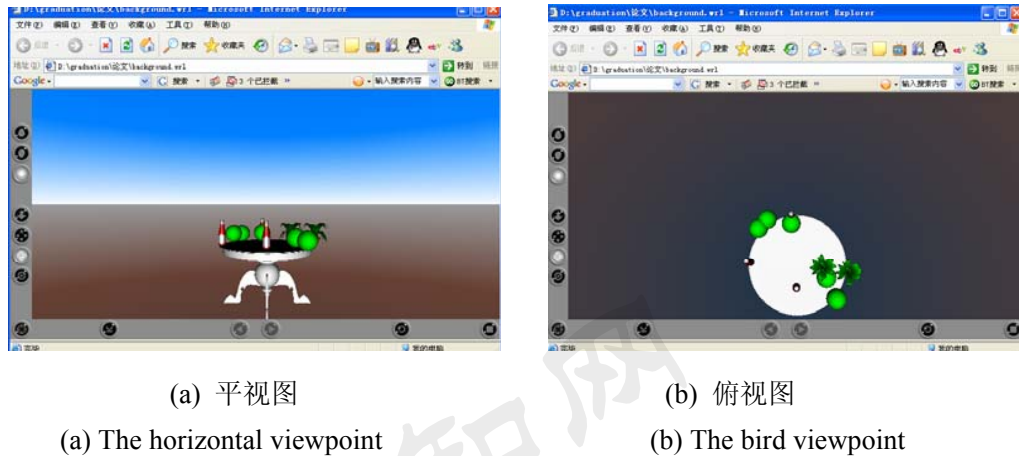


图 4-8 第一组空间关系的非优化摆放场景
Figure 4-8 The non-optimized scene of the first input

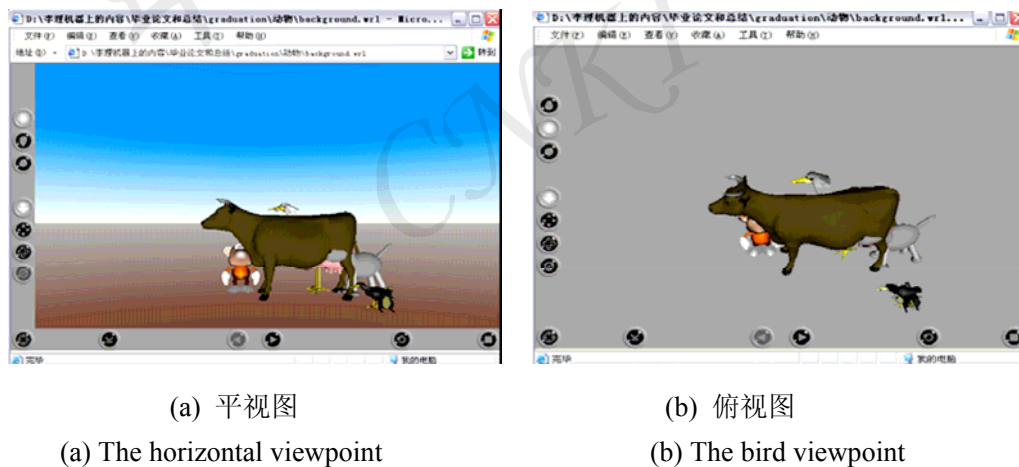


图 4-9 第二组空间关系的非优化摆放场景
Figure 4-9 The non-optimized scene of the second input

比较图 4-6和图 4-8、图 4-7和图 4-9的两组场景组，不难看出，图 4-6和图 4-7显示的场景没有重叠发生，整体布局较令人满意，而图 4-8和图 4-9显示的场景中，有明显的物体重叠现象，整体布局过于拥挤，且不美观，因此，基于最优保持简单遗传算法的、满足 5 个约束的摆放方法可以实现较满意的场景。

4.6.2.3 性能比较 目前,还没有关于摆放问题的标准评价平台,因此,这里我们把几个和摆放问题相关的、具有代表性的几个系统,从年代、方法、算法细节等角度,对它们的性能进行比较,见表 4-4所示。

表 4-4 相关摆放系统的综合比较

Table 4-4 Comparative synthesis of our study with related works

项目	DEM ² ONS ORANOS	Multi Formes4	ARCHi PLAN	MultiCAD -GA	DEM ² ONS -GA	DEM ² ONS Manhattan	This paper
文献	[92]	[104]	[108]	[49]	[109]	[29]	本文
年代	1997-1998	1999	2001	2002	2000-2003	2003	2003-2005
方法	构造	构造	构造	迭代	迭代	构造	构造+迭代
结果	不能保证 找到结果	能找到结果	能找到 结果	不能保证 找到结果	不能保证 找到结果	能找到结果	全局最优 解
物体	3D 包围盒	3D 包围盒	2D 三 角形	3D 包围 盒	3D 包围盒	3D 包围盒	3D 包围 盒
度	9	6	4	6	7	5	7
朝向	24 种正交	1 种朝向	0 或者 $\pi/2$	24 种正交	沿高轴任 意旋转	24 种正交	沿高轴任 意旋转
约束	是	否	否	否	是	否	是
算法	DHN-CSP DHN 过 滤;增量的 FC 或 RFL;动态 启发(基于 FFP)	CSP 增量的 FC 或 PL; BJ; AC5 过滤; 动态启发 (基于 FFP)	Branch and bound (最优化);动 态空间 排序启 发	SGA 处理 20 个 体;约束 编程(生成 初始群 体);简单 遗传算法	SGA 处理 300 个体;域过 滤;简单遗 传算法	S-CSP FC 或 RFL;过 滤的 SG -AC 3D; 动态启发 (基于 FFP);随机 枚举	OMSGA 基于空间 本体库的 摆放区域 确定(得到 约束范围) + 最优保 持简单遗 传算法

注: CSP=约束满足问题(Constraint Satisfaction Problem); S=空间的(Spatial); DHN=动态层次数字化的(Dynamic Hierarchical Numerical); T=后跟踪(Backtracking); FC=前检验(Forward Checking); RFL=真实全向前看(Real-Full-Lookahead); PL=部分向前看(Partial Lookahead); BJ=后跳(Backjumping); CBJ=矛盾控制(Conflict directed Backjumping); FFP=错误优先规则(Fail First Principle); AC = 弧度一致性(Arc Consistency).

4.6.3 图形检索实验

为了检验基于受限语义图形检索方法的有效性,检索性能与语义库相关性,我们分别在《同义词词林》、《同义词词林(扩展版)》和知网三个语义库上做实验。其中《同义词词林》的概述见2.2.2.2节,知网的概述见2.2.2.1节,《同义词词林(扩展版)》介绍见2.2.2.3节。

基于不同的语义库,对应的语义相似度具体计算方法有差异。基于知网语

义库的时候,首先,使用公式(2-1)计算概念相似度,即语义相似度;在基于《同义词词林》的时候,使用公式(4-14);在基于《同义词词林(扩展版)》的时候,仍然使用公式(4-14),只是Classof(c)函数返回词语c的语义代码集合中的没有语义代码用5级表示。

图形库中含有402个从网上下载的图形文件,每个文件都自动匹配文件名;检索词语来自伊索寓言前100个高频名次。

根据经验,权重 α 的值取0.5;权重 β 的值取0.1;相似度阈值 θ 取1,因为,我们认为同义词才符合检索要求。

自动检索结果,首先,由五位研究生分别人工评价;然后,综合五个人的结果,如果,有一名同学认为这个检索是错误的,那么,错误数量增加1,否则,正确数量增加1;最后使用正确率指标评价检索结果,具体结果见表4-5。

表4-5 基于不同语义库的图形检索结果

Table 4-5 The result on different knowledge database

语义库	同义词词林	同义词词林(扩展版)	知网2005
正确率	0.91	0.97	0.9795

我们分析造成错误率的主要原因有两个方面,一、语义库定义的同义词和评价者理解的同义词有差别,例如,在知网2005中,“小车”和“马车”被看作是同义词,而评价者不认为它们是同义词;二、通过模糊相似度得到的检索结果,有时不被评价者接受,例如:在知网2005中,不含有“铁锤”这个词语,通过模糊相似度找到“铁链”作为其同义词,但是不被评价者接受。

4.7 本章小结

本章在空间本体库和空间关系抽取基础上进行了物体自动摆放研究。首先,基于空间本体库确定摆放区域。利用空间本体库中关于方位的本体概念,给出现代汉语方位词的空间区域计算模型,以及方位消歧规则。确定的摆放区域大大降低了后期摆放的搜索空间,提高了速度。其次,基于最优保持简单算法实现物体摆放。在确定的空间区域的约束下,对无重叠约束、尺寸约束、位置约束以及朝向约束进行了最短基因编码,基于层次约束给出适应度函数,以及体现最优保持的遗传算子定义。此方法保证场景布局的合理性,和同等功能系统比较,有较明显的优势。最后,基于受限语义检索图形。提出结合具体情况,词语相似度有两层含义:语义相似度和模糊相似度。语义相似度用于那些包含于语义库的词语,而模糊相似度用于那些没有被包含于语义库的词语。同时词

语相似度由两个部分组成：中心词相似度和辅助词相似度。这个方法在保证正确率升高的前提下，降低检索失败率，从而降低了空场景出现的概率，满足场景生成的要求。

物体摆放是空间关系可视化的关键问题。本章在空间本体库和机器学习结合基础上实现了三维多个物体的自动摆放。事实上，本文提出的技术路线，不仅可以用于汉语文本自动场景生成，还可以平行移植到其它语言，而且，还可以为动态场景提供初始场景布局。这种自动的摆放方法增强了场景生成的自动性，大大降低了对场景构建人员的素质要求和成本，为场景自动生成提供了一套可行的解决方案。

