

# 基于主题森林结构的对话管理模型<sup>1)</sup>

邬晓钧 郑方 徐明星

(清华大学智能技术与系统国家重点实验室 语音技术中心 北京 100084)

(E-mail: [wuxj, fzheng, xumx]@sp. cs. tsinghua. edu. cn)

**摘 要** 现有许多对话系统都是面向任务的,但很少考虑对话过程涉及的多主题、主题切换、主题间的信息共享,以及不同信息重要程度不同等问题. 该文提出基于主题森林结构的对话管理模型,较好地处理了上述问题,能实现对话过程的混合主导. 该模型能保证交互过程中对话内容的一致性,对上下文的理解、对话控制和应答决策灵活自然. 其推理机应用了多种与领域无关的策略,具有较好的可移植性.

**关键词** 对话系统,对话模型,对话管理,主题森林

**中图分类号** TP319.42

## Topic Forest Based Dialog Management Model

WU Xiao-Jun ZHENG Fang XU Ming-Xing

(Center of Speech Technology, State Key Laboratory of Intelligent Technology and Systems,  
Tsinghua University, Beijing 100084)

(E-mail: [wuxj, fzheng, xumx]@sp. cs. tsinghua. edu. cn)

**Abstract** There are many task-oriented dialog systems, but few have taken into account the issues on multiple topics, topic shifts, information sharing among topics, and difference in importance for information items. A dialog management model based on the topic forest structure is presented here, which can well handle those issues with mixed-initiative dialog control. This model is consistent during the dialog transaction, so that it can understand the user according to the context. Its dialog control and response decision are both natural and flexible. With several domain-independent strategies, its reasoning engine is portable too.

**Key words** Dialog systems, dialog models, dialog management, topic forest

## 1 引言

人机口语对话技术属于综合应用语音技术和语言理解技术领域. 目前一个人机口语对

1) 国家自然科学基金(19871045)和清华基础研究基金(JC1999042)资助

Supported by the National Natural Science Foundation of P. R. China(19871045) and the Basic Research Foundation of Tsinghua University(JC1999042)

收稿日期 2001-06-08 收修改稿日期 2001-12-20

Received June 8, 2001; in revised form December 20, 2001

话系统通常是面向任务的,它通过不断的对话,理解用户的意图,并最终提供用户所希望的服务.对话管理是对话系统的核心,它根据前级语言理解部分分析得到的用户需求,更新或查询对话系统内部维护的知识库状态,然后根据对话进展情况按一定策略生成应答.对话模型是对话管理的理论基础,形式化地描述了语义和语用,以及两者之间的相互关系.对话模型主要有语法模型和规划模型两种.语法模型把对话看作一个状态机,状态有先后次序和限制条件,用类似语法描述的方法来表示.规划模型假设人们在和别人或计算机系统对话时,头脑中早有一些目标和规划,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,因此对话由规划和规划识别构成.

对话管理模型是对话模型理论的具体实现.相应于两种对话模型,目前很多实用的口语对话系统采用的对话管理模型也主要有两种,即有限状态模型和自组织模型.在有限状态模型中,系统事先定义好所有可能的对话状态和状态间的转移,组成有限状态网络,这些将决定所有合法的对话过程<sup>[1]</sup>.由于对话路径事先确定,当用户提供的信息比要求的多时,系统将不能处理;如果考虑任一情况下用户所有可能的回答,则状态数激增,给对话管理带来很大负担;如果任务的各个参数之间有复杂的关系,需要用户和系统一起协商,则模型不能适用.在自组织模型中,对话过程事先并不确定,完全依据当前对话状态的动态发展而变化,这避免了有限状态模型中事先定义状态和转移的工作,便于处理复杂的任务,对话控制较为灵活,易于实现对话的人-机混合主导<sup>[2~5]</sup>.自组织模型将特定的对话任务划分为多个小目标,控制对话过程逐步实现它们,并最终完成整个对话任务.这个模型表达对话状态的能力强,但需要寻求一个好的数据结构以及相应的对话管理策略.

对话系统的评价目前并不规范,客观标准和主观标准都有.对话管理模型的评价则主要从一致性、灵活性和可移植性三个方面来进行,即能否保证对话上下文中主题内容的一致性和对话多个主题之间内容的一致性;基于对话上下文的理解、应答和对话控制是否灵活;是否便于领域任务的移植.

