

对话管理和可定制 对话系统框架的研究

(申请清华大学工学博士学位论文)

培 养 单 位 : 清华大学计算机科学与技术系

学 科 : 计算机科学与技术

研 究 生 : 邬 晓 钧

指 导 教 师 : 吴 文 虎 教 授

副 指 导 教 师 : 郑 方 副 教 授

二〇〇三年十月

Research on Dialogue Management and Customized Framework for Dialogue Systems

Dissertation Submitted to
Tsinghua University
in partial fulfillment of the requirement
for the degree of
Doctor of Engineering

by
Xiao-jun WU
(**Computer Science and Technology**)

Dissertation Supervisor: Professor Wen-hu WU

Associate Supervisor: Associate Professor Fang ZHENG

October, 2003

关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。

(保密的论文在解密后遵守此规定)

作者签名：_____

导师签名：_____

日 期：_____

日 期：_____

摘要

信息咨询的口语对话系统是当前对话系统研究的热点，对话管理是对话系统的核心技术之一。论文工作在对话管理的基础理论、具体方法和实际应用等方面都进行了深入的研究和积极的实践，成果如下：

1、提出对话管理的心理模型。现有多种对话管理方法借用传统自然语言理解技术中的对话模型作为理论指导，关注于对话的语言表层，不能广泛适用于各种领域任务，无法获得良好的系统性能。所提出的对话管理的心理模型吸取了心理语言学关于对话心理的理论经验，由知识、交互历史和对话策略三个部分组成，输入为对话语义，输出为应答语义。模型的三个组成部分相互联系相互影响，完成知识管理、交互历史管理和应答推理三方面的对话管理任务。

2、提出主题森林的知识表示方法。已有多种对话管理方法在领域概念和图式表示上或者过于偏重概念表示，或者过于偏重图式表示，使得相应的对话管理方法无法广泛适用。主题森林的知识表示方法，能应用于包含多个主题和复杂主题的领域任务，能表示主题中信息项的不同重要程度和主题间的共享信息。主题森林的表示方法将领域概念表示和图式表示有机地结合在一起，还可直接表示对话交互知识和对话交互历史，可替代现有对话管理中的多种表示方法，适用于不同的领域任务。

3、提出基于主题森林的对话管理方法，给出了知识管理、交互历史管理和应答推理的算法。背景任务的对话系统 EasyFlight 应用了所提出的对话管理方法，具有良好的系统表现。大规模的系统评测结果说明，本文所提出的对话管理方法在一致性、灵活性和可移植性各个方面都达到了较高的水平。

4、提出一个支持多路对话的可定制对话系统框架。该框架的核心是基于主题森林的对话管理器，其自然语言理解模块采用了基于语义类的上下文无关增强方法。多个项目的应用实践表明，利用该框架可快速有效地建立信息咨询的汉语口语对话系统。

关键词：对话系统，对话管理，对话管理模型，可定制对话系统框架

Abstract

Spoken dialogue systems for information retrieval are very popular in the research of dialogue systems. Dialogue management is one of the kernel technologies to set up such systems. In this dissertation, in-depth study and active practice are carried out on the fundamental theory, the concrete methods, and the practical application of dialogue management. The achievements are as following:

1. A psychological model for dialogue management is proposed. The current dialogue management methods focus on the surface layer of dialogue utterances according to the dialogue model theories that come from the technology of natural language understanding. These methods cannot be applied to tasks in different domains and the system performances are not good enough. With the reference to the theories of dialogue mentality in psycholinguistics, a psychological model for dialogue management is proposed. With utterance semantics as input and response semantics as output, this model is made up of three parts: knowledge, history and strategies, which interact and influence each other. The dialogue management can be divided into knowledge management, history management, and response reasoning.

2. A knowledge representation method is proposed as Topic Forest. There are different representations for domain notions and schema in the current dialogue management methods. Some of them emphasize on representation of notions while others on representation of schema. That is the radical reason of their poor portability. The knowledge presentation method named Topic Forest is thus proposed, which is applicable for domain tasks with multiple topics and complex topics. This method is powerful to represent the items with different importance to the topic and shared information between topics. Because the representation of domain notions are combined together with the representation of domain schema, Topic Forest can be directly used to represent the dialogue knowledge and dialogue history. This representation method can replace the current ones in different domain tasks.

3. The dialogue management method based on Topic Forest is further proposed, which includes algorithms of knowledge management, history management and response reasoning. The dialogue management method has been applied to a background dialogue system called *EasyFlight*, and the system performance is excellent. A large-scale system evaluation was executed, which demonstrated good performance on consistency, flexibility, and portability for the proposed dialogue management.

4. A customized framework for dialogue systems is proposed, which is capable of management for multiple concurrent dialogues. The core of the framework is the dialogue management module applied with the dialogue management method based on Topic Forest. In the natural language understanding module, an enhanced kind of context free grammar, which is based on semantic classes, is applied. It is concluded with the practice of several projects that the framework is very useful to set up Chinese spoken dialogue systems for information retrieval rapidly and effectively.

Keywords: dialogue systems, dialogue management, dialogue management model, customized frameworks for dialogue systems

目 录

摘 要	I
ABSTRACT (英文摘要)	II
第一章 对话系统概述	1
1.1 对话系统的发展	1
1.1.1 自然语言理解的发展	1
1.1.2 语音技术的发展	2
1.1.3 信息咨询的口语对话系统研究	3
1.2 对话系统研究现状	4
1.2.1 对话系统组成	4
1.2.2 对话系统介绍	4
1.2.3 对话系统的评价	6
1.2.4 口语对话系统研究中的几个主要问题	7
1.3 对话管理	8
1.3.1 对话管理的概念	8
1.3.2 对话管理的设计和评价准则	9
1.4 研究工作概述	10
1.5 论文组织	12
第二章 对话管理的心理模型	13
2.1 对话管理方法	13
2.1.1 基于对话语法模型的方法	14
2.1.2 基于对话规划模型的方法	15
2.1.3 基于对话互动模型的方法	18
2.1.4 现有对话管理方法总结	19
2.2 心理语言学的相关理论	20
2.2.1 语言理解和心理表征	21
2.2.2 概念储存模型	21

2.2.3 图式	22
2.2.4 思维模型	24
2.2.5 语言生成理论	24
2.2.6 现有对话管理方法的心理本质分析	25
2.3 对话管理的心理模型	26
2.3.1 对话管理模型的组成	26
2.3.2 对话管理模型的运作	28
2.3.3 对话管理模型的实用指导意义	30
第三章 主题森林	33
3.1 对话系统中的知识表示	33
3.1.1 知识表示的内容	33
3.1.2 对现有对话系统中的知识表示方法的分析	34
3.2 主题森林	37
3.2.1 主题树的结构	38
3.2.2 主题树的建立	39
3.2.3 主题树的概念与图式功能	40
3.2.4 共享信息索引	42
3.2.5 与其它表示方法的比较	43
3.3 主题森林特点总结	44
第四章 基于主题森林的对话管理方法	46
4.1 知识管理	46
4.1.1 领域概念和图式管理	46
4.1.2 数据库管理	47
4.1.3 交互知识管理	47
4.2 交互历史管理	50
4.3 应答推理	51
4.3.1 与主题森林结构相关的应答生成	51
4.3.2 应答焦点的推理算法	52
4.4 基于主题森林的对话管理流程	53

4.5 对话举例	55
4.6 系统评测	56
4.6.1 对话场景设计	57
4.6.2 用户满意度调查问卷	57
4.6.3 客观指标与系统日志	58
4.6.4 评测方法	58
4.6.5 主观评测结果	59
4.6.6 客观评测结果	63
4.6.7 系统评测总结	67
4.7 对话管理特点总结	68
第五章 可定制对话系统框架	69
5.1 对话系统的建立	69
5.1.1 WOZ 方法与开发-更新方法	69
5.1.2 对话系统开发工具	70
5.1.3 对话系统开发工具的需求及技术分析	71
5.2 可定制对话系统框架	72
5.2.1 对话系统框架组成	73
5.2.2 各组成模块的输入输出	74
5.2.3 自然语言理解模块的定制	75
5.2.4 对话管理器的定制	78
5.2.5 数据库管理和自然语言生成模块的定制	80
5.2.6 对话系统的建立	81
5.3 可定制对话系统框架应用实例	82
5.3.1 绿色食品信息查询系统	82
5.3.2 足球竞猜系统	86
5.4 可定制对话系统框架总结	87
第六章 论文工作总结	89
参考文献	91
致谢及声明	97

目 录

附 录	98
个人简历、在学期间的研究成果及发表的论文	111

第一章 对话系统概述

对话系统是人机对话系统的简称，是与人进行信息交互的计算机系统。信息交互的方式可以是文字，可以是语音，也可以是人的眼神、姿态、动作。早期的对话系统是基于自然语言理解技术的纯文字人机接口；随着语音技术的逐渐发展，出现了人机语声接口的对话系统；当触摸屏技术、图像处理技术发展起来之后，又出现了多模式（multi-model）的对话系统。

本章首先介绍自然语言理解技术和语音技术发展对对话系统研究产生的影响，分析对信息咨询的口语对话系统进行研究的重要性和必要性；接着对口语对话系统的研究现状进行总结，指出口语对话系统研究的重点和难点；然后介绍对话系统的核心——对话管理的概念；最后给出本论文工作的具体内容和创新之处。

1.1 对话系统的发展

对话系统的发展与自然语言理解的发展和语音技术的发展密切相关。

1.1.1 自然语言理解的发展

所谓自然语言，是相对人造语言而言的。自然语言是指人类语言集团的本族语^[1]，如汉语、英语、法语等，而人造语言指世界语或计算机的各种程序语言。语言作为思维的载体，是人际交流的重要工具。在信息化社会中，语言信息处理的技术水平和年处理信息量已成为衡量一个国家现代化水平的重要标志之一。

自然语言理解作为语言信息处理技术的一个高层次的重要方向，一直是人工智能界所关注的核心课题之一。创造和使用自然语言是人类高度智能的表现，对自然语言的研究有助于我们探究人类语言、语言规律、语用规律，以及人的思维、思维机制、思维过程和思维规律^[2]。如果计算机能够理解自然语言，那么人与计算机间的信息交流就能以人们所熟悉的本族语来进行。所以对自然语言理解的研究在理论方面和应用方面都有重大的意义。

自然语言理解的发展大体上经历了三个时期：60年代以关键词匹配为主流的早期，70年代以句法语义分析为主流的中期，以及80年代后走向实用化和工程化的近期^[1]。

60年代开发的自然语言理解系统，主要依靠关键词匹配技术来识别输入句子的意义，著名的系统如美国麻省理工学院的 SIR 系统、STUDENT 系统和 ELIZA 系统。70年代采用句法语义分析技术后，在语言分析的深度和难度方面都比早期有了很大的进步，著名的系统有美国 BBN 公司设计的世界上第一个可以用普通英语同计算机数据库对话的人机接口 LUNAR、美国麻省理工学院设计的通过人机对话进行搭积木游戏的 SHRDLU 系统、美国斯坦福大学的 MARGIE 系统。80年代以来，出现了一大批商品化的自然语言人机接口，如美国人工智能公司生产的英语人机接口系统 Intellect、美国弗雷公司生产的 Themis 人机接口、日本富士通公司开发的 ATLAS 英日英机译系统等等。

1.1.2 语音技术的发展

语音技术可大致分为语音识别技术和语音合成技术。

大规模的语音识别研究开始于70年代初期^[3]，到80年代出现了语音命令技术，进入90年代后，出现了说话人识别、人名拨号、数字串识别、关键词检出等多方面的技术。随着大词汇量、非特定人、连续语音识别技术的日趋成熟，以及大规模集成电路和超大规模集成电路的迅速发展，语音识别系统性能不断提高，系统的实时性成为可能。

语音合成技术的研究可以追溯到公元17世纪，那时人们用机械装置来模拟人发音。70年代以后，大规模集成电路和超大规模集成电路技术的发展，特别是数字信号处理与实验语音学、现代音韵学的交叉发展和互相促进，推动了语音合成技术的长足进步。到80年代末，基音同步叠加法（Pitch-Synchronous Overlap Add，即 PSOLA）的提出，使语音波形编码合成技术得到了很大发展和广泛应用。

在自然语言理解的基础上，语音技术的进步使得人们渴望已久的人机语声接口这一具有重大意义的研究课题成为可能。人机语声对话系统的研

究涉及计算机技术、信号处理、自然语言处理、语言学、心理学、行为科学、生理学等诸多学科领域的知识。它将语音识别、自然语言理解、知识表示、对话管理以及语音合成等各项技术紧密地联系起来，彼此之间相互指导相互促进，最终达成以人机对话的形式进行信息交流的目的。

1.1.3 信息咨询的口语对话系统研究

信息咨询的口语对话系统是当前对话系统的研究重点，如电话语音查号、航班信息咨询、旅游景点信息咨询、食品订购服务等。这类对话系统有两大特点，一是面向特定任务（task-oriented）或某一特定领域的（domain-dependent）^[4]，二是采用口语作为人机交互方式。

选择具有特定面向的系统作为研究对象有两方面的原因。其一，从实际应用的角度看，现代社会是信息社会，信息咨询是人们生活中不可缺少的部分。这类对话系统应用于实际，可以为社会节省大量的资源，可以提高信息服务的质量。其二，从研究策略的角度看，广义的对话系统非常复杂且难以获得所需的心理和行为的计算模型，而信息咨询的对话系统似乎是一种比较简单、易于获得计算模型、系统实际所需工作量较少的对话系统。同时，这类对话系统也具有一定的代表性，是进一步研究广义对话系统的基础。

语言是人类最方便使用最多的交流媒介。用口语作为人机交互的方式，对用户来说最方便，而且可以适应多数的场合。语言的表达能力也是其它交流方式所无法比拟的。Bill Gates 在 2003 年 2 月“微软新一代技术展望”大会上指出，随着通讯技术和网络技术的发展，几乎无限的海量信息已不可能在有限的屏幕上展现，人们要求能随时随地获取信息，就会抛弃键盘、手写笔、摄像头之类的工具，最终的人机接口只能是具有自然语言理解能力的语音接口。

由此可见，研究信息咨询的口语对话系统有其重要性和必要性。

近十年来，发达国家投入了大量的人力、物力、财力来研究信息咨询的人机口语对话系统，美国有 DARPA 的 Communicator 计划，欧洲有 ARISE（自动铁路信息研究）计划、REWARD 计划和 VERBMOBIL 计划，日本文

部省也曾于 1990-1991 年间以巨资支持人机对话系统的研究。许多著名的学府和研究机构也竞相开展这项研究，例如美国麻省理工学院的 SLS 实验室^[5]、卡耐基梅隆大学的 ISL 实验室、Lucent-Bell 实验室、日本的 ATR 实验室、德国的 Erlangen-Nuremberg 大学和 Philips 公司等。在国内，人机口语对话工程也已紧锣密鼓地展开。中科院自动化所^[6]、清华大学^[7]、香港中文大学^[8]、台湾大学^[9]等，都投入了相当大的精力进行这方面的研究。

1.2 对话系统研究现状

1.2.1 对话系统组成

通常一个人机口语对话系统主要由以下几个功能模块组成：语音识别模块、自然语言理解模块、对话管理模块、自然语言生成模块和语音合成模块。其中，对话管理模块往往与一个后端数据库相联，存取有关的信息，如图 1-1 所示。

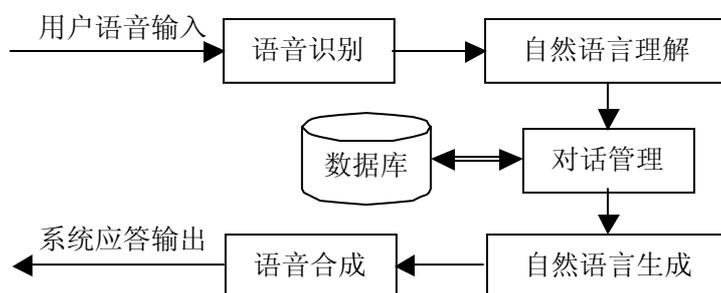


图 1-1 人机口语对话系统的主要模块

在信息咨询的人机口语对话系统中，用户通过自然语言表达自己的思想，与计算机就某一领域的内容进行信息交互。计算机识别并理解用户的言语，了解用户的意图，并根据相关的数据库信息做出适当的应答。以上过程不断重复，最终达到交换信息和解决问题的目的。

1.2.2 对话系统介绍

根据不同应用，对话系统可以构建于不同的平台，有着不同的表现形式。

嵌入式平台：Huginin 等人^[10]设计了一个基于 Microsoft Excel 的嵌入式电子表格系统。该系统采用人机对话的方式进行电子表格的自动设计与填充。与使用鼠标键盘的手工输入方式和简单屏蔽鼠标键盘的语音命令方式相比，语音对话的方式提高了效率，而且使用户更加轻松。

WWW 平台：Issar^[11]设计了一个用于在 WWW 网页上填充表格的语音软件。该软件使用 Java applet 作为用户接口，采用名为 Sphinx-II 的语音识别器和基于框架的语义分析器，以 plug-in 程序的方式处理语音输入输出。Issar 认为，这种基于表格的语音接口是探索分布式自然语言系统的重要一步。

机器人平台：Jijo-2^[12]是一个可移动的办公室机器人，能通过语音对话的方式，完成人员查询、引路、接通特定人电话、给特定人发电子邮件等任务。这类机器人平台的系统面临的主要问题是实际使用环境中的噪音，以及系统响应的实时性。

电话平台：随着大量公有信息的出现（订票、信息查询等）以及电话的普及，基于电话的对话系统越来越多。欧洲的 ARISE 计划^[13]下有法语、荷兰语、意大利语等若干系统，Els 等研究人员对各系统进行了横向比较，这有助于找到不同方法的优缺点和提高研究水平。基于电话平台的系统应用前景广阔，有很好的社会效益和经济效益，其技术挑战主要在于电话信道的窄带特性、信道之间的差异，以及现实生活中的噪音问题。

以下是国内外一些对话系统的简介。

一、麻省理工学院的 GALAXY 系统^[5]。这是一个通过口语对话获取旅游信息的系统，有大约 1500 个词的词汇量，能够提供大约 750 个城市的天气预报和大约 250 个城市的航班情况。它的语音识别器 SUMMIT，采用基于分段（Segment-Based）的识别方法，建立了 Anti-Phones 模型，词识别率为 83.9%；它直接采用了另一个对话系统 TINA^[14]的自然语言理解模块，用语义框架的结构来描述语义；自然语言生成为 GENESIS；语音合成采用 Off-the-shelf 的硬件和软件。该系统的第二代 GALAXY-II^[15, 16, 17]采用了 client-server 体系结构，成为美国 DARPA Communicator Program 的第一个

参考体系结构。GALAXY-II 系统作为发展人类语言技术的试验平台，在其基础上，已经开发了许多不同领域、不同语言的系统，如电话天气预报查询系统 JUPITER^[18]，航班订票系统 MERCURY^[19]。

二、德国的 VERBMOBIL 系统。这个对话系统用于会议的安排，可以识别并翻译大量的不同口语表达。它通过一个动态建立的上下文模型和一个建立在语料库之上的随机模型，可以预测对话某一点的下一句将会是什么。

三、由英德法意等国共同开发的 SUNDIAL 系统^[20]。这是一个提供航班和火车时刻信息的电话口语对话系统。它的词汇量为 1000 词左右，是非特定人的系统，而且具有很好的对话管理功能，通过电话进行的对话成功率达到 96%。

四、中国科学院自动化所模式识别国家实验室的 LODESTAR 系统^[6]。该系统向用户提供旅游信息，并且可以根据用户的要求计划旅游路线。它采用了大词表连续语音识别的技术，识别结果经过语义项的匹配得到有关的语义概念。它实现了对话的人机混合主导，基于模板生成系统应答，整个系统的应答准确率达到 90.9%。

五、清华大学智能技术与系统国家重点实验室语音技术中心的 EasyNav 系统^[21]。该系统向用户提供友好的清华大学校园导游服务，包括校园内的建筑物信息和交通信息。它考虑了口语中的省略指代现象，能处理上下文相关的对话。当信息查询结果为空时，该系统还会主动放宽某些约束条件，提供用户可能关心的信息。

1.2.3 对话系统的评价

在人工智能界，人们普遍认为可以采用著名的图灵（Turing）试验（Turing 于 1950 年在《Mind》杂志上发表了名为《计算机与智能》的论文，提出了如何判别一台计算机是有智能的实验方法）来判断计算机是否“理解”了某种自然语言，这主要用于评价文本输入的自然语言理解系统。

对于一个特定领域内用于完成特定任务的人机口语对话系统，目前并没有公认和统一的评价标准。常用的客观标准包括：交互时间、对话回合数、修正修补回合占对话总回合数的比率、交互成功率等，可直接由计算机统计得到；常用的主观标准包括：使用是否方便、应答是否清楚自然友好、系统是否鲁棒、系统误解率、主观交互时间等，一般要求用户填写调查表，再进行统计得到。

然而，由于各个对话系统的领域任务不同，相同任务的系统评测时所用对话场景不同，而且各系统采用的应答策略不同，自然语言生成也不同，所以各系统评测结果可比性不大。

1.2.4 口语对话系统研究中的几个主要问题

当前在口语对话系统的研究中，有以下几个重要的问题。

一、口语问题。由于采用语音作为交互方式，系统接收的是（或十分接近）日常生活中的口语，发音比较随意，给语音识别带来较大困难。与书面语相比，口语中包含大量的口语现象，如停顿、重复、更正、无意义词等，而且口语句式比较随意，给自然语言理解带来较大困难。

二、对话上下文问题。对话是一个有特定背景连续交互过程，在这一过程中经常出现上下文省略和指代的情况。一句话的意义有时要结合对话上下文或者相关的背景（如对话场所、对话时间）才能确定，而现有的自然语言理解主要基于上下文无关文法，因此对话系统需要维护对话上下文，包括交互的语义和特定对话的背景。

三、领域任务的知识表示问题。知识表示一直就是人工智能领域的重要课题，也是对话系统提供信息服务的基础。对话系统相关的领域任务可能有复杂的组成，牵涉很多的因素，只有了解这些因素的关系和相关的含义，才能与用户做到真正意义上的交流。

四、对话策略问题。对话策略涉及很多方面，其中最主要的有对话的主导方式，对话信息的确认机制。对话主导方式可以分为用户主导（user initiative）、系统主导（system initiative）和混合主导（mixed initiative）三

类。在口语对话中，往往认为对话中的提问者就是主导对话进行者。作为一个对话的积极参与者，任何一方都应该主动配合对方，同时在必要时主导对话进行，形成混合主导的方式。早期的对话系统是由用户主导或系统主导的，在当前的对话管理研究中，系统应答的目标是自然友好积极，在不会发生问题的情况下，让用户尽可能自主，实现对话的混合主导。当前的语音识别技术和自然语言理解技术都有一定的错误率，为了保证对话信息的正确处理，系统要有一套确认机制。确认方式有显示和隐式两种，多项信息可以在一个或多个回合中进行确认。不同的确认策略会影响系统的易用性和友好性。

1.3 对话管理

1.3.1 对话管理的概念

对话管理，又称对话控制，其目的是在人机对话系统中指导产生计算机这一方的所有话语^[22]。对话管理需要引导对话的发展并且实现特定的任务，解决对话中出现的识别和理解错误，关系到上述对话系统研究中后三个重要问题，是整个对话系统的核心。

智能的对话系统应该不仅能回答用户提出的问题，而且在某些情况下能主动提出问题来澄清一些模糊的概念，这些情况包括识别错误、用户提供信息不足、上下文语义不一致等。因此对话管理要设计一定的交互策略来应付多种可能出现的对话情况，如怎样将大量的信息分散在一些短小的句子中表达出来，用户要求的信息没有该怎么办。对话管理应该能够提供基于对话上下文的有用信息来解决用户疑难，而不是简单地回答是否或简单地罗列信息。

对话管理不仅决定了系统的应答语义（自然语言生成模块决定具体的应答语句），还可以预测对话的进行方向，从而为语音识别和自然语言理解模块提供一定的指导。

1.3.2 对话管理的设计和评价准则

为了使系统能够完成目标任务，并且具有较好的性能，对话管理的设计应该遵循一些原则。这些原则也是对话管理的评价标准。

1.3.2.1 一致性原则

在通常的对话过程中，往往出现这样的情况：用户起初对自己的目标只有一个模糊的认识，在与对方的对话过程中，才逐步地明确和完善这一目标，最后通过对方的帮助来完成目标任务。以订机票为例，起初用户可能只确定了终点城市，连要哪一天的航班都没有确定（因为航班并不一定每天都有）。随着对话的深入，用户最后会决定要哪一天的航班，还会确定要什么时候起飞的，甚至提出要乘坐某一特定型号的飞机。即使用户先前有确定的计划，也可能因为售票情况、天气因素、时间关系等而改变初衷。在用户确定目标的多次反复过程中，以及根据上下文从其他领域的主题获取信息的过程中，保持知识的前后一致性是非常重要的问题。

在对话过程中主要存在两类一致性问题。

其一，语音识别系统不可避免的存在识别错误，导致内容上的不一致性。如果是单句中出现了不一致情况，则在语义分析阶段可能检查出来；如果导致的是上下文间的不一致，则需要对话管理来解决。

其二，不管是用户还是系统，都有可能主动在不同主题间切换，这时就要保证主题间相关内容的一致性。当系统主动切换时要保持上下文内容的一致，当用户主动切换时，系统首先要检测一致性，再采取相应动作。

1.3.2.2 灵活性原则

对话系统在与用户进行交互时应该具有一定的灵活性。

首先，当对话主题切换时，系统应该能够明白不同主题间相同内容的共享关系。例如在订票时，用户经常会考虑天气问题。当用户询问几天内天气情况时，系统应该意识到用户是有意于在这几天内出行，生成的应答也应该体现出这一点。

其次，系统产生的对话应该有一定的灵活性，表现在对话过程的控制和应答句式的变化上。当对话过程中需要用户在多个选择中决定一个，或者出现不一致情况时，系统最好能尽量延长要求用户决定的时间，以免对话频繁地进入确认环节。系统在相同状态下，也可以生成不同句式的回应，使对话显得更加灵活和智能。

1.3.2.3 可移植性原则

现有很多实用的对话系统，但绝大多数是为某特定领域的特定任务而建立的。当要在其他领域建立一个对话系统，甚至只是在相近领域建立另一个功能相似的对话系统时，不得不另起炉灶，重新设计和搭建系统。

现有的自然语言理解系统的功能有限，只能处理自然语言在某个领域内的子集，所以当对话系统要移植到新的工作领域时，首先要更新句法语义分析模块，这引起了与对话管理模块接口的变化。不同领域的知识结构不同，即对话管理的推理基础不同，而且数据库也有所不同，所以移植系统时，还要更新对话管理的基本结构以及和数据库的接口。不同领域使用的语言有自己的特色，因此还要改变对话生成部分。

上述原因使得对话系统向不同领域移植时，要作较大的改动和补充。目前，如何设计对话管理使得系统移植时尽可能减少工作量，是研究人员普遍关注的课题。只有当新对话系统的建立不需要太多的经验和技巧，也不需要太大的工作量的时候，人机对话系统才能更快更容易地普及开来，为人们的生活提供各种便利服务。

1.4 研究工作概述

本文的研究工作始于清华大学 985 项目的 EasyFlight 系统，该系统是一个通过电话提供航班信息查询和订票的汉语口语对话系统。

EasyFlight 系统由语音识别、自然语言理解、对话管理、自然语言生成、语音合成等几个部分组成。其中语音识别部分接收电话语音，采用关键词检出技术进行识别，输出形式为词网格；自然语言理解部分采用了基于语义类的上下文无关增强文法和文法与语义的合一分析，直接在词网格上进

行操作，给出以语义框架形式表示的用户语义信息；对话管理部分根据对话模型，参考用户语义和对话的上下文，获取用户模型，并根据需要访问数据库，同时生成应答焦点，即应答所包含的主题信息项；自然语言生成部分根据应答焦点，参照对话状态和用户模型，生成合适的文本应答；最后语音合成部分将文本应答转化为语音反馈给用户。整个系统的功能模块及数据流见图 1-2。