中国知网
CNKI

第5章 基于自然语言描述的空间概念建模原型

5.1 引言

目前概念建模系统可以分为两大类：基于脚本的原型系统和基于自然语言的原型系统。基于固定格式的短语、类高级语言的建模系统属于基于脚本的原型系统。基于对话或者基于篇章的建模系统属于基于自然语言的系统。

早期的基于人机对话控制积木块微观世界的SHRDLU系统^[110,111,112,113]、从一个自然语言(日文)写的故事到计算机动画的MULTRAN翻译器^[114,115]、从日语文本摆放物体的YAMADA^[44]、Mukerjee在文献^[52,116,117,118]阐述得从描述郊区花园场景文本生成场景模型系统、从描述场景文本生成静态场景的WordsEye系统^[3,119]、从故事到动画片全过程自动生成的天鹅系统^[2,120,121,122,123,124,125,126]以及从交通事故报告恢复事故现场场景的Carsim系统^[4,127,128,129]都是基于自然语言概念建模系统的典型代表。

而从意大利语书写的、描述简单空间关系短语建模的NALIG^[50]、可以从描述复杂空间关系英语短语生成3D场景的Put系统^[51]、直接从约束脚本摆放几十、几百个物体的DEM²ONS系统^[92]、MultiCAD系统^[49]、MultiFormes系统^[104,130]等属于基于脚本的原型系统。

基于脚本的建模系统对于前端的自然语言理解问题，以及离散数据到连续数据映射都涉及较浅或者不涉及，而主要着力解决后端的场景建模问题，这主要是由其处理的脚本格式清楚、含义明朗决定的。

基于自然语言的建模系统主要涉及三个方面的问题：前端的自然语言理解技术、后端的场景生成技术以及自然语言离散数据到图形场景连续数据的映射。到目前为止，自然语言理解问题还主要处于句子级别的自然语言处理技术，没有达到篇章级别处理，更没有达到理解层次。这主要是因为研究者们对处理的自然语言加以限制，例如只含有主谓宾的简单句、每次最多描述一个空间关系等。场景生成问题被看作是约束编程问题，使用构造方法和迭代方法解决，具体的内容在4.2节阐述。自然语言离散数据到图形场景连续数据的映射应该是基于自然语言建模系统的核心问题，也是目前的瓶颈问题，主要的解决方法是基于知识的方法。例如：WordsEye系统利用手工标注图形库空间标记实现映射、天鹅系统依赖大量常识库实现映射，这两种方法过于依赖人力，消耗工作量。

从上面阐述可以看出单纯依赖知识的方法有一定的缺点，同时机器学习领域提供了很多可以学习到知识的算法。因此，本文在小规模空间本体库的基础上，结合机器学习方法构建基于自然语言的建模系统。使用该方法既保持了对知识的依赖性，又充分利用了机器学习算法的自动性、快速性等特点，为解决离散信息到连续信息的映射提供了一种新的策略。

本章利用空间本体库和机器学习算法相结合模型实现一个概念建模原型系统，而且提出了基于 t 检验的对该系统类型的定量评价策略。通过评价验证了该模型的有效性和实用性。第 2 节介绍了建模系统的总体框架，分析了本体库和机器学习算法结合模型的特点；第 3 节给出了系统空间信息数据结构；第 4 节提出基于 t 检验的评价策略，并且给出该原型系统的评价。最后对本章内容进行了总结。

5.2 基于自然语言描述空间概念建模原型框架

图 5-1 中给出了一个基于空间本体库和机器学习算法相结合模型的概念建模系统框架。左部为空间本体获取过程，包括本体概念、关系等体系定义、实例化等模块；中部是获取的空间本体库，包含知网等语义知识库、方位本体、实体本体以及图形本体；右部是概念建模的实现过程，包含汉语语言分析模块、空间关系抽取模块以及物体摆放模块等。

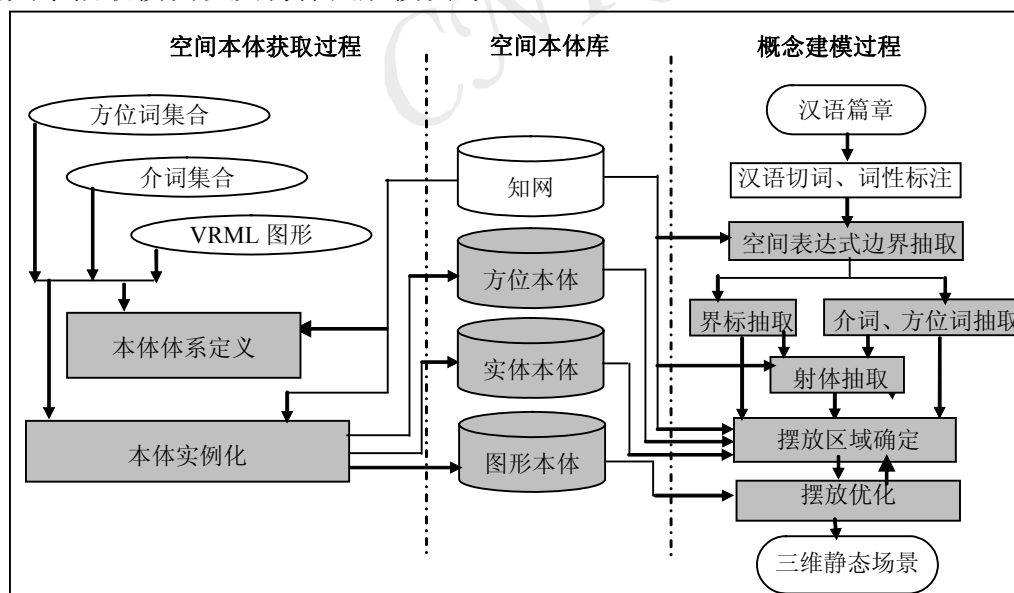


图 5-1 基于本体库的概念建模系统框架

Figure 5-1 Framework of conceptual model system based on spatial ontology

本文前几章研究的主要内容就是围绕着本体论和机器学习方法,构建小规模空间本体库、抽取空间关系以及摆放空间物体,其目的是在有限知识背景下,合理的、高效的利用机器学习方法,全自动实现文本到场景的转换。综合前几章的工作,就可以建立起一个全自动从本文出发的场景建模系统。

概念建模过程使用了汉语分析和场景可视化部分。其中,汉语分析部分使用了本实验室开发的机器翻译系统MTS2000的汉语词性标注子模块,其采用基于隐马尔可夫模型的统计词性标注方法^[131]。场景可视化部分采用了Cortona SDK®工具包,利用VRML Automation技术。Cortona解释器是目前最常用的VRML解释器之一,它为用户提供了应用程序接口(API),用户可以通过该接口应用多种编程工具,包含C/C++, Visual Basic, Java Script等,把VRML嵌入到自己的应用程序中,并可进行控制^[132]。本课题具体策略是,使用Cortona SDK作为VRML解释器,在Visual C++编程界面下,利用Cortona ActiveX控件以及VRML Automation技术来访问和控制VRML场景,这样,增加高级语言与VRML虚拟现实场景的交互能力,从整体上降低系统开发的难度。

与现有概念建模系统相比,本文的自动概念建模系统在技术路线上有以下优点:

(1) 与传统的基于脚本、自然语言的建模系统相比,本文所处理的对象是完全真实的自然语言文本。在基于脚本中的建模系统中,需要人工把描述场景的文本转换为固定格式的脚本,或者直接按照固定的格式书写脚本,总之,需要具有一定素质的人员完成脚本,浪费成本和时间。在基于自然语言的建模系统中,研究者都或多或少的限制了句法格式、规定单一的空间关系表述方式,任何修改都会影响场景生成有效性。而本文处理对象是客观世界中,不依赖于本课题存在的真实文本,能够处理复杂的空间关系组合,因此,可以没有任何成本、充分利用已有资源,提高人机接口的人性化。

(2) 与单一依赖知识库的建模系统相比,本文提出的小规模空间本体库和机器学习方法结合的方法更加具有鲁棒性。在完全依赖知识库的系统中,知识、常识是整个系统的瓶颈问题,而且知识的获取过分依赖于专家、专业人员的特性,导致系统建设周期长、不够强健。而由于机器学习算法的参与,使定性和定量信息直接的转换,更加快速、高效,大大降低对专家的依赖,缩短系统建设周期。

(3) 与单一基于构造方法或者迭代方法的建模系统相比,本文分成的摆放区域和摆放优化两个模块比一个模块更加灵活,有的放矢。基于单一方法的建模系统在建模过程中,把定性信息和定量信息放在一个步骤中解决,加大了任

务量以及难点的密度。而在使用两步法进行建模的时候，由于定性信息定量化、定量信息最优化两个问题逐一攻破，大大减低了工作量，降低了问题的难度，因此，两步法有利于提高系统的整体效率。

(4) 这种基于本体库的建模系统具有较好的可移植性。只要得到基本空间概念知识，就可以通过本文的方法获取所有空间概念定量信息，因此，容易升级、更新，将最新知识、发现应用到系统中。

5.3 空间概念建模过程的实现

在已知汉语词性标注结果的基础上，概念建模过程是一个抽取、定量化的过程。图 5-2 给出了一篇篇章生成对应摆放场景的建模流程，同时给出了建模过程中所使用的本体情况。

对于给定的待生成场景的篇章，首先截成子串得到子串集合；然后针对每个子串循环识别空间表达式边界，得到空间表达式边界集合。在每个空间表达式边界内分别识别界标、方位介词以及方位词的抽取，至此完成空间表达式的抽取，且得到空间表达式集合。针对每个空间表达式分别识别其对应的全部射体，得到空间关系，且形成空间关系集合。针对每个空间关系，首先进行方位词空间语义消歧；然后计算摆放对应射体们的区域范围；最后基于遗传算法得到每个射体的位置以及朝向。依次循环处理每个空间关系，得到所有实体的坐标点以及朝向。使用 Cortona SDK 工具包生成篇章对应的静态 3D 场景。