当前对话系统研究开发的重点,从名址录、日程安排等简单对话,逐渐转向票务服务、自动售货等更为复杂的对话.在这些复杂项目中,对话涉及的内容信息项较多,而且信息项之间有复杂的逻辑关系.在对话服务过程中,用户的要求可能从不明确到明确,可能根据实际情况多次更改,而且由于现有语音识别和自然语言理解技术的局限,有时不能正确把握用户语义.同时,我们希望在多回合的对话中,系统应答尽可能自然友好.这样的对话系统需要有一个强大的对话管理模块来支持.

本文针对当前对话系统的要求,根据对话的规划模型,提出了基于主题森林结构的自组织对话管理模型.该模型允许用户在对话涉及的多个主题间自由切换,并能对单个复杂主题的多种信息不断适应用户的要求.

## 2 主题森林

对话任务涉及的一些基本内容称为信息,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.现有的对话系统大多局限于单一主题,然而实际情形要复杂得多.例如,一个航班订票的对话可能会涉及到航班信息、天气状况、地区时差、订票改票与退票、送票服务等许多主题;不同的主题可能共享某些信息,引起对话主题切换时的省



略现象.此外,一个复杂主题可能包含多项信息且重要程度不同,用户关心的内容也不同.例如航班订票中,日期、时间、起止城市是大家都关心的必要信息,而航空公司、飞机机型则不是,只是主题中的可选信息.

这里涉及到多主题、主题切换、主题间的信息共享,以及信息的不同重要程度等问题,目前的许多对话系统没有考虑这些因素.文献[2~4]都只处理单个主题,其中文献[3,4]只处理主题的必要信息,而文献[2]对下面两种不同的情形:用户已确定 A 信息正与系统协商 B 信息和用户已确定 B 信息正与系统协商 A 信息,采取同一个对话动作.文献[5]介绍了一种处理多主题的对话模型,可允许主题切换,但没有考虑主题间的信息共享.本文提出的对话模型,可以有效地解决上述问题.下面以航班订票对话系统 EasyFlight 为例,说明这一对话管理模型的主要数据结构——主题森林结构,以及运行时的多种策略.图 1 是 EasyFlight 系统中主题森林的一部分,只列出了“航班信息主题”、“航空公司信息主题”和“时差信息主题”三个主题.