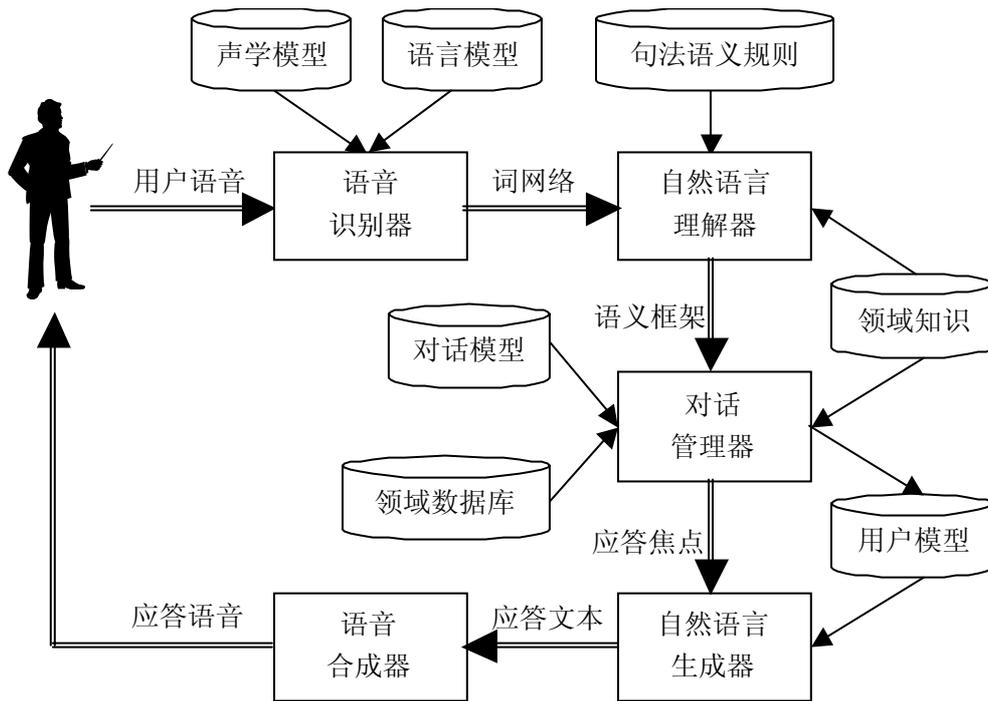


图 1-2 EasyFlight 系统的功能模块及数据流图

作者的研究任务是该口语对话系统中的对话管理技术。作为对话系统的核心，对话管理涉及前述口语对话系统研究中第二、三、四个重要问题。

在对 EasyFlight 项目进行研究开发的实践基础上，本文作者深入研究了对话系统中对话管理的行为。通过对国内外已有对话系统中对话管理方法进行研究比较，并查阅和参考了大量语言学和心理语言学中关于语言理解和对话生成的理论，作者提出了对话管理的心理模型理论。该模型描述了对话过程中思维模型的建立和演变，这正是对话管理的基础；该模型能够说明在对话系统中对话管理应具有的功能，即在对话过程中进行知识管

理、交互历史管理和应答推理，可指导实际对话系统的建立；该模型能够解释现有不同对话管理方法各自的适用条件和应用局限。

在对话管理模型理论的指导下，在具体的对话管理实践中，作者提出了主题森林的知识表示方式，适用于现有的多种信息咨询的对话系统。主题森林可用来表示复杂的领域概念和知识，表示领域任务的执行过程，同时主题森林可用作对话过程中的信息积累的框架。

作者提出了基于主题森林的对话管理方法，应用该方法的对话管理器在知识管理、对话历史管理，及应用多种交互策略进行应答推理等方面，都具有良好的性能。

在对话管理研究的基础上，作者进一步研究了对话系统各模块的协作关系，分析比较不同对话系统中的各组成模块的系统功能（共性）与领域任务特点（个性），将与领域无关的关键技术及与领域相关的特殊操作进行划分，设计并实现了可定制对话系统框架，用于建立文本或语音接口的信息咨询对话系统。该框架以基于主题森林的对话管理器为核心，提供了专用于处理汉语口语的可定制自然语言理解模块，给出了数据库操作和自然语言生成模块的定制方法。实践表明，利用该框架可快速而有效地建立多个领域的对话系统。

1.5 论文组织

本文的第二章，分析现有的对话管理方法和心理学中与对话有关的理论，提出对话管理的心理模型；第三章介绍主题森林的知识表示方法，第四章介绍基于主题森林的对话管理方法；第五章介绍可定制对话系统框架；第六章总结全文。

第二章 对话管理的心理模型

作为对话系统的核心，对话管理的重要性显而易见。在自然语言理解发展的初期，所处理的文句较简单，对对话管理要求不高。随着自然语言处理能力的提高，盲目地进行对话管理已不能满足系统要求。自然语言理解理论中的对话模型，描述了对话在语义层面的特点，不能揭示更深层面的心理规律，因此以对话模型为理论基础的各种实用的对话管理方法都只是依据对话语句的表层语义分析进行应答推理，无法触及这一过程的心理本质，因而各有其适用性和限制性。只有研究对话过程的心理规律，才能从根本上解释各种对话管理方法为什么可适用于某些领域任务，为什么不适用于某些领域，才能真正解决如何根据领域任务选择合适的对话管理方法的问题。

本章首先介绍已有的对话系统中应用的多种对话管理方法，考察它们的理论根据和内容特点，分析比较它们的优劣，指出应在对话的心理规律上加强研究。其次根据心理语言学中关于语言理解和语言生成的理论，对现有对话管理方法的心理本质进行分析。在此基础上，提出对话管理的心理模型，给出模型的组成，描述了在对话过程中该模型的运作情况，说明在实际对话管理器的设计中该模型的指导意义。

2.1 对话管理方法

对话管理的实现可以非常简单，例如在对话系统中存储一系列规则，对用户输入进行规则匹配，若输入与某规则匹配就生成相应的回答。这种对话管理的方法不考虑对话的上下文，处理能力极其有限。

在已有的对话系统中，用得较多的对话管理方法都是以对话模型理论为基础的。对话模型形式化地描述了语义和语用，以及两者之间的相互关系[23]。对话模型可大致分为三类，语法模型（dialogue grammars）、规划模型（plan-based models）和互动模型（joint action models），以不同的对

话模型为理论依据的对话管理方法分别称为有限状态方法、基于规划的方法和互动的方法。

2.1.1 基于对话语法模型的方法

常见的对话一般都是一问一答地进行，而且对话的内容有一定的顺序，对话的语法模型正是以此为出发点。语法模型理论认为：对话可以看作一个状态机，状态有先后次序和限制条件，可用类似语法描述的方法来表示；在对话的语法中，终结符是对话动作（speech acts），非终结符是对话阶段（dialogue stages）；系统根据当前的对话动作，进行适当的状态转移，产生相应的对话动作，并且可以预测对话发展。其中，对话动作指语句中包含的给出信息、提问、要求确认等语义。

这一对话模型比较假定总是可以从用户语句中分析得到对话动作，然后根据这一动作决定对话状态转移的方向，并且一次转移只有一个可能的状态。但是在实际对话过程中，一句话可能包含多个对话动作，也可能多句话组合在一起是一个对话动作，这是语法模型的不合理之处。实际对话中，对于同一个语句的对话动作，往往可以有多个合理的应答，对话的语法模型无法解释这一现象，而且这一对话模型没有指出对话过程中不同的对话动作之间有何联系。

对话管理的有限状态方法以对话的语法模型为依据，在许多对话系统中都有应用。该对话管理方法将对话分为多个状态，每个状态与某个问题相联系，用户的应答是状态的转移条件。对话状态和状态转移构成一个对话状态的网络，决定了系统在对话过程中任一状态下的应答，即决定了整个对话过程。

一个典型的例子是 Furui^[24]介绍的一个查询某地区或某商行资料的人机对话系统。这一系统采用触摸屏和语音作为输入，系统先向用户提出问题，根据不同的回答采取不同的对策。对话过程中，用户所有可能的选择都已经事先定义为不同状态下的转移条件，系统可确定对话过程的下一个状态，并且已事先定义好每一状态下的应对策略。

这一对话管理方法有很大的局限性。首先，由于对话过程已由对话系统事先确定，用户只能配合而不能主导对话的进行。如果用户在对话的某回合中提供的信息多于系统事先定义的内容，系统按既定功能处理时会忽略多余信息，但在后面的对话过程中却会向用户询问那些被忽略信息；在对话过程中，用户有时会改变原来的意图，更改某些信息，有时没有听清系统应答会要求重复，而这时系统已无法退回到之前的状态，也就无法正确回应用户。

实际系统中为了解决上面提到用户信息过量问题，经常多定义一些对话状态及状态转移条件，同时斟酌系统应答的语句，引导用户较好地配合系统，只回答系统提问的信息。很多系统采用逐步确认各项信息的策略，尽量保证用户在对话过程中不会要求对话状态的回退。这些做法，都是以对话状态转移网络的复杂化为代价的。对于复杂的领域任务，由于信息多，对话状态的网络结构本身就比较复杂，采用这些解决方法会导致对话状态和状态转移条件数目飞快膨胀，以至于对话状态难以列举清楚，状态转移条件也难以分析明确。

其次，有些对话任务是需要系统和用户进行协商的，对话相关的各项信息彼此约束相互影响，用户不可能在对话一开始就确定所有信息，只能通过和系统的交流讨论才能完成。这时对话管理的有限状态方法无法适用。

最后，领域任务可能会涉及多个相关主题，在对话过程中可以在多处自然地出现主题转移。有限状态方法难以将这些情况都事先定义到对话状态网络中。

可见，对话管理的有限状态方法只适用于信息结构相对简单、对话状态数目有限的对话任务，或者是具有单一菜单式结构和少数应答选项的对话任务。常见的应用有限状态的对话管理方法实现的对话系统主要提供名录查询^[25]，进行问卷调查，或提供简单的列车时刻查询^[26]等。

2.1.2 基于对话规划模型的方法

对话的规划模型建立在这样的假设之上：人们在对话的时候，通常头脑中早有一些大大小小的目标和实现目标的规划，对话的作用只是逐步确

定或实现这些目标和规划，最后得到所期望的结果。该模型认为，对话语句总是暗含一定的对话动作，而对话动作是人内心目标的反映，因此对话是由规划（plan）和规划识别（plan-recognition）构成的，即对话一方根据另一方的对话动作，识别对方内心的规划，以此修改或确定自己的规划并表达出来；在一个对话语句中可以包含多个规划。

对话的规划模型要求从对话的每一语句中分析和识别出规划。但是实际上，规划的识别应考虑对话的语句、对话上下文和对话者当时的精神状态等多个因素。而且与对话动作相类似，规划与对话语句也有一对多和多对一的关系，规划模型有不合理之处。另外，规划本身可能与对话任务相关，也可能只和对话本身相关，需要区别对待，应该建立多层次的规划结构，在不同层次上进行识别和处理，这是规划模型理论中比较薄弱有待充实的部分。可见在实际对话过程中，规划的识别并不是容易的事，而且有时可识别的规划还具有不确定性。

规划模型提到了对话者的心理目标，但是没有解释人的思想状态与规划到底有何联系，对话一方究竟如何根据心理目标和对话情况确定当前回合的规划；同样没有说明对话另一方如何才能正确识别出对方的规划。

基于规划的对话管理方法以对话的规划模型为依据，与前述有限状态方法相比，应用范围较广，是目前比较流行的对话管理方法。基于规划的方法有时也称自组织方法。

与有限状态方法不同，基于规划的方法并不事先确定对话过程，而是将特定的对话任务划分为多个小目标，依据当前对话状态动态地决定对话的进展方向，控制对话过程逐步实现这些小目标，最终完成整个对话任务。这一方法避免了有限状态方法中事先定义状态和转移的工作，表达对话状态的能力强，便于处理复杂的任务，且对话控制较为灵活，易于实现对话过程的人机混合主导。

依据对话的规划模型实现对话管理的对话系统比较多，而且具体表现形式多种多样，有表形（见图 2-1^[27]）、树形（见图 2-2^[28]）、事件驱动的^[29]等等。

	D_ST	D_TIME	A_ST	A_TIME	DATE	NUM	
S1	—	—	—	—	—	—	Welcome()
S2	—	X	X	X	X	X	Ask(D_ST)
S3	X	—	X	—	X	X	Ask(D_TIME)
...							...

图 2-1 某火车订票系统的对话管理

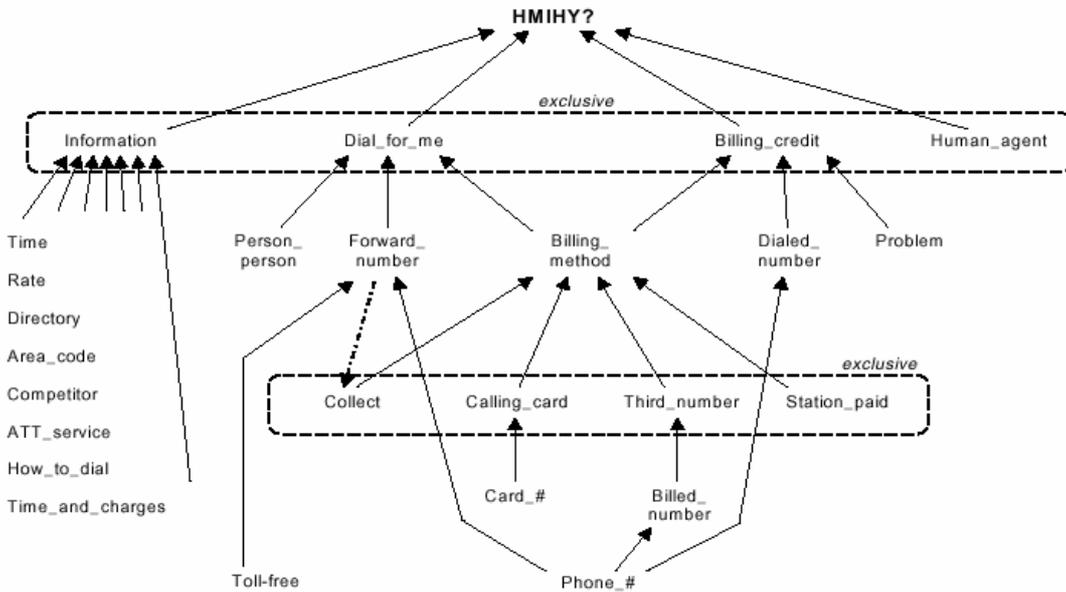


图 2-2 某电话服务系统的对话管理

类似图 2-1 所示的表格形式的对话管理，出现在很多系统中。这类表格的实质是列举领域任务中所有的信息状态的组合，并且给每一组合规定一个应答。在有限状态方法中，如果取消所有转移条件，将原对话状态表示为以前所有状态与状态转移的总和，就可以得到该表格的内容。因此这种表格状的对话管理方法只是省去了有限状态方法中定义转移条件的困难，当领域任务复杂时，仍然难以列举清楚所有的信息状态组合。

图 2-2 所示的树状对话管理，其结点处是对话系统用户的动作。根据用户依次采取的动作，系统可以确定用户需要的服务。然而很多提供信息咨询服务的对话系统，没有办法总结出用户采取的动作。而且对于一个需要用户和系统进行协商的领域任务，用户事先对目标的细节可能并不清楚，

多个用户动作间会有不定次数的反复，这一用于对话管理的树状结构将难以构建。

同是基于规划的对话管理方法，上述表格状和树状的方法所处理的内容显然不同，而且不能相互转化，这说明两个领域任务中的规划本身不同。那么怎么确定某个领域任务的规划类别，选择规划合适的表现形式呢？这其实也是规划识别的问题，基于规划的对话管理方法无法从对话的规划模型理论上找到这个问题的答案。因此对于一个实际对话任务，依据对话的规划模型设计实现该任务可用的对话管理需要一定的技巧，而且对于一个复杂的领域任务，可能是非常困难的事情。基于同样的原因，寻求具有较大适用范围和具有较好移植性的规划方法，是目前对话管理研究的一大热点。

由于理论上不能说明在对话过程中应该如何进行规划，所以基于规划的对话管理方法必须经过系统测试才能证明其应答的合理性。在对话应答设计中，对于某一对话状态经常可以有不同语义的应答，无法依据对话模型理论来判断哪种应答更好，只能通过与用户满意度相关的测试来比较。

2.1.3 基于对话互动模型的方法

相对以上两种对话模型，对话的互动模型提出得较晚，它试图解释为何对话中常有强调、重复、请求确认等对话动作，为何对话能够继续。该模型假定，对话双方都试图使对方在对话内容上与自己取得一定的共识，保持对话的连续性，从而使对话能够进行下去。

对话的互动模型摆脱了语法模型和规划模型的常规思路，可以解释对话中重复、指代、确认等一部分对话现象产生的原因，也能给出对话系统处理的内容和处理的原则。该模型目前正在发展之中，有待进一步丰富和提高。如果能在特定面向的对话系统中，体现普遍的互动理论，就可以解释很多的对话现象了，也许可为智能系统间相互交流提供某种规范。

对话管理的互动方法以对话的互动模型为依据。为了保证对话的连续性，互动的方法根据当前对话状态得到多个可能的对话进行方向，然后根

据一定的策略选择最佳的方向。参考文献^[30]介绍了应用该方法的一个对话系统，图 2-3 为该系统的应答策略示意图。

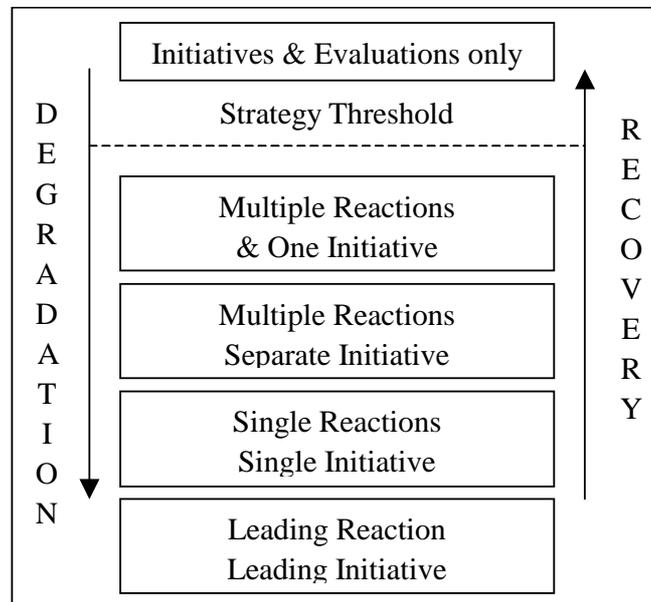


图 2-3 互动模型系统应答策略示意图

这一对话管理方法仍需要识别对话动作。由于应答只注重对话动作的连续性，忽视了不同领域任务的知识对系统应答的影响，因此系统的主动性有限，而且很难与任务内容结合。应用该对话管理方法的实际系统在实现时加上了大量与特定任务相关的限制，大有喧宾夺主之嫌。由于对话的互动模型出现得比较晚，现有系统所应用的互动方法的对话管理又不易移植或推广，应用该对话管理方法的系统极少。

2.1.4 现有对话管理方法总结

目前对话系统的研究是一大热点，各种实验系统或实用系统很多，研究人员各有各的观点，算法也多种多样，呈现百家争鸣的景象。但是总的来说，目前对话管理的研究水平并不高，主要表现为没有针对对话管理本身的理论模型，现有各类对话管理方法都还不成体系。

现有的对话管理方法，借鉴传统的自然语言理解中的对话模型理论，拘泥于语言表层。各种对话管理方法，都不可避免需要识别语句中的对话

动作或规划，希望从中得到应答的依据。然而如前所述，在现实对话中语句与对话动作或规划并不是一一对应的，这已经注定了这些方法得不到广泛应用或者表现不尽如人意。

对话是一个交互的心理和行为过程，单纯以语言为对象进行处理，不考虑对话双方的心理过程，其成效必然有限。同时，领域任务的特点也会影响对话时的心理运作，因此还应该研究具体的领域知识对于心理的影响。

上述三种对话管理方法，目前应用最多最成功的是基于规划的方法。对话的规划模型指出领域知识或领域任务在心理上有某种结构，对话者的行为服从这一结构所体现出来的心理趋势。基于规划的对话管理方法注重不同领域任务的表示形式，表征领域知识的能力强，应答时具有一定的心理倾向（规划），正是相对较优的原因。当然，现有对话的规划模型还只是简单提及了心理因素与对话行为的关系，并没有真正揭示对话双方的心理运作规律。

根据以上分析，对话管理的研究需要在领域知识表征和心理规律两个方面有所加强，具体说来就是要研究领域知识在对话者心理中的组织结构，以及对话时的心理运作规律。

2.2 心理语言学的相关理论

对话管理归根结底是一个心理过程，仅仅从研究语言表象的传统自然语言理解理论中寻求解决之道是无法真正成功的。对话管理需要建立心理模型，来解释说明对话过程中理解对方和确定应答的心理过程，从而指导具体对话管理方法的实现。

建立对话管理的心理模型，可以借鉴心理语言学的相关理论。

心理语言学，也称语言心理学，是本世纪 50 年代在心理学和语言学之间发展出来的一门新兴的学科^[31]。它的主要任务是研究三个方面的问题：言语的产生、言语的理解和儿童的语言获得^[32]。对语言的心理学研究，“为人类心理过程的研究提供了一个非常有利的视角”（乔姆斯基语），不仅

对探讨语言现象的本质和规律有重要意义，而且对提示人类认知系统的本质和规律，也有重要的意义^[33]。

2.2.1 语言理解和心理表征

心理语言学理论指出，语言理解是指从语言的表面结构中提取出深层命题结构的一种积极的推理的过程，或者说是从语言的表层结构去建构意义的过程。它包括语言知觉、词汇通达、句子与话语理解等。人们理解语言先要接受由外部输入的语言刺激，然后在心理词典中进行搜索，获得单词的意义，再经过句法分析和语义分析，揭露句子和话语的意义。理解不仅依赖于外部输入的信息，而且依赖于人们自己的知识组织、认知结构，依赖于内部的心理图式。因此，语言理解过程又是一个自下而上加工和自上而下加工交互作用的过程。

在语言理解过程中，外部语言材料所负载的信息经过感知和理解，被转化成一种心理形式，输入人脑并储存起来。信息在人脑中存在的方式或形态，称为心理表征。理解的过程就是建立心理表征的过程。一方面，通过理解建立起来的对语言材料的心理表征被储存起来，成为语言记忆的主要内容；另一方面，只有当外部信息纳入到记忆的已有信息网络中，心理表征才得以建立，理解才得以完成。所以理解和记忆有着密切的相互关系；理解是良好记忆的前提，记忆中已有的知识又是理解的基础。

我们平时在听到或读过一段话语以后，如果没有特别的要求，过一段时间就只能记住话语的意思，而不是它的原话。这说明理解时建立了一种被我们称之为“意义”的心理表征，这种心理表征是对句子内容深层次的理解，而不是句子表层形式的理解。

2.2.2 概念储存模型

概念储存模型是描述人脑中词的意义和概念结构的模型。理解句子或更复杂的语言材料，是以对单词意义的掌握为基础的。单词有其一般性的意义，不随句子语境的变化而变化，这在心理上表现为相应的概念。只有以概念为建筑材料才能建立句子的命题表征。换句话说，在理解句子时，必须从记忆库中提取句子中单词的意义，即从心理词典中查找出单词的意

义，才可能建构命题表征。单词所表示的概念在头脑中有其固有的、内在的联系，形成一定的结构，人们是按照这种内部认知结构来储存所输入信息的，这也是长久稳固的记忆所要求的。

2.2.3 图式

对单词意义的理解，即概念系统中相应部分的激活，只是理解语句或文章的基础。在此基础上还必须运用更高级的心理组织，将整个文章或语句的已有的知识结合起来，才能最后完成对文句的理解。这些心理组织称为图式。

图式有各种各样的具体形态，但它们在基本的结构与功能上是共同的。常见的图式有事件图式、场景图式、角色图式与故事图式。以事件图式为例，一个事件图式保存着两方面的信息，一是清单性信息，即某种情境下将发生哪些子事件，一是结构性信息，即各子事件间有何关系。作为事件图式，其基本的内部结构是一致的，可以用上饭馆吃饭为例来说明事件图式的结构特点，见图 2-4^[33]。

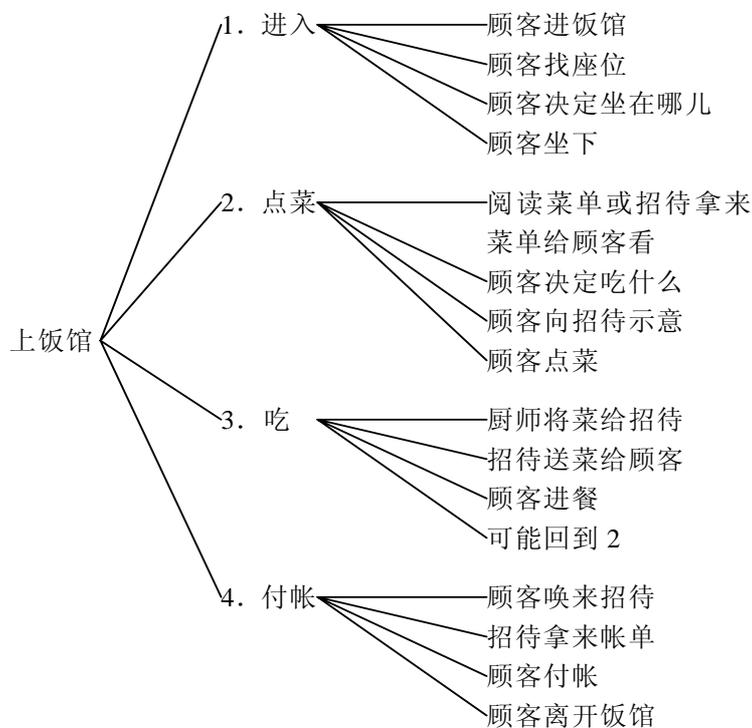


图 2-4 上饭馆的事件图式

上饭馆的事件一般都按上述图式所示的过程进行，人们熟知这一过程并把它作为一种知识结构储存在头脑中，形成一个上饭馆的事件图式。在这里，知识的单元是事件，一个事件是由若干个下一层次的小事件组成，而每个小事件又由若干个更细小的事件构成。因此，事件图式是以事件为单元的等级层次结构，上下层事件是整体和部分的关系，下层事件是上层事件不可或缺的一个组成部分。在事件图式中，同一级的事件由于被包含在同一事件的整体中而获得心理上的联系，它们之间具有时间上的连续性，构成一个事件的子事件是按顺序相继发生的，这种顺序一般不可更改，因为子事件之间是因果相承的，前者为后者的条件。

由于图式集合了关于某一定型事物的具体构成知识，它能为人们理解语言材料提供一种积极的准备状态。当图式被激活后，人们对即将要叙述的内容便会产生一种预期。由图式提供的信息不一定能清晰地出现在意识中，但它们的可利用性是明显的。当预期同材料所叙述的内容一致时，图式将促进对材料的迅速理解。相反，当图式的预期同实际的描述不一致时，图式将阻碍对材料的理解。

在有些情况下，句子可能省略事件过程中一些“不言而喻”的部分，这时只要相应的图式得到唤醒，在理解时所建立的意义表征就会包含关于这些部分的信息。心理语言学实验证明，内部图式的这种补充作用能在记忆结果中直接表现出来，以至于我们会误以为这些信息是从外部材料中输入的。

某个图式一旦被有关的线索激活，就会为信息的加工储存提供一种框架。能被图式组织进来的信息将获得巩固的记忆，而与图式无关、难以组织进来的信息则很容易遗忘。在有些情况下，不能由图式组织起来的信息会很快丢失。

可见，图式现象充分体现了人类已有的认知结构在处理外界信息时的主动作用。这种作用贯穿于信息加工的全过程。此外，在对话理解中，图式本身也往往需要根据上下文所提供的线索进行主动的分析、判断和修正。

2.2.4 思维模型

思维模型是头脑中的一种数据结构，它表现了真实世界或虚拟世界的部分情景^[34]。当对话者根据已有的概念和图式结构，加入对话过程中获得的信息，就在头脑中形成了一定的思维模型。经典实验研究了与空间顺序、空间大小相关的语言理解，证明人在语言理解过程中确实建立了相应的心理表征。基于这样的心理表征，人们对文句中直接描述的空间关系或大小关系的判断，与对可由文句意义推理出的空间关系或大小关系的判断，两者同样迅速。这也从另一方面证明人们理解记忆的是更深层次的意义，而不是文句表面的意义。

思维模型认为人脑中的理解和记忆是以一些标记（token）为中心的，当出现新的描述时，在旧有标记之间查找是否有唯一符合该描述的标记，如果有则确定为对旧有标记的指代，如果没有则要建立新的标记，如果有多个则需要更多信息来确认。默认值只有根据上下文才能决定。