可以看出，整个概念建模过程都是围绕着空间关系进行的，因此，设计一个能够表达篇章级空间关系的数据结构是十分重要的。我们把篇章中的所有空间关系抽象成一个森林，拥有共同界标的空间关系是其中的一棵树，一个空间关系中界标、介词、方位词以及射体分别是节点。5.1 给出了该数据结构的类 C 语言描述。

数据结构 5-1:

```
(1) class CTTSNode{           //节点数据结构
(2) public:
(3) string m_Preposition;      //介词
(4) Object m_Landmark;        //界标
(5) Object m_Landject;        //射体
(6) SpaceWord m_Area;         //方位词，以及对应的区域
(7) CTTSNode* m_LeftBrother; //当前节点的左兄弟
(8) CTTSNode* m_RightBrother; //当前节点的右兄弟
```

```

(9)  CTTSNode* m_LeftMostChildren; //当前节点的最左儿子
(10) CTTSNode* m_Parrent;          //当前节点的父亲
(11) public:
(12) void SetLandMarkName(string); //设置界标实体本体信息
(13) void SetLandJectName(string); //设置射体实体本体信息
(14) void SetObjectInf(Object*,bool); //设置相关图形本体信息};
(15) class CTTSNodeList{           //森林数据结构
(16) public:
(17) string    m_Text;              //存储一篇输入文章
(18) CCilexpand m_VRML_inExpand; //抽取各个实体对应图形本体信息
(19) CGiveZone   m_GiveZone;       // 各个实体对应的摆放区域
(20) CTTSNode *m_ForestBegin;     //森林中第一棵树的根节点
(21) CTTSNode *m_ForestEnd;      //森林中最后一棵树的根节点
(22) vector<CTTSNode*> m_AllNode;  //缓存所有的节点，以便建立森林
(23) public:
(24) void SetText();               //存储一篇输入文章
(25) void SetCache2();            //缓存空间表达式识别后数据
(26) void WriteObjInfo();         //得到实体本体信息
(27) void SetZone();              //得到各个实体对应的摆放区域
(28) void TextToVector();         //从文本中得到所有的节点
(29) void AddNodeToVector(CTTSNode*); //加入一个 node 到 vector 中
(30) void CreatForest();          //建立森林
(31) void AddToNode(CTTSNode*,CTTSNode*); //子树中添加节点
(32) void AddRoot(CTTSNode *);    //把森林中树的根节点建立起来
(33) void AddToForest(CTTSNode *); //把节点加入到森林中 };

```

5.4 空间概念建模场景的评价

5.4.1 TTS 系统的评价研究概述

TTS(text-to-scene)的含义是文本到场景的转换。本文研究的空间概念建模系统是文景转换系统的一个分支，因此，把整个文景转换系统评价研究在这一节进行概述。

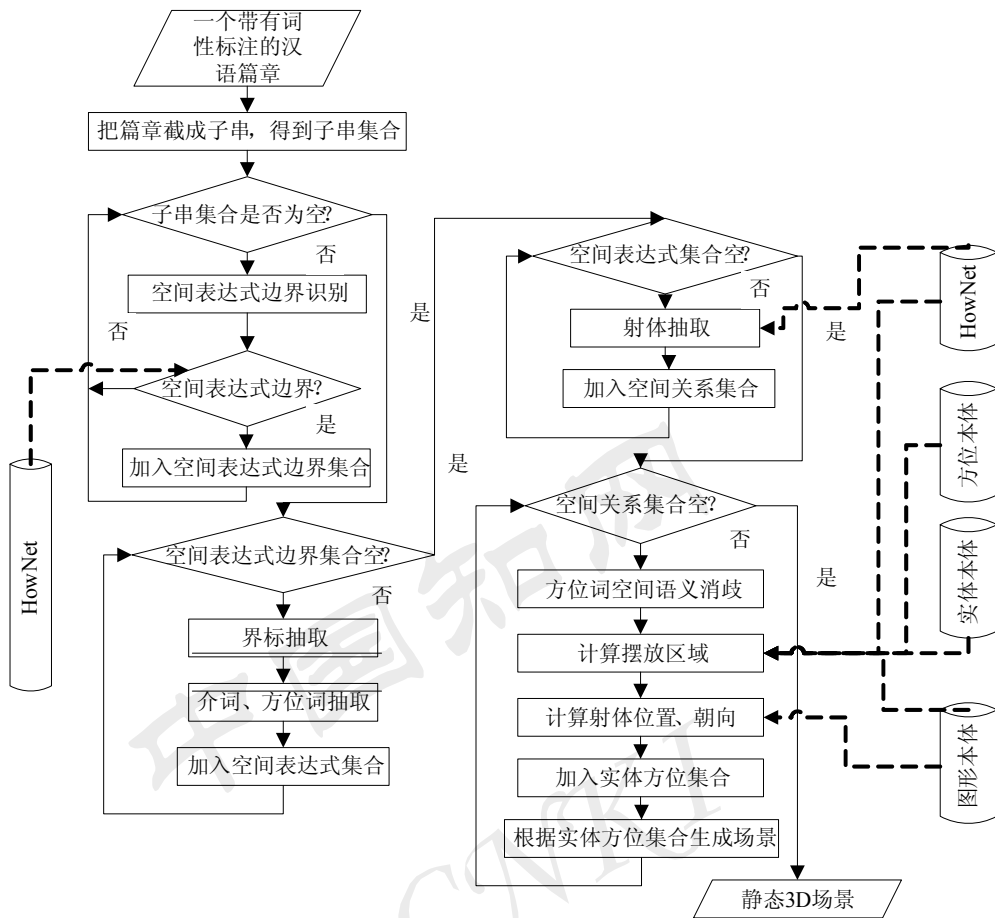


图 5-2 基于本体库的建模过程

Figure 5-2 Modeling procedure based on spatial ontology

目前, 对于基于高级描述的场景生成系统没有公共的评价体系, 从文献上看, 评价方法多集中于目测^[1,2,3,4,44,92,104]。目测就是指, 把系统生成的几个结果, 即场景, 展现给评价者, 然后, 评价者给出主观的优劣评价。本文又根据主观优劣评价的复杂度, 把目测评价策略分为纯粹目测评价策略和统计目测评价策略两大类。纯粹目测评价策略的特点是, 最终的评价结果就是直接使用评价者的主观上的“好”、“坏”判断, 而没有进一步的分析。文献[1]、[2]、[3]、[44]和[92]等中都使用这种方法。统计目测评价策略的特点是, 首先, 对判断标准提出严格的说明, 例如: 等级说明, 或者分值说明, 评价者依据这个标准给出主观评价; 然后, 从统计角度, 给出主观评价统计数字。文献[129]中是这种方法的实例。下面较详细的介绍Carsim系统的评价策略。

CarSim 系统的评价策略属于统计目测评价策略。其评价策略的具体方案如下:

(1) 评价者。由于这个系统的目标是把交通事故报告描述的场景恢复成 3 维动画片, 给交通局或者保险公司使用, 所以, 其把从交通局请来的四位警官作为评价者。

(2) 评价文本。使用从网络新闻报纸收集来的 50 篇文章作为系统的输入, 也就是评价文本, 每篇长度在 36 到 541 个单词之间。

(3) 评价等级。评价等级就是评价者给系统生成的动画质量进行分类的标准。该系统使用 3 个分值表示不同的等级, 0 分表示错误、1 分表示多少还正确、2 表示很好。

(4) 统计手段。该系统使用统计手段, 对评价者给出的评价等级, 即分值, 进行统计分析。使用的统计手段包含二次权重加权 k 系数(Weighted k Coefficient with Quadratic Weights)^[133]、标准偏差(Standard Deviation)。统计的信息包含成对内部评价一致性(Pairwise Inter-rater Agreement)、每个文本的标准偏差以及评价的成对互联系性(Pairwise Correlation)。

使用该评价方案, 该系统得到的统计数据是: 成对内部评价一致性的值是 0.73, 通常大于 0.70 的值就被认为具有很好的一致性; 每个文本的标准偏差值是 0.45; 评价的成对互联系性值是 0.75。

从上面阐述可以, 在文景转换领域的评价手段有如下的弱点:

(1) 没有公认的评价资源。这里的评价资源, 主要是指用于评测的公认资源, 包含高级描述和场景。在这个领域中, 还没有如同其它领域, 比如自然语言处理领域, 有众多的国际公认的评价平台。造成这个问题的主要原因是, 图形、图像的统一化、公认化较其它表述方法难; 而且三维, 乃至二维计算机图形技术发展成熟不是很久以前的事情, 资源建建需要一定的时间。

(2) 没有公认的评测方案。也就是, 在作评测的时候没有可以遵循的优劣标准。优劣的评测还处于各自的主观判断。

为了加快文景转换方向评价体系的标准化, 受机器翻译领域评价体系的启发^[134,135], 本文提出基于 t 检验的评价模型。实践证明, 可以一定程度的减少主观因素对评价的影响。

5.4.2 基于 t 检验的评价模型

t 检验是统计学上的术语, 用于假设检验, 或者称显著性检验的一种方法。把这个方法引入概念性建模系统评测系统, 目的是为了量化系统的显著性。

基于 t 检验的评价模型包含评价资源、评价标准和评价指标三个部分。

在给出这三个部分具体阐述之前，首先给出评价类型。我们把评价分为封闭评价和开放评价两种类型。封闭评价是指，把用于训练的资源作为评价资源；开放评价是指，用于评价资源和用于测试的资源无关。