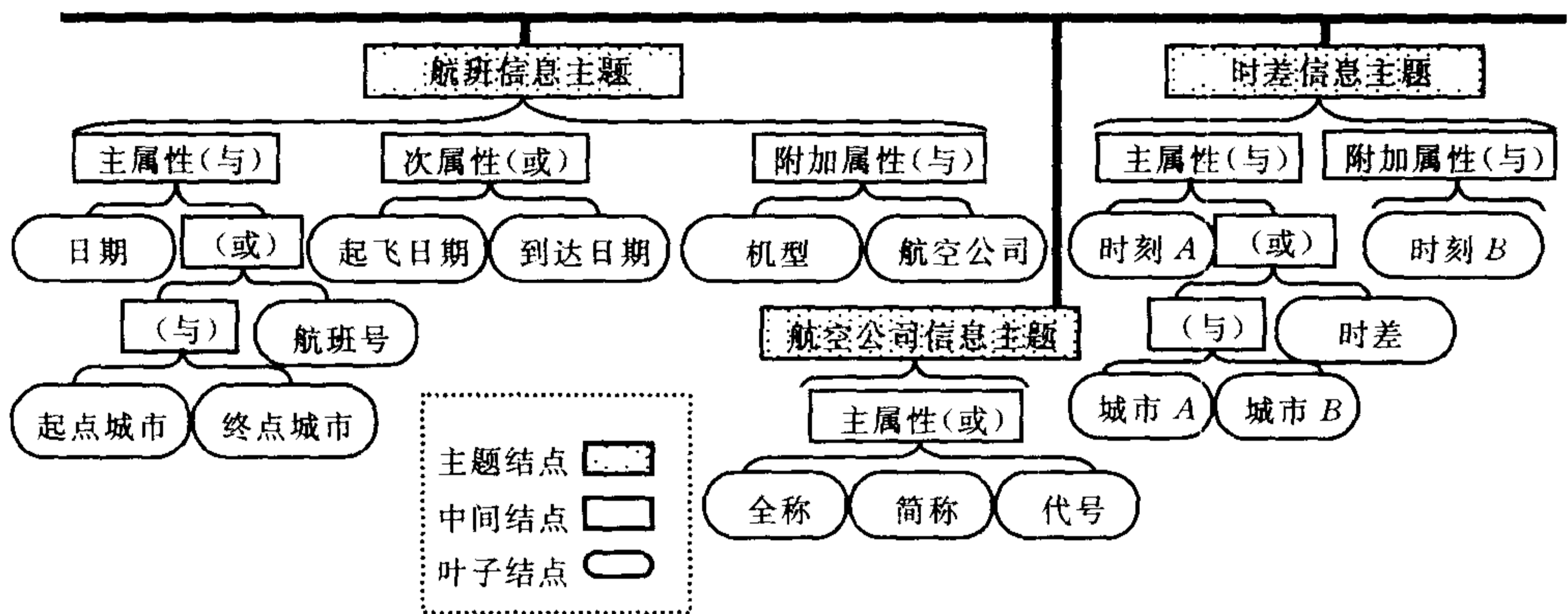


图 1 部分主题森林  
Fig.1 Part of topic forest

### 2.1 主题树结构

主题树是用以描述单个主题所包含的各项信息及信息之间的关系,记录对话过程中主题信息的知识状态的树结构.

主题树有三类结点:主题结点、中间结点和叶子结点.1)每个主题结点都是一棵主题树的根,标明该主题的类型和相关知识库.2)中间结点给出它的子结点间的逻辑关系,包括“与”和“或”.主题树中有三类特殊的中间结点,即主属性结点、次属性结点和附加属性结点.顾名思义,主属性结点下存放该主题最关键的信息,它们的取值情况决定了系统是否查询知识库;次属性结点下存放对主题的最后确定起重要作用的细节信息,在对话过程中也是不可缺少的交互内容,系统将会主动与用户讨论并确定这些信息项,从而完全确定本主题内容;附加属性结点下存放主题的其他信息,如果用户不提及该信息,系统也不会主动提及.按确定主题的重要程度将信息项分类,系统就能区别对待它们并且结合用户兴趣控制对话过程.3)主题树的叶子结点用来存储主题的某项信息,与某一特定语义相联系,其取值状态就是该信息项的知识状态.一棵主题树所有叶子结点状态的总和,就表示了一个对话主题的知识状态.

每一结点都有一个二值的有效状态标志.叶子结点的有效状态标志表明该结点的信息



是否有效,即是否已经得到用户确认;中间节点的有效状态标志取值为所有子节点有效状态标志的逻辑运算值;主题节点只有当该主题所包含的多种信息都唯一时,也就是主题可被完全确定时,才为真.

每一个结点都有一个对话生成函数与之对应.在不同的系统状态下,调用该函数会得到不同的应答输出,例如给出信息项的内容,要求用户提供有关信息,提示用户确认有关信息等.

### 2.2 主题森林和共享信息索引

对某个特定任务的若干主题都建立主题树,就构成了该对话任务的主体森林.事实上,主题森林是主题树的集合,集合的元素,即主题树,可多可少,完全取决于对话任务涉及的内容.在主题树以及主题森林上分别定义有各自不同的操作.

由于每棵主题树都包含了该主题的完整信息,为了处理主题间的共享信息,本文提出共享信息索引结构.共享信息索引是从特定语义到叶子结点的一对多的映射.在主题森林建立后,对话管理的推理机会遍历所有叶子结点,将对应于同一特定语义的叶子结点都记录下来.当对话过程中要获取某主题中某项信息的知识状态时,推理机会根据叶子结点对应的特定语义查找并参考其他主题的共享信息.

共享信息索引表示了不同主题间具有同类信息的关系,将各棵主题树联系在一起,成为一个相互关联的整体.图 2 给出了共享信息索引的示例.