2.2.5 语言生成理论

语言生成就是人们利用语言表达思想的心理过程。

口语的生成可分为三个阶段：构思阶段、转换阶段和执行阶段。在构思阶段，说话者根据自己的目的在头脑中产生所要表达的思想，确定说话的内容。在转换阶段，说话者运用句法规则将所要表达的思想转换成言语信息，这是将句子的潜在结构当作一个整体并转换成它的表层结构。在执行阶段，说话者将头脑中的言语信息变成口头语言。

说话内容的确定要受到许多因素的影响，其中有两个是主要的。第一，听话者的知识状态。说话者只有考虑到听话者的认知状态才能成功地进行交谈。第二，参与对话者要共同遵守的一些谈话原则，称为“合作原则”。例如，说话者要按照交谈的需要给予适量的信息，信息不要过多和过细；说话者必须给出真实的信息；说话内容要与交谈的主题有关。

说话者的说话内容（即思想或意义）在说出之前必须在头脑中加以表征（至少有一部分被表征），表征这些信息的心理代码必须囊括将要出现在言

语中的全部信息。这种内部表征称为意图（intention）。在头脑中用来表征意义信息的是一种抽象的命题表征系统。这种潜在于句子中的表征系统是以命题和概念之间的关系为基础的。说话者的这种抽象的命题表征借助一定的语言结构成为最终所要表达的言语信息。

2.2.6 现有对话管理方法的心理本质分析

如上所述，心理语言学的理论指出：对话者已有的心理结构（包括概念和图式）是语言理解的基础，是理解得到的意义存储的框架，也是确定说话内容时要考虑的因素；确定说话内容还要考虑心理结构以外的其它一些因素。

将心理语言学理论与已有的对话管理方法相对照可以知道，有限状态方法中的状态控制比较近似于事件图式的组织，因此有限状态方法用以处理有序的事件任务比较合适，例如进行问卷调查。但是其它心理结构，如概念模型和场景图式，并不适合用状态转移网络来表示。基于规划的对话管理方法采取的多种表现形式，实质是不同心理结构的表现。图 2-1 的表结构表示的是领域任务的概念组成，图 2-2 的树结构表示的是领域任务的事件图式，所以它们有各自的表现形式且不能相互替代。但它们各自都适合领域任务的表示，因此都能较好的完成领域任务。而且树结构的表示方法也可以看成是一个有限状态网络，用有限状态方法来进行对话管理应该也能完成领域任务。可见，有限状态方法和基于规划的方法都是对话管理所采用的表面形式，其本质都是对个别心理结构的表示，当这种表示适合领域任务时，对话管理方法就能完成对话系统任务；反之，当这种表现形式不适合领域任务时，生搬硬套这种对话管理方法，系统表现当然不好。至于对话管理的互动方法，它只强调了确定说话内容（即生成应答）时要考虑“合作原则”，却没有考虑领域任务的心理结构，所以不易在实际系统中移植和推广。

已有的多种对话管理方法，忽视了对话过程中已交流的信息在固有心理结构上的积累，即思维模型在对话过程中的建立和演变，而思维模型却

是应答生成的基础。可见现有对话管理方法的应答决策，只是一种无意识的自发行为。

2.3 对话管理的心理模型

根据以上分析，有限状态方法和基于规划的方法都是不同心理结构在对话管理上的反映，只有依据对话管理的心理模型（以下简称为对话管理模型），而不是自然语言理解中的对话模型，才能判断对话管理的具体表现是否合理。同时对话管理模型还有助于说明在实际的对话系统中，对话管理模块应具有哪些功能，这也是对话模型理论所不能胜任的。

2.3.1 对话管理模型的组成

对话管理模型的组成可用图 2-5 说明。

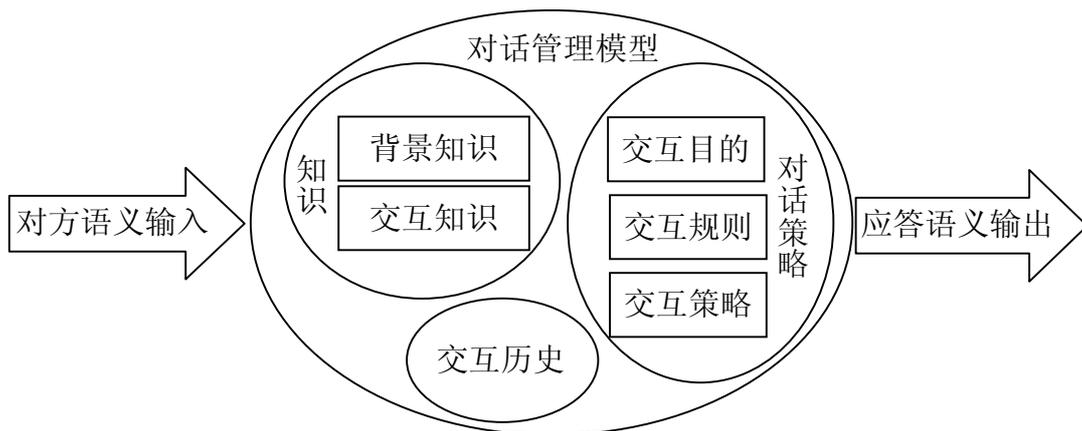


图 2-5 对话管理模型组成

对话管理模型由三个部分组成：知识、交互历史和对话策略。模型的输入是对方语句的语义，模型的输出是应答语义。

知识包括两个部分：背景知识和交互知识。背景知识指对话者的常识和所掌握的信息，即固有的心理结构和具体的知识内容。心理结构包括各种概念的结构和相互关系，包括各种定型事物的图式，是深层次语义理解的基础，是交互知识的存储框架，是对话过程控制的依据。具体的知识内容是各种概念的外延，相当于概念后备的数据库。在对话过程中，某些概

念被激活的同时，会有一些相关的具体知识也被激活。交互知识指对话过程中积累的各种信息的总和，它以固有的心理结构为基础，是对话过程中建立的思维模型，是对话中所领会的浅层和深层意义的心理表征。

背景知识和交互知识是从相对静态和动态的角度加以区分的。虽然在对话过程中概念和图式都可能被更新，具体知识也可能被添加、修正或删除，但对话结束时这些知识都仍然存在，在下一次对话过程中仍然会起作用。而交互知识只是当前的对话状态，它不仅每个回合都在改变，而且对话结束的同时也就失去了意义。

交互历史是以前多个回合双方交流的语义的历史。交互历史与交互知识的不同之处在于：交互知识中以前多回合的语义以固有的心理结构为框架进行组织，受到固有心理结构的影响，其意义已超过了语义的简单相加，而且交互知识中还包括被激活的具体知识。交互历史只是以前语义的简单记录。例如在图 2-6 所示的对话中，在第二回合 A 改变了原来的信息，则交互知识是：A 明天去圆明园。而交互历史包括了 A 和 B 每个回合给出的语义，包括第一回合 A 说明天去颐和园，第一回合 B 问是不是去划船，第二回合 A 更正为去圆明园，第二回合 B 说圆明园不能划船。

A:	我明天要去颐和园玩。
B:	去颐和园划船吗？
A:	噢，说错了，我明天要去的是圆明园。
B:	那里可划不了船。

图 2-6 对话 1

对话策略包括三个部分：交互目的（动机）、交互规则和交互策略。思维本身是由动机激发的^[Lie 1997]，交互目的决定对话双方的合作态度、交流内容等因素；交互规则指双方为了保证对话过程顺利进行所应遵循的规则，包括心理语言学中提到的对话的“合作原则”，还包括领域中可能的特殊规划；交互策略包括对话的主导方式、确认机制、语义的组织等策略。对话策略决定了如何根据当前的交互知识和交互历史构思确定应答的语义。

对话管理模型的输入为对方语义，包括对话对方语句中的各个概念、概念间的关系（即命题），以及某些具体知识。文句语义的获得，其实也要依靠固有的概念结构，是复杂的心理过程，但是这一心理过程与对话管理的心理过程无关。对话管理模型的输出为应答语义，它不是最终应答所用的句式和文字，而是指应答所应包含的概念、命题和具体知识。至于应答最终采用哪种句式，用什么词汇，都是属于自然语言生成的任务，而不是对话管理所关心的。

2.3.2 对话管理模型的运作

对话管理模型的运作是对话过程中心理规律的体现，是从得到对方语义输入，到推理产生应答语义的过程。对话管理模型的具体运作可分为三个方面的任务：知识管理、交互历史管理和应答推理。

（一）知识管理

在对话过程中，得到新的语义输入后，首先要确定新语义所适用的心理结构。有的语义表示的概念属于某个固定的概念结构，该结构给出概念的组成以及与其它概念的关系，有的语义表示的概念还与图式中的概念相关。确定语义适用的心理结构应该考虑已有的交互知识，这有助于在多个可能的心理结构中进行选择。有时文句本身具有歧义，即使结合交互知识也不能够确定。心理学实验告诉我们，人们遇到这种情形时，一般都假设某种解释是正确的并进行下一步的处理，当后来发现该解释不正确时，才重新选择另一种解释进行处理，而不是从一开始就同时对两种解释情况分别进行处理。

新语义所适用的心理结构很多情况下就是上回合的心理结构。新的心理结构会被加入到交互历史中作为容纳新语义的框架。将新语义加入到交互历史的框架中，也就更新了交互知识。新的心理结构可能蕴含一些领域任务或对话场景的默知识（不言而喻的信息），这些信息也一并加入交互知识，体现了语言的深层次理解。这一过程有时会激发某些具体的知识，它们也会被加入到交互知识中，并且通过应答传递给对话的另一方，改变另一方的交互知识。

由此可见，更新了交互知识中可能包含原来所没有的具体知识，这些知识满足一定条件时会加入到自身的背景知识中。交互知识还可能导致原来所应用的概念或图式结构不再适用，需要进行修正，这也会影响背景知识中原概念或图式结构的改变，甚至生成新的概念和图式结构。以上就是信息的获取和知识的习得过程。

在知识管理过程中，有某些策略或原则在起作用，前面提到的歧义处理就是一种策略。相对固定的背景知识不是轻易就会被更新的，而是要满足一定的条件才行，这也是一种策略。在交互知识更新时，有时旧信息会失去价值，这些信息是否保留、保留多长时间，也是一种策略。在图 2-6 所示的对话中，B 第二回合的话说明在 B 的交互知识中，虽然圆明园的信息代替了颐和园的信息，但之前由颐和园概念激发的划船信息并没有立即从交互知识中去除。

（二）交互历史管理

交互历史记录了对话双方的语义交流历史，也起着对话模型中输入语义到交互知识的缓冲作用。实际对话中，人们对对话回合历史会留有一些记忆，知道哪些信息是对方给出的，哪些信息是自己给出的。对话过程中会出现“刚才说的不算”这样的语句，这需要将交互知识恢复到“不算”的那个回合之前。这时仅有交互知识是无法实现“不算”的心理过程的，可见交互历史管理是对话管理不可缺少的一个功能。

在人们的实际对话中，由于人的记忆特点，交互历史往往只保留较近一些回合的信息，或者某些特殊的回合信息。对于计算机来说，交互历史可以从对话一开始就保存下来，但如果对话的另一方是人的话，保留所有的交互历史并不一定有意义，至少并不会有利于图灵试验的结果。因此如何保留交互历史，也是可研究的策略之一。

（三）应答推理

对话管理的根本目的是应答，根据交互历史和交互知识，应用不同的对话策略，可以得到不同的应答语义，这一过程称为应答推理，也就是对话者构思应答语义的过程。

应答推理的基础是交互知识的结构，也就是固有的概念和图式结构。这些结构决定了应答的表现，好像是用对话管理的有限状态方法进行对话控制，或是用基于规划的对话管理方法进行控制。对话策略中的交互目的决定了应答推理的大方向，例如在信息咨询的对话中，咨询信息者要尽量给出相关的信息条件，提供信息者要尽量提供所求的信息，双方均要积极配合完成信息咨询的任务。根据交互知识（含概念和图式结构）和交互目的，明确了已知信息和未知信息，确定了推理的方向，就可以决定可能的应答语义。对话策略中的交互规则保证对话语义易为对方所接受，使得对话能顺利进行，领域任务能更好地完成。应答语义的组织应该考虑这个因素，例如在报告火警时，应该给出火灾地点及火势的必要且精确的信息，其它细节信息可省略。满足上述条件后，仍然有多种可能的应答，交互策略可用来决定当前应答选用哪一种语义组合，即选择哪些语义在当前回合的应答中输出，甚至预先决定未来回合的应答语义组合。当然，对话策略在对话过程中也是可变的，直接影晌应答推理的结果。

以上三个方面的任务是对话管理应该具有的功能，且这三个功能彼此联系，缺一不可。

2.3.3 对话管理模型的实用指导意义

在对话系统中，对话管理模块常被称为对话管理器(dialogue manager)，自然语言理解模块为对话管理器提供语义输入，对话管理器的应答语义输出由自然语言生成模块转换成具体的应答语句。

根据前面的分析，已有的对话管理方法不管是有限状态方法，还是基于规划的方法，都具有一定的盲目性。它们自身的控制方式既定，只适合具有相应心理结构的领域任务。对话管理器应根据领域任务自然地选用不同的对话管理方法和具体形式。

对话系统所要解决的领域任务决定了对话可能涉及的各个概念和可能用到的各种图式，这些概念和图式可以称为领域相关概念和图式。由于领域任务的局限性，领域相关概念和图式的表示并非广泛意义上的知识表示，

以目前的技术水平是可以实现的。对话系统中任务相关的具体知识，可以用数据库来表示和存储。这些是对话管理所需的背景知识的准备。

图 2-7 是一个抽象的对话管理器模块组成图，其中知识管理、交互历史管理和应答推理三个子模块分别负责对话管理的三个任务，这三个子模块由一个总控模块进行协调。知识管理子模块又可细分为三个更小的子模块：领域概念和图式管理、数据库管理和交互知识管理。

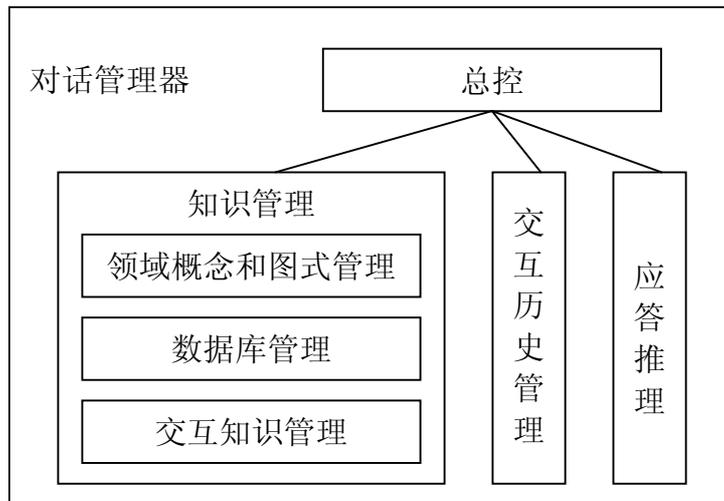


图 2-7 对话管理器模块组成

当对话管理器得到语义输入后，领域概念和图式管理子模块首先判断这些语义属于哪个概念结构或图式结构，这一判断要结合交互知识和交互历史才能进行。确定了语义所对应的概念或图式之后，由交互知识管理子模块根据新的语义信息进行交互知识的更新，同时交互历史管理模块也进行交互历史的更新。这样应答推理模块就可根据交互历史和交互知识进行推理。在推理过程中，可能要通过数据库管理子模块存取具体知识，也可能要通过领域概念和图式管理子模块改变原有的领域概念和图式。推理过程最后得到要应答的语义，这些语义也进入交互历史。

在上述过程中，各个模块都有如何选择所用策略的问题，这些策略有的可以根据领域任务要求来确定，有的可以根据系统表现来选择。例如应答推理相关的交互目的、交互规则由领域任务决定，交互策略可根据系统

期望达到的性能和系统用户的要求来决定，也可以比较不同策略下的系统表现从而选择较好的策略。

从现有的技术水平来讲，对话管理器的具体实现关键要解决好两方面的问题。第一是领域概念和图式的表示问题，这是整个对话管理的基础。领域任务的特点决定了不同的概念和图式结构，应选用适当的数据结构来表示，领域概念和图式管理子模块、交互知识管理子模块和应答推理模块都需要这一数据结构上进行操作。第二是应答推理的策略实现问题。概念和图式的结构不同，应答推理实现同样策略的操作也可能不同，而且对于多种策略还存在选优的问题。有些策略选择可以事先决定，固化在应答推理的实现中，有些策略选择却需要根据具体系统表现才能确定，甚至需要根据对话的具体情况来动态选择。如果有一种较好的数据结构可以表示很多领域任务的概念和图式，同时在这一数据结构上实现了较好的应答推理机制，那么相应的对话管理器就具有较好的可移植性，可以用于建立那些领域任务的对话系统。

第三章 主题森林

对话者真正理解对方语句并进行应答推理，是以头脑中已有的领域概念和图式为基础的。用计算机来表示人头脑中的概念和图式，属于知识表示的范畴。知识表示一直是人工智能领域的难题，至今没有得到较好的解决。但是对于信息咨询的对话系统来说，由于涉及的信息内容较少，信息间的关系相对较简单，有可能找到比较好的知识表示方法。

现有的对话系统多应用有限状态网络、表格、树形等结构来表示领域任务涉及的概念和图式，应用局限性大。作者提出了主题森林的知识表示方法，适用于信息咨询的对话系统，能表示领域任务涉及的多个对话主题和复杂主题；在同一主题中，按信息的重要程度分类，可在对话过程中记录用户兴趣，进行不同的处理；考虑了多主题间共享信息的问题；与领域数据库相联系；可表示对话过程中的交互知识和交互历史，用于实现智能的对话管理。

本章首先提出信息咨询的对话系统中的知识表示问题，然后分析现有的一些表示方法存在的问题，随后提出主题森林表示方法，说明其结构组成与表示功能，最后与现有表示方法进行比较，并总结其特点。

3.1 对话系统中的知识表示

对话管理的心理模型指出，领域任务涉及的概念和图式是对话管理的基础，对话过程中交互知识管理和应答推理都是在领域概念和图式的基础上进行的。领域概念和图式的表示，在本质上是一个知识表示问题。对于信息咨询的对话系统，涉及的概念和图式范围较小，有可能找到较好的知识表示方法。

3.1.1 知识表示的内容

对话任务涉及的一些基本内容，我们称为信息，如名址录中的姓名和地址。一些相互间有较紧密关系的信息，组成对话中的主题。对于一个主题来说，可将具体的一项信息称为信息项。

在信息咨询的对话系统中，概念表示就是要表示出这些信息项如何构成主题的概念，这包括信息项和信息项之间的关系，以及不同信息项对主题的重要程度；还要表示不同主题间的关系是由哪些信息项体现的。

在信息咨询的对话系统中，图式表示一般只涉及事件图式和场景图式，即咨询任务一般有固定的咨询方式，或者咨询的内容本身是一些信息项的简单组合。

3.1.2 对现有对话系统中的知识表示方法的分析

现有的对话系统多用有限状态网络，或者表形树形的结构来支持对话管理。根据对话管理的心理模型，这些表示方法在本质上是对领域概念和图式的表示。

（一）有限状态网络

有限状态网络的方法是把任务相关的各信息项按某种顺序列举，没有明显的概念表示；图式功能蕴含在状态转移中，同时每一状态也蕴含着部分的交互历史信息。由于没有明显的概念表示，如果信息项间有复杂的关系，或者存在可有可无的信息项，用有限状态网络来表示比较困难；当任务涉及多个主题时，无法将不同主题的信息项安排入同一个网络中。由于图式功能蕴含在网络中，当任务涉及信息项较多时，网络中状态数目和状态转移的数目会快速增长，以至于该有限状态网络的建立非常困难；若任务目标的达成需要用户与系统反复协商，这时有限状态网络无法适用。

例如在一个航班信息查询系统中，如果知道了航班号和日期就可以查询相关航班，如果知道了起点城市、终点城市和日期也可以查询相关航班。如果用有限状态网络来表示这样的图式功能，则采用系统主导方式，由系统主动问用户将给出航班号还是起点城市和终点城市，然后根据不同的答案进入不同的对话过程。有的用户可能对航班的起飞时刻或到达时刻有要求，可能对航空公司或航班机型有要求，因此有限状态网络中应设定询问用户是否给出这些要求的图式功能，而且对于不同的回答都要进入不同的状态。仅以上两种情况的组合，就需要网络中包含数倍于信息项的对话状态和转移数目。在航班查询中，用户还可能会问到机型或航空公司的基本

信息、起点城市和到达城市两地的时差等，属于不同的主题，这无法在航班信息查询的有限状态网络中加入。

在图式功能表示上，如果用户在回答系统提问时，给出了比问题本身更多的信息，例如在系统提问起点城市时用户还给出了终点城市，那么一般情况下，更多的信息会被忽略，并且在对话进入特定状态时，由系统再次提问这些信息。如果要实现较人性化的对话控制，那么需要在每一可能的状态下，都对这种“多余”信息进行分析，加入转移条件。如果用户对查询得到的航班不满意，要改变日期或航班号等信息，系统还要加入适当的状态转移，由于这种信息的修改是不定次数的且在很多对话状态下都可能发生，所以需要大量的状态转移。

可见对于有限状态网络方法来说，表示航班信息这一主题相关的概念和图式就已经非常复杂，难以实现，而且该方法无法将任务相关的其它多个主题的概念和图式加入到同一网络表示中。

（二）表格形表示

表格形表示也是用得较多的表示方法，例如图 2-1，这种方法其实是对有限状态网络表示的一种压缩。表格中每一条记录包含两个部分，一部分是信息项状态的某种组合，另一部分是在该组合下系统应采取的动作。信息项集合就是所在主题的概念表示，信息项状态组合与系统动作的对应就是图式功能的表示。

用表格来表示航班信息主题，表中至少要包括航班号、起点城市、到达城市、日期、起飞时刻、到达时刻、航班机型、航空公司这 8 项信息，不同的信息项状态组合可达 28 种，即使通过某些办法将一些状态合并（例如忽略某些信息项的状态，使得不同信息项状态组成对应相同的动作），仍然需要列出很多记录，实现比较困难。

由于表格形表示相当于将有限状态网络的状态进行压缩，丢失了对话过程的信息，所以无法记录哪些信息项分别是在哪些回合中得到的。例如在参考文献^[27]中，对于下面两种不同的情况：用户已确定 A 信息正与系统

协商 B 信息，用户已确定 B 信息正与系统协商 A 信息，系统将采取同一个对话动作。

表格形表示可以处理多个主题，即为每个主题建立一个表格，但是不同主题间的信息关系却无法表示。共享信息体现了不同主题之间的关系，也是人们概念组成的一部分，因此在领域概念和图式的表示中，不应忽略。在对话过程中，两个含有共享信息的主题间会产生自然的切换，而且主题切换过程中存在上下文省略现象。例如在航班查询的对话过程中，查询得到某航班后，用户直接问航班票价，其实是从航班信息主题转换到了航线信息主题，两个主题的共享信息为起点城市和到达城市。这时航线信息查询条件已满足（根据上下文，起点城市和到达城市的信息已确定），系统应该给予回答。但是这种主题间的共享信息关系，无法加入到表格中，对话管理不能意识到主题间的共享信息，不能很好地处理这样自然的主题转移，无法合理地处理主题切换时的上下文省略。

（三）树形表示

树形表示有不同的表示方法，图 2-2 的实质就是有限状态网络方法。参考文献^[35]中的树形表示方法，如图 3-1 所示，同时表示了概念和图式。

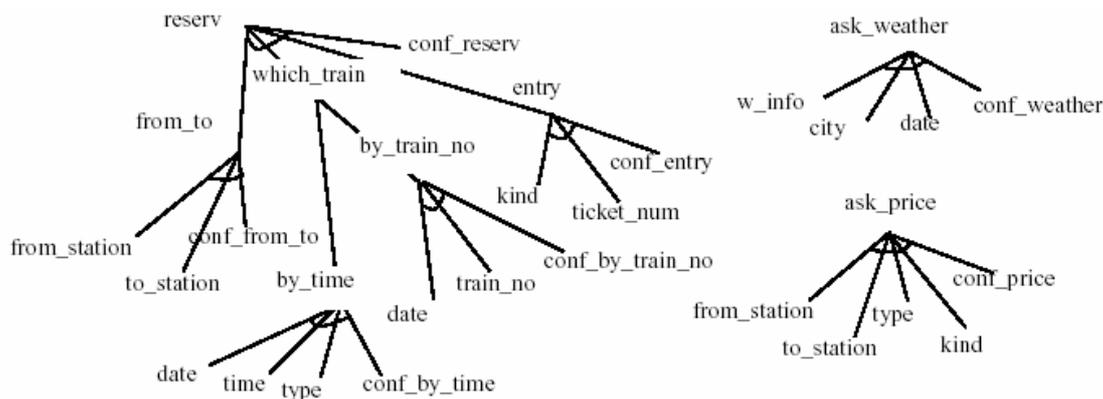


图 3-1 概念和图式的树形表示

在图 3-1 右上方名为 ask_weather 的树中，w_info、city 和 date 是该主题的信息项，表示该主题的概念组成，而 conf_weather 与对话的确认有关，表示图式功能。这种树形表示才是与上述有限状态网络和表格形表示所不

同的表示方法，本文以下提到的树形表示都是特指这种意义的树形表示方法。

树形表示的图式功能还在于它不直接给出各信息项的状态，而且在对话过程中动态地对各信息项的状态进行归纳总结，并依据树形结构进行应答推理。由于信息项的状态是在对话过程中动态得到的，所以树形表示省去了事先定义状态和转移的工作，设计实现较简单，可以用于复杂的主题，但是应该配合适当的交互历史管理。

树形表示可用于涉及多主题的区域任务，如图 3-1 所示，能够用来处理主题间的转移。

然而现有的树形表示方法仍有缺陷。首先，树形结构中各信息项的地位都是一样的，但是事实上，组成主题的各信息项的重要程度可能不同，这会引发对话过程的不同。用这种树形表示方法来表示航班信息主题时，如果将航班机型和航空公司信息项加入树形结构中，那么系统就会总是要求用户给出相关信息，如果不加入这两个信息项，则用户无法给出这两项信息的限制条件。两种情况都不能令人满意，用这样的树形表示来处理航班信息主题并不合适。其次，虽然树形表示支持多个主题，但是仍然没有表示主题间的共享信息，因此对话管理仍然不能很好地处理自然的主题转移和由此产生的上下文省略。

根据以上分析，在现有信息咨询的对话系统中，对于领域概念和图式的表示以树形表示方法最为优越，适用于较复杂的系统。但是应该加强复杂主题中信息项的重要程度的表示，加强多主题共享信息的表示，支持更自然的主题切换。同时，根据对话管理的心理模型，应注意以领域概念和图式的表示结构为基础表示交互知识，并且结合交互历史的管理。

3.2 主题森林

通过对现有对话系统中知识表示方法的分析，作者提出了主题森林的表示方法，用来表示领域概念和图式，能够克服现有表示方法的一些缺点，较好地表示出复杂主题、主题各信息项的不同重要程度、多主题和主题间

的共享信息。同时主题森林可用于对话过程中交互知识的表示，并且直接记录交互历史。

主题森林由若干棵主题树以及相关的共享信息索引组成。

3.2.1 主题树的结构

主题树是表示对话单个主题相关的概念和图式的树形结构，包括构成该主题的各项信息及信息项之间的关系，可用来表示对话过程中该主题的交互知识，记录该主题相关的交互历史。

主题树由三类结点组成：主题结点、中间结点和叶子结点。

（一）主题结点

主题树的根结点称为主题结点，该结点给出主题的名称、主题的类型，以及与该主题相关的数据库关系表。一个主题结点可以有三个分支，分别称为主属性分支、次属性分支和附加属性分支，后两个分支可以为空。每一分支都是以一个中间结点开始的。

顾名思义，主属性分支用来表示该主题最重要的信息项，这些信息项的交互知识状态决定系统是否应该执行数据库操作。次属性分支用来表示该主题的其它一些重要的细节性信息项，也是对话过程中不可缺少的交互内容。附加属性分支用来表示剩下的一些信息项，对话过程中可能不会出现关于这些信息项的对话交互。

（二）中间结点

中间结点是连接主题结点和该主题树所有叶子结点的结点，中间结点的子结点可以是叶子结点或其它中间结点。中间结点给出它所有直接子结点间的逻辑关系，包括“与”和“或”两种关系。

直接与主题结点相连的中间结点是主属性分支、次属性分支或附加属性分支三棵子树的根结点，相应的结点又称为主属性结点、次属性结点和附加属性结点。其它中间结点，又称为一般性中间结点。

（三）叶子结点

叶子结点分为两种，一种是普通叶子结点，用于表示主题的单项信息，包括该项信息所对应的语义、在数据库表中对应的字段名称、关于该信息项的交互知识，以及关于该信息项的交互历史。另一种是记录状态的叶子结点，只用于辅助领域图式的表示。