评价资源，这里不把着眼点放在得到一个含有具体的高级描述和场景的评价资源库，而是着重给出获取用于评价的资源方法。资源包含高级描述和相应场景。获取资源方法是：

(1) 获取高级描述。从语料库中随意抽取不小于 1% 的描述作为系统输入。对于封闭评价，则从训练语料库中选取，而对于开放评价，则从测试语料中选取。

(2) 获取场景。场景是指和高级描述相应的、反映描述内容的图形、图像。通过三种途径获取每个描述对应的三个场景，一、主观获取，是指人，包含用户、专业人员等谙熟描述的人员，在理解高级描述基础上，描绘出相应场景，这样得到的场景称为主观场景，作为评价标准的上限；二、系统获取，是指系统本身，利用现有体系、方法，处理高级描述，生成相应的场景，这样得到的场景称为系统场景，是评价的核心；三、随意获取，是指，利用任意手段，生成高级描述的相应场景，这样得到的场景称为随意场景，作为评价标准的下限。

通过这个获取方法，得到用于评价的三个集合：主观场景集合、系统场景集合和随意场景集合，作为评价资源。

评价标准，就是指将生成场景质量分为不同的等级，并且详细说明等级范畴。我们提出的评价标准包含以下 6 个等级，具体说明如下：

- A. 场景准确、流畅地传达了原高级描述的信息，除个别因审美差异造成的偏差外，无需修改；
- B. 场景传达了原高级描述的信息。不用参照原高级描述，就能通过场景推测出原文的意思，但是场景在方位、动作、情节等方面的精确性有些问题，需要修改。不过这种修改无需参照原高级描述也能有把握地进行，且修改也较容易；
- C. 场景大致表达了原高级描述的意思，局部与原高级描述有出入，一般情况下需要参照原高级描述才能改正。有些情况即使无需参照原高级描述也能猜测到原高级描述的意思，但译文的不妥明显是由于系统程序的缺陷造成的；
- D. 场景有一部分符合原高级描述的一部分意思，整体场景没有匹配上，不过原高级描述的实体都孤立地生成了，对人工后编辑有点用处；
- E. 看了译文不能得到与原高级描述相符的意思，不过总是生成了一些实

体;

F. 完全没有得到场景。

以上各等级在计算平均值时分别对应于 95、80、65、45、20 和 0 分。

评价指标包含平均数、统计量t、统计量t的两尾概率以及单尾概率四个具体的评价值。评价数是指场景评价等级的算术平均值，计算公式如(5-1); 统计量t的计算公式是(5-2); 统计量t的两尾概率以及单尾概率可以查数学表得到。

$$\bar{x}_k = \frac{\sum_{i=1}^m \left(\frac{\sum_{j=1}^{n_i} x_{ij}}{n_i} \right)}{m} \quad (5-1)$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{S_{x_1-x_2}}$$

其中 $S_{x_1-x_2} = \sqrt{\frac{\sum (x_1 - \bar{x}_1)^2 + \sum (x_2 - \bar{x}_2)^2}{(n_1 - 1) + (n_2 - 1)} \times \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$;

自由度 $df = (n_1 - 1) + (n_2 - 1)$ (5-2)

其中, x_{ij} 表示第j个评价者给第i个场景的评价分值; n_i 表示第i个场景的评价者的个数; \bar{x}_k 表示第k类场景的样本平均数。

5.4.3 评价及分析

利用本章提出的基于本体论和机器学习方法结合的概念建模系统, 以及基于t检验评测方法进行了一组空间概念建模实验, 包含封闭空间概念建模实验和开放空间概念建模实验。在封闭实验中, 使用3.6节中训练语料库的随机6篇文章作为封闭测试语料。在开放实验中, 使用3.6节中测试语料库的随机2篇文章作为开放测试语料。

为了尽量保证目测评价的公正性, 降低主观因素的影响, 在评测中, 我们保证以下几点要求:

- (1) 在评测结果文件中, 把主观获取、系统获取以及随机获取的场景进行了随机排序处理, 即不同途径获取的场景的先后顺序是不确定的;
 - (2) 将不同途径获取的场景并列在一起, 能够一目了然地看出相对差别;
 - (3) 评测被告知所有的场景都是系统自动生成的;
 - (4) 评测者之间在评测过程中, 不交流和评测有关的任何内容;
 - (5) 评测者对每个场景都是凭着第一印象打分, 即不需要繁复周密的思考;
- 下面详细说明评价资源获取的方案。

在获取主观场景的时候，首先，由本课题组中即理解空间关系具体含义，又善于场景建模的研究人员阅读、理解作为评价用的篇章；然后，根据理解，使用Cosmo World2.0®建立一幅反映该篇章的静态三维场景，场景中的实体都来自符合2.4.4.2节阐述的图形库；

在获取随意场景的时候，首先，把评价篇章进行词性标注，并且把其中名次类别词语编号；然后，由一名同学投掷两个骰子得到一个数字，根据这个数字选取词语，其中选取词语的个数由评价篇章中的方位词个数决定；最后，使用Cosmo World2.0®，根据选取的词语和篇章中方位词建立一幅静态三维场景，场景中的实体亦都来自符合2.4.4.2节阐述的图形库。

在此基础上，由3名评测者同时对封闭测试和开放测试的场景进行评价。表5-1给出了这组实验的三个不同场景集合的样本平均值以及得分比例；表5-2给出了这组实验的基于t检验的显著性评价；表5-3给出了用于封闭评价和开放评价的篇章名称，由于篇幅，就不在这里详列其内容；图5-3显示了《乌鸦与狐狸》的系统场景。《乌鸦与狐狸》文本内容是：“有只乌鸦偷到一块肉，衔着站在大树上。路过此地的狐狸看见后，口水直流，很想把肉弄到手。他便站在树下，大肆夸奖乌鸦的身体魁梧、羽毛美丽，还说他应该成为鸟类之王，若能发出声音，那就更当之无愧了。乌鸦为了要显示他能发出声音，便张嘴放声大叫，而那块肉掉到了树下。狐狸跑上去，抢到了那块肉，并嘲笑说：“喂，乌鸦，你若有头脑，真的可以当鸟类之王。”^[89]。

表 5-1 样本平均值以及分布

Table 5-1 The averages and propositions on samples

评价条件	场景集合	A	B	C	D	E	F	平均分
封闭测试	主观	0%	44.44%	55.56%	0%	0%	0%	71.67**
	系统	0%	16.67%	50%	27.78%	5.56%	0%	59.45
	随意	0%	0%	33.33%	55.56%	11.11%	0%	48.89**
开放测试	主观	0%	16.67%	66.67%	16.67%	0%	0%	64.16 ^{ns}
	系统	0%	16.67%	50%	33.33%	0%	0%	60.83
	随意	0%	0%	16.67%	50%	33.33%	0%	40*



图 5-3 《乌鸦与狐狸》的系统场景

Figure 5-3 The system scene of “Crow and Fox”

表 5-2 成对双样本均值 t 检验分析结果

Table 5-2 The results of t-test on pairwise samples averages

评价条件	场景 1	场景 2	统计量 t	单尾 $P(T \leq t)$	t 单尾临界界 $\alpha=0.05$	t 单尾临界界 $\alpha=0.01$	双尾 $P(T \leq t)$	t 双尾临界界 $\alpha=0.05$	t 双尾临界界 $\alpha=0.01$
封闭测试	主观	系统	3.6082	0.001	1.7396	2.5669	0.002	2.1098	2.8982
	系统	随意	2.7391	0.007	1.7396	2.5669	0.01	2.1098	2.8982
开放测试	主观	系统	0.4589	0.33	2.015	3.3649	0.67	2.5705	4.0321
	系统	随意	2.7916	0.02	2.015	3.3649	0.04	2.5705	4.0321

表 5-3 评价用篇章名称

Table 5-3 The texts for the evaluation

评价条件	篇章 1	篇章 2	篇章 3	篇章 4	篇章 5	篇章 6	文献
封闭测试	寒鸦与狐狸	口渴的鸽子	乌龟与老鹰	乌鸦与羊	乌鸦与狐狸	一只眼睛的鹿	[89]
开放测试	人与狐狸	哲学家、蚂蚁和赫耳墨斯					[89]

从以上实验结果可以看出：

(1) 在封闭测试以及开放测试中，系统场景评价样本均值都高于随意场景，而且双尾概率 $P(|t| \geq 2.7391)$ 的值都是在 0.01 和 0.05 之间，可见处理效应是十分明显的，说明基于线性分类器抽取空间关系的方法，以及基于小规模本体库实现定性信息到定量信息转换的方法，很好的理解了自然语言描述的空间概念、量化了其信息；

(2) 在封闭测试中，主观场景评价样本均值高于系统场景评价样本均值，而且双尾概率 $P(|t| \geq 3.6082)$ 的值都远远小于 0.01，在 0.002 附近，说明自动化概念建模的效果和人类审美理解还是有相当的差距；

(3) 在开放测试中，主观场景评价样本均值高于系统场景评价样本均值，但是双尾概率 $P(|t| \geq 0.4589)$ 的值都远远大于 0.01，在 0.67 附近，说明实验存在相当一定因素的误差，失去了统计意义。导致这个结果的原因，很可能是因为统计样本的数量过少。而且，从中不能对这两个场景样本做出任何判断。

为了使评测更加客观、具体，在完成目测评价后，我们和评价者进行了口头交流。评价者口头评价内容可以归纳为以下几点：

(1) 图形不够真实。这个评价是指，用于评价所有场景中的三维图形都不够美观、赏心悦目，也就是离人们想象中的图形有很大的差距。而且，评价者指出，这是影响评价等级的主要因素之一；