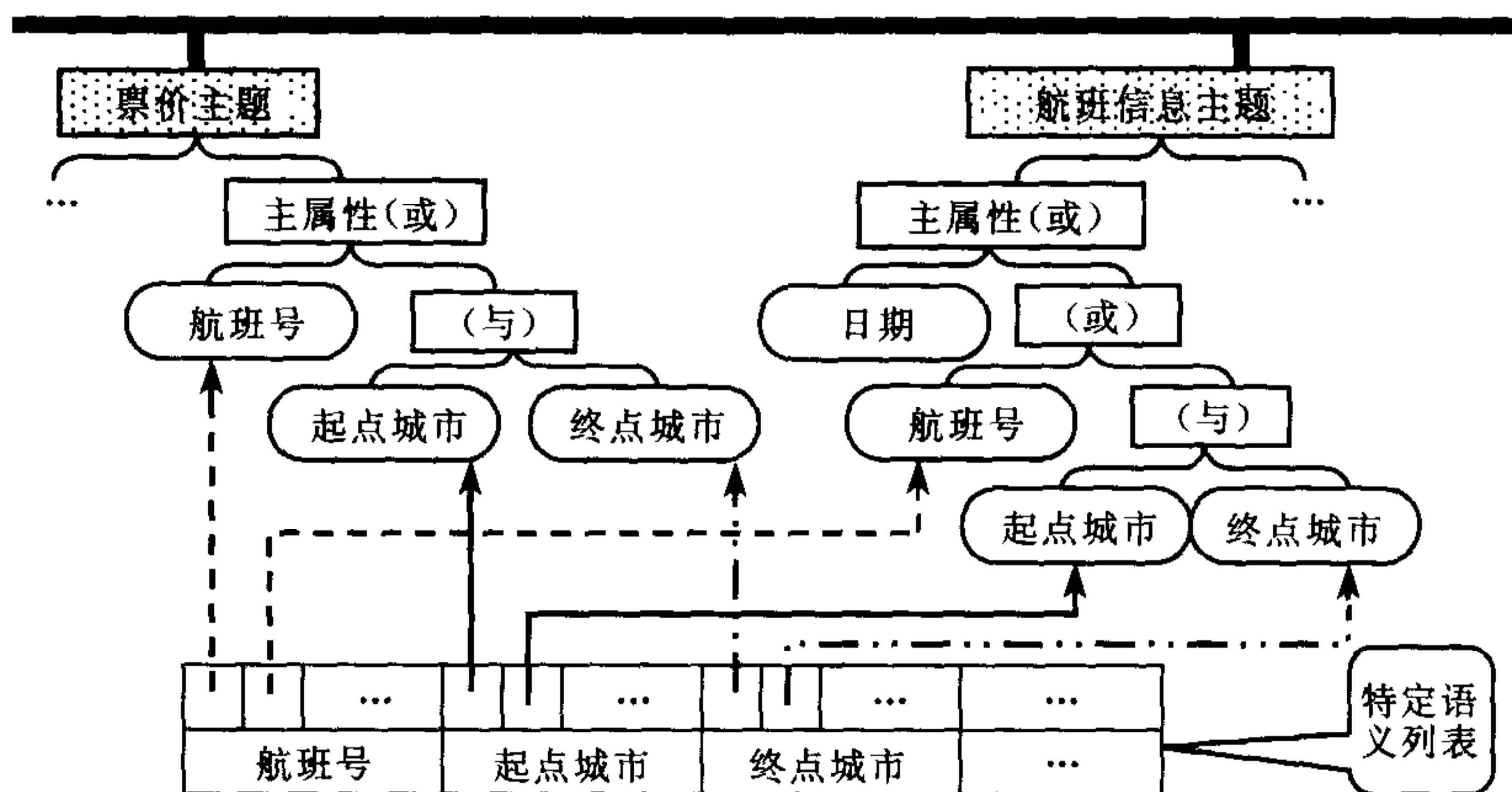


图 2 主题树与共享信息索引

Fig. 2 Topic trees and shared information index

### 2.3 其它数据结构

主题森林结构是该对话管理模型中最主要的数据结构,此外还有一些辅助的结构,用于对话的语义理解、流程控制和应答生成.

直接上下文信息和最近使用信息,分别用于对话过程中一问一答间的省略处理,以及对话多回合间的省略处理.

对话辅助信息包括对话的当前主题、对下回合用户语句所属主题的预测、系统应答标志、知识库查询状态、对话确认状态等信息.