3.2.2 主题树的建立

下面就以提供航班信息查询的对话系统为例，说明如何用主题树来表示对话涉及的航班信息主题。

首先建立主题结点，给出主题名称（“航班信息主题”）和该主题相关的数据库关系表的名称。

其次考虑该主题的多个信息项分别应属于哪一个属性分支。对于实际的航班信息查询来说，一般的用户都会给出起点城市和终点城市，以及航班日期，这三项信息在航班信息查询中是最重要的，应属于主属性分支。有些用户会直接给出航班号和日期，这两项信息的组合也满足航班信息查询的条件，因此航班号也是该主题的关键信息，也属于主属性分支。对于这两种情况，日期都是必不可少的，起点终点城市或航班号可以任选，因此用中间结点“(与)”连接“起点城市”结点和“终点城市”结点，再用中间结点“(或)”与“航班号”结点连接，最后与“日期”结点连接在主属性结点“主属性(与)”下。用户查询最关心的是航班的起飞时刻和到达时刻，根据时刻信息用户可确定乘坐哪一个航班，在实际对话中系统和用户将会协商确定用户需要的航班时刻，因此这两项细节信息应属于次属性分支，它们的关系可用“或”来表示。其它与航班信息有关的机型和航空公司不是每个用户都会关心并提及的，它们应属于附加属性分支，用“与”关系连接。

最后填写每个叶子结点的内容，给出各项信息所对应的语义以及在数据库表中对应的字段名称。每项信息可以有默认的限制条件，例如在用于北京地区的航班信息查询系统中，可以为起点城市添加默认值为“北京”。

经过以上三步，可得到航班信息主题树，如图 3-2 所示。

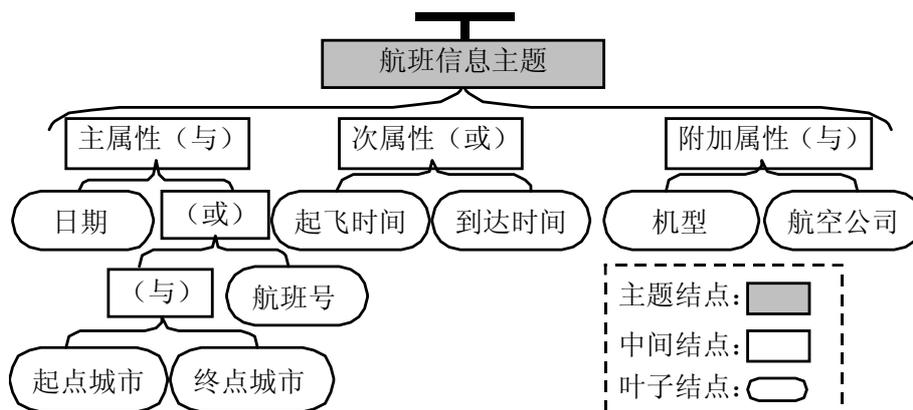


图 3-2 航班信息主题树

3.2.3 主题树的概念与图式功能

由主题树的建立过程可知，主题树包含了组成某主题的所有信息项，并且用“与或”关系表示信息项之间的关系。每项信息是否已在对话过程中获得某种限制条件，可以用逻辑真假来表示，通过中间结点标志的“与或”逻辑关系，可以计算得出每个属性分支的信息状态。主题树所有结点的信息状态，可以在对话过程中动态计算，用于交互知识的管理。主题的交互历史分散地记录在表示各信息项的叶子结点上，用于交互历史的管理。

由于信息项按重要程度分属不同的属性分支，因此主题树的分支结构暗含了图式的信息：当主属性分支的信息状态为真时，系统可以进行数据库操作，否则认为关键信息不充分，应通过对话获得更多信息；在主属性分支的信息状态为真、次属性分支的信息状态为假的情况下，系统应主动与用户关于次属性分支下的信息项进行交互；不管附加属性分支的信息状态如何，系统都不会主动与用户讨论该分支下的各项信息。

在 3.2.2 中建立的主题树没有使用记录状态的叶子结点，所表示的图式功能只限于主题概念组成暗含的合理的对话交互方式。与前面介绍的树形表示方法相关，由于描述了信息项对主题的不同重要程度，图式功能更完善。对于航班信息查询所需的对话管理来说，这种增强的领域概念和图式表示可以满足对话管理的要求。

如果对话系统要在航班信息查询的基础上，进一步提供航班订票服务，则订票过程应该用一个事件图式来描述，需要将记录状态的叶子结点加入到主题树中。

如图 3-3 所示，在原航班信息主题的附加属性分支中加入订票相关的信息项叶子结点“剩余票数”、“欲订票数”和“身份证号码”，并加入记录状态的叶子结点“对话状态”。

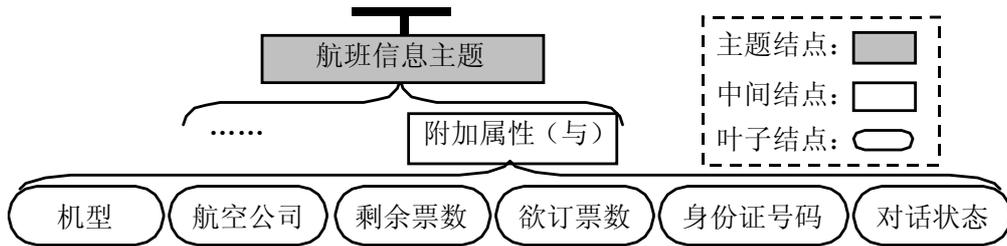


图 3-3 加入新结点的附加属性分支

加入和订票相关的信息项叶子结点后，主题树表示航班信息查询的概念和图式功能不受影响。记录状态的叶子结点可取某些特定的值，表示不同的对话状态。由于该主题的对话交互知识是由主题树表示的，所以根据树中各结点（包括记录状态的叶子结点本身）的知识状态，可以得知对话状态的变化，并按合理的对话过程进行交互。可用状态转移网络形象地表示相应的图式，见图 3-4。当对话处于非起始状态和结束状态时，如果用户改变航班信息，则对话状态应转移到起始状态，为清楚起见，这些状态转移在图中省略了。

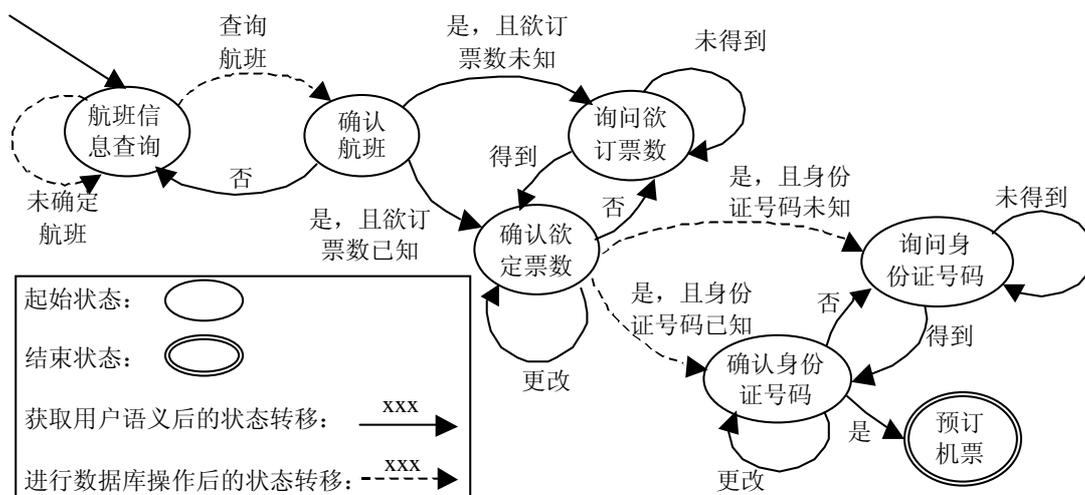


图 3-4 用状态转移网络表示对话状态的改变

对话管理的心理模型指出，在对话过程中除了对话交互外，访问数据库也会引起对话交互知识的改变，从而引起状态的改变。在图 3-4 中用实线和虚线来表示两种不同原因引起的状态转移，与现有对话系统中有限状态网络表示方法相比更加精细和准确。

在主题树中加入记录状态的叶子结点，只用于表示纯事件的、与概念组成无关的图式功能。如图 3-4 所示，状态转移网络主要表示订票的事件流程，而之前的航班信息查询只占一个状态。实际对话时，可能在大部分对话回合中系统都处于这一状态。该状态下的对话控制由之前建立的主题树结构确定。

对话系统任务相关的每一个主题都可以用一棵主题树来表示。由于每棵主题树都表示了该主题的概念和图式，当对话在不同主题间切换时，都有相应的概念和图式作为交互基础。

3.2.4 共享信息索引

共享信息索引是主题树的附属结构，用于表示不同主题间的共享信息。共享信息索引是从特定语义概念到主题树叶子结点的一对多的映射。

在 3.1.2 中提到，航班信息主题和航线信息主题共享起点城市和终点城市两项信息，在主题森林中这种关系通过共享信息索引来表示，如图 3-5 所示。

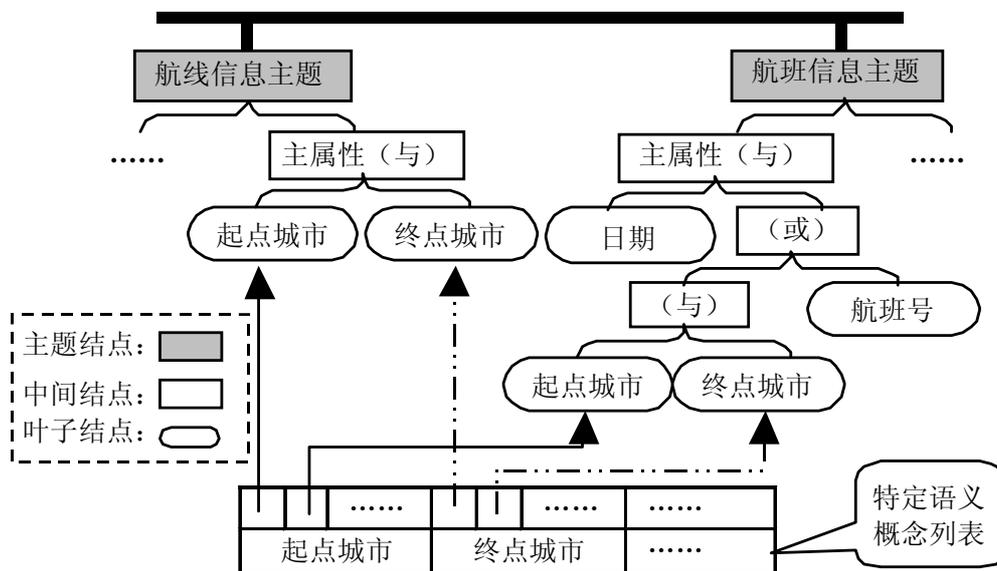


图 3-5 共享信息索引

在对话过程中，不同主题的共享信息虽然来自人脑中同一概念，但由于分属不同主题，在思维模型中要分别建立标记。在主题森林中，每棵主题树都表示相应主题的所有信息项，便于在对话过程中模拟思维模型，进行交互知识的管理。共享信息索引表示出不同主题间共享信息的同源关系，使得各棵主题树不是孤立地存在，而是相互关联，组成主题森林这一有机整体。

3.2.5 与其它表示方法的比较

主题森林作为一种知识表示方法，表示对话系统领域概念和图式，与已有的一些知识表示方法相比有其特点。

（一）与其它用于对话管理的表示方法的比较

现有对话系统用有限状态网络、表格形、树形等表示方法来支持对话管理。这些系统的对话管理以对话模型为基础，局限于对话语言表层，没有意识到所用的表示方法实质上用于对领域概念与图式的表示。这些表示

方法或者偏重于概念表示，或者偏重于图式表示，没有将两者很好地结合在一起，因此不能广泛适用于信息咨询的对话系统。

主题森林方法将领域概念与图式在同一结构中表示，可替代已有的多种表示方法，因此在主题森林的基础上实现对话管理，可兼有现有不同对话系统对话管理控制的优点，适用范围更广。具体来说，对于复杂主题和多主题，由于主题森林的表示采用树形结构，各信息项状态在对话过程中动态归纳，克服了有限状态方法和表格形方法列举信息项状态的难题；主题森林可直接用于对话交互知识的表示，而有限状态方法和表格方法在对话过程中，都需要有另外的数据结构来记录已有的各项信息条件；利用分支结构和记录状态的叶子结点，结合信息项状态的动态归纳，主题森林能表示领域任务的事件图式；利用共享信息索引将各主题有机的联系在一起，克服了已有的树形表示方法在控制对话状态和理解主题转换两方面的弱点。

（二）与一般知识表示方法的比较

一般性的知识表示是人工智能多年来的难题，虽然一些知识表示方法已有广泛应用，但仍存在这样或那样的缺点。对于信息咨询的对话系统来说，领域任务涉及的概念较少，关键是要将概念与图式有机地结合在一起，便于用于对话管理。主题森林的表示方法是已有的多种知识表示方法在对话系统背景下的结合与特例化，比一般的知识表示方法更适合对话系统的对话管理，也较容易在计算机上实现。

3.3 主题森林特点总结

综上所述，主题森林是用于对话系统领域概念和图式的知识表示方法，其结构与领域无关。

主题森林可表示领域中多个主题的概念组成。对于含多项信息的复杂主题，该方法不仅可表示各项信息间的关系，还可表示不同信息项对主题的不同重要程度。由于采用树形方式表示主题所有信息，对话状态可在对话过程中动态计算，与概念组成相关的状态及状态转移不必事先确定，可

适用于含多主题和复杂主题的领域任务。主题森林通过共享信息索引表示多个主题间的共享信息，将领域任务相关的各个主题概念表示联系在一起。

主题森林的分支结构以及记录状态的叶子结点，可表示领域的事件图式功能。这样的图式表示，通过各项信息状态的动态归纳来确定状态和转移条件，分支结构只用于表示由于主题概念组成及不同重要程度的信息项引起的图式功能，记录状态的叶子结点只用于表示与领域概念无关的纯事件顺序的图式功能。这样的分散表示方法，使得领域图式的表示更容易。

主题森林可代替有限状态方法、表形、树形等现有的多种表示方法，具有广泛的适用性，可直接用于对话过程中交互知识管理和交互历史管理，进行应答推理。主题森林采用了树形结构，易于在计算机上实现各种操作，具有较好的可移植性。

主题森林在表示领域概念和事件图式上较为成功，但对于场景图式表示能力有限，这使得主题森林结构不宜使用在地图查询类对话系统中。

第四章 基于主题森林的对话管理方法

作为对话系统的核心，对话管理是以一定的领域概念和图式表示为基础的。领域概念和图式的表示能力，关系到对话管理的性能和实现方法。已有对话系统的概念和图式表示存在这样或那样的缺陷，使得对话管理往往要使用辅助的数据结构才能实现必要的管理功能。本文作者提出的主题森林结构，是较好的领域概念和图式表示方法，为对话管理提供了强大的支持。在 EasyFlight 系统中已实现了基于主题森林的对话管理器。

对话管理的心理模型给出了对话管理的三大任务：知识管理、交互历史管理和应答推理，本章将依次介绍基于主题森林的对话管理方法如何实现这任务，说明在这些任务中可应用的策略；接着以 EasyFlight 系统实际运行时的一个对话例子说明所实现的对话管理器的性能；然后介绍对 EasyFlight 系统进行评测及评测结果，特别对系统的对话管理功能进行了评价；最后总结基于主题森林的对话管理方法所具有的特点。

4.1 知识管理

根据对话管理的心理模型，在一个实际的对话管理器中，知识管理可细分为领域概念和图式管理、数据库管理和交互知识管理三个子任务。

4.1.1 领域概念和图式管理

领域概念和图式是对话的基础，对于信息咨询的对话系统来说，可以认为领域概念和图式是固定的心理结构，不会在对话过程中被修改。主题森林结构具有广泛的适用性和较好的可移植性，为了使得对话管理也具有同样的特点，应实现主题森林结构在对话过程中的动态载入。

当对话系统的领域任务已知，可根据任务确定相关的各个主题，根据主题中各信息项的关系建立相应的多棵主题树。这是一个离线设计的过程，可按某种遍历方式（例如前序遍历）将主题树结点信息写成对话管理器的配置文件。由于每棵主题树都给出了该主题的完整信息，共享信息索引不必在配置文件中给出。对话管理器从文件中读入各棵主题树后，再遍

历各个叶子结点，归纳出主题间的共享信息。对于主题树中记录状态的叶子结点，可用程序形式来表示复杂的状态和状态转移。

当对话管理器得到新的用户语义输入时，要根据所含的语义内容确定当前主题。这有两种情况，一是根据用户语义可确定所属主题，即含有某主题的特定语义，或者只有某主题才涵盖所有的语义信息；另一种情况是有多个主题可涵盖所有语义信息，根据对话交互的连续性原则，即对话的上下文句子应该相关，可认为当前主题仍为前一回合主题。如果当前语句是对话的第一句，可认为属于该对话场景下某一默认主题。

只要确定了当前主题，就可找到该主题对应的主题树，将用户语义填入到主题树结构中。

4.1.2 数据库管理

由于主题树的主题结点和叶子结点中给出了数据库相关的表和字段名称，对话管理遍历整棵树后，可根据各信息项状态，给出数据库操作所需的条件。记录状态的叶子结点的内容也作为数据库操作的条件之一，用来确定在不同对话状态下的操作类型，例如在航班查询订票的对话过程中（参考第三章中图 3-4），当状态为“航班信息查询”时，执行数据库查询操作，当状态为“预订机票”时，执行数据库添加记录操作，当状态为其它各状态时，可省略实际的数据库操作。

数据库查询的结果是相应各字段的可能取值，对话管理可以将取值填入到相应的主题树结构中。

4.1.3 交互知识管理

对话过程中的交互知识，可直接填入主题森林中，模拟对话中建立的思想模型。

对话管理获得当前用户语义后，可确定所属的主题（如 4.1.1 所述），然后将各项语义填入到相应的主题树中。填入方式有两种：添加和替换。添加指加入新的语义信息，即用户给出的信息项的限制条件，替换指用新

的语义信息代替已有的信息。添加和替换实际上是对话过程中信息的更新方式，其效果如同对话时发生的记忆与遗忘现象。

如何填入用户语义的策略，就是交互知识管理中记忆与遗忘的策略。当某项用户语义相关的信息项还没有限制条件时，或者限制条件是来自于该项信息的默认值时，对话管理将执行语义信息的添加操作。如果该信息项已有限制条件，首先判断相关主题的类型属于查询型主题，还是记录型主题。所谓查询型主题，指该主题用于信息查询，例如查询电话号码、天气预报、查询火车信息等。由于查询型主题只用于查询功能，旧的查询条件没有保留的必要，可以采用替换操作。所谓记录型主题，指主题的信息条件可能会加入数据库，例如航班订票、约会提醒、股票买卖等。对于记录型主题，要考虑用户语句是陈述句还是疑问句。如果用户语句为陈述句，可认为用户是明确给出信息项的限制条件，应采用替换操作；如果用户语句为疑问句，表示用户只是查询，并没有确定当前的限制条件，应采用添加操作，即系统仍“记忆”着原来的信息条件。

图 4-1 所示为航班订票对话的一个片段，其中 U 表示用户语句，S 表示系统语句。U2 语句为疑问句，用户想知道后天有没有航班，以便在明后天之间取舍，对此系统采用添加操作比较合适。U3 语句为陈述句，用户决定航班日期，对此系统采用替换操作比较合适。

U1: 明天有去乌鲁木齐的航班吗?
S1: 有。
U2: 那后天有吗?
S2: 后天也有。
U3: 那我要后天的吧。

图 4-1 记忆与遗忘策略

在 4.1.1 中提到，确定用户语句的主题有时要参考上下文。当已确定当前要操作的主题树并填入用户语义时，也需要考虑上下文省略。例如在航

班信息查询时，用户给出一个时刻条件“八点二十”，这时仅凭该语句无法确定是指起飞时刻还是到达时刻，必须考虑上下文，这就需要对话管理使用交互历史记录来确定确切的语义。当交互历史记录也无法确定当前确切语义时，可认为该项语义是对话场景下的默认语义，例如在航班信息查询的对话中，用户一开始就说“我要去上海，查一下八点的飞机。”时刻信息“八点”无法依据对话上下文判断是起飞时刻还是到达时刻，可以默认在这一对话场景下是指起飞时刻。

当对话过程中出现主题切换且切换前后的主题具有共享信息时，也会出现上下文省略的现象。根据共享信息索引，对话管理能意识到可能存在的共享信息省略，从而将共享信息从一棵主题树复制到另一棵主题树中，使得新主题保持最新的交互知识状态。

用户语句中包含的语义信息和数据库操作得到的数据信息最终都存储到主题森林结构中，即交互知识直接利用主题森林来表示。只要遍历主题森林，就可知道各主题所处的状态，从而确定对话状态。

在对话中有时会发生信息更正的现象，产生原因可能是说话者改变了原本的意图，也可能说话者意识到听话者误解了先前的语句。对于口语对话系统来说，由于目前语音识别技术和自然语言理解技术水平有限，这两种产生信息更正的情况都可能发生。用户更正原有信息一般有两种方式。一种方式是直接给出正确的信息，根据前述的策略，这时对话管理将采取替换方式更新信息；另一种方式是用户给出诸如“刚才一句不算”之类表示取消前一回合信息的句子，这时对话管理要依据交互历史才能恢复到先前的对话状态。

用户语义中除了包含信息项的限制条件外，还可能包含直接对信息项进行某种操作的语义。例如在航班信息查询中会有“返程”的语义，表示将起点城市和终点城市互换。这时对话管理的操作等价于执行两次填入操作。在实际对话系统应用中，一般还提供重新开始对话的“超级命令”，即用户可以要求当前对话中止并重新开始，这时对话管理将把主题森林中所有信息项的限制条件清除。

用户语义中还可能含有与领域任务无关的一般性对话语义，例如寒暄语“你好”、“再见”，要求重复的“请再说一遍”、“我没听清”等等。这些语义不影响对话交互知识，对话管理直接进行相应的处理。

4.2 交互历史管理

对于信息咨询的对话系统来说，交互历史管理主要是对与领域任务相关的信息交互历史进行管理，也可利用主题森林结构进行。每回合对话管理将用户语义填入到主题森林的各个叶子结点时，将回合编号与该信息项条件同时记录下来，就构成交互历史。

在 2.3.2 中提到，交互历史的记录可以采用不同的策略。第一种策略，保留从对话开始到当前的所有交互历史。对话管理可在主题森林每一叶子结点处维护一个栈结构，每当有新的语义信息就入栈，如果用户要求回退到之前的对话回合，则将语义信息出栈，并重新填入主题森林，维护对话交互知识。第二种策略，只保留最近若干回合内的历史记录，可对不同的信息项设置不同的交互历史记录长度，模拟实际对话中人们对不同信息项所能记住的内容。对话管理可在每一叶子结点处维护不同深度的栈，当栈满时，删除栈底记录。

4.1.3 中提到，有时用户语义由于上下文的原因，要借助交互历史才能确定。为了能够快速处理这类情况，省去在不同叶子结点查找交互历史的麻烦，可以建立一个数据结构，表示当前根据上下文可取的确切语义，如图 4-2 所示。

主题 1			主题 2		
语义 1	语义 2	语义 1	语义 2
确切语义 1	确切语义 2	确切语义 1	确切语义 2

图 4-2 根据上下文可取的确切语义

例如在航班信息查询主题中，语义 1 可为时刻，确切语义 1 可为起飞时刻或到达时刻，语义 2 可为城市，确切语义 2 可为起点城市或终点城市。

如果上下文中还没有出现过该语义，相应的确切语义可以为空。如果系统有默认的确切语义，可以在对话开始时就设置。由于对话过程中可能出现主题转换，所以该数据结构同时记录所有主题的上下文确切语义，并在每一回合更新。这一数据结构可在对话管理载入主题森林时自动建立。

图 4-2 的结构可用于处理同一说话者所表达的语义省略，而有时语义省略会发生在对话双方的一问一答中，例如系统问用户：“您想要大概什么时候起飞的航班？”用户答：“下午一点左右。”对话管理应记录提问的确切语义，才能正确处理用户语义。为此，交互历史管理还应记录如下回合信息：当前回合编号、当前主题、对话主导方、系统应答焦点、系统应答文本、是否列举查询内容、当前列举的值、用户提问的类型、系统提问的信息项、系统是否请求用户确认。回合信息有些是在应答推理过程中逐渐得到的，有些是在应答生成后才设置的，用于下一回合对话处理。这些回合信息也用栈结构记录，当用户要求对话回退时可从栈中恢复。

4.3 应答推理

应答推理是对话管理最重要的任务。对话管理器根据当前对话回合得到的用户语义信息，考虑对话交互知识和交互历史，必要时进行相关的数据库操作，决定应该回答的语义。应该回答的语义称为应答焦点。

4.3.1 与主题森林结构相关的应答生成

信息咨询的对话过程是问与答的交替进行，对于一个积极参与对话的对话系统来说，采取混合主导的对话控制方式比较好，即当用户主动提问时，系统尽量给予回答，否则系统主动向用户提问。

由于主题内所有信息项都在主题树中用叶子结点来表示，当系统应答的语义与某主题整体或某信息项有关时，只要给出主题结点或相应的叶子结点，再结合当前的回合信息，就可确定应答语义。因此对话管理在维护回合信息的基础上，只要推理出应答相关的结点，就得到了应答焦点，可简单表述为应答焦点是该相关结点。在主题森林的配置文件中，如果为主题结点和叶子结点附加一个编号，则应答焦点相当于该编号。

对话系统中实际生成应答文本的是自然语言生成模块，该模块的输入就是应答焦点。当应答焦点为某个结点编号时，根据回合信息中对话主导方是用户还是系统，可确定应该生成问题的答案还是生成相关的问题。如果用户提问，则根据用户提问的类型，可确定回答的内容；如果系统提问，则根据结点的取值情况可确定提问的内容。

4.3.2 应答焦点的推理算法

根据主题森林所表示的领域图式，考虑合作的对话交互原则，可以应用如下算法得到应答焦点：

- (1) 确定用户语义所属主题，即确定进行应答推理的主题树。
- (2) 将用户语义信息填入主题树。
- (3) 如果主属性分支信息状态为假，
- (4) 则系统主导对话，找主属性分支下某个未知信息的叶子结点作为应答焦点，结束。
- (5) 否则进行数据库操作。
- (6) 如果数据库操作失败，
- (7) 则应答焦点为该主题数据库操作失败，结束。
- (8) 否则将操作结果信息填入主题树。
- (9) 如果本回合用户主导对话（主动提问），
- (10) 则应答焦点为答案所在结点（主题结点或叶子结点），结束。
- (11) 否则，系统主导对话，按主属性分支、次属性分支、附加属性分支的先后顺序，查找“记忆”了多个信息限制条件的叶子结点。
- (12) 如果有这样的结点，
- (13) 则应答焦点为该叶子结点，结束。
- (14) 否则，如果数据库操作结果唯一，

- (15) 则应答焦点为主题结点，结束。
- (16) 否则按次属性分支、主属性分支、附加属性分支的先后顺序，查找具有信息限制条件且得到多个数据库操作结果的叶子结点。
- (17) 如果有这样的结点，
- (18) 则应答焦点为该叶子结点，结束。
- (19) 否则按次属性分支、主属性分支、附加属性分支的先后顺序，查找没有信息限制条件且得到多个数据库操作结果的叶子结点。
- (20) 应答焦点为找到的叶子结点（一定存在），结束。

从该算法可以看出，这一应答推理过程确实是以主题树的结构特点为基础的。首先，算法步骤(3)根据主属性分支的信息状态决定是否进行数据库操作；其次，算法步骤(11)按信息项的重要程度向用户提问，要求在多个条件中确认；最后，算法步骤(16)和(19)都优先考虑次属性分支下的细节信息向用户提问。

该算法在基于主题树结构进行推理时，还考虑了不同用户的兴趣。算法步骤(16)与步骤(19)区别在于，系统将会先与用户讨论曾得到用户语义的信息项，因为交互历史表明用户对该项信息的兴趣较大。

算法步骤(1)其实是领域概念与图式管理的内容，当用户转移主题时，系统可选择相应的主题进行应答推理。由于每棵主题树都存储了相应主题完整的对话交互知识（包括对话状态）和交互历史，所以当主题转移并再次转移回来时，应答推理仍能顺利地进行。