(2) 场景不够丰富。这个评价是指，用于评价所有场景都不能够很充实的反映篇章故事情节，即一幅静态场景是不够的，应该是多幅的动态场景，这是没有任何场景被评价为 A 级的原因；

(3) 空间关系体系较完整。这个评价主要是针对系统场景集合来说的，评价者一致认为场景展现的空间关系能够很好的体现原篇章对空间关系的描述。

从以上口头评价可以看出，自然语言描述的空间关系三维可视化研究仅仅是文景转换系统的一个子模块。但是，结合目测评价可以看出，这个研究是文景转换系统的核心、基础部分，也就是空间关系展现的可接受性是整个场景可接受的必要条件。

需要指出的是，本文探索基于本体论和机器学习的自动建模技术，但是并不否定人类认知、审美的重要作用。事实上，本文在各项技术的研究中都尽量利用了现有的人类认知、审美的研究成果，如在空间关系抽取研究中，只有人、事物或者图形也可充当空间关系成分这一人类认知观点是特征的主要思想来源；在空间本体库概念定义中，充分依据“目之所及”为“前”、“触之可得”为“近”等空间认知观点等等。本文认为，自动化理解自然语言、建立场景模

型和人类理解自然语言、建立场景是相符相承的，人与机器要适当配合，才能发挥各自所长，更好的推动概念建模研究的发展。实验证明，本文提出的方法有效的实现了自然语言描述的空间关系三维可视化，大大缩小了本体库的规模，提高了概念建模的自动化程度。由于目前所依赖的自然语言处理技术比较单一，只包含词性标注技术，以及图形库不够完善，其中图形不够精细，空间关系抽取不可避免的还存在一些错误，三维场景也不可避免的不能满足人类视觉要求。如果在精度更高的自然语言处理技术，以及更加精细的图形基础上，将会进一步提高空间关系抽取的精度，以及生成场景的美感，最终达到提高概念建模质量的目的。

5.5 本章小结

本章综合前面几章内容建立了一个基于本体库和机器学习方法的概念建模原型系统，并提出基于 t 检验的评价方法。虽然该系统的基础部分没有离开知识，但是这些知识是通过机器学习方法自动获取的，而且规模很小。与传统的基于构造、基于迭代概念建模系统相比，该系统在知识处理方面具有表达直观、易于理解、概括性强、获取效率高、可移植性好等优点。通过对生成场景的评价验证了基于 t 检验评价方法的有效性，同时也证明了本文提出的本体论和机器学习方法相结合模型的可行性及重要价值。

结论

空间概念建模是自然语言理解和文景转换的基础问题。探索基于自然语言描述的空间概念建模方法具有重要的理论研究意义和广阔的应用前景。本文在 本体论和机器学习理论基础上,研究了空间本体库构建技术以及篇章级空间关系定性到定量转化技术。具体内容包含基于归并的空间本体半自动实例化方法、基于线性分类器的空间关系抽取方法、基于空间本体库和遗传算法的物体空间摆放方法。在这些研究基础上,构建了一个基于小规模空间本体库的自然语言空间概念建模原型系统,且提出基于 t 检验的目测评价体系。通过对生成场景的评价,验证了基于自然语言描述的空间概念建模的有效性。本文提出的空间概念理解、量化的方案有效地提高空间本体库实例化效率、缩短三维静态场景构建周期、增强了空间本体知识的扩充性,为自然语言描述空间概念建模这一基础问题提供了一套可行的方案

具体地说,本文的创新成果包括如下几个方面:

1. 提出了基于规则、聚类等学习策略的半自动空间本体实例化方法。与传统的手工实例化方法相比,避免了陷入专家权威的困境,实现了自动空间本体实例化,大大提高了本体库的构建效率和鲁棒性。

2. 提出了一种基于线性分类器抽取篇章级空间关系的方法。首次在计算语言学领域中,提出空间关系的形式化描述。把空间关系的抽取抽象为一个二值分类问题,然后在线性分类器基础上实现了篇章级空间关系的抽取。与现有信息抽取方法相比,该方法实现了从句子级别关系抽取到篇章级别关系抽取,以及句法关系抽取到空间认知关系抽取的进步。而且,只利用词性标注实现了空间认知关系的抽取,避开了自然语言语义、甚至语用分析的难题,同时有效地保证了空间关系抽取的正确率。

3. 提出了基于小规模本体库确定摆放区域方法。摆放区域作为物体摆放的第一个步骤,大大缩小了物体摆放搜索空间,降低了摆放约束条件,同时保证了摆放区域确定的合理性。该方法在小规模空间本体库的基础上,结合聚类等机器学习方法,把空间关系的定性信息转换为三维场景所需要的定量信息,实现自然语言离散系统和场景连续系统的平滑过渡,使物体摆放只能处理脚本描述的现状跨越到完全可以处理自然语言描述。

4. 提出了基于 t 检验的目测评价体系。该体系给出了明确的评价资源获取

策略、以及定量的评价标准和评价指标，使该研究中评价朝定量化、统一化迈进了一步。

虽然取得以上阶段性的研究成果，但是本文研究还存在很多需要进一步改进的工作，主要包括：

1. 进一步提高空间关系抽取的正确率。空间关系抽取是空间关系可视化的基础环节，直接影响到生成场景和原自然语言文本的匹配度。本文在显示的空间关系概念基础上，仅利用词性标注工具抽取空间关系。在今后的工作中可以尝试扩充空间关系概念，使其能够包含文本中的隐含的空间关系，提供自然语言描述空间关系理解的完备性。

2. 进一步提高图形表达的精确度。三维图形的形式化细致程度是空间关系可视化不容忽视的问题，直接影响生成场景的视觉可接受性。本文使用单层包围盒表达三维图形。在今后的工作中可以尝试多级包围盒表达方法，使其能够配合部分与整体的空间关系的可视化，提高场景描述的精度。

3. 深一步实现基于语言的动态场景生成。动态场景更具有实用性和交流性，是文景转换的最理想的结果。本文基于自然语言实现了空间概念建模，构建了静态场景。在今后的工作中完全可以结合上面两个方面的改进，引入时序分析，扩充本体库，实现基于语言的动态场景生成。

参考文献

- 1 A. Mukerjee, K. Gupta and S. Nautiyal. Conceptual Description of Visual Scenes from Linguistic Models. *Image and Vision Computing* 18(2), 2000: 173~187
- 2 陆汝铃, 张松懋. 从故事到动画片——全过程计算机辅助动画自动生成. *自动化学报*, 2002, 28(3): 322~348
- 3 B. Coyne and R. Sproat. WordsEye: An Automatic Text-to-scene Conversion System. *Proceedings of the SIGGRAPH 2001 Annual Conference on Computer Graphics*, Los Angeles, CA USA, 2001: 487~496
- 4 R. Johansson, D. Williams and A. Berglund. A System to Visualize Written Road Accident Reports as Animated 3D Scenes. In Hirst, G. and Nirenburg, S., editors, *ACL2004: Second Workshop on Text Meaning and Interpretation*, Barcelona, Spain, 2004: 57~64
- 5 J. Haiman (ed). *Iconicity in Syntax*. John Benjamins, 1985
- 6 沈家煊. 汉语特点研究和语言共性研究. 中国语文研究四十年纪念文集. 北京语言学院出版社. 1993
- 7 M. Ma and P. Kevitt. Semantic Representation of Events in 3D Animation. In *Proceedings of the Fifth International Workshop on Computational Semantics (IWCS-5)*, Tilburg, The Netherlands, 2003: 253~281
- 8 M. Ma and P. Kevitt. Visual Semantics and Ontology of Eventive Verbs. *The First International Joint Conference on Natural Language Processing (IJCNLP-2004)*. Hainan, China. 2004: 187~196
- 9 J. Dhiraj, Z. James and W. Li. The Story Picturing Engine: Finding Elite Images to Illustrate a Story Using Mutual Reinforcement. *ACM Transactions on Multimedia Computing, Communications and Applications*. 2006, 2(1): 68~89
- 10 S. Yusuke, T. Takenobu and T. Hozumi. Kairai - Software Robots Understanding Natural Language. *Third International Workshop on Human-Computer Conversation*. Bellagio, Italy, 2000: 158~163
- 11 S. Yusuke, T. Takenobu and T. Hozumi. Processing of 3-D Spatial Relations for Virtual Agents Acting on Natural Language Instructions. *Second Workshop*

- on Intelligent Virtual Agents, 1999: 67~78
- 12 T. Tokunaga, T. Koyama and S. Saito. Bridging the Gap between Language and Action. Intelligent Virtual Agent - 4th International Workshop IVA 2003. LNAI. 2003, 2792: 127~135
 - 13 T. Hozumi. Natural Language Understanding and Action Control. Technical Report. Department of Computer Science, Tokyo Institute of Technology. 2001: 1~38
 - 14 W. Terry. Procedures as a Representation for Data in a Computer Program for Understanding Natural Language. MIT AI Technical Report 235. MIT, 1971: 10~30
 - 15 W. Terry. Understanding Natural Language. Academic Press, 1972: 4~8
 - 16 R. Simmons. The CLOWNS Microworld. Proceedings of the workshop on Theoretical Issues in Natural Language Processing TINLAP 75. Morristown, NJ, USA. 1975: 17~19
 - 17 K. Kahn. Creation of Computer Animation from Story Descriptions. Ph.D. Thesis, AI Tech. Report 540, AI Lab, MIT, Cambridge, MA, 1979: 1~104
 - 18 Ulysse. <http://www.info.unicaen.fr/~nugues/research.html>
 - 19 N. Pierre and B. Olivier. A Conversational Agent to Help Navigation and Collaboration in Virtual Worlds. Virtual Reality, 1998, 3(1):71~82
 - 20 Persona. Chapter on Persona, J. Bradshaw (ed). Software Agents, MIT Press, 1997: 56~58
 - 21 Nautilus. <http://www.aic.nrl.navy.mil/~severett/vr.html>
 - 22 S. Stephanie, W. Kenneth and A. Manuel. Spoken Language Interface to a Virtual Reality System. Fifth Conference on Applied Natural Language Processing (ANLP 97). 1997: 36~37
 - 23 B. Webber. Instructing Animated Agents: Viewing Language in Behavioral Terms. Lecture Note in Computer Science, Multimodal Human-Computer Communication, Systems, Techniques, and Experiments. Berlin: Springer-Verlag. 1998, 1374: 89~100
 - 24 N. Badler, C. Erignac, and Y. Liu. Virtual Humans for Validating Maintenance Procedures. Comm. of the ACM. 2002, 45(7): 56~63
 - 25 J. Allbeck and N. Badler. Representing and Parameterizing Agent Behaviors. In "Life-like Characters: Tools, Affective Functions and Applications." Helmut