用户模型有两方面,一方面是对主题信息的兴趣模型,已经体现在主题树结构中;另一方面是用户的用语习惯模型,记录下来用于应答的生成,例如用户对于日期的表述(可



能表述为星期几、几月几日或明后天)。

### 3 对话管理

本节介绍对话管理的离线和在线操作,包括主题森林和共享信息索引的建立,以及对话管理所应用的多种策略。

#### 3.1 准备——主题森林及共享信息索引的建立

给定某领域的特定任务,根据所有可能涉及的多项信息间的亲疏关系,可以确定任务包含的多个主题,并为每一主题建立相应的主题树。每棵主题树的建立以该主题包含的各信息项的性质和相互关系为依据。下面以图 1 中“航班信息主题”的建立为例说明。

先建立主题结点,并标明与该主题相关的知识库。然后考虑存放信息项的叶子结点应分别属于主属性结点、次属性结点还是附加属性结点。对于实际的航班信息来说,每个用户都知道起点城市和终点城市,对航班的日期有所要求,这三项信息是航班信息查询的关键,因此属于主属性结点。有的用户会直接要求航班号和日期,因此航班号也可能是确定主题的关键信息,也属于主属性结点。对于这两种情况,日期都是必不可少的,起终点或航班号可以任选,因此用“(与)”结点连接“起点城市”结点和“终点城市”结点,再用“(或)”与“航班号”结点连接,最后与“日期”结点连接在“主属性(与)”下。用户可能不了解确切的起飞时间和到达时间,但是这对于最后确定哪一次航班是重要的,实际对话中系统和用户协商决定的也多是这两项信息,因此这两项信息应归入次属性结点,它们的关系显然可用“或”来表示。与航班信息有关的机型、航空公司不是每个用户都会提及和关心的,它们属于附加属性结点,并用“与”关系连接。接着填写每个叶子结点的内容,标明它们所对应的特定语义(可以有默认值)、对应的知识库中的字段名称、存储的值类型等。最后将每个结点所对应的对话生成函数连接到结点上,就完成了整棵主题树的构建。

同样方法可建立与任务相关的其他主题树,这些主题树的集合就是该对话系统的主题森林。若将该对话模型应用于名址录等简单的对话任务,则主题森林可能退化为只有一棵树的情形。

当领域任务确定后,主题森林的建立只是一个离线操作,完成后可保存供对话系统启动时加载。系统载入主题森林后,遍历所有结点,根据各结点相关的特殊语义,自动生成共享信息索引。这一过程几乎不耗时间。

#### 3.2 运行——对话管理策略

对话管理的推理是指根据当前的对话状态,决定系统应采取的动作。对话系统的推理机与所采用的对话管理模型的结构和表达能力是密不可分的,是整个对话管理的核心。我们设计和实现的推理机以主题森林为基础。

推理机的输入是对用户话语进行语音识别和语义分析的结果,包括语义框架和该语句所属主题。语义框架由一系列语义槽组成,每个语义槽包含某一特定语义,将会被填充到相应的叶子结点中。用户语句的其他非主题信息,也包括在语义框架中。推理机维护了一些系统状态参数和辅助变量,用以表示除主题信息外的其余一些对话状态,记录一些对话的上下文以指导语义分析并帮助推理。

推理机的输出是系统应采取的动作,具体表现为对某个对话生成函数的调用。对话生成函数将会参考推理机维护的系统状态参数和辅助变量,给出合理的应答。对话系统的领域知



识体现在语义框架和语义槽的定义、主题森林的设计、结点的对话生成函数和领域的知识库上,与推理机无关。

### 3.2.1 记忆与遗忘策略

主题的各项信息存储在各个叶子结点中,信息的来源可以是输入的语义框架、其他主题树中的相关叶子结点或知识库查询结果,分别模拟对话中主题信息的三个来源:用户信息、与其他主题的共享信息和自身掌握的信息。

把主题信息项填入到相应主题树的叶子结点中的操作,称之为主题树的填充,分为添加和替换两种。添加操作丰富主题的信息,帮助确定主题;替换操作用新的信息代替以前的信息。添加和替换的策略,就是对话过程中系统记忆和遗忘的策略。在图 3 所示的对话过程中,U2 的意图是想在两个日期中间进行选择,应该采用添加操作;U3 则应采用替换操作。