4.4 基于主题森林的对话管理流程

正如对话管理的心理模型所描述的，对话管理的三大任务相互关联，密不可分。对话管理的流程如图 4-3 所示。

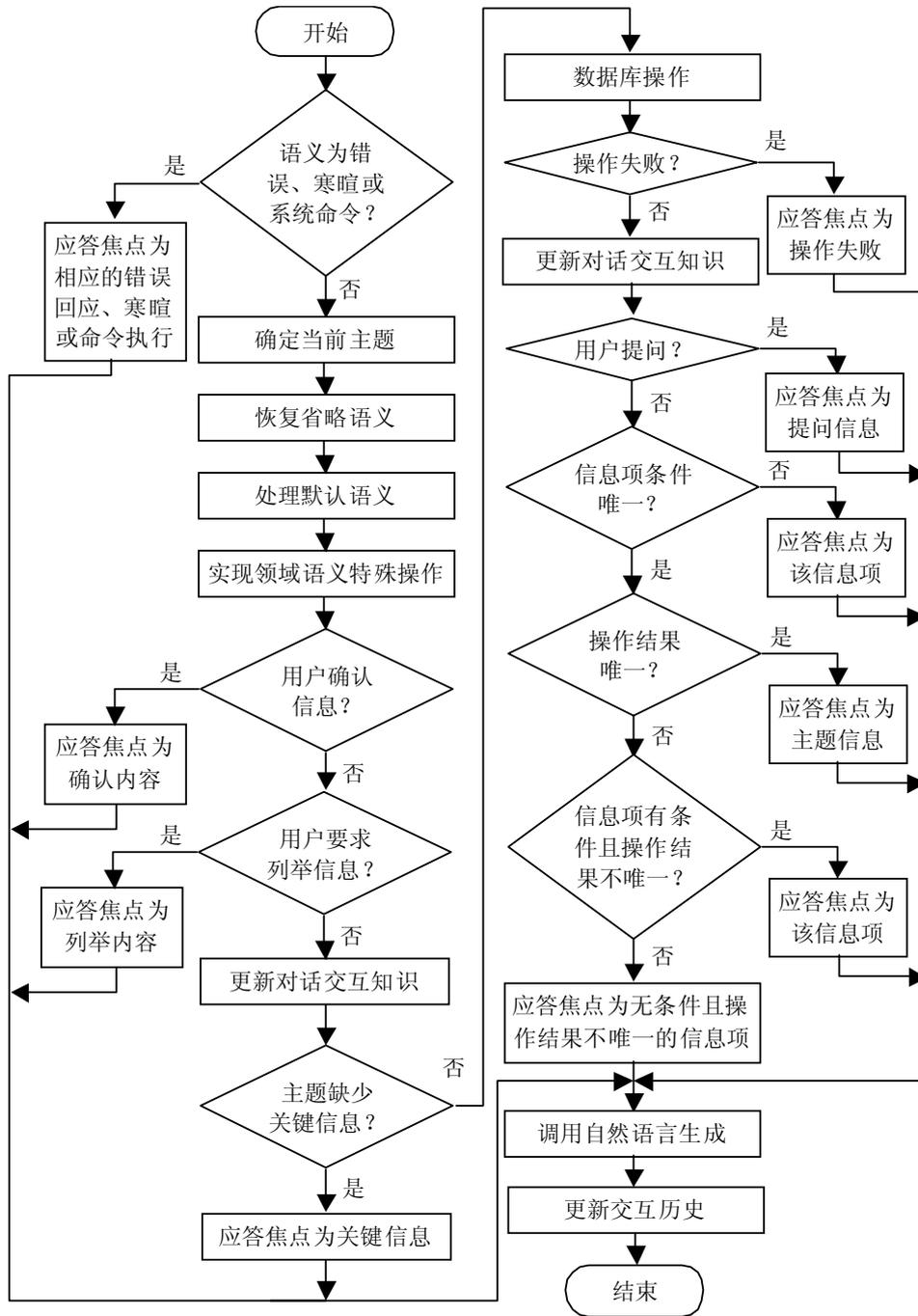


图 4-3 对话管理流程图

4.5 对话举例

作者在 EasyFlight 系统的对话管理器中应用了上述基于主题森林的对话管理方法，达到了较好的对话控制效果。图 4-4 所示是一个典型的对话例子（U 表示用户语句，S 表示系统语句）。

S1: 您好，欢迎使用本航班查询订票系统！

U1: 我要两张去乌鲁木齐的机票。

S2: 从北京到乌鲁木齐。您想问哪一天的？

U2: 后天有吗？

S3: 后天。有，查到 4 次不同的航班。

U3: 都是什么时间的？

S4: 起飞时刻为 11:25、13:50、19:15 和 19:25。

U4: 什么时候到？

S5: 到达时刻为 15:20、17:40、22:55 和 23:05。

U5: 那我要下午到的。

S6: 您要 9 月 20 日 11:25 从北京出发，15:20 到达乌鲁木齐的 CJ6901 次航班的机票吗？

U6: 好的。

S7: 您要 2 张票是吗？

U7: 票价是多少？

S8: 从北京到乌鲁木齐的机票每张 2140 元。

U8: 嗯，我要 2 张票。

S9: 请问您的身份证号码是多少？

U9: 123456789012345678。

S10: 您的身份证号码是 123456789012345678 吗?
U10: 没错。
S11: 您订了 2 张票, 该记录已加入数据库。
U11: 谢谢, 再见!
S12: 谢谢使用本系统, 再见!

图 4-4 典型的对话例子

在上述对话中, S2 的应答表明主题森林将北京作为起点城市的默认值, 由于日期是航班查询的重要信息, 所以系统主动提问。S6 一句, 系统进行航班数据库查询时, 得到唯一的航班记录, 所以问用户是否要订票。S7 表明 U1 语句中的“两张票”信息已作为交互知识记录下来, 所以系统不问要订多少张票, 而是直接请求用户确认票数。U7 一句, 用户将对话主题转换为航线信息查询, 并询问票价信息。这时根据用户语句包含的所有语义, 可确定该语句所属主题, 对话管理器意识到发生了主题转换, 转而在新的主题树上进行应答推理。根据主题森林的共享信息索引, 对话管理器意识到新主题的主属性分支信息状态为真, 可以进行数据库操作并且回答用户问题。U8 一句, 用户将主题转换为航班订票, 并确认票数。S9 一句, 对话管理器也回到航班订票主题树上进行应答推理, 继续询问用户身份证信息。S12 是对寒暄语“再见”的应答。

4.6 系统评测

进行系统评测有助于了解对话系统是否达到设计要求, 是否胜任领域对话任务, 有助于定性地比较不同对话系统的性能^[36, 37]。进行系统评测还可以比较不同的算法或策略对于系统性能的影响^[38, 39]。进行系统评测也是发现系统中深藏的错误, 从而改进系统的有效方法。

为了真正了解 EasyFlight 系统的性能, 特别是对话管理的性能, 我们设计组织了以下的系统评测。

4.6.1 对话场景设计

这次系统评测为对话者设计了 5 个对话场景，每一场景都给出一段文字描述说明场景任务。每个场景都附有一张表，表中已填入场景任务已知的信息条件，要求测试者通过对话交互获取要求的信息。图 4-5 所示是第一个对话场景。

对话场景一： 假定您在深圳工作，有个北京的朋友要来看您。他坐 6 月 22 日（星期六）的 X2 117 次航班来深圳，您需要知道飞机的到达时间以便去接他。	
对话场景一	
航线	北京→深圳
日期	6 月 22 日（周六）
航班号	X2 117
到达时刻	

图 4-5 对话测试场景一

在五个对话场景中，场景一、二只涉及航班信息查询，比较简单；场景三要求订单程机票；场景四要求订往返机票，且要求航班机型为几种可能机型之一；场景五要求订往返机票，有日期限制（该航线不是每天都有航班），且要求获取票价信息。因此在对话任务的难易度方面，场景一和二难度较小，场景三难度适中，场景四和五难度较大。

这五个场景描述的是实际生活中常见的航班查询或订票需求，任务给出的数据是精心准备的，保证在对话顺利的情况下可以获得唯一的信息结果，这将有利于计算客观评价指标。

4.6.2 用户满意度调查问卷

对话系统的评价指标分为主观指标和客观指标两类。常用的主观指标包括系统的理解能力、对话合作能力、正确应答能力等，对系统综合性能的评价称为用户满意度。主观指标的评价一般采用调查问卷的方式，问题涉及对话交互的各个方面，由测试者填写，再汇总统计。

为了得到主观评价结果，作者设计了一份用户满意度调查问卷，要求测试者在完成对话后填写。该问卷共有七个问题：

- 1、在对话中，系统理解你说的话吗？
- 2、对话任务可以很容易完成吗？
- 3、对话进展合理吗，有没有感到跳跃？
- 4、在对话中，你知道系统可以接受哪些话，不能接受哪些话吗？
- 5、系统响应及时吗？
- 6、系统的响应和你所预料的一样吗？
- 7、今后你愿意用这样的口语对话系统进行航班查询或订票吗？

对于每个问题，问卷都给出了从最好评价到最坏评价共五个级别的答案供选择。

4.6.3 客观指标与系统日志

对话系统评测常用的客观指标包括：回答正确率、任务完成率、对话回合数、对话时间、系统平均响应时间、出错信息率、平均句子长度等等。

为了计算这些客观评价指标，我们为 EasyFlight 系统加入了日志功能，记录内容包括：用户语音、每回合语音时刻、用户语音识别结果、分析得到的用户语义、每回合系统应答、对话回合数。

4.6.4 评测方法

我们找了 80 个大学本科学生，把他们分成两组。第一组的学生完成对话场景一和三，可选做场景五；第二组的学生完成对话场景二和三，可选做场景四。每个测试者完成对话场景任务后，填写用户满意度调查问卷。场景四和五作为选做场景是因为在实际生活中，虽然这两个场景任务也是常见的，但相对其他三个场景来说出现的比例较低。由于任务难度会影响系统性能，所以适当减少这两个任务的对话，可以使评测结果更接近实际应用的情形。

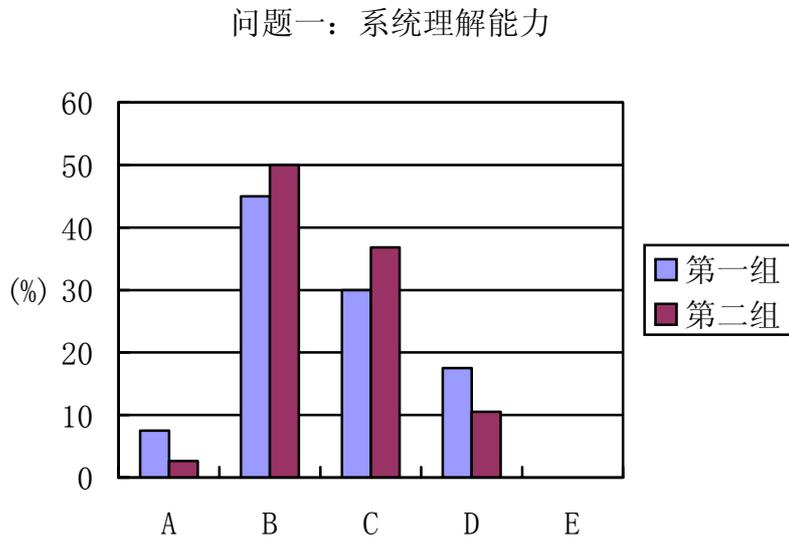
进行分组的目的在于：如果两组的评测结果有明显差异，说明场景任务对系统性能影响较大，分组的评测结果可信度不高，否则可以认为评测结果能够基本反映系统的整体性能。

由于系统评测的复杂性，对话系统的评测规模一般都不大，例如[38]中测试者为 32 人，[40]中测试者为 42 人。相对来说，我们组织的这次系统评测，规模已经相当大了。

在进行对话系统评测时，测试者通过麦克风与系统对话，系统应答显示在计算机屏幕上。第一组实际收到 40 份调查问卷，其中有 15 人选做了场景五。第二组实际收到 38 份调查问卷，其中 8 人选做了场景四。这 78 份调查问卷成为计算主观评价指标的依据。由于操作的问题，系统日志中实际只记录到 125 个对话，作为计算客观评价指标的依据。

4.6.5 主观评测结果

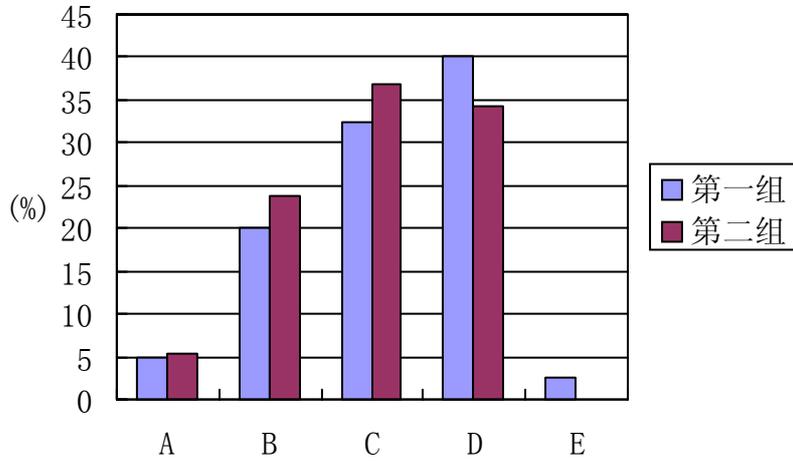
图 4-6 至 4-12 为调查问卷中各问题答案的柱状分布图。



A: 绝大部分; B: 大部分; C: 一般;
D: 一小部分; E: 几乎不能

图 4-6 用户满意度调查问题一的答案结果

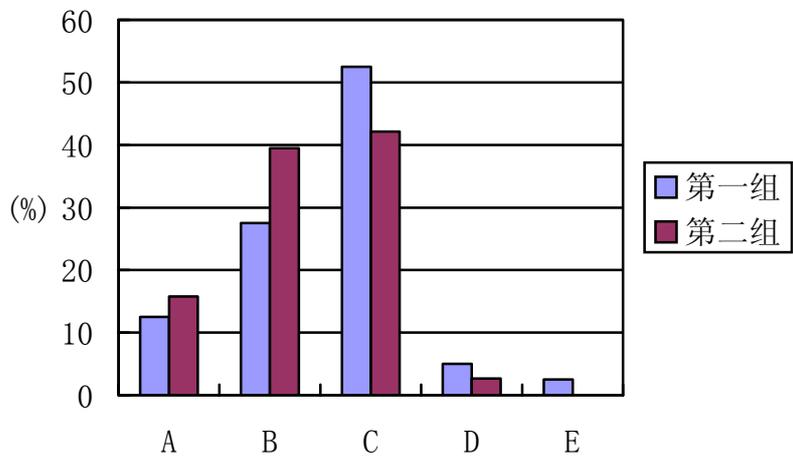
问题二：对话任务难易度



A: 非常容易; B: 没问题; C: 不算难;
D: 有点难; E: 根本无法完成

图 4-7 用户满意度调查问题二的答案结果

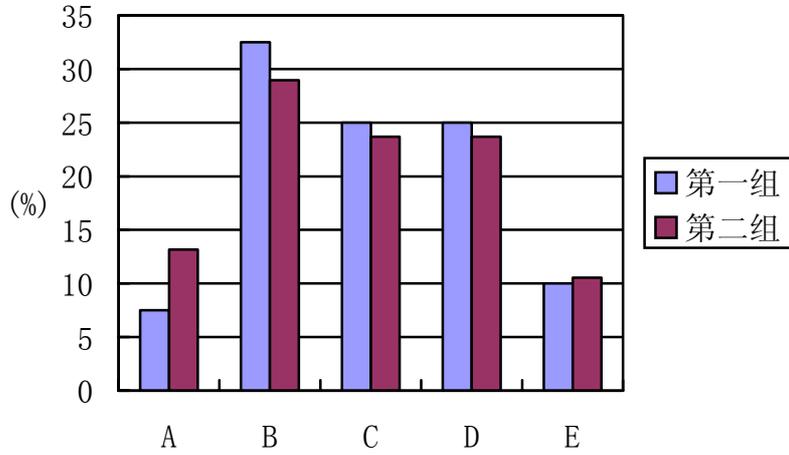
问题三：对话过程合理性



A: 非常合理; B: 感觉有时有点跳跃;
C: 会有些跳跃但还可以接受;
D: 我不喜欢这样的对话; E: 我讨厌这样的对话

图 4-8 用户满意度调查问题三的答案结果

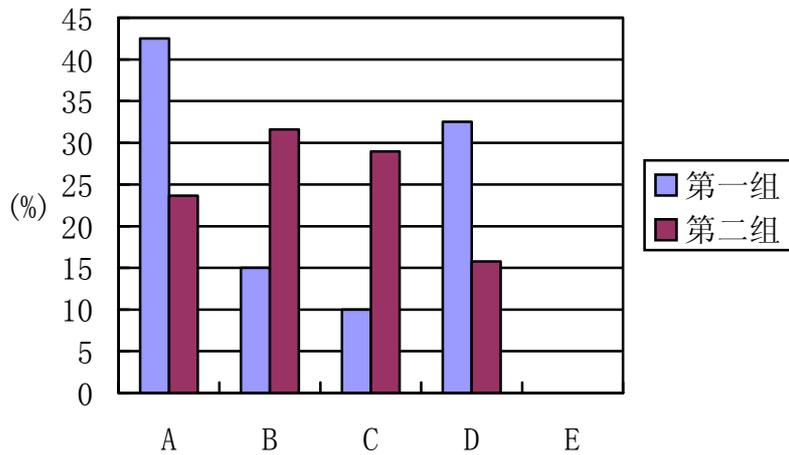
问题四：输入语句的随意性



- A: 从没考虑过，应该都能接受吧；
- B: 个别情况下只接受某些话；
- C: 不同对话情况下，都能接受几种不同内容的话；
- D: 个别情况下可接受几种不同内容的话，一般不行；
- E: 系统总是只接受对话过程中特定的话

图 4-9 用户满意度调查问题四的答案结果

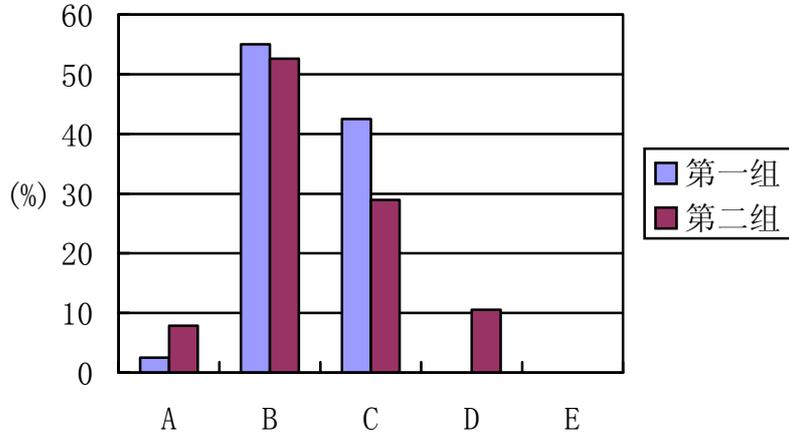
问题五：系统响应实时性



- A: 反应总是很快；
- B: 我有时要等；
- C: 还行；
- D: 有点慢；
- E: 太慢了

图 4-10 用户满意度调查问题五的答案结果

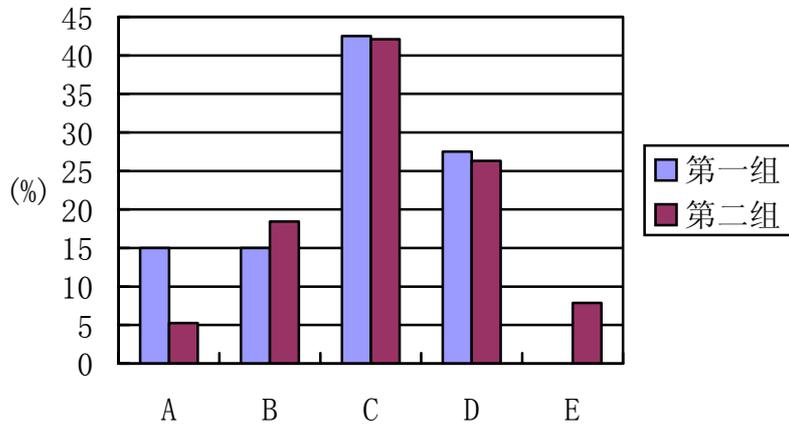
问题六：系统响应可预料性



- A: 总是在我预料中; B. 有时出乎意料, 但可以理解;
 C: 时有意料之外的, 但不算过分; D. 在预料之中的回答不多;
 E: 我根本无法预计系统会说什么

图 4-11 用户满意度调查问题六的答案结果

问题七：系统的实际可用性



- A. 这个系统比航班订票点的服务好, 我非常愿意;
 B. 这样的系统和航班订票点的服务差不多, 我愿意;
 C. 如果航班订票点电话打不进去, 我会用的;
 D. 如果航班订票点电话打不进去, 事情又急, 我会用的;
 E. 这样的系统根本没有用, 我不会浪费时间去用的

图 4-12 用户满意度调查问题七的答案结果

从上述直方图中可以看到，两组测试者除了对问题五的答案分布差别稍大外，其余问题的答案分布都差不多，说明场景任务分配对用户满意度测试没有明显影响，两组测试结果都可认为是可信的。

问题七是对整个系统可用性的综合评价，从答案中可以看到大部分测试者认可这一系统的服务性能（该题各答案选项如下：A. 这个系统比航班订票点的服务好，我非常愿意；B. 这样的系统和航班订票点的服务差不多，我愿意；C. 如果航班订票点电话打不进去，我会用的；D. 如果航班订票点电话打不进去，事情又急，我会用的；E. 这样的系统根本没有用，我不会浪费时间去用的）。

根据用户对场景任务是否成功的主观判断结果，第一组（场景任务一、三、五）对话成功率为 80.1%，第二组（场景任务二、三、四）对话成功率为 76.0%，基本相等。由于在实际评测过程中，测试者并不是非常合作，系统仍达到了这样的性能令人非常满意（[41]中介绍的系统只提供航班信息查询服务，功能不如 EasyFlight，对话成功率为 50%）。

4.6.6 客观评测结果

由于语音识别性能可用别的方法进行评测，本次对话系统客观指标评测主要是测试自然语言理解与对话管理的性能，因此对日志记录进行处理时，以用户语音的识别结果当作用户语句的输入，用户语音本身只用于判断真正的输入语句。

我们对系统日志记录的 125 个对话进行了分析，统计了如下几项数据：对话总回合数、用户语句中领域外的句子数、用户语句包含的语义数、系统正确理解的语义数、没有理解的语义数、错误理解的语义数、以及系统正确的应答回合数。其中所统计的语义包括：起飞城市、到达城市、日期条件、起飞时刻条件、到达时刻条件、航班机型、航空公司、订票张数、航班号，对以上各项的疑问，以及对系统问题的确认。

（一）用户语句输入

所谓领域外句子指用户语句既不包含以上语义，也不是寒暄语。由于 EasyFlight 系统的自然语言理解模块采用了针对汉语口语特点的增强文法，只要用户语句中含有特定语义，即使整个句子不合常规，也认为是领域内的句子。在日志记录的对话中，共统计有 775 条用户语句，其中领域外句子为 172 句，占 22.2%。

分析对话上下文可知，领域外句子较多的原因是目前所采用的关键词语音识别算法性能不好，例如在多个对话过程中都将用户确认时所说的“是”识别成“十”。

（二）用户语义的理解

在统计的各项数据中，语义是否被正确理解指自然语言理解模块分析出的语义是否正确。所谓没有理解的语义，指用户语句包含但没有被系统发现的语义；所谓错误理解的语义，指用户语句包含但被系统错误理解的语义，或者用户语句不包含但被系统误认为含有的语义。对于口语对话系统来说，用户语句包含的语义没有被理解与被误解成别的语义，是两种差别很大的错误。语义没有被理解，一般只导致对话回合数的增加，而语义被误解，还会产生应答不合理的现象。

在日志记录的 $775 - 172 = 603$ 句领域内用户语句输入中，共含各类语义 752 项。其中被正确理解的语义共 653 项，正确率为 $653/752 \times 100\% \approx 86.8\%$ ；没有理解的语义共 79 项，缺失率为 $79/752 \times 100\% \approx 10.5\%$ ；错误理解的语义共 35 项，错误率为 $35/752 \times 100\% \approx 4.7\%$ 。

从以上数据中可以发现，平均每回合用户语句所含的语义只有 $752/603 \approx 1.25$ 项，由于起点城市和终点城市经常在同一个回合中给出，所以基本上其它语义都单独出现在一个语句中。这一现象说明对话管理与系统应答较好地控制了用户语句所包含的语义数，使得用户语句比较简单，有利于对话任务的完成。

仔细检查那些含有没有理解语义和错误理解语义的用户语句，发现这些现象主要发生在日期和时刻语义的理解上。自然语言理解模块可能存在程序错误，现象一：“六月二十三日”在有的语句中被正确理解，而在有

的语句中没有理解；现象二：“六月二十二号早上”被正确分析出含有日期和时刻语义，但具体数据错误，导致语义的误解。此外，自然语言理解模块不能理解由“对”和“好”单个字构成的句子。这两个问题可以通过排除程序错误和扩充自然语言理解所用的领域文法来解决。

（三）对话应答

系统正确应答的回合数，直接反映对话管理的性能。我们采用表 4-1 所示的标准来判定某个系统应答是否正确。

表 4-1 系统应答是否正确的判定标准

用户语句中的语义项	应答情况	是否正确
有，多个	多个语义全对	是
	部分语义对	是
	其它	否
有，单个	单个语义对	是
	其它	否
无	无	是
	有	否

在对话过程中，只要回合间语句的语义保持一致，对话就可顺利进行。因此当用户语句含多项语义信息时，只要系统应答表示理解某项语义，且不含错误理解的语义，就可认为应答是正确的。没有理解的语义可以在以后回合中交互得到，整个对话过程仍保持流畅。如果系统应答包含了错误理解的语义，即使其它多项语义都理解正确，仍会使用户不得不纠正该错误，打断正常的交互过程，因此认为应答错误。

根据以上标准，共有 693 个回合的系统应答是正确的，应答正确率为 $693/775 \times 100\% \approx 89.4\%$ 。同样采用表 4-1 的标准，对于自然语言理解模块的回合理解正确率可进行如下估计：领域外语句占总回合的 22.2%，理解正确

率（即没有得到任何语义）不大于 100%；领域内语句占总回合的 77.8%，理解正确率小于语义正确理解率 86.8%。所以自然语言理解模块的回合理解正确率小于 $22.2\% + 77.8\% \times 86.8\% \approx 89.7\%$ 。可见对话管理模块的响应正确率接近 100%。

（四）回合时间

根据系统日志记录的用户每回合语音时刻，可以计算每回合对话的时间，用以计算平均回合时间。由于实际评测时，测试者只是处在模拟场景下，有时因翻看场景介绍或者其它原因，不能及时与系统进行对话，所以将每回合对话时间简单平均，并不能得到可信的结果。

为此我们编写程序统计不同时间段内的对话回合数，并计算平均回合时间，见表 4-2。

表 4-2 不同时间段内的对话回合数和平均回合时间

时间（秒）	回合数	平均时间（秒）
<15	139	11.9
<20	311	15.0
<25	443	17.1
<30	519	18.6
<35	564	19.7
<40	596	20.6
<50	615	21.3
<60	629	22.1
<80	638	22.8
<100	644	23.4
< ∞	657	43.3

从表中可以看到，简单取所有回合时间的平均值不能正确反映实际的情况。根据上表数据，计算不同时间间隔内的对话回合数，可用图 4-13 表示。

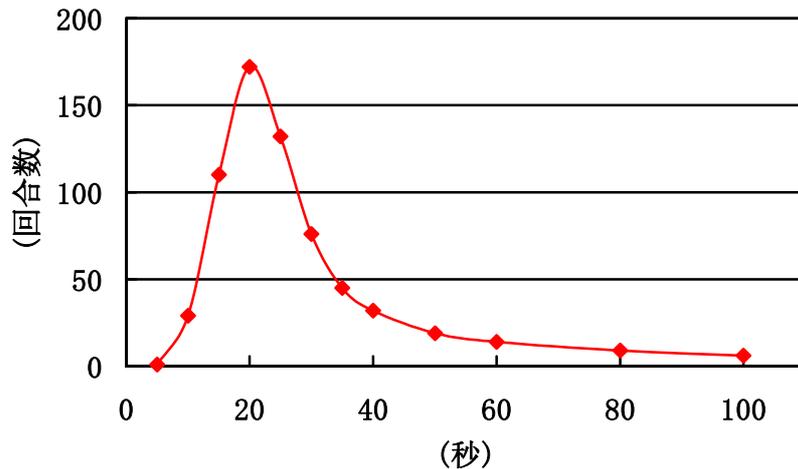


图 4-13 不同时间间隔内的对话回合数

根据图 4-13，平均的对话回合时间应该在 20 秒左右。进一步分析表明，平均对话回合时间稍长主要是因为语音识别速度较慢，自然语言理解也占用一些时间，对话管理几乎不占时间。由于 EasyFlight 没有使用网上航班数据库，而是直接将航班数据放在内存中，所以数据库操作时间很短。如果该对话系统用于实际，连接网上的航班信息数据库，平均回合时间会更长。总之，对话管理占用时间较少，不影响系统的实时性。