-
- Prendinger and Mitsuru Ishizuka, Ed. Springer, Germany, 2003: 34~40
- 26 H. Tanaka, T. Tokunaga and Y. Shinyama. Animated Agents that Understand Natural Language and Perform Actions. Proceedings of PRICAI-02 Workshop on Liflike Animated Agents. 2002: 89~ 94
- 27 K. Barnard and D. Forsyth. Learning the Semantics of Words and Pictures. Proceedings International Conference on Computer Vision, 2001: 408~415
- 28 D. Brown and B. Chandrasekaran. Design Considerations for Picture Production in a Natural Language Graphics System. ACM SIGGRAPH Computer Graphics. 1981, 15(2): 174~207
- 29 O. Le and V. Gaildrat. Constraint-based 3d Isothetic Object Layout for Declarative Scene Modeling. The 2003 International Conference on Imaging Science, Systems, and Technology, CISST'03. Las vegas, USA. 2003: 320~326
- 30 S. Diomidis. On the Declarative Specification of Models. IEEE SOFTWARE. Published by the IEEE Computer Society. 2003, 20(2): 94~96
- 31 R. Adiga, et al. An Overview of the BlueGene/L Supercomputer. Supercomputing, ACM/IEEE 2002 Conference. 2002: 60~68
- 32 G. Kjeldsberg, F. Catthoor, and E.J. Aas. Data Dependency Size Estimation for Use. Memory Optimization. Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions. 2003, 22(7): 908~921
- 33 S. Furber. Future Trends in SoC Interconnect. 2005 IEEE VLSI-TSA International Symposium. 2005: 295~298
- 34 Y. Kubo and A. Takahashi. Global Routing by Iterative Improvements for Two-layer Ball Grid Array Packages. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. 2006, 25(4): 725~733
- 35 K. Fukunage and S. Yamada. Placement of Circuit Modules Using a Graph Space Approach. Proceedings of the 20th Design Automation Conference. 1983: 465~473
- 36 F. Liu. Fuzzy Pole Placement Design with H_{∞} Disturbance Attenuation for Uncertain Nonlinear Systems. IEEE Conference on Control Applications. 2003, 1: 392~396
- 37 G.M. Huang, J.S. Lei and A. Abur. A Heuristic Approach for Power System Measurement Placement Design. International Symposium on Circuits and Systems. 2003, 3: 407~410

- 38 M. Grohe and N. Schweikardt The Succinctness of First-order Logic on Linear Orders. Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science. 2004:438~447
- 39 Y. Chou and Y.L. Lin. Effective Enforcement of Path-delay Constraints in Performance-driven Placement. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. 2002, 21: 15~22
- 40 R. Hentschke and R. Reis. Plic-Plac: a Novel Constructive Algorithm for Placement. Proceedings of the 2003 International Symposium on Circuits and Systems. 2003, 5: 461~464
- 41 P. Maidee, C. Ababei and K. Bazargan. Timing-driven Partitioning-based Placement for Island Style FPGAs. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. 2005, 24(3): 395~406
- 42 A. Kennings and K.P. Vorwerk. Force-Directed Methods for Generic Placement. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. 2006, 25(10): 2076~2087
- 43 A. Deshpande and J.M. Hellerstein. Decoupled Query Optimization for Federated Database Systems. 18th International Conference on Data Engineering. 2002: 716~727
- 44 A. Yamada. Studies on Spatial Description Understanding Based on Geometric Constraints Satisfaction. PhD thesis, Kyoto University, Kyoto, 1993: 1~108
- 45 A. Abdullah and A. Hussain. Heuristics and Meta-heuristics for Bandwidth Minimization of Sparse Matrices. IEEE International Conference on Engineering of Intelligent Systems. 2006: 1~6
- 46 H. Mori. Applications of Meta-heuristics to Power Systems in Japan. Transmission and Distribution Conference and Exhibition 2002: Asia Pacific, IEEE/PES. 2002, 1:661~663
- 47 王知人. 一种改进的模拟退火算法. 高等学校计算数学学报. 2006, 28(1): 15~19
- 48 N. Fu and J. Yu. Fuzzy Tabu Search Algorithm for the VLSI Placement Problem. IEEE 2002 International Conference on Communications, Circuits and Systems and West Sino Expositions. 2002: 2(29): 1146~1150
- 49 N. Vassilas, G. Miaoulis and D. Chronopoulos. MultiCAD-GA: A System for the Design of 3D Forms Based on Genetic Algorithms and Human Evaluation.

- Methods and Applications of Artificial Intelligence, Second Hellenic Conference on AI, SETN 2002. Thessaloniki, Greece. Lecture Notes in Computer Science 2308. 2002: 203~214
- 50 G. Adorni , M. Manzo and F. Giunchiglia. Natural Language Driven Image Generation. Proceedings of COLING 84, Stanford, California. 1984: 495~500
- 51 S. Clay and J. Wilhelms. Put: Language-based Interactive Manipulation of Objects. IEEE Computer Graphics and Applications, 1996: 31~39
- 52 A. Mukerjee. Conceptual Description of Visual Scenes from Linguistic Models. Journal of Image and Vision Computing. 2000, 18(3): 1~13
- 53 WordsEye. <http://semantclight.com>
- 54 R. Lu, S. Zhang and Z. Wei. Generate Computer Animation from Natural Language Stories. Proceedings of Pacific Asian Conference on Expert Systems. Los Angeles, USA, 1999
- 55 F. Annika. Design and Use of Ontologies in Information-providing Dialogue Systems. Dissertation, Linköping Studies in Science and Technology, Thesis No. 874, School of Engineering at Linköping University. ISBN: 91-7373-947-2. Sweden, 2004: 15~19
- 56 K. Mahesh and S. Nirenburg. A Situated Ontology for Practical NLP. Proceedings of IJCAI'95 Workshop on Basic Ontological Issues in Knowledge Sharing. Nontreal, Canada. 1995: 1~10
- 57 S. Russell and P. Norvig. Artificial Intelligence: a Modern Approach, (2nd international edition), Prentice Hall, Upper Saddle River, N.J., chapter 10: Knowledge Representation. 2003: 320~340
- 58 T. Gruber. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal of Human-Computer Studies. 1995, 43(5/6): 907~928
- 59 B. John and F. Scott. Spatial Ontology Baseline. ONTOSPACE Projecto Report, Deliverable D2, University of Bremen, Germany. 2005: 1~105
- 60 I. Niles and A. Pease. Origins of the Standard Upper Merged Ontology: A Proposal for the IEEE Standard Upper Ontology. Working Notes of the IJCAI-2001 Workshop on the IEEE Standard Upper Ontology. URL: <http://projects.tecknowlege.com/IJCAI01/index.Html>. 2001: 37~42
- 61 R. Casati and A. Varzi. Parts and Places: the Structures of Spatial

- Representation, MIT Press (Bradford Books). Cambridge, MA and London, 1999: 89~104
- 62 B. Smith. Mereotopology: a Theory of Parts and Boundaries. *Data and Knowledge engineering*, 1996(20): 287~303
- 63 N. Guarino. Formal Ontology, Conceptual Analysis and Knowledge Representation. *International Journal of Human-Computer Studies*, 1995, 43(5/6), 625~640
- 64 D. Lenat and R. Guha. *Building Large Knowledge-based Systems: Rrepresentation and Inference in the CYC Project*. Addison-Wesley Publishers, New York. 1989: 201~209
- 65 陆汝钤, 石纯一, 张松懋等. 面向 Agent 的常识知识库. *中国科学(E 辑)*, 2000, 30(5): 453~463
- 66 C. Masolo, S. Borgo and A. Gangemi. The WonderWeb Library of Foundational Ontologies: Preliminary Report, WonderWeb Deliverable D17, ISTC-CNR, Padova, Italy. 2002: 23~70
- 67 C. Masolo, S. Borgo and A. Gangemi. Ontologies Llibrary (final), WonderWeb Deliverable D18, ISTC-CNR, Padova, Italy. 2003: 34~67
- 68 董振东, 董强. 知网. <http://www.keenage.com>. 1999
- 69 梅家驹. 同义词词林. 上海辞书出版社, 1996
- 70 同义词词林扩展版. <http://ir.hit.edu.cn/>
- 71 R. Langacke. *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Stanford University Press, Stanford, Calif., 1987, I: 78~98
- 72 廖秋忠. 空间方位词和方位参考点. *中国语文*. 1989(1): 9~16
- 73 邹韶华. 现代汉语方位词的语法功能. *中国语文*. 1984(3): 173~178
- 74 方经民. 现代汉语方位成分的分化和语法化. *世界汉语教学*. 2004, 68(2): 4-15
- 75 董振东, 董强. HowNet2005 版本使用说明书
- 76 张辉. 论空间概念在语言知识构建中的作用. *解放军外语学院学报*. 1998, 21(1): 19~22
- 77 刘宁生. 汉语如何表达物体的空间关系. *中国语文*. 1994(3): 169~179
- 78 L. Talmy. Figure and Ground in Complex Sentences. G. Jeseoph (ed). *Universals of Human Language (Vol. 4)*. Stanford Un ivers ity Press, 1978: 347~408
- 79 蓝纯. 从认知角度看汉语的空间隐喻. *外语教学与研究*. 1999(4): 7~15