U1:明天有去上海的飞机吗?	
S1:有.	
U2:后天呢?	(添加操作)
S2:有.	
U3:那我要后天的票.	(替换操作)

图 3 对话示例 1

Fig. 3 Dialogue example

判断应采用添加操作还是遗忘操作,首先要考虑用户的语句是陈述句还是疑问句。陈述句在提供一定信息的同时,含有较强的肯定态度;疑问句提供信息的同时,往往含有待定的意味。其次,应用添加操作还是遗忘操作还和主题的类型有关。在 EasyFlight 系统中,对于时差信息主题,只采用替换操作就可以了,而且替换比添加更合适,因为主题功能本身就是查询,主题信息没有多选的可能;对于航班信息主题,则添加操作和替换操作都是必需的。最后,将知识库查询结果填充到主题树中的时候,一般采用替换操作,用查询结果值(可能多个)替换原有值。

### 3.2.2 用户兴趣的适应策略

决定系统的应答焦点时,推理机根据对话历史和对话状态,考虑各信息项之间的制约关系和信息重要程度。如果仍有多个主题信息项可作为应答焦点,则会优先选择那些曾被用户提及的信息项,因为这些信息项是用户最可能感兴趣的。这一策略是下面的应答推理策略的组成部分。

### 3.2.3 应答推理策略

根据用户输入和其他主题共享信息获取某主题的信息后,推理机进行推理,最后归结到某个结点采取相应动作,调用它的对话生成函数,或调用固定的系统应答函数,生成合理的对话应答输出。应答推理策略如图 4 所示(其中  $Lx$  表示第  $x$  行伪代码),其中  $L3$  寻找主属性结点下一个无效的结点,实际上是寻找该主题关键信息中的未知项; $L11$  中寻找次属性结点下的某结点,应用用户兴趣的适应策略,其目的是主动与用户交流协商有关信息,引导对话的发展; $L12$  再次检查主属性结点下所有结点的信息状态,使得当用户多次改变主题的关键信息时,系统也能提示用户确认最终的信息; $L14$  的操作也应用了用户兴趣的适应策略。

上述的推理策略中, $L8$  和  $L9$  使得系统会优先回答用户提问(这时是用户主导);当系统有必要主导对话进行时,推理机根据主题树的结构,能容易地把握当前主题下各项信息不同重要程度,结合对话状态,决定最合适的应答焦点( $L3, L11, L13$ ),从而实现对话的混合主导。

### 3.2.4 主题切换策略

对话过程中的多个主题既相互联系,又相对独立。当用户确实切换了主题时,用户语句

L1:	将语义框架信息填入主题树
L2:	IF 主属性无效
L3:	THEN 找主属性下某个无效结点作为应答焦点
L4:	ELSE 组织数据库查询
L5:	IF 查询失败
L6:	THEN 应答焦点为主题查询失败
L7:	ELSE 将查询结果信息填入主题树
L8:	IF 本回合用户提问
L9:	THEN 应答焦点为答案所在结点
L10:	ELSE IF 查询结果不唯一
L11:	THEN 找次属性下某结点为应答焦点
L12:	ELSE IF 主属性下所有结点取值都唯一
L13:	THEN 应答焦点为主题结点
L14:	ELSE 主属性下某取值不唯一的结点为应答焦点

图 4 推理策略

Fig. 4 The reasoning strategy

必有明显的语义标志(与领域知识有关),这时自然语言理解模块可以知道切换后的主题.当对话管理从语义框架中得知主题切换,则将所有对话管理策略应用于切换后的主题树,在新的主题中与用户进行交互.如果有必要,系统也可以主动切换主题.这时要通知自然语言生成器生成有关的应答,并修改相应的对话状态,为下一回合的主题处理做好准备.

不管当前对话是什么主题,有何意义,推理机只是基于事先设计的主题森林结构和各结点的信息状态进行推理,而不涉及结点内容实际的语义,因而与领域无关.

## 4 对话举例

图 5 为一个实际的对话例子.