4.6.7 系统评测总结

这是一次规模较大的对话系统评测，结果显示整个系统表现良好，对话任务完成率较高，达到一定的实用程度，而且用户满意度水平也较高。评测结果表明，系统的语音识别性能相对较差，是系统整体性能提高的瓶颈；自然语言理解还存在一些小问题，但可以改进；对话管理的表现非常好。

4.7 对话管理特点总结

基于主题森林的对话管理方法以对话管理的心理模型为理论指导，以主题森林的领域概念和图式表示为基础，在对话系统中能较好地完成对话知识管理、交互历史管理和应答推理的任务。

由于主题森林结构所支持的领域概念和图式表示功能，涵盖了有限状态网络和表形、树形等基于规划的表示功能，基于主题森林的对话管理方法可用于有限状态的对话管理方法和基于规划的对话管理方法所胜任的领域任务，而且由于主题森林表示能力增强，能获得更好的对话管理表现。

由于单个主题的信息都存在单棵主题树中，对话管理每回合都收集该主题所有信息，并且根据主题树表示的图式进行推理，所以能够保持单个主题上下文信息的一致性。由于主题森林包含共享信息索引结构，使得对话管理能意识到主题间的共享信息，在处理主题转移时也能维护主题间信息的一致性。

基于主题森林对领域概念和图式的表示，该对话管理方法实现了对话控制的混合主导，对于复杂主题可以区别处理不同重要程度的信息，对于多主题可以自然地处理主题转移，使得对话过程比较灵活，用户自由度大。该对话管理方法速度很快，有利于系统实时应答。

主题森林是对话系统领域概念和图式的知识表示方法，适用范围较广。基于主题森林的对话管理方法在进行操作时，只根据主题森林的结构和内容进行，不涉及真正的语义，所以对话管理的操作与领域无关。因此该对话管理方法也具有广泛的适用性和较好的可移植性。

第五章 可定制对话系统框架

当今社会是信息社会，人们日益增长的信息获取要求决定了口语对话系统具有广阔的应用前景。口语对话系统集成语音识别、自然语言理解、对话管理、自然语言生成、语音合成等多项技术，目前建立对话系统的方法主要是针对单个领域任务进行设计和实现，所需时间长，投入的人力多，难以适应市场应用需求。要解决这一问题，使得对话系统的建立快速而有效，需要研究可定制对话系统框架。

本章先介绍对话系统建立的一般方法以及已有的对话系统开发工具，对当前对话系统开发工具进行需求和技术上的分析；然后提出可定制对话系统框架，说明该框架的组成及定制方法；最后以应用实例说明这个可定制对话系统框架的实用性。

5.1 对话系统的建立

如图 1-1 所示，对话系统由多个功能模块组成，系统的建立是一个复杂的过程。一般认为，由于当前对话系统所涉及的多项技术水平都不够高，当用户知道是与计算机进行交互时，会表现得和与真人对话时不同，即人-机交互时具有与人-人交互不同的特点^[42, 43, 44]。对话系统的建立要考虑这种差别。

5.1.1 WOZ 方法与开发-更新方法

Wizard of Oz 方法，简称 WOZ 方法，是一种获取人-机对话特点并测试比较不同对话策略的方法^[45]。该方法常用于对话系统建立过程的初期，具体做法是：告诉用户要与计算机进行对话，完成一定的场景任务，而实际上对话另一方并不是计算机，而是一名领域专家。该专家根据计划要在对话系统中实现的对话策略，模拟计算机与用户进行交互，从而获取较为真实的人-机交互资料。利用 WOZ 方法，可以在建立对话系统之前，就比较不同的对话策略，从而为实现对话系统提供指导，少走弯路^[46, 47, 48]。但是这一方法需要所谓专家能够表演好计算机的角色，由于当前语音识别、自

然语言理解技术水平不高，专家的表现往往比真实系统好很多，自然实际系统的性能水平会低于预计水平。

开发-更新方法与 WOZ 方法相反，希望尽快建立一个可运行的系统，然后让用户与该系统交互，根据交互结果，对系统各模块进行更新^[49]。开发-更新的过程反复进行，最后使整个系统达到一个较高的性能水平。该方法可得到较好的人-机交互数据，使系统性能稳步提升。但是这一方法在每次系统更新后都要组织用户与系统进行交互，工作量大，最初建立的系统可能与最后的系统差别很大，人力物力浪费较多。

以上两种方法可在一定程度上结合起来，以求较快较好地建立对话系统，具体做法是先估计各模块的性能、算法及策略需求，实现一个原型系统，然后经过少数几个开发-更新周期的改进并最终完成对话系统。

5.1.2 对话系统开发工具

为特定对话任务建立新的原型系统一般需要很长时间。因为不同对话系统领域任务不同，适用场合不同，应用的策略不同，组成对话系统的各个模块需要针对具体任务特制。如果把各模块中与领域无关的功能提炼出来，并提供与领域相关功能的定制，形成对话系统开发工具，就可以大缩短建立新的对话系统的时间。对话系统开发工具最重要的内容是提供一个对话系统框架，可以有效协调系统各组成模块的运行，并且可以为领域任务定制各模块。此外，对话系统开发工具应附带一些应用程序，帮助对话系统建立者定制各模块。

目前口语对话系统中的语音识别模块一般采用统计方法，领域定制过程就是收集领域语料进行训练的过程。系统中的自然语言理解模块如果也采用统计方法，就也要进行收集语料和训练的工作，如果采用规则文法的方法，则需对领域任务设计特定的文法。对于这两个模块来说，基本算法都是与领域无关的，易于形成开发工具。对话系统的自然语言生成模块与数据库操作模块完全与领域任务相关，而语音合成模块与领域任务基本无关，这些组成模块定制简单，易于集成到对话系统框架中。

对话管理与领域任务直接相关，对话管理的定制方法是现有对话系统框架最重要的核心技术。据我们所知，目前已有的可定制对话系统框架中大都采用有限状态的对话管理方法^[50, 51, 52]，定制新系统的对话管理的过程是建立适用于该领域任务的有限状态网络的过程，其中著名的有 CSLU 对话系统开发工具^[26, 53, 54]、EUROPA 框架^[55, 56]和 Nuance 开发包中的 DialogBuilder^[57]。有限状态网络是抽象的知识表示方法，可定制对话系统框架采用有限状态的对话管理，优点是易于描述和定义。然而本文第二章已指出，有限状态的对话管理方法难以适用于复杂的系统，已有的这些对话系统开发工具已不能适应目前对话系统建立的需求。

5.1.3 对话系统开发工具的需求及技术分析

目前信息咨询的口语对话系统需求广泛，简单应用如公交换乘查询、地点查询、气象查询等，对话可能只有一个回合，复杂应用如车船订票、旅馆预订、娱乐信息咨询等，对话涉及多主题和复杂主题，一般都需要多回合交互。一个实用的对话系统，还需要具有同时处理多路对话的能力。

由于现有语音识别水平不高，特别是在具有背景噪音及电话信道的条件下，只适于极简单的运用。而另一方面，由于互联网的繁荣和无线通讯的发展，通过网页传输和手机短信等方式进行文本交互的对话系统越来越受到人们的欢迎。但是，传统的网页一般只提供关键词接口的查询界面，用户难以表达复杂的查询要求，只能通过简单查询得到大量信息后再人工进行挑选。现有的手机短信服务都是依靠发送特定的功能代码和固定格式的一组关键词来获取信息，功能代码及关键词格式记忆困难，也不能自由地表达特定需求，使用不便。

综上所述，为了适应当前的应用需求，对话系统开发工具要求能适用于不同的领域任务，即适用于不同复杂度的对话任务，而且所建立的对话系统能支持并发的多路对话。作为对话系统开发工具的核心，可定制对话系统框架应提供文本接口，文本内容为自然语言，而不是特定代码和固定格式的关键词组。

可定制对话系统框架最大的难点在于对话管理。对话管理的基础在于领域概念和知识表示，用有限状态网络或表格形式来表示，易于实现对话管理的定制，但适用性有限；用树形方式来表示，适用性好，但难以将与领域无关的因

素从整个对话管理中提炼出来，也就难以提供定制功能。由于对话管理的核心地位，只有确定了对话管理定制化的形式，才能进一步确定系统框架中各组成模块的接口以及相互协调的调用关系。

不管对话系统采用语音接口（通过语音进行对话）还是文本接口（通过文本输入输出进行对话），自然语言理解都是一个难点。传统的自然语言理解注重于句子的合法性检查，以得到句子结构为目标，而对话系统中的自然语言理解以得到句子语义为目标。一般性的语义表示是与知识表示类似的难题，目前只有限制其表示内容与领域相关，才能较好的实现。目前各种语音识别技术的识别结果中都存在不少插入错误、删除错误和替换错误，识别正确率都不高。由于语音输入时用户语音的自发性和随意性，识别结果中不可避免含有停顿、重复、修正等多种口语现象^[58]。对于文本输入来说，实际用户的汉语文本输入在很大程度上接近生活口语，而不是书面语，句子往往比较短小，同一个意义可以有多种不同的说法，各个语义成分的顺序比较随意，且各成分间会出现一些可有可无的助词。从技术上说，对语音识别结果和用户文本输入进行自然语言理解有很大的相似性。自然语言理解要能够分析这些汉语口语句子，得到所包含的正确语义，是一件很困难的事。

目前的自然语言理解技术可大致分为统计方法和规则方法两类。统计方法要求收集大量语料进行标注和训练，而且由于汉语口语的特点，训练语料的需求量会比书面语大得多。规则方法一般由专家给出领域内的规则，所需时间少，而且一般说来性能好于统计方法，但是需要专家有丰富的经验。两者比较，规则方法的投入比较小，更有利于快速定制领域任务相关的自然语言理解模块。

此外在自然的对话过程中存在上下文省略现象，基于上下文无关文法的自然语言理解还需要结合其它方法才能有效处理。在对话系统中，对于跨回合的上下文省略的理解，对话管理发挥着重要的作用。

5.2 可定制对话系统框架

本文提出的主题森林表示方法，具有领域概念和图式的表示功能，能够适用于复杂的对话系统；基于主题森林的对话管理方法具有广泛的适用

性和良好的可移植性。通过对不同对话领域任务的分析比较，作者对所提出的基于主题森林的对话管理方法做了进一步提炼和适当扩充，形成了基本的可定制对话系统框架。以此框架为核心，可设定对话系统其它组成模块的功能与接口。

清华大学智能技术与系统国家重点实验室语音技术中心的燕鹏举博士，针对对话系统中用户语音自发性、随意性，以及语音识别输出中插入错误、删除错误和替换错误，提出了适用于汉语口语分析的基于语义类的上下文无关增强文法，以及相应的 *Marionette* 分析器，具有较高的理解成功率^[59, 60, 61]。作者对该文法及相应分析器做了局部的改动，将其纳入到对话系统框架中，形成可定制的自然语言理解模块。

根据对话系统框架的定制接口，作者将数据库操作、自然语言生成模块也加入到该框架中，并给出各自的定制方法。

为了实际应用时更快更好的定制各模块，作者设计实现了几个辅助的应用程序。这些应用程序与可定制对话系统框架组成了对话系统开发工具。

5.2.1 对话系统框架组成

作者提出并实现的可定制对话系统框架如图 5-1 所示。

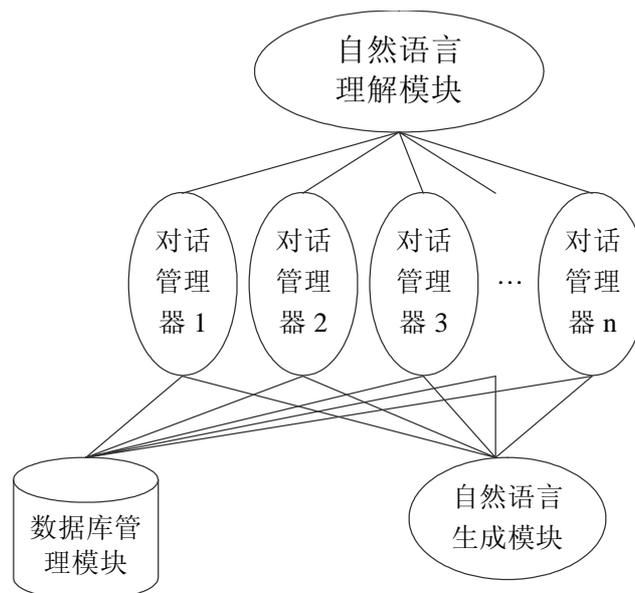


图 5-1 可定制对话系统框架

从图中可以看到，该框架只包含一个自然语言理解模块、一个数据库管理模块、一个自然语言生成模块，以及多个对话管理器。在系统运行时，每一路对话都由一个对话管理器负责对话控制，但这些对话可以共享自然语言理解模块、数据库管理模块和自然语言生成模块。

该对话系统框架中不包含语音识别模块和语音合成模块，有以下几点原因。一是因为目前很多对话系统都只要求文本输入和输出，用不到语音识别或语音合成；二是因为语音识别和语音合成在语音接口的对话系统中，与其它模块的耦合度小，经常可简化成语音识别→文本接口的对话系统→语音合成的串行模式，即容易在文本接口的对话系统上增加语音输入输出功能；三是因为以目前的语音技术水平，为某个具体任务定制时经常要选择合适的算法并进行训练或自适应，提供一个固定的算法模块难以适应不同对话系统的需求，而提供多种算法的可定制语音技术模块实现起来又太复杂，其难度不小于可配置对话系统框架。

5.2.2 各组成模块的输入输出

可定制对话系统框架中，各组成模块间数据流的输入输出如图 5-2 所示。

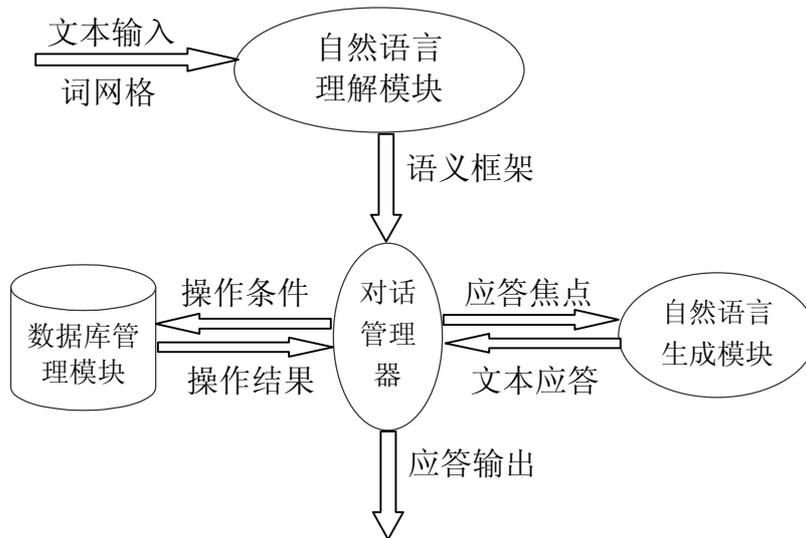


图 5-2 各组成模块的输入输出

自然语言理解模块可以接收文本输入或者语音识别得到的词网格，输出用语义框架形式表示的用户语义。对话管理器接收每回合的语义框架，视需要进行数据库操作，生成应答焦点，然后输出由自然语言生成模块产生的真正文本应答。从图 5-2 的数据流向也可以看到，对话管理器确实是对话系统的核心。

5.2.3 自然语言理解模块的定制

自然语言理解模块的定制，主要分为领域关键词、领域语法和语义解释三方面的工作。

（一）定制领域关键词

领域关键词是对话过程中最基本的语义单位，是领域语法的终结符。当自然语言理解模块的输入为文本句子时，先要根据领域关键词对句子进行分词处理。

领域关键词采用配置文件的方式定制。将领域关键词按所属语义类型分成不同的关键词类，每一类用一个段（segment）表示。段中每一个关键词都给出文本形式和汉语拼音，分别用于输入为文本和语音识别的词网格两种情况。有些关键词类所表示的语义是领域信息项，在自然语言理解模块对句法分析结果进行语义解释时，要将关键词转化成某种格式的语义表示，可以在配置文件中将这些关键词类标记出来。具体的关键词文件格式参见附录一。

合法的领域关键词可以包含汉字、数字和英文字母，数字和英文字母用特殊的字符串来代替拼音。这样定义关键词，使得文本输入中包含数字和英文字母时，自然语言理解模块也能处理。

在设置关键词配置文件时，绝大多数关键词都可以从数据库中直接导出。本文作者设计了一个应用程序，作为对话系统开发工具的一部分，可以将数据库输出的关键词文本，转换成关键词配置文件可用的格式。

（二）定制领域语法

领域语法给出了对话过程中合法句子的结构，是对输入进行句法分析的依据，也是进行语义解释的基础。

自然语言理解模块采用了基于语义类的上下文无关增强文法，其特点如下：

- 1、文法规则分为若干优先级，每一优先级的规则可以指定为词法分析的或非词法分析的；
- 2、规则有五种可能的类型，分别是苛刻型规则、跳跃型规则、长程型规则、无序型规则和交叉型规则。苛刻型规则就是传统方法的规则，表示连续句法成分组成更高级成分的关系。跳跃型规则允许组成更高级成分的各句法成分之间有少量的其它成分，用以解决口语中出现的停顿和无意义语气词现象。长程型规则允许组成更高级成分的各句法成分之间有任意数目的其它成分，用以解决口语中出现的重复和修正现象。无序型规则允许组成更高级成分的各句法成分先后顺序任意，用以解决汉语口语语序随意的的问题。交叉型规则允许组成更高级成分的各句法成分在句子中占位交叉但不冲突，且各句法成分先后顺序任意，用以解决汉语中某些特定句型，如“有……吗？”和“是……吗？”。

领域语法也采用配置文件的方式定制，说明各优先级是否是词法分析的，然后列出每一优先级的规则。具体的文法文件格式参见附录二。

（三）定制语义解释

自然语言理解模块进行句法分析的结果是将句子划分成一棵句法成分树，然后依据成分归约所应用的规则调用相应的语义解释函数，得到句子的语义表示。

语义解释的定制采用程序接口，即为每一规则定义一个语义解释函数，并为那些要用特定格式表示语义的关键词类定义一个语义解释函数。每个规则对应的语义解释函数都负责把该规则的语义表示出来。规则右边各子成分是根据别的规则归结得到或者就是某个关键词类，可利用自然语言理

解模块所提供的接口，调用该子成分归结时实际所用规则对应的语义解释函数，或者调用关键词类的语义解释函数。当自然语言理解模块按句法分析结果调用最顶级规则对应的语义函数时，就可根据句法成分树依次调用各规则的语义解释函数，直至关键词类的语义解释函数，最终得到整个句子的语义解释结果。句子所含语义用语义框架的形式表示，可以满足对话过程中回合语义的表示要求，并作为对话管理器的输入。

本文作者设计了一个应用程序，作为对话系统开发工具的一部分，可为语义解释的定制自动生成一组函数模板，直接面向每一规则的语义解释定制，并且支持定制内容的增量式更新。

可定制对话系统框架面向多种可能的应用，而领域语义解释可能非常复杂，提供这样的程序接口工作量小，表达能力强。

（四）自然语言理解模块算法简介

根据以上三个部分的定制内容，对话系统框架可得到一个专用于特定领域的自然语言理解模块。

自然语言理解模块在初始化时读入配置的关键词文件和文法文件，就可以对输入进行句法分析。

当输入为文本句子时，自然语言理解模块采用双向的关键词最大优先匹配算法对句子进行分词。在分词过程中对出现的连续数字可按单个数字串或某个数值进行不同的分词处理。分词结果是一个关键词网格，用于下一步的句法分析。当输入为语音识别的词网格时，直接进行句法分析。

句法分析根据优先级的高低，先应用最高优先级的规则进行归结分析，然后对分析结果进行剪枝优化，再应用下一优先级的规则进行归结分析，如此直至所有优先级规则都已应用过。由于采用五种不同类型的规则，句法分析算法比较复杂。在每次对分析结果进行剪枝优化时，要根据当前优先级规则是词法分析的还是非词法分析的，进行不同处理。如果当前优先级规则是词法分析的，则只有应用该优先级中最顶级规则归结得到的句法成分才被保留下来，参与下一优先级规则的分析，而其它部分分析结果都

被删除。如果当前优先级规则是非词法分析的，则所有分析结果都被保留。对于被保留的分析结果，其子成分都不再参与以后的分析。每次应用某一优先级的规则进行分析之后，对于相互冲突的不同分析结果，考虑汉语口语中可能的重复和修正现象，考虑口语趋向于结构简单、句子短小的特点，考虑语音识别的声学得分，将部分非最优结果剪枝去掉。规则分优先级、保留成分的子成分不参与后续分析、每次优先级分析后的剪枝，都有利于自然语言理解模块提高句法分析的速度。

在建立实际的对话系统时，语义解释函数与可定制对话系统框架在同一工程中被编译和链接。自然语言理解模块会根据句法分析结果，自动调用相应的语义解释函数。

5.2.4 对话管理器的定制

对话管理器的定制主要分为主题森林的定制、领域相关回合信息的定制、其它领域相关操作的定制三方面的工作。

(一) 主题森林的定制

第三章已介绍了主题森林结构，该结构可广泛应用于信息咨询的对话系统中领域概念和图式的表示。

按 3.3.2 所述确定主题森林的结构，然后按固定格式以前序遍历方式给出所有结点及其内容，即可生成领域任务的主体森林配置文件。主题森林包含的共享信息索引由对话管理器读入主题森林配置文件后自动生成，不必写入配置文件。主题森林的配置文件中要加入与数据库操作和自然语言生成相关的信息。具体的主题森林文件格式参见附录三。

如主题森林中含有记录状态的叶子结点，则相关的状态判断和状态转移以程序方式定制。当对话管理器根据主题的交互知识状态认为需要进行数据库操作时，会给出当前主题所有信息项的交互知识，这时可在进行数据库操作前进行对话状态的判断和转移。当数据库操作完成之后，在将查询结果返回给对话管理器时，可再一次进行对话状态的判断和转移。由于对话管理的下一步操作就是将数据库操作结果填入主题树中，所以状态信

息会写入记录状态的叶子结点中。从实际程序代码上看，状态判断和转移的程序接口与数据库操作的接口可以合二为一。

（二）领域相关回合信息的定制

4.2 中已介绍了基于主题森林的对话管理所维护的回合信息。除了这些必要的回合信息外，不同领域还可能需要特殊的回合信息。例如在一些含有日期信息的领域对话中，用户可能用几月几日、星期几或明后天等不同方式表达日期信息，虽然自然语言理解可以将不同表达都解释成同样的语义，但是为了符合用户的用语习惯，自然语言生成可以使用与用户相同的表达方式。这时就需要加入领域相关的回合信息，表明当前用户使用的表达方式，或者直接表明自然语言生成时应采用的表达方式。对话管理器为领域相关回合信息的定制提供了程序模板。

在对话过程中由于识别或理解错误，或者用户改变了意图，可能需要对话过程回退或重置。可定制的对话管理器提供了对这类对话控制的支持，只要定制自然语言理解模块使之能正确理解对话回退或重置的“高级命令”，再按预定义的格式填写语义框架，对话管理器就可执行相应的对话控制操作。对话管理器还提供了设置对话可回退回合数的函数接口。

（三）其它领域相关操作的定制

对话管理器提供了定制其它可能的领域相关操作的程序接口，主要是以下各项：

1、 设置领域默认语义。

在本文第四章中指出，在对话过程中有些语义表达要结合领域任务场景才能正确理解，涉及的语义包括领域任务默认的主题和主题各信息项默认的限制条件。

2、 设置领域特殊的语义相关操作。

在特定领域任务的交互规则中，某些语义信息的更新意味着另一些语义信息的取消。例如在电话区号查询的对话中，用户改变了省的信息，则市、区县的语义也失去了意义。在特定领域任务的对话中，还有一些语义不

表示信息项的限制条件，而表示对信息项条件进行操作，例如航班信息咨询对话中会出现“返程”的语义。在对话系统框架中实现这些领域特殊的语义相关操作，具体方法是利用提供的程序接口，在对话管理器得到自然语言理解输出的语义框架后，往语义框架中添加语义操作信息。当对话管理器将回合语义填入到主题森林中时，就会同时执行这些操作。

3、 设置领域特殊的语义转换操作。

在信息咨询的对话交互过程中，数据库查询结果经常被用来作为信息项的限制条件，因此存在从查询结果变为限制条件的语义转换。可定制的对话管理器提供这样的接口，使得用户能够自然地与系统对话而不必重复系统反馈的查询结果，对话交互知识真正为对话双方共享。

（四）对话管理算法简介

根据以上三个部分的定制内容，对话系统框架可得到一个专用于特定领域任务的对话管理器。该对话管理器可用于单个对话交互的对话控制。

对话管理器在初始化时读入主题森林配置文件，然后自动建立共享信息索引。在对话过程中，对话管理器采用图 4-3 所示的对话管理算法流程。

对话系统框架提供了对话管理器相关的生成、应答和删除三个可调用函数，用于定制实现可处理多路对话的系统。当系统得知有一个新的对话开始（例如接收到某手机发来的短信，或者接到某个电话，或者已建立新的网络连接等情况），就调用生成函数得到一个新的初始化了的对话管理器；当对话进行时，系统调用应答函数将用户输入传给相应的对话管理器，即可得到应答输出，反馈给用户；当系统判断对话结束（例如发生电话挂断，用户反应超时等情况），就调用删除函数将相应的对话管理器删除。由于每个对话管理器都维护各自所需的对话信息，而对话系统框架中其它模块或者与对话进程无关，或者从对话管理器获得对话相关信息，所以上述算法可以保证多路对话同时进行，互不干扰。

5.2.5 数据库管理和自然语言生成模块的定制

数据库管理与自然语言生成都是完全与领域相关的。

可配置对话系统框架要求在主题森林配置文件中给出与主题及各信息项相关的数据库信息，包括数据库中表和字段的名称，在调用对话管理器的生成函数时给出数据库管理的对象指针。在对话管理器运行时，会根据对话情况，自动调用数据库操作。该操作采用程序接口，给出当前回合的对话交互知识，以及数据库操作要求的结果。该接口与定制对话状态判断和转移的程序接口相同，两者可合二为一。

自然语言生成模块的定制就是根据对话管理器给出的对话交互知识、交互历史和应答焦点，生成真正的应答文本，相应的定制接口为程序接口。具体来说，要按程序接口的规定生成相应的系统帮助信息、系统寒暄信息、系统无法理解用户语义时的反馈信息、数据库操作失败信息、用户提问的确认信息，以及主题或语义项相关的信息。

当自然语言理解模块无法理解用户输入时，对话管理器给出当前已连续多少回合无法理解，并要求自然语言生成模块给出相应的信息反馈。自然语言生成模块可根据领域任务特点，设定不同连续回合数下的应答信息，使应答更人性化。

对于主题或语义项相关的信息，自然语言生成模块要先根据回合信息决定生成的是陈述句（用户主动提问）还是疑问句（系统主动提问）。如果是陈述句，则根据用户的具体问题给出答案，或者按用户要求列举某些信息。如果是疑问句，则根据主题或信息项的取值状态，结合当前主题的对话状态，生成相应的问句，或询问限制条件，或请求用户确认，或请求用户选择。

在生成具体应答后，自然语言生成模块应该通知对话管理器更新某些回合信息。例如当系统主动向用户提问缺失的关键信息时，用户的回答可能是承接问题的省略语句，这时必须有回合信息才能确定确切语义。

5.2.6 对话系统的建立

利用可定制对话系统框架建立新领域的对话系统可以采取以下步骤：

- 1、定制适用该领域任务的自然语言理解模块。

- 2、 定制适用该领域任务的对话管理器。
- 3、 定制该领域任务的数据库管理模块。
- 4、 定制该领域任务的自然语言生成模块。
- 5、 根据领域任务需要实现语音识别模块。
- 6、 根据领域任务需要实现语音合成模块。
- 7、 建立对话系统工程，利用对话系统框架提供的接口将各模块组合在一起，根据领域任务需要实现多路对话控制。