- 80 R. Langacker. *Foundations of Cognitive Grammar. Volume I.* Stanford, CA.: Stanford University Press. 1987: 207~210
- 81 廖秋忠. 现代汉语篇章中空间和世界的参考点. *中国语文*, 1983, (4): 257~263
- 82 方经民. 现代汉语方位参照聚合类型. *语言研究*. 1987(2): 34~40
- 83 方经民. 论方位参照的构成要素. *中国语学*. 东京, 1993(240): 168~174
- 84 方经民. 空间方位参照的认知结构. *世界汉语教学*. 1999, 50(4): 32~38
- 85 K. Church. A Stochastic Parts Program and Noun Phrase Parser for Unrestricted Text. *Proceedings of the Second Conference on Applied Natural Language Processing (Association for Computational Linguistics)*. Morristown, NJ, 1988: 136~143
- 86 M. Collins. *Head-driven Statistical Models for Natural Language Parsing*. PhD thesis, University of Pennsylvania, Philadelphia, PA, 1999
- 87 N. Littlestone. Learning Quickly when Irrelevant Attributes Abound: A new Linear-threshold algorithm. *Machine Learning* 1988, 2:285~318
- 88 T. Mark. Using Siscourse Focus, Temporal Focus, and Spatial Focus to Generate Multisentential Text. *Proceedings of the Fifth International Workshop on Natural Language Generation, acl-1990*. Pittsburgh, PA. 1990: 70~78
- 89 伊索寓言. <http://www.white-collar.net/child/fable/yisuo/index.html>
- 90 Machine Translation Laboratory of Harbin Institute of Technology. The Final Report of the MTS2000 Project. 2001, PIII-A: 52
- 91 Y. Chen and M. Yang. A Lexicalized Second-order-HMM for Ambiguity Resolution in Chinese Segmentation and POS Tagging. *The Highest Letter (English)*. 2005, 11(4): 346~350
- 92 G. Kwaiter, V. Gaildrat and R. Caubet. DEM²ONS: a High Level Declarative Modeler for 3D Graphics Applications. *Proceedings of the International Conference on Imaging Science Systems and Technology, CISST'97*. Las Vegas, 1997: 149~154
- 93 J. Frazer. *Evolutionary Architecture*. Architectural Association. London, UK, 1995: 232~235
- 94 P. Bentley. *Evolutionary Design by Computer*. Morgan Kaufmann, San Francisco, USA, 1999: 102~103
- 95 J. Gero and V. Kazakov. *Evolving Design Genes in Space Layout Planning*

- Problems. *Artificial Design in Engineering*. 1998, 12(3): 163~176
- 96 P. Coates and D. Makris. Genetic Programming and Spatial Morphogenesis. *Proceedings AISB'99: Symposium on Creative Evolutionary Systems*. University of Edinburgh, Edinburgh, UK, 1999: 105~114
- 97 D. Makris. *Evolutionary Design Environments (EDEN)*. MSc thesis. School of Architecture, University of East London, London, UK, 1999: 60~70
- 98 T. Fuhr, G. Socher and C. Scheering. A Three-dimensional Spatial Model for the Interpretation of Image Data. *IJCAI-95 Workshop on the Representation and Processing of Spatial Expressions*. Montreal, Canada, 1995: 93~102
- 99 P. Olivier and J. Tsujii. Quantitative Perceptual Representation of Prepositional Semantics. *Artificial Intelligence Review*. 1994: 147~150
- 100 K. Gapp. Basic Meanings of Spatial Relations: Computation and Evaluation of 3d Space. *Proceedings of the twelfth national conference on Artificial intelligence*. Seattle, Washington, United States. 1994, 2: 1393~1398
- 101 J. Lyons. *Semantics (Vol. 2)*. Cambridge University Press. 1977: 690
- 102 E. Gibson. *Principles of Perceptual Learning and Development*. New Jersey, Prentice Hall Inc. 1969: 37
- 103 恽为民,席裕庚. 遗传算法的全局收敛性和计算效率分析. *控制理论与应用*. 1996, 13(4):45~460
- 104 P. Bonnefoi and D. Plemenos. Object Oriented Constraint Satisfaction for Hierarchical Declarative Scene Modeling. *International Conferences in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG'99)*, Plzen, Czech Republic, 1999:194~215
- 105 <http://multimedia.lycos.com>
- 106 Donquix. Mastering the Retrieval of Google V4.0. <http://www.se-express.com/article/article13-2.htm>
- 107 王斌. 汉英双语语料库自动对齐研究. 博士论文, 中国科学院计算技术研究所, 1999, 7
- 108 B. Medjdoub and B. Yannou. Dynamic Space Ordering at a Topological Level in Space Planning. *Artificial Intelligence in Engineering*. 2001, 15: 47~60
- 109 S. Sanchez, O. Roux and H. Luga. Constraint-based 3D Object Layout using a Genetic Algorithm. *The Sixth International Conference on Computer Graphics and Artificial Intelligence (3IA'03)*. Limoges, France, 2003

- 110 W. Terry. Procedures as a Representation for Data in a Computer Program for Understanding Natural Language. MIT AI Technical Report 235, 1971
- 111 W. Terry. Understanding Natural Language. Academic Press, 1972
- 112 W. Terry. <http://hci.stanford.edu/winograd/shrdlu/code>
- 113 W. Terry. [http:// hci.stanford.edu/ ~winograd/ shrdlu/ download/ java-shrdlu.zip](http://hci.stanford.edu/~winograd/shrdlu/download/java-shrdlu.zip)
- 114 T. Noma, K. Kai and J. Nakamura. Translating from Natural Language Story to Computer Animation. Proceedings of the First Singapore International Conference on Interligent System (SPICIS'92). 1992: 475~480
- 115 Y. Takashima. Story Driven Animation. Proceedings of SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface (CHI+GI'87). 1987: 149 ~153
- 116 A. Mukerjee. Modeling Visual Objects Based on Accidental Alignments. IEEE Conference on Systems, Man and Cybernetics. Los Angeles, 1990: 539~542
- 117 A. Mukerjee. Neat vs Scruffy: a Review of Computational Models for Spatial Expressions. P. Olivier and K. Gapp, editors, Computational Representation and Processing of Spatial Expressions. Lawrence Erlbaum Associates, Mahwah, NJ. 1998: 1~32
- 118 A. Mukerjee, R. Agrawal and N. Tiwari. Qualitative Sketch Optimization. Journal of Artificial Intelligence in Engineering Design Analysis and Manufacturing (AI-EDAM). 1997, 11(4): 311~323
- 119 R. Sproat. Inferring the Environment in a Text-to-scene Conversion system. Proceedings of the K-CAP'01. 2001: 1~8
- 120 R.Q. Lu. 知识科学与计算科学. 清华大学出版社. 2003: 3~10
- 121 S. Zhang. Story grammar — a Failed Approach or Apromising Technique. Proceedings of 14th International Conference on AI, KBS, Expert Systems and Natural Language. Paris, France, 1994
- 122 张松懋. 相关的弱优先故事分析文法. 计算机学报. 1994, 17(6): 477~480
- 123 张松懋. 含无序产生式的故事分析文法的研究. 软件学报. 1994, 5(1): 11~18
- 124 R.Q. Lu and Z. Jin. Automatically Multi-paradigm Requirements Modeling and Analyzing: An Ontology-based Approach. 中国科学 F 辑 (英文版) SCIENCE IN CHINA (INFORMATION SCIENCES). 2003, 46(4): 279~297

- 125 R.Q. Lu and S.M. Zhang. Automatic Generation of Computer Animation: using AI for movie animation. Berlin. 2002: 1~198
- 126 R. Lu, S. Zhang and H. Shi. Solving the Quantification Problem in Animation Generation. Australian Journal of Intelligent Information Processing Systems, 1995, 2(4): 35~45
- 127 R. Johansson, A. Berglund and M. Danielsson. Automatic Text-to-scene Conversion in the Traffic Accident Domain. Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI-05). Edinburgh, Scotland, 2005: 1073~1078
- 128 R. Johansson and P. Nugues. Automatic Conversion of Traffic Accident Reports into 3D Animations. The Annual SIGRAD Conference Special Theme: Mobile Graphics (SIGRAD-05). Lund, Sweden, 2005: 340~345
- 129 R. Johansson. Natural Language Processing Methods for Automatic Illustration of Text. Lund Institute of Technology, Lund University, Licentiate Thesis, 2006
- 130 W. Ruchaud and D. Plemenos. MultiFormes: a Declarative Modeller as a 3D Scene Sketching Tool. International Conference ICCVG'2002, Zakopane (Poland), 2002: 1~6
- 131 吕亚娟. 双语语料库对齐及翻译知识自动获取技术研究. 哈尔滨工业大学博士学位论文. 2003: 1~99
- 132 金辉,石敏. 基于 Cortona 解释器的 VRML Automation 软件开发技术. 系统仿真学报.2002, 14(9): 1179~1182
- 133 J. Cohen. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement. 1960, 20:37~46
- 134 俞士汶, 姜新, 朱学锋. 机器翻译译文质量评价的实践与分析. 中文电脑国际会议论文集. 1994:26~32
- 135 俞士汶, 段慧明. 英汉机器翻译译文质量测试大纲. 计算机世界. 1998.4. 13:10~11