U1: 我要一张去纽约的机票.	(用户主动,确定终点城市)
S1: 您要哪一天的?	(系统主动,提问起飞时间)
U2: 星期三.	(用户应答)
S2: 有好几次航班,您要大概什么时候起飞的?	(系统主动,提问起飞时间)
U3: 10:00 之后吧.	(用户应答)
S3: 您要 10:10 起飞,还是 16:30 起飞的航班呢?	(系统主动,提问起飞时间)
U4: 16:30 起飞的什么时候到?	(用户主动,提问到达时间)
S4: 到达时间为 19:10.	(系统应答)
U5: 那时是北京时间几点?	(用户主动,切换主题,询问时差)
S5: 08:10.	(系统应答)
U6: 那就要 16:30 起飞的吧.	(用户主动,切换主题,确定起飞时间)
S6: 您要 6 月 14 日 16:30 从北京出发,19:10 到达纽约的 CA 981 次航班的机票吗?	(系统主动,请求确认航班信息)
U7: 是的.	(用户应答)

图 5 对话示例 2

Fig. 5 Dialogue example 2

根据 U1 的语义分析结果,推理机填充图 1 中“航班信息主题”主题树的“终点城市”结点.这时主属性结点无效,系统找到“日期”结点,调用对话生成函数向用户提问.得到 U2 输



入后,主属性结点有效(“起点城市”结点有默认值),推理机搜索主题树生成查询条件.由于查询结果不唯一,推理机在次属性结点下选取“起飞时间”结点作为应答的结点,该结点的应答函数生成 S2.得到 U3 输入后,推理机的动作与 U2 输入后一样,这次“起飞时间”结点的应答函数生成 S3. U4 用户提问,推理机在查询后,找到回答问题的结点为“到达时间”结点,相应对话生成函数回答 S4. U5 用户切换主题,推理机在填充“时差信息主题”的时候,根据共享信息索引从“航班信息主题”中取得城市的信息,使得主属性结点有效可以进行查询并应答. U6 用户再次切换回主题,这次查询结点唯一且主属性结点下的结点信息都确定,推理机调用主题结点的对话生成函数生成应答 S6.

在这段对话中,用户没有主动提及起飞时间、机型和航空公司,而系统主动与用户协商航班的起飞时间,对机型和航空公司只字不提,正是区别处理主题不同类型信息的结果.只有当用户主动提及机型和航空公司时,系统才会与用户进行这些信息的交流.

## 5 总结与评价

本文提出的基于主题森林结构的对话管理模型,是对话规划模型的一种具体实现.主题森林以信息项为最小的意义-规划单元,通过“与”和“或”关系构成最大为主题的不同层次的意义-规划单位,其中一个主题为对话任务的规划;不同的主题之间通过共享信息索引,建立了规划之间的联系.

由于每棵主题树都记录了对话过程中主题的所有知识状态,推理机总是根据整棵树的状态进行应答决策,所以该模型可以保证上下文中主题内容的一致性;共享信息索引的引入,使得主题之间的一致性也得到了保证.

这一对话管理模型具有较好的灵活性,体现在:能处理涉及多个主题和复杂主题的对话;允许用户自由地切换主题;对单个主题的不同信息能根据用户兴趣区别处理;考虑了对话一问一答及多回合间的省略情况,正确理解用户语义;能根据用户用语习惯指导自然语言的生成;基于主题森林结构的推理机根据对话发展,动态地决定应答内容,实现了对话控制的人-机混合主导.

与领域有关的知识主要体现在自然语言理解、语义框架的定义、领域知识的构成、知识库的组织和查询,以及自然语言生成上用户用语模型和某些与应答有关的对话状态,也与领域有关.但是这一模型的应答推理,只应用普遍策略,不涉及真正语义,有利于系统移植.

## References

- 1 Michael F McTear. Modeling spoken dialogues with state transition diagrams; Experiences with the CSLU Toolkit. In: Proceedings of 5th International Conference on Spoken Language Processing, v4. Sydney Australia; Australian Speech Science and Technology Association (ASSTA), 1998. 1223~1226
- 2 Lin Yi-Chung, Chiang Tung-Hui, Wang Heui-Ming, Peng Chung-Ming, Chang Chao-Huang. The design of a multi-domain mandarin Chinese spoken dialogue system. In: Proceedings of 5th International Conference on Spoken Language Processing, v2. Sydney Australia; Australian Speech Science and Technology Association (ASSTA), 1998. 41~44
- 3 Jennifer Chu-Carroll. Form-based reasoning for