利用对话系统开发工具中的一些应用程序，可有效地加速模块的定制化过程，从而更快速地建立整个对话系统。

5.3 可定制对话系统框架应用实例

在燕鹏举博士对汉语口语的自然语言理解研究和本文作者对对话管理的研究基础上，我们用了近三年的时间，建立了 EasyFlight 系统，并进行了如 4.6 所述了对话系统评测。研究并实现了可定制对话系统框架后，作者花了一周的时间，利用该框架实现了 EasyFlight 系统中从文本输入到文本输出部分的功能。

为了进一步验证所实现的可定制对话系统框架的有效性，作者结合实际项目，利用该可定制框架，快速建立了绿色食品信息查询系统和足球竞猜系统。

5.3.1 绿色食品信息查询系统

随着经济的不断发展，人民生活水平的不断提高，越来越多的人开始关注环保，关注环境与健康的关系。绿色产品逐渐被公众认可和重视。经过市场调查发现，在各种绿色产品中，绿色食品的种类最丰富，与人们生活最息息相关，人们对绿色产品的认知度也高于其它绿色产品。据国家中心简报载，截止 2003 年 6 月底，全国有效使用绿色食品标志的企业总数达 1929 家，产品总数达 3427 个。建立绿色食品信息查询系统，有助于人们了解、选择和鉴别绿色食品。一些典型的查询如：

龙乾春绿茶是不是绿色产品？山东有没有绿色西瓜？穗穗牌绿豆的编号是 LB-06-0201040141 吗？

绿色食品信息查询系统采用手机短信方式进行对话，可利用可定制对话系统框架实现手机短信内容的分析与应答。

（一）自然语言理解模块的定制

一种绿色食品的信息包括品牌、产品名称、绿色食品编号、生产单位名称和所在（省市自治区、市县），其中品牌、产品名称、单位名称、单位所在都会有不同的称呼。这些关键信息每项都对应一个关键词语义类，加上表示这些信息项抽象概念的关键词语义类，组成短语和句子的其它语义类，共得到关键词语义类 42 个。附录一给出了定制自然语言理解模块所用的关键词文件，其中 `mat_brand1` 类（表示第一组品牌）、`mat_province` 类（表示省市自治区）、`mat_city` 类（表示市县）、`mat_company` 类（表示单位名称）、`mat_product1` 类（表示第一组产品名称）和 `mat_product2` 类（表示第二组产品名称）中的关键词较多，附录中只给出前面几个。这些关键词可直接从数据库中得到，并利用对话系统开发工具提供的应用程序，转换成可用的格式。其它语义类所含的关键词较少，而且很大一部分是汉语口语中常用的，在附录中全部列出。

以定义的关键词类为终结符，设计领域文法，思路是：

- 1、 设计规则用于分析表示关键信息项的短语，如具体的标志号、省市名称、产品名称等。
- 2、 设计规则用于分析表示某种抽象语义的短语，如泛指的企业、产品、标志号等。
- 3、 设计规则用于分析表示不同类型问题的短语。
- 4、 设计规则用于分析各类特殊疑问句。
- 5、 设计规则用于分析各类一般疑问句。

- 6、 设计底限规则，当用户语句不能用已有规则分析归结时，可利用底限规则来获取语句中可能存在的信息项。
- 7、 让所有规则最后都归结为文法初始符号。

根据以上思路设计绿色食品信息查询系统的所用文法，共有 228 条规则，见附录二。

利用对话系统开发工具生成文法相应的语义解释模板，实现各规则的语义解释函数。

以上操作完成了绿色食品信息查询系统中自然语言理解模块的定制。

(二) 对话管理的定制

绿色食品信息查询只涉及一个主题，且主题所含信息项比较少，主题森林可设计如图 5-3 所示，相应的配置文件见附录三。

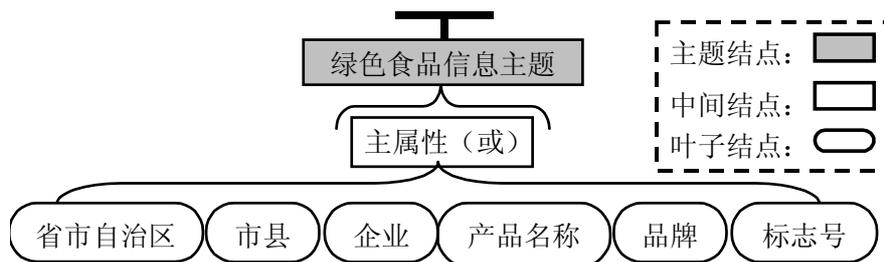


图 5-3 绿色食品信息查询系统的主题森林

绿色食品信息查询时，不必设置特殊的回合信息。

由于绿色食品信息查询的对话只包含一个主题，所以该主题就是领域默认主题，该主题的信息项没有默认的限制条件。

绿色食品信息中省市区、市县、企业几项语义之间有限制关系，需要进行领域特殊的语义操作。如当前回合信息中含企业信息，则应把主题中原有的省市区和市县信息清除，否则如当前回合信息中含市县信息，则因清除省市区信息。

对于绿色食品信息查询来说，信息项的限制条件只有“等于”或“不等于”两种，将查询结果变为限制条件的语义转换比较简单。

以上操作完成了绿色食品信息查询系统中对话管理器的定制。

（三）数据库管理和自然语言生成模块的定制

我们用 Microsoft Access 存储绿色食品信息，并实现数据库管理。主题森林配置文件中主题结点和各叶子结点给出的数据库中表和字段名称，其实是面向对话管理的虚拟数据库信息。实际数据库中对绿色食品信息进行了划分，建立了多张关系表。定制数据库管理模块的内容包括：1、实现数据库操作前的初始化（包括数据库连接）；2、将对话管理器给出的面向虚拟数据库的操作转换为面向真正数据库的 SQL 语句，并将操作结果返回对话管理器。

由于绿色食品信息查询系统采用手机短信方式进行对话，所以系统应答中除了用户提问的答案外，还应给出相关的一些信息，以减少对话回合数。例如用户问：“穗穗牌绿豆是不是绿色产品？”系统回答：“穗穗牌绿豆是绿色产品，标志号是 LB-06-0201040141。”这一应答既给出用户所需的答案，又提供了绿色食品标志号作为证据，用户不必再发短信问标志号了。这样的智能应答可以分为 5 类，分别用于回答关于绿色食品各项信息的特殊疑问句或一般疑问句。只要根据当前回合用户给出的语义，就可以确定应答的类型。

（四）对话系统的建立

完成以上各模块的定制，就可以建立一个文本输入文本输出的绿色食品信息查询系统了。这个系统的建立（不包括数据收集和数据库建立的工作），共用了二周的时间。

根据短信服务运行商的需求，可将已建立的绿色食品信息查询系统与手机短信服务平台相连，根据短信服务平台提供的手机号码及发送时间，决定如何调用对话管理器相关的生成、应答和删除函数。

短信服务运行商给出了 100 句句子作为绿色食品查询测试，其中有些句子非常口语化，例如：“鱿鱼丝什么的好？”“干燥裙带菜，大连有吗？”“豆浆哪的好？”“查一下山西省壶关县辛寨老陈醋厂。”“通化是哪的？”

“新疆屯河投资股份有限公司做什么的？”由于加入了底限规则，即使用户输入的语句不符合汉语口语习惯，自然语言理解也能抽取相应的语义。已建立的绿色食品信息查询系统对用于测试的 100 句句子都能正确应答。

在 CPU 为 Intel Pentium 4 1400MHz，内存为 256MB，操作系统为 Windows 2000 Professional 的计算机上，该绿色食品信息查询系统每秒可处理 70 条短信，回合时间主要消耗在访问 Access 数据库上。

5.3.2 足球竞猜系统

足球竞猜系统是与中央电视台合作的项目，用于央视的足球比赛有奖竞猜活动。由于实际参加竞猜的用户有上万人，高峰时收到的手机短信数目可达每秒近千条，而且大多数用户输入并不严格按中央电视台给出的标准格式，这样巨大数量和多变内容的短信无法在有限时间内靠人工来处理。利用计算机建立一个处理系统，充分发挥计算机处理速度快的优点，可保证竞猜活动顺利进行。

足球竞猜系统的输入为手机短信，包括手机号码和短信内容，系统输出为用户投票比分的的支持率。利用可定制对话系统框架建立该系统，关键要正确理解用户投票比分，然后进行数据库操作。足球竞猜系统还应易于修改用于不同的比赛，可定制对话系统框架提供的定制方式（配置文件方式），可以满足这一需求，而不必重新编译链接生成新的应用程序。

在实际足球比赛直播时，中央电视台在屏幕上给出足球竞猜的方式。例如 2003 年 8 月 2 日，中国龙之队在北京对阵西班牙皇家马德里队，央视要求短信竞猜格式为“1mA:B”，表示龙之队得分为 A，皇家马德里队得分为 B。但是实际收到的用户竞猜短信中，很多短信没有“1m”字样或者用“ml”字样，很多短信用了全角冒号而不是半角冒号，很多短信用“比”、“.”、“—”、“*”来代替冒号，有些短信中只有两个连续数字，有些短信索性用文字来表示两队的比分，也有一些短信内容与足球竞猜无关。

分析足球竞猜涉及的语义，共得到 11 个关键词语义类，放入领域关键词配置文件中。为了分析多种多样的用户短信，先编写了一个函数对输入进行初步的格式化处理，统一全角和半角符号，删除重复符号（如“——”、

“**”)，然后在领域文法配置文件中设计了 42 条规则，其中 0~4 优先级的规则都是词法分析的。在手机短信中会出现连续的数字，例如两个连续数字，应理解为竞猜比分；如果存在三个或更多的连续数字，可认为该短信无效；但连续 11 个数字又应视为有效的手机号码。通过调整规则的优先级并设定词法分析，这三种情况被依次处理，消除了分析时产生的歧义。

实现语义解释时，对于表示“龙马 3 比 2”、“龙 3 马 4”、“龙 3 比 2 胜马”、“马 4 比 3 胜”、“龙 3 比 0”、“3 比 2 龙胜”等短语的不同规则，都要正确给出两队各自的得分。

由于足球竞猜不涉及对话上下文，严格说来不属于对话过程，所以不需要定制对话管理器，只要根据自然语言理解模块得到的语义框架，直接进行数据库操作。

由于在足球竞猜过程中，只需记录竞猜投票，更新各比分的投票支持率，赛后才根据实际比赛结果计算中奖情况，所以数据库管理在程序内部实现就可以了，不必用专门的数据库管理软件。

系统的应答焦点是投票更新后该比分的支持率，自然语言生成模块只用一个应答模板就可以实现。

足球竞猜系统的建立用了三天时间。由于领域关键词数目较少，领域文法简单，数据库操作在系统内部实现，系统的处理速度非常快。在 CPU 为 Intel Pentium 4 1400MHz，内存为 256MB，操作系统为 Windows 2000 Professional 的计算机上，该足球竞猜系统每秒可处理 2500 条短信。

5.4 可定制对话系统框架总结

现有许多对话系统都是花了很长时间，投入了很多人力物力才建立起来的，然而传统的建立对话系统的方法已不能满足社会需求。已有对话系统开发工具，采用有限状态网络的对话管理方法，可定制领域任务较少。

基于主题森林的对话管理方法具有广泛的适用性和良好的可移植性，作者以此为核心设计并实现的可定制对话系统框架能够用来快速建立目前大多数的信息咨询对话系统。该框架的自然语言理解模块采用了基于语义

类的上下文无关增强文法，可定制处理不同领域的汉语口语。可定制的对话管理器与自然语言理解模块是这个可定制对话系统框架最重要的组成部分。

可定制对话系统框架以及用于模块定制的一些应用程序，构成了对话系统开发工具。据我们所知，目前国内还没有任何对话系统开发工具，而国外也没有应用非有限状态对话管理的可定制对话系统框架。

我们利用可定制对话系统框架建立了多个实用对话系统，所花时间较少，系统性能良好。

第六章 论文工作总结

对话系统的研究由来已久，但是随着自然语言理解技术和语音技术的发展，随着现代社会人们对信息需求的发展，对话系统中的对话管理显得日益重要，已成为对话系统的核心技术。本文的工作重点是对信息咨询的对话系统中对话管理技术的研究。

本文的研究工作和贡献主要分为以下三个方面。

（一）提出对话管理的心理模型。

传统的对话管理方法以自然语言理解技术中不同的对话模型理论为指导，在理论上存在这样那样的先天不足。由于对话模型注重在对话的语言层面，较少触及对话的心理层面，相应的对话管理方法在实际应用中领域局限性较大，难以在复杂领域任务中实现灵活自然的对话控制。作者从已有的研究思路中解脱出来，从心理语言学理论中汲取营养，加以归纳整理和提炼，提出了对话管理的心理模型。

对话管理的心理模型是从对话应答心理活动的角度提出的，模型组成分为知识、交互历史和对话策略三部分，其中知识与对话过程中的心理表征有关，交互历史与对话过程中的记忆有关，对话策略与对话过程中应答的心理取向有关。该模型描述了对话过程中应答生成的心理运作。对话管理的心理模型说明了已有对话管理方法的心理本质，可以解释已有对话管理方法的优缺点和适用性。该理论指出，实用对话系统中对话管理应完成三大任务：知识管理、交互历史管理和应答推理。

（二）提出主题森林结构和基于主题森林的对话管理方法。

根据对话管理的心理模型，领域概念和图式的知识表示是对话交互的基础，关系到对话者的语义理解能力和应答内容。在实际对话系统中，领域概念和图式的表示能力决定了对话管理的性能特点。

作者提出了主题森林结构，用于信息咨询的对话系统中，表示领域概念和图式。主题森林结构将概念的树形表示和图式的有限状态网络表示结

合在一起，可适用于含多主题和复杂主题的领域任务。由于表示了不同主题间的共享信息，可实现对话中主题转换的深入理解。由于表示了单个复杂主题中不同重要程度的多项信息，实现了与概念组成有关的图式表示。

基于主题森林的对话管理方法，以主题森林为基础，应用了信息咨询的对话所应采取的最基本的交互策略，可较好地完成知识管理、交互历史管理和应答推理三大任务。该对话管理方法可协调数据库操作和自然语言生成模块的运作。在提供航班信息查询和订票的 EasyFlight 系统中，对话管理器应用了基于主题森林的对话管理方法。根据一次大规模的对话系统评测结果，对话管理可较好的协调系统各组成模块，灵活自然地进行对话控制，完成领域任务，对话管理器的应答正确率接近 100%。

(三) 提出并实现了可定制对话系统框架。

由于基于主题森林的对话管理方法可广泛适用于信息咨询的对话系统，有较好的可移植性，以应用该对话管理方法的对话管理器为核心，作者设计并实现了可定制对话系统框架。该框架协调了对话管理器、自然语言理解模块、领域数据库操作模块和自然语言应答模块，可用于建立信息咨询的汉语口语对话系统，支持多路并发的对话控制。基于该对话系统框架，已快速度高质量地建立了多个汉语口语对话系统。

参考文献

- [1] 石纯一, 黄昌宁, 王家廛. 人工智能原理. 北京: 清华大学出版社, 1993
- [2] 王开铸. 自然语言理解—计算机能思维吗?. 哈尔滨: 哈尔滨工业大学出版社, 1996
- [3] 杨行峻, 迟惠生, 等. 语音信号数字处理. 北京: 电子工业出版社, 1995
- [4] 戴礼荣. 人机语声对话特点及系统设计. 安徽大学学报 (自然科学版), 1997. 21(4): 52~56
- [5] Goddeau D., Brill E., Glass J., et al. GALAXY: A Human Language Interface to Online Travel Information. In: Proceedings of the 3rd International Conference on Spoken Language Processing (ICSLP'94). Yokohama, Japan, 1994. 707~710
- [6] Huang C., Xu P., Zhang X., et al. LODESTAR: A Mandarin Spoken Dialogue System for Travel Information Retrieval. In: Proceedings of the 6th European Conference on Speech Communication and Technology (EuroSpeech'99). Budapest, Hungary, 1999. v3, 1159~1162
- [7] Huang Y.F., Zheng F., Xu M.X., et al. Language Understanding Component in Chinese Dialogue System. In: Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP2000). Beijing, China, 2000. v3, 1053~1056
- [8] Meng H., Tsui W.C. Comprehension across application domains and languages. In: Proceedings of the 2nd International Symposium on Chinese Spoken Language Processing (ISCSLP2000). Beijing, China, 2000. 69~72
- [9] Lin B.S., Lee L.S. Computer-aided Design/Analysis for Chinese Spoken Dialogue Systems. In: Proceedings of the 2nd International Symposium on Chinese Spoken Language Processing (ISCSLP2000). Beijing, China, 2000. 57~60
- [10] Hugunin J., Zue V. On the Design of Effective Speech-based Interfaces for Desktop Applications. In: Proceedings of the 5th European Conference on Speech Communication and Technology (EuroSpeech'97). Rhodes, Greece, 1997. 1335~1338
- [11] Issar S. A Speech Interface for Forms on WWW. In: Proceedings of the 5th European Conference on Speech Communication and Technology (EuroSpeech'97). Rhodes, Greece, 1997. 22~25
- [12] Asoh H., Matsui T., Fry J., et al. A Spoken Dialog System for a Mobile Office Robot. In: Proceedings of the 6th European Conference on Speech Communication and Technology (EuroSpeech'99). Budapest, Hungary, 1999. v3, 1139~1142

- [13] Os E.D., Boves L., Lamel L., et al. Overview of the ARISE project. In: Proceedings of the 6th European Conference on Speech Communication and Technology (EuroSpeech'99). Budapest, Hungary, 1999. v4, 1527~1530
- [14] Seneff S. TINA: a Natural Language System for Spoken Language Applications. Computational Linguistics, 1992, 18(1): 61~86
- [15] Seneff S., Hurley E., Lau R., et al. Galaxy-II: A Reference Architecture for Conversational System Development. In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sydney, Australia, 1998. 931~934
- [16] Seneff S., Lau R., Polifroni J. Organization, Communication, and Control in the GALAXY-II Conversational System. In: Proceedings of the 6th European Conference on Speech Communication and Technology (EuroSpeech'99). Budapest, Hungary, 1999. 1271~1274
- [17] Polifroni J., Seneff S. Galaxy-II as an Architecture for Spoken Dialogue Evaluation. In: Proceedings of the International Conference on Language Resources and Evaluation (LREC2000). Athens, Greece, 2000. 725~730
- [18] Zue V., Seneff S., Glass J., et al. JUPITER: A telephone-based conversational interface for weather information. IEEE Transaction on Speech and Audio Processing, 2000, 8(1): 100~112
- [19] Seneff S., Polifroni J. Dialogue Management in the Mercury Flight Reservation System. In: Proceedings of the ANLP/NAACL 2000 Workshop on Conversational Systems. Seattle, USA, 2000. 1~6
- [20] Simpson, A., Fraser, N. Black Box and Glass Box Evaluation of the SUNDIAL System. In: Proceedings of the 3rd European Conference on Speech Communication and Technology (EuroSpeech'93). Berlin, Germany, 1993. 1423~1426
- [21] 黄寅飞, 郑方, 燕鹏举, 等. 校园导航系统 EasyNav 的设计与实现. 中文信息学报, 2001, 15(4): 35~40
- [22] Schmid P., Seneff S., Polifroni J. BIANCA: A Dialogue Management Engine for PEGASUS. In: research reports. MA: The MIT Lab for Computer Science, 1998. 23~25
- [23] Cole R. A., Mariani J., Uszkoreit H., et al. Survey of the State of the Art in Human Language Technology. Pittsburgh, PA: Center for Spoken Language Understanding CSLU, Carnegie Mellon University, 1995. 227~280
- [24] Furui S., Yamaguchi K. Designing a Multimodal Dialogue System for Information Retrieval. In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sydney, Australia, 1998. v4, 1191~1194

- [25] Córdoba R., San-Segundo R., Montero J.M., et al. An Interactive Directory Assistance Service for Spanish with Large-Vocabulary Recognition. In: Proceedings of the 7th European Conference on Speech Communication and Technology (EuroSpeech2001), Aalborg, Denmark. v2, 1279~1282
- [26] McTear M.F. Modelling spoken dialogues with state transition diagrams: experiences with the CSLU toolkit. In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sydney, Australia, 1998. 1223~1226
- [27] Lin Y.C., Chiang T.H., Wang H.M., et al. The Design of a Multi-Domain Mandarin Chinese Spoken Dialogue System. In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sydney, Australia, 1998. 41~44
- [28] Wright J., Gorin A., Abella A. Spoken Language Understanding Within Dialogs Using a Graphical Model of Task Structure. In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sydney, Australia, 1998
- [29] Wang K. An Event Driven Model for Dialogue Systems. In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sydney, Australia, 1998. v2, 393~396
- [30] Heisterkamp P., McGlashan S. Units of Dialogue Management: an Example. In: Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP'96). Philadelphia, USA, 1996. 200~203
- [31] 桂诗春. 新编心理语言学. 上海: 外语教育出版社, 2000
- [32] 朱曼殊. 心理语言学. 上海: 华东师范大学出版社, 1990
- [33] 彭聃龄. 语言心理学. 北京: 北京师范大学出版社 1991
- [34] Garnham A. Mental models as representations of discourse and text. Chichester: E. Horwood, 1987
- [35] Lin B.S., Wang H.M., Lee L.S. Consistent Dialogue across Concurrent Topics Based on an Expert System Model. In: Proceedings of the 6th European Conference on Speech Communication and Technology (EuroSpeech'99). Budapest, Hungary, 1999. v3, 1427~1430
- [36] Bennacef S., Devillers L., Rosset S., et al. Dialog in the RAILTEL Telephone-Based System. In: Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP'96). Philadelphia, USA, 1996. 550~553
- [37] Constantinides P.C., Hansma S., Tchou C., et al. A Schema Based Approach to Dialog Control. In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sidney, Australia, 1998. 409~412

- [38] Derillers L., Maynard H.B. Evaluation of Dialog Strategies for a Tourist Information Retrieval System, In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sidney, Australia, 1998, 1187~1190
- [39] Lin B.S., Lee L.S. Computer-Aided Design/Analysis for Chinese Spoken Dialogue Systems. In: Proceedings of the 2nd International Symposium on Chinese Spoken Language Processing (ISCSLP2000). Beijing, China, 2000, 57~60
- [40] Hanrieder G., Heisterkamp P., Brey T. Fly with the EAGLES: Evaluation of the "ACCeSS" Spoken Language Dialogue System. In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sidney, Australia, 1998, v2, 503~506
- [41] Ipsic I., Mihelic F., Dobrisek S., et al. A Slovenian Spoken Dialog System for Air Flight Inquiries. In: Proceedings of the 6th European Conference on Speech Communication and Technology (EuroSpeech'99). Budapest, Hungary, 1999. v6, 2659~2662
- [42] Bell L., Gustafson J. Interaction with an Animated Agent in a Spoken Dialogue System. In: Proceedings of the 6th European Conference on Speech Communication and Technology (EuroSpeech'99). Budapest, Hungary, 1999. v3, 1143~1146
- [43] Gustafson J., Lindberg N., Lundeberg M. The August Spoken Dialogue System. In: Proceedings of the 6th European Conference on Speech Communication and Technology (EuroSpeech'99). Budapest, Hungary, 1999. v3, 1151~1154
- [44] Kumamoto T. An Analysis of Dialogues with Our Dialogue System through a WWW Page. In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sydney, Australia, 1998. v4, 1219~1222
- [45] Fraser, N.M. Gilbert, G.N. Simulating Speech Systems. *Computer Speech, and Language*, 1991. 5(1): 81~99
- [46] Pirker H., Loderer G., Trost H. Thus Spoke the User to the Wizard. In: Proceedings of the 6th European Conference on Speech Communication and Technology (EuroSpeech'99). Budapest, Hungary, 1999. v3, 1171~1174
- [47] Relano G.J., Tapias D., Villar J.M., et al. Flexible Mixed-Initiative Dialogue for Telephone Services. In: Proceedings of the 6th European Conference on Speech Communication and Technology (EuroSpeech'99). Budapest, Hungary, 1999. v3, 1179~1182
- [48] Wyard P., Churcher G. A Realistic Wizard of Oz Simulation of a Multimodal Spoken Language System. In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sydney, Australia, 1998. v4, 1247~1250

- [49] Seneff S., Hirschman L., Zue, V.W. Interactive Problem Solving and Dialogue in the ATIS Domain. In: Proceedings of the 4th DARPA Speech and Natural Language Workshop. Pacific Grove, California. Defense Advanced Research Projects Agency, Morgan Kaufmann
- [50] Pargellis A., Kuo J., Lee C.H. Automatic Dialogue Generator Creates User Defined Applications. In: Proceedings of the 6th European Conference on Speech Communication and Technology (EuroSpeech'99). Budapest, Hungary, 1999. v3, 1175~1178
- [51] Brndsted T., Bai B., Olsen J. The REWARD Service Creation Environment. An Overview. In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sydney, Australia, 1998. 1175~1178
- [52] Nouza T., Nouza. J. Graphic Platform for Designing and Developing Practical Voice Interaction Systems. In: Proceedings of the 7th European Conference on Speech Communication and Technology (EuroSpeech2001). Aalborg, Denmark, 2001. v2, 1287~1290
- [53] Sutton S., Cole R.A., de Villiers J., et al. Universal Speech Tools: The CSLU Toolkit. In: Proceedings of the 5th International Conference on Spoken Language Processing (ICSLP'98). Sydney, Australia, 1998. v7, 3221~3224
- [54] Cole R., de Villiers J., Shobaki K. Demonstrations of Dialogue Design Tools in the CSLU Toolkit. In: Proceedings of the ESCA Tutorial and Research Workshop, Interactive Dialogue in Multi-Modal Systems, 1999, Kloster Irsee, Germany
- [55] Sasajima M., Yano T., Kono Y. EUROPA: A Generic Framework for Developing Spoken Dialogue Systems. In: Proceedings of the 6th European Conference on Speech Communication and Technology (EuroSpeech'99). Budapest, Hungary, 1999
- [56] Sasajima M., Yano T., Shimomori T., et al. MINOS-II: A Prototype Car Navigation System with Mixed Initiative Turn Taking Dialogue. In: Proceedings of the 7th European Conference on Speech Communication and Technology (EuroSpeech2001). Aalborg, Denmark, 2001. v2, 1311~1314
- [57] <http://foto.hut.fi/research/TargetJr/manpages/Viewing/IVAdditions/DialogBuilder.html>
- [58] Zue V. Conversational Interfaces: Advances and Challenges. In: Proceedings of the 5th European Conference on Speech Communication and Technology (EuroSpeech'97). Rhodes, Greece, 1997. KN9~18
- [59] Yan P.J., Zheng F., Xu M.X. Robust Parsing in Spoken Dialogue Systems. In: Proceedings of the 7th European Conference on Speech Communication and Technology (EuroSpeech2001). Aalborg, Denmark, 2001. v3, 2149~2152

参 考 文 献

- [60] Yan P.J., Zheng F., Sun H. et al. Spontaneous Speech Parsing in Travel Information Inquiring and Booking Systems. *Journal of Computer Science and Technology*, 2002, 17(6): 924~932
- [61] 燕鹏举. 对话系统中的自然语言理解研究: [博士学位论文]. 北京: 清华大学计算机科学与技术系, 2002

致 谢

衷心感谢导师吴文虎教授和副导师郑方副教授。他们引领我进入学术研究的殿堂，不断给予我精心的指导，并以严谨求实的治学作风影响着我。他们的言传身教将使我终生受益。

特别感谢课题组中黄寅飞和燕鹏举两位师兄。在与他们合作研究的过程中，我学习到了不少知识，留下了快乐的回忆。

感谢智能技术与系统国家重点实验室语音技术中心全体老师和同学们对我的热情帮助和支持！感谢父母多年来对我的不断鼓励和鞭策。

声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名：_____日 期：_____

附录一 领域关键词配置文件

本附录给出绿色食品信息查询系统中用于定制自然语言理解的关键词配置文件。其中 mat_brand、mat_province、mat_city、mat_company、mat_product1 和 mat_product2 几类因所含关键词较多，只列出了一部分。

为了同时支持对文本输入和语音识别词网格输入，文件格式规定每一类中的关键词在文本或无调拼音上不能相同，因此有些相同语义的关键词实际分成不同的关键词类，如 mat_product1 和 mat_product2。

每一类关键词以[关键词类名]开始，后跟语义解释方式（Null、Copy 或 Custom），每一个关键词给出文本和拼音串（为节省篇幅，下面以双栏方式给出文件内容）。