附录

附录 1: 方位词词形实例集合

SPG = {上, 下, 前, 后, 东, 西, 南, 北, 东面, 西面, 南面, 北面, 东头, 西头, 南头, 北头, 上头, 下头, 前头, 后头, 里头, 外头, 左, 右, 里, 外, 中, 内, 旁, 之上, 之下, 之前, 之后, 之东, 之西, 之南, 之北, 之外, 之内, 之中, 之间, 上边, 下边, 前边, 后边, 东边, 西边, 南边, 北边, 左边, 右边, 里边, 外边, 上面, 下面, 前面, 后面, 左面, 右面, 里面, 外面, 以东, 以西, 以南, 以北, 以内, 以外, 上下, 前后, 南北, 左右, 里外, 内外, 底下, 顶上, 头里, 当中, 当间, 中间, 内中, 内里, 面前, 跟前, 头前, 背后, 旁边, 边上, 面上, 对面, 这儿, 这里, 那儿, 那里, 以远, 以近, 附近, 周围, 上部, 下部, 前部, 后部, 东部, 西部, 南部, 北部, 中部, 内部, 东端, 西端, 南端, 北端, 左端, 右端, 上端, 下端, 前端, 后端, 东侧, 西侧, 南侧, 北侧, 左侧, 右侧, 外侧, 内侧, 东南边, 东北边, 西南边, 西北边, 东南面, 东北面, 西南面, 西北面, 东南部, 东北部, 西南部, 西北部, 东南方, 东北方, 西南方, 西北方, 东南角, 东北角, 西南角, 西北角, 左上头, 左下头, 右上头, 右下头, 左上部, 左下部, 右上部, 右下部, 左上方, 左下方, 右上方, 右下方, 前上方, 后上方, 左上角, 左下角, 右上角, 右下角, 一边, 一面, 一头, 两边, 两头, 两旁, 两旁边, 四面, 四边, 四外, 四周, 前后左右, 边, 头, 尾, 面, 顶, 外边, 侧, 际, 四围, 四周围, 中心, 东南, 东北, 西南, 西北, }

攻读博士学位期间发表的论文

1. 李晗静, 李生, 赵铁军. 汉语中方位参考点恢复研究. 计算机研究与发展. (已录用, 预 2007 年 2 月发表, EI 检索源)
2. 李晗静, 李生, 赵铁军. 基于自然语言理解的实体自动摆放的研究. 电子与信息学报. (已录用, 预 2007 年发表, EI 检索源)
3. Hanjing Li, Tiejun Zhao, Sheng Li, Yanhai Han. The Extraction of Spatial Relationships from Texts Based on Hybrid Method. 2006 IEEE International Conference on Information Acquisition (ICIA2006), Weihai, China, Aug. 2006: (ei indexed)
4. Hanjing Li, Tiejun Zhao, Sheng Li. Graphic Retrieval Based on Limited Semantics. The IEEE International Conference on Natural Language Processing and Knowledge Engineering (IEEE NLP-KE 2005), Wuhan, China, 2005: 535-539 (Accession Number: 8884615)
5. 李晗静, 赵铁军. EH 实时计算机图像生成系统设计. 哈尔滨工业大学学报, 第 37 卷, 第 6 期. 2005, 37(6): 760-762 (EI 检索源. 2005-ei-0482)
6. 李晗静, 郭恒业. 合并纹理图像技术. 计算机工程与应用. 2003(29): 90-91, 124 (2003-GNLW-0895)
7. Hanjing Li, Sheng Li, Tiejun Zhao. The Extraction of the Trjectory from Text Based on the Linear Classification. (投递到 International Journal of Computational Linguistics and Chinese Processing.)
8. 李晗静. VRML 格式转换成专用格式的软件实现. 电脑学习. 2002(3): 60-62
9. 李理, 李晗静, 赵铁军. 基于遗传算法空间物体最优摆放. 第 10 届 机器学习会议. 海口, 2006, 10
10. 赵纪元, 李晗静, 赵铁军. 汉语空间表达式中射体识别研究与分析. 第三届学生计算语言学联合学术会议. 沈阳, 2006, 8

哈尔滨工业大学博士学位论文原创性声明

本人郑重声明：此处所提交的博士学位论文《基于自然语言描述的空间概念建模研究》，是本人在导师指导下，在哈尔滨工业大学攻读博士学位期间独立进行研究工作所取得的成果。据本人所知，论文中除已注明部分外不包含他人已发表或撰写过的研究成果。对本文的研究工作做出重要贡献的个人和集体，均已在文中以明确方式注明。本声明的法律结果将完全由本人承担。

作者签字： 日期： 年 月 日

哈尔滨工业大学博士学位论文使用授权书

《基于自然语言描述的空间概念建模研究》系本人在哈尔滨工业大学攻读博士学位期间在导师指导下完成的博士学位论文。本论文的研究成果归哈尔滨工业大学所有，本论文的研究内容不得以其它单位的名义发表。本人完全了解哈尔滨工业大学关于保存、使用学位论文的规定，同意学校保留并向有关部门送交论文的复印件和电子版本，允许论文被查阅和借阅，同意学校将论文加入《中国优秀博硕士学位论文全文数据库》和编入《中国知识资源总库》。本人授权哈尔滨工业大学，可以采用影印、缩印或其他复制手段保存论文，可以公布论文的全部或部分内容。

作者签名： 日期： 年 月 日

导师签名： 日期： 年 月 日

哈尔滨工业大学博士学位涉密论文管理

根据《哈尔滨工业大学关于国家秘密载体保密管理的规定》，毕业论文答辩必须由导师进行保密初审，外寄论文由科研处复审。涉密毕业论文，由学生按学校规定的统一程序在导师指导下填报密级和保密期限。

本学位论文属于 保密 ，在 年解密后适用本授权书
不保密

(请在以上相应方框内打“√”)

作者签名： 日期： 年 月 日

导师签名： 日期： 年 月 日

致谢

值此论文完成之际，心中感慨良多。这篇论文能够得以顺利完成，使我既体会到了辛勤劳动后的喜悦，又深深地感到它与大家的热心帮助是分不开的。

首先向我尊敬的导师李生教授致以深深的谢意。李老师开阔的视野、严谨的学风以及严以律己、宽以待人的高尚品质无不是学生学习的楷模。是李老师高瞻远瞩的考虑得以这个题目顺利敲定，论文的每一步进展都离不开李老师的悉心指导和亲切关怀。

感谢赵铁军教授多年来的大力支持和帮助，是赵老师把我领进了自然语言处理研究的大门。他渊博的知识、严谨的治学态度和勤恳的工作作风深深感染了我，始终激励着我不断地把工作做得更好。在本文撰写过程中，每一步进展都和赵老师的指导和支持分不开，再次向赵老师表示深深的谢意。

感谢中科院的陆汝钫院士和微软亚洲研究院的黄昌宁教授的启发和帮助。是这两位老师启发，作者的研究工作得以深入。

感谢美国宾夕法尼亚大学人体建模仿真中心的 Norm Badler 教授，以及瑞典 Linköping 大学自然语言处理实验室的 Lars Ahrenberg 教授的关心和帮助。作者在两处分别访问期间得到两位教授对本文部分工作的诚恳建议。两位教授渊博的知识、缜密的思维使本人受益匪浅。

感谢杨沐昀、于浩、郑德权、徐冰几位同事的鼓励和支持，感谢一起学习与奋斗过的薛永增、韩习武、张春祥、梁颖红、谭红叶、骆婕、韩延海、李理、叶利军、李士奇、赵纪元、季伟、郭键以及研究室的所有成员。这篇论文的顺利完成是与他们的热心帮助和密切合作分不开的。

感谢本人工作的由王宇颖老师率领计算机语言基础教研室的所有同事，感谢他们的关心、支持和鼓励。

感谢培育我的哈尔滨工业大学，感谢计算机系的所有老师们。

感谢国家 863 高科技项目基金、国家自然科学基金对本课题研究的支持。

感谢对本论文进行评审、提出宝贵意见的各位专家。

向养育我长大成人的父母致以崇高敬意，感谢他们赋予我生命、培养我性格、历练我思维，感谢他们支持我、激励我、鼓舞我。我的成功属于他们。

感谢所有曾经给予我关心、支持和帮助的人们。

个人简历

李晗静，女，汉族，1974年2月生。

教育背景：

1993.9-1997.7 吉林大学计算机科学系，获理学学士学位。

2000.9-2002.7 哈尔滨工业大学计算机系，获工学硕士学位。

2002.3- 至今 哈尔滨工业大学计算机系，博士研究生。

工作经历

1997.7- 2003.7 哈尔滨工业大学计算机系，助教。

2003.7- 至今 哈尔滨工业大学计算机系，讲师。

研究方向： 自然语言理解、文景转换

主要科研工作及成果：

1. 参加国家“973”项目，独立完成虚拟场景建模系统人机交互开发；
2. 参加国家“863”项目，独立完成基于实例汉日翻译原型系统的开发；
3. 参加国家自然科学基金项目，名称是基于ontology的自然语言描述的空间概念三维可视化研究（60575041），排名第二；
4. 参加国家自然科学基金项目，名称是面向特定领域基于Ontology跨语言信息检索技术研究（60302021），排名第四；
5. 参加242项目，名称是基于内容和链接分析的Blog信息获取关键技术研究，排名第四；
6. 参加多项国家自然科学基金项目和863项目，负责自然语言理解的相关研究；2002年获得国家教育部公派访问学者项目，作为访问学者于瑞典Linkoping大学访问，从事本体论研究；
8. 2003年获得哈尔滨工业大学双语教学项目，作为访问学者于美国宾夕法尼亚大学短期访问，从事自然语言指令驱动动画研究；
9. 先后在国内外会议和期刊上发表论文10多篇；