[ato_0_9]Copy	馨禾 -> xin1 he2
8 -> ba1	园艺 -> yuan2 yi4
2 -> er4	
9 -> jiu3	[mat_brand3]Custom
0 -> ling2	伊娜可 -> yi1 na4 ke3
6 -> liu4	MUTUINACO -> e_m0 e_u0 e_t0 e_u0 e_i0
7 -> qi1	e_n0 e_a0 e_c0 e_o0
3 -> san1	INACO -> e_i0 e_n0 e_a0 e_c0 e_o0
4 -> si4	
5 -> wu3	[ato_brand]Null
1 -> yi1	牌 -> pai2
	牌子 -> pai2 zi3
[ato_a]Copy	
A -> e_a0	[mat_province]Copy
	安徽 -> an1 hui1
[ato_lb]Copy	北京 -> bei3 jing1
LB -> e_l0 e_b0	重庆 -> chong2 qing4

[mat_brand1]Copy	
阿珍 -> a1 zhen1	[mat_province2]Copy
艾丽曼 -> ai4 li4 man4	陕西 -> shan3 xi1
爱绿森 -> ai4 lv4 sen1	
.....	
	[ato_province]Null
[mat_brand2]Copy	省 -> sheng3
冬梅 -> dong1 mei2	自治区 -> zi4 zhi4 qu1
绿裕 -> lv4 yu4	市 -> shi4
三鹿 -> san1 lu4	

附录

[mat_city]Copy 鞍山 -> an1 shan1 巴东 -> ba1 dong1 白城 -> bai2 cheng2	[tag_exist]Null 有 -> you3 [tag_not_exist]Null 没有 -> mei2 you3
[ato_city]Null 盟 -> meng2 市 -> shi4 县 -> xian4	[tag_exist_q]Null 有吗 -> you3 ma0 [tag_exist_or_not]Null 有没有 -> you3 mei2 you3 好不好 -> hao3 bu4 hao3
[mat_company]Copy 阿坝藏族羌族自治州红原县奶粉厂 -> a1 ba4 zang4 zu2 qiang1 zu2 zi4 zhi4 zhou1 hong2 yuan2 xian4 nai3 fen3 chang3 阿城金都绿色食品厂 -> a1 cheng2 jin1 dou1 lv4 se4 shi2 pin3 chang3 阿珍名特绿色食品有限公司 -> a1 zhen1 ming2 te4 lv4 se4 shi2 pin3 you3 xian4 gong1 sil	[tag_is]Null 是 -> shi4 属于 -> shu3 yu2 [tag_isnot]Null 不是 -> bu4 shi4 [tag_is_or_not]Null 是不是 -> shi4 bu4 shi4
[mat_product1]Copy 阿魏菇 -> a1 wei4 gu1 阿魏菇罐头 -> a1 wei4 gu1 guan4 tou0 阿珍小米 -> a1 zhen1 xiao3 mi3	[tag_is_q]Null 是吗 -> shi4 ma0 [tag_what]Null 哪 -> nei3 哪个 -> nei3 ge4 哪些 -> nei3 xie1 什么 -> shen2 me0
[mat_product2]Copy 白瓜子 -> bai2 gua1 zi3 菜子烹调油菜子油 -> cai4 zi3 peng1 tiao2 you2 cai4 zi3 you2 菜子色拉油 -> cai4 zi3 se4 la1 you2	[tag_at]Null 在 -> zai4
[ato_chanpin]Null 产品 -> chan3 pin3 类产品 -> lei4 chan3 pin3	[tag_where]Null 哪儿 -> na3 er2 哪里 -> na3 li3 哪的 -> na3 de0 哪有 -> na3 you3
[tag_may_i_ask]Null 查一下 -> cha2 yi1 xia4 告诉我 -> gao4 su4 wo3 请问 -> qing3 wen4 问一下 -> wen4 yi1 xia4 想问 -> wo3 xiang3 wen4 想知道 -> wo3 xiang3 zhi1 dao4 我想找 -> wo3 xiang3 zhao3	[tag_how_many]Null 多少 -> duo1 shao3 几 -> ji3 [tag_question_mark]Null 吗 -> ma0

附录

[tag_sheng]Null 省 -> sheng3	卖 -> mai4 [tag_chanpin]Null 产品 -> chan3 pin3 商品 -> shang1 pin3
[tag_shi]Null 市 -> shi4 地方 -> di4 fang1	[tag_lb]Null 标志号 -> biao1 zhi4 hao4 编号 -> bian1 hao4 标号 -> biao1 hao4
[tag_qiye]Null 公司 -> gong1 si1 厂 -> chang3 厂家 -> chang3 jia1 单位 -> dan1 wei4 企业 -> qi3 ye4 生产厂家 -> sheng1 chan3 chang3 jia1 生产商 -> sheng1 chan3 shang1 厂商 -> chang3 shang1 商家 -> shang1 jia1	[tag_chandi]Null 产地 -> chan3 di4 [tag_pai]Null 牌 -> pai2 品牌 -> pin3 pai2 牌子 -> pai2 zi0
[tag_lvsechanpin]Null 绿色产品 -> lv4 se4 shan3 pin3 绿色食品 -> lv4 se4 shi2 pin3	[tag_de]Null 的 -> do0 [tag_lvse]NULL 绿色 -> lv4 se4
[tag_produce]Null 出 -> chu1 产 -> chan3 生产 -> sheng1 chan3 出产 -> chu1 chan3 有 -> you3 做 -> zuo4	[tag_the]Null 这个 -> zhe4 ge4 那个 -> na4 ge4 这 -> zhe4 那 -> na4

附录二 领域文法配置文件

本附录给出绿色食品信息查询系统中用于定制自然语言理解的领域文法配置文件。

文件格式如下：先说明哪些优先级的规则为词法分析的，然后列出所有规则。每条规则的格式为：

左端非终结符 优先级 规则类型 -> 右端成分列表

其中右端成分列表可包含任意项的终结符或非终结符（关键词类）。优先级为从 0（最高）到 9（最低）的整数，可以缺省，默认为 5。规则类型如下表：

表：各类规则的符号和意义

规则类型	符号	意义
苛刻型	*	传统意义下的规则
跳跃型	无	该规则各右项间可跳过一定数目的垃圾词
长程型	~	该规则各右项间距离可以任意长
无序型	@	该规则各右项不考虑顺序，且右项不能有重复
交叉型	#	1、该规则各右项间所覆盖成分可以是重叠的 2、该规则各右项不考虑顺序，且右项不能有重复

文法中不允许出现无意义的无穷递归规则，例如 `name -> name`。交叉型规则会在同优先级规则处理完之后再处理，所以应适当安排交叉型规则与其他规则的优先级。

文件中可以加注释，注释格式与 C++ 语言相同。

[Lexical Analysis]

0

[Rules]

// 标志号

lb_number1 0*-> ato_lb ato_0_9 ato_0_9 ato_0_9 ato_0_9 ato_0_9 ato_0_9
ato_0_9 ato_0_9 ato_0_9 ato_0_9 ato_0_9 ato_0_9

lb_number 0*-> lb_number1

lb_number 0*-> lb_number1 ato_a

lb_number 0*-> lb_number1 ato_a ato_a

lb_number 0*-> tag_lb lb_number1

// 品牌

brand1 1-> mat_brand1

brand1 1-> mat_brand2

brand1 1-> mat_brand3

brand 2-> brand1

brand 2-> brand1 tag_the ato_brand

brand 2*-> brand1 ato_brand

// 省

province 2-> mat_province

province 2-> mat_province2

province 2*-> mat_province ato_province

// 市

city 2-> mat_city

city 2*-> mat_city ato_city

// 企业

company 2-> mat_company

// 产品

product1 1-> mat_product1

// 具体产品 1

product1 1-> mat_product2

// 具体产品 2

product 2-> product1

product 2-> tag_lvse product1

// 绿色 xx

product 2-> tag_lvsechanpin product1

// 绿色产品 xx

product 2*-> product1 ato_chanpin

// xx(类)产品

// 指某一概念

ref_province -> province

// xx 省

ref_city -> city

// xx 市

ref_cify -> ref_province city

// xx 省 xx 市

ref_company -> company

// xx 公司

ref_company -> ref_province company

// xx 省 xx 公司

ref_company -> ref_city company

// (xx 省)xx 市 xx 公司

附录

ref_brand -> brand // xx 牌
ref_brand -> ref_province brand // xx 省 xx 牌
ref_brand -> ref_city brand // (xx 省)xx 市 xx 牌
ref_brand -> ref_company brand // (xx 省)(xx 市)xx 公司 xx 牌

ref_product -> product // xx 产品
ref_product -> ref_province product // xx 省 xx 产品
ref_product *-> ref_province tag_de product // xx 省的 xx 产品
ref_product -> ref_city product // (xx 省)xx 市 xx 产品
ref_product *-> ref_city tag_de product // (xx 省)xx 市的 xx 产品
ref_product -> ref_company product // (xx 省)(xx 市)xx 公司
// xx 产品
ref_product *-> ref_company tag_de product // (xx 省)(xx 市)xx 公司
// xx 产品
ref_product -> ref_brand product // (xx 省)(xx 市)(xx 公司)
// xx 牌 xx 产品

ref_lb -> tag_lb // 标志号
ref_lb -> ref_product tag_lb // (xx 省)(xx 市)(xx 公司)
// (xx 牌)xx 产品标志号
ref_lb -> ref_brand tag_lb // (xx 省)(xx 市)(xx 公司)
// xx 牌标志号

ref_chandi -> tag_chandi // 产地
ref_chandi -> ref_brand tag_chandi // (xx 省)(xx 市)(xx 公司)
// xx 牌的产地
ref_chandi -> ref_product tag_chandi // (xx 省)(xx 市)(xx 公司)
// (xx 牌)xx 产品的产地

// ref_company 指含有具体名称的公司, ref_qiye 指不含具体名称的企业
ref_qiye1 -> tag_qiye // 企业
ref_qiye2 -> ref_brand tag_qiye // (xx 省)(xx 市)(xx 公司)
// xx 牌企业
ref_qiye2 -> ref_product tag_qiye // (xx 省)(xx 市)(xx 公司)
// (xx 牌)xx 产品企业
ref_qiye -> ref_qiye1
ref_qiye -> ref_qiye2
ref_qiye #-> tag_produce ref_qiye2 // (xx 省) 生产 xx 的企业

// ref_product 指含有具体内容的产品, ref_chanpin 指不含具体内容的产品
ref_chanpin1 -> tag_lvsechanpin // 绿色产品
ref_chanpin1 -> tag_chanpin // 产品
ref_chanpin -> ref_province ref_chanpin1 // xx 省的产品
ref_chanpin -> ref_city ref_chanpin1 // (xx 省)xx 市的产品
ref_chanpin -> ref_company ref_chanpin1 // (xx 省)(xx 市)xx 公司
// 的产品

附录

```
ref_chanpin -> ref_brand ref_chanpin1 // (xx 省)(xx 市)(xx 公司)
// xx 牌的产品

// 提问类型
mark_q_is -> tag_is_or_not // 是不是
mark_q_is ~-> tag_is tag_question_mark // 是……吗
mark_q_is -> tag_is_q // 是吗
mark_q_is *-> tag_exist_or_not tag_is // 有没有是
mark_q_exist -> tag_exist_or_not // 有没有
mark_q_exist ~-> tag_exist tag_question_mark // 有……吗
mark_q_exist -> tag_exist_q // 有吗
mark_q_what -> tag_what // 什么
mark_q_what *-> tag_is tag_what // 是什么
mark_q_what *-> tag_exist tag_what // 有什么
mark_q_how_many -> tag_how_many // 多少
mark_q_how_many *-> tag_is tag_how_many // 是多少
mark_q_how_many *-> tag_exist tag_how_many // 有多少
mark_q_tell_me -> tag_may_i_ask // 告诉我
mark_q_where -> mark_q_is tag_where // 是不是哪儿
mark_q_where -> tag_where // 哪儿
mark_q_where -> tag_at tag_where // 在哪儿
mark_q_where -> tag_is tag_where // 是哪儿
mark_q_where -> tag_exist tag_where // 有哪儿
mark_q_where ~-> tag_at tag_question_mark // 在……吗

// 特殊疑问句
// 问哪个省
ask_sheng 8-> mark_q_what tag_sheng // 哪个省
query_sheng 8-> ask_sheng
query_sheng 8-> ref_brand ask_sheng // xx 牌是哪个省的
query_sheng 8-> ref_brand ask_sheng tag_produce // xx 牌是哪个省
// 生产的
query_sheng 8-> ask_sheng tag_produce ref_brand // 哪个省生产了
// xx 牌
query_sheng 8-> tag_produce ref_brand ask_sheng // 生产 xx 牌的是
// 哪个省
query_sheng 8-> ref_product ask_sheng // xx 产品是哪个省的
query_sheng 8-> ref_product ask_sheng tag_produce // xx 产品是哪个省
// 生产的
query_sheng 8-> ask_sheng tag_produce ref_product // 哪个省生产了
// xx 产品
query_sheng 8-> tag_produce ref_product ask_sheng // 生产了 xx 产品的
// 是哪个省
query_sheng 8-> ref_chandi ask_sheng // xx 牌/xx 产品的
// 产地是哪个省
```

// 问哪个市	
ask_shi 8-> mark_q_what tag_shi	// 哪个市
ask_shi 8-> tag_is ref_province tag_what tag_shi	// xx 省哪个市
query_shi 8-> ask_shi	
query_shi 8-> ref_brand ask_shi	// xx 牌是哪个市的
query_shi 8-> ref_brand ask_shi tag_produce	// xx 牌是哪个市生产的
query_shi 8-> ask_shi tag_produce ref_brand	// 哪个市生产了 xx 牌
query_shi 8-> tag_produce ref_brand ask_shi	// 生产了 xx 牌的是
	// 哪个市
query_shi 8-> ref_product ask_shi	// xx 产品是哪个市的
query_shi 8-> ref_product ask_shi tag_produce	// xx 产品是哪个市
	// 生产的
query_shi 8-> ask_shi tag_produce ref_product	// 哪个市生产了 xx 产品
query_shi 8-> tag_produce ref_product ask_shi	// 生产了 xx 产品的是
	// 哪个市
query_shi 8-> ref_chandi ask_shi	// xx 牌/xx 产品的产地
	// 是哪个市
query_shi 8#-> mark_q_exist ref_company	// 有 xx 企业吗
// 产地是哪里	
query_chandi 8-> ref_chandi mark_q_where	// xx 牌/xx 产品的产地
	// 是哪里
query_chandi 8-> mark_q_where tag_produce	// 哪里产的
query_chandi 8-> ref_product mark_q_where ref_chanpin1	// xx 产品是哪里的产品
query_chandi 8-> ref_brand mark_q_where tag_de ato_brand	// xx 牌是哪里的牌子
query_chandi 8-> ref_brand mark_q_where	// xx 牌是哪里的
query_chandi 8-> ref_product mark_q_where	// xx 产品哪儿有
// 问哪个企业	
ask_qiye 8-> mark_q_what tag_qiye	// 哪个企业
ask_qiye 8-> tag_is ref_province tag_what tag_qiye	// xx 省哪个企业
ask_qiye 8-> tag_is ref_city tag_what tag_qiye	// (xx 省)xx 市哪个企业
query_qiye 8-> ask_qiye	
query_qiye 8-> ref_brand ask_qiye	// xx 牌是哪个企业的
query_qiye 8-> ref_brand ask_qiye tag_produce	// xx 牌是哪个企业
	// 生产的
query_qiye 8-> ask_qiye tag_produce ref_brand	// 哪个企业生产了 xx 牌
query_qiye 8-> tag_produce ref_brand ask_qiye	// 生产了 xx 牌的是
	// 哪个企业
query_qiye 8-> ref_product ask_qiye	// xx 产品是哪个企业的
query_qiye 8-> ref_product ask_qiye tag_produce	// xx 产品是哪个企业
	// 生产的
query_qiye 8-> ask_qiye tag_produce ref_product	// 哪个企业生产了
	// xx 产品
query_qiye 8-> tag_produce ref_product ask_qiye	// 生产了 xx 产品的是
	// 哪个企业

附录

```
// 问品牌(问有什么, 问有多少)
ask_pinpai 8-> mark_q_what tag_pai // 哪些牌子
ask_pinpai 8-> mark_q_what tag_pai tag_de ref_chanpin1 // 哪些牌子的产品
ask_pinpai 8-> mark_q_what tag_pai tag_de product1 // 哪些牌子的 xx 产品
ask_pinpai 8-> mark_q_how_many tag_pai // 多少牌子
ask_pinpai 8-> mark_q_how_many tag_pai tag_de // 多少牌子的产品
ref_chanpin1
ask_pinpai 8-> mark_q_how_many tag_pai tag_de product1 // 多少牌子的 xx 产品
query_pinpai 8-> ask_pinpai
query_pinpai 8-> ask_pinpai tag_is ref_chanpin1 // .....是绿色产品?
query_pinpai 8-> query_pinpai
query_pinpai 8-> ref_product query_pinpai // xx 产品哪些牌子的
query_pinpai 8-> ref_province query_pinpai // xx 省有哪些牌子
query_pinpai 8-> ref_city query_pinpai // (xx 省)xx 市有哪些
// 牌子
query_pinpai 8-> ref_company query_pinpai // (xx 省)(xx 市)xx 企业
// 有哪些牌子

// 问产品
ask_chanpin1 *-> tag_produce tag_what // 生产什么
ask_chanpin1 8-> mark_q_what ref_chanpin1 // 哪些产品
ask_chanpin1 8-> mark_q_exist ref_product // 有没有 xx 产品
ask_chanpin1 8-> mark_q_what ref_product // 哪些 xx 产品
ask_chanpin1 8-> mark_q_tell_me ref_product // 我要找 xx 产品
ask_chanpin1 8-> mark_q_how_many ref_chanpin1 // 多少产品
query_chanpin1 8-> ask_chanpin1
query_chanpin1 8-> ask_chanpin1 tag_is ref_chanpin1 // 哪些产品是绿色产品
query_chanpin 8-> query_chanpin1
query_chanpin 8-> ref_province query_chanpin1 // xx 省有哪些产品
query_chanpin 8-> ref_city query_chanpin1 // (xx 省)xx 市有哪些
// 产品
query_chanpin 8-> ref_company query_chanpin1 // (xx 省)(xx 市)xx 企业
// 有哪些产品
query_chanpin 8-> ref_brand query_chanpin1 // (xx 省)(xx 市)(xx 企业)
// xx 牌有哪些产品
query_chanpin 8-> lb_number mark_q_what // xx 标志号是什么
query_chanpin 8#-> lb_number mark_q_exist // xx 标志号有吗
query_chanpin 8#-> ref_brand mark_q_exist // xx 牌有吗?
query_chanpin 8-> ref_product query_chanpin1 // xx 产品有什么

// 问标志号
query_lb 8-> ref_lb // .....标志号
query_lb 8-> ref_lb mark_q_what // .....标志号是什么
query_lb 8-> ref_lb mark_q_how_many // .....标志号是多少
query_lb 8-> ref_chanpin ref_lb // 产品编号
query_lb 8@-> lb_number ref_company // xx 标志号是 xx 公司的
```

附录

```
// 一般疑问句
query_is_lb1 #-> mark_q_is lb_number // 是 LB-xx-xxxxxxx 吗
query_is_lb 8@-> ref_lb query_is_lb1 // .....标志号是 xxxx 吗
query_is_lb 8@-> ref_brand query_is_lb1 // xx 牌是 xxxx 吗

query_is_lvse1 #-> mark_q_is ref_chanpin1 // 是绿色产品吗
query_is_lvse 8-> ref_product query_is_lvse1 // xx 产品是绿色产品吗
query_is_lvse 8-> ref_brand query_is_lvse1 // xx 牌是绿色产品吗
query_is_lvse 8-> ref_chanpin query_is_lvse1 // xx 产品是绿色产品吗

query_exist_lvse1 #-> mark_q_exist ref_chanpin1 // 有绿色产品吗
query_exist_lvse1 #-> mark_q_exist ref_product ref_chanpin1 // 有 xx 绿色产品吗
query_exist_lvse 8-> ref_chanpin // 辅助下面三条规则
query_exist_lvse 8-> query_exist_lvse1 // 有绿色的吗
query_exist_lvse 8-> ref_province query_exist_lvse1 // xx 省有绿色产品吗
query_exist_lvse 8-> ref_city query_exist_lvse1 // (xx 省)xx 市有
// 绿色产品吗

query_exist_lvse 8-> ref_company query_exist_lvse1 // (xx 省)(xx 市)xx 企业
// 有绿色产品吗

query_exist_product 8@-> ref_product ref_city mark_q_exist // xx 产品 xx 市有吗
query_exist_product 8@-> ref_product ref_province // xx 产品 xx 省有吗
mark_q_exist
query_exist_product 8#-> ref_product mark_q_exist // xx 产品有吗

query_exist_qiye 8@-> ref_qiye ref_city mark_q_exist // xx 企业 xx 市有吗
query_exist_qiye 8@-> ref_qiye ref_province mark_q_exist // xx 企业 xx 省有吗

query_is_sheng1 -> ref_province tag_de // xx 省的
query_is_sheng2 -> ref_province tag_de tag_chanpin // xx 省的产品
query_is_sheng2 -> ref_province tag_de tag_lvsechanpin // xx 省的绿色产品

query_is_sheng3 #-> mark_q_is query_is_sheng1 // 是 xx 省的吗
query_is_sheng4 #-> mark_q_is query_is_sheng2 // 是 xx 省的绿色产品吗
query_is_sheng5 8-> query_is_sheng3
query_is_sheng5 8-> query_is_sheng4
query_is_sheng 8-> query_is_sheng5
query_is_sheng 8-> ref_product query_is_sheng5 // xx 产品是 xx 省的
// 绿色产品吗

query_is_sheng 8-> ref_brand query_is_sheng5 // xx 牌是 xx 省的
// 绿色产品吗

query_is_sheng 8-> city query_is_sheng3 // xx 市是 xx 省的吗
query_is_sheng 8-> ref_company query_is_sheng3 // (xx 市)xx 企业是
// xx 省的吗

query_is_shi1 -> ref_city tag_de // xx 市的
query_is_shi2 -> ref_city tag_de tag_chanpin // xx 市的产品
```

附录

query_is_shi2 -> ref_city tag_de tag_lvsechanpin	// xx 市的绿色产品
query_is_shi3 #-> mark_q_is query_is_shi1	// 是 xx 市的吗
query_is_shi3 #-> mark_q_where ref_city	// 在 xx 市吗
query_is_shi4 #-> mark_q_is query_is_shi2	// 是 xx 市的绿色产品吗
query_is_shi5 8-> query_is_shi3	
query_is_shi5 8-> query_is_shi4	
query_is_shi 8-> query_is_shi5	
query_is_shi 8-> ref_product query_is_shi5	// xx 产品是 xx 市的 // 绿色产品吗
query_is_shi 8-> ref_brand query_is_shi5	// xx 牌是 xx 市的 // 绿色产品吗
query_is_shi 8-> ref_company query_is_shi3	// xx 企业是 xx 市的吗
query_is_company1 -> ref_company tag_de	// xx 企业的
query_is_company1 -> ref_company tag_produce tag_de	// xx 企业生产的
query_is_company2 -> query_is_company1 tag_chanpin	// xx 企业的产品
query_is_company2 -> query_is_company1 tag_lvsechanpin	// xx 企业的产品
query_is_company3 #-> mark_q_is query_is_company2	// 是 xx 企业的产品吗
query_is_company 8-> ref_brand query_is_company3	// xx 牌是 xx 企业的 // 产品吗
query_is_company 8-> ref_product query_is_company3	// xx 产品是 xx 企业的 // 产品吗
query_is_company 8-> query_is_company4	
query_is_company4 -> ref_company tag_produce product	// xx 企业生产 xx 产品吗
// 底限句子	
query_all 8#-> ref_product	
query_all 8#-> ref_company	
query_all 8#-> lb_number	
// 最高级句型	
query 9-> query_sheng	// 问省
query 9-> query_shi	// 问市
query 9-> query_chandi	// 问产地
query 9-> query_qiye	// 问企业
query 9-> query_pinpai	// 问品牌
query 9-> query_chanpin	// 问产品
query 9-> query_lb	// 问标志号
query 9-> query_is_lb	// 问标志号是不是
query 9-> query_is_lvse	// 问是不是绿色的
query 9-> query_exist_lvse	// 问有没有绿色的
query 9-> query_is_sheng	// 问是不是 xx 省
query 9-> query_is_shi	// 问是不是 xx 市
query 9-> query_is_company	// 问是不是 xx 企业
query 9-> query_exist_product	// 问有没有产品
query 9-> query_exist_qiye	// 问有没有企业
query 9-> query_all	// 底限句子

附录三 主题森林配置文件

本附录给出绿色食品信息查询系统中用于定制对话管理的主题森林配置文件。

文件以前序遍历方式依次给出主题森林中各结点的内容。每个结点都以该结点所属的结点类型开始，每类结点都有一些固定的属性。

文件中可以加注释，注释格式与 C++ 语言相同。

```
[TopicForest]
NumberOfTrees=1 // 只含一个主题

[TopicNode] // 主题结点
Name=GreenFood
FunctionIndex=0
TopicType=2 // 查询型主题
DatabaseTable=GreenFood // 数据库表名称
QueryFailedFunctionIndex=1
AnswerTopicFunctionIndex=2
PrimaryTree=Yes // 只含有主属性分支
SecondaryTree=No
AdditionalTree=No

[PrimaryTree] // 主属性分支
Name=
FunctionIndex=3
Relation=OR // “或”关系
NumberOfSons=6 // 含 6 项信息

[LeafNode] // 叶子结点，省市自治区
Name=Province
FunctionIndex=4
SlotName=Province // 语义槽信息
SlotType=0
ColumnName=Sheng // 数据库表中字段名称
DataType=STRING // 数据类型为字符串型
Default=No // 无默认值

[LeafNode] // 叶子结点，市县
Name=City
FunctionIndex=5
SlotName=City // 语义槽信息
```

SlotType=0
ColumnName=Shi // 数据库表中字段名称
DataType=STRING // 数据类型为字符串型
Default=No // 无默认值

[LeafNode] // 叶子结点, 企业
Name=Company
FunctionIndex=6
SlotName=Company // 语义槽信息
SlotType=0
ColumnName=QiYe // 数据库表中字段名称
DataType=STRING // 数据类型为字符串型
Default=No // 无默认值

[LeafNode] // 叶子结点, 产品名称
Name=Product
FunctionIndex=7
SlotName=Product // 语义槽信息
SlotType=0
ColumnName=ChanPin // 数据库表中字段名称
DataType=STRING // 数据类型为字符串型
Default=No // 无默认值

[LeafNode] // 叶子结点, 品牌
Name=Brand
FunctionIndex=8
SlotName=Brand // 语义槽信息
SlotType=0
ColumnName=PinPai // 数据库表中字段名称
DataType=STRING // 数据类型为字符串型
Default=No // 无默认值

[LeafNode] // 叶子结点, 标志号
Name=Label
FunctionIndex=9
SlotName=Label // 语义槽信息
SlotType=0
ColumnName=LB // 数据库表中字段名称
DataType=STRING // 数据类型为字符串型
Default=No // 无默认值

个人简历、在学期间的研究成果及发表的论文

个人简历

1976年9月21日出生于江苏省苏州市，1995年9月免试保送清华大学计算机科学与技术系，1999年7月本科毕业并获得工学学士学位，同年9月免试保送清华大学计算机科学与技术系攻读博士至今。

已发表论文

- [1] Wu X.J., Zheng F., Xu M.X. Topic Forest: A Plan-based Dialog Management Structure. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2001). Salt Lake City, USA, 2001. v1, 617~620
- [2] 邬晓钧, 郑方, 徐明星. 基于主题森林结构的对话管理模型. 自动化学报, 2003. 29(2):275~283
- [3] 邬晓钧, 燕鹏举, 徐明星. 口语对话系统中日期和时刻的处理. 见: 第六届全国人机语音通讯学术会议论文集. 深圳, 2001. 107~110
- [4] Wu X.J., Xu M.X, Wu W.H. Preparing for Evaluation of a Flight Spoken Dialogue System. In: Proceedings of the 3rd International Symposium on Chinese Spoken Language Processing (ISCSLP2002). Taipei, China, 2002. 315~317
- [5] Wu X.J., Zheng F., Wu W.H. A Hybrid Dialogue Management Approach for a Flight Spoken Dialogue System. In: Proceedings of the 1st International Conference on Machine Learning and Cybernetics (ICMLC2002). Beijing, China, 2002. 824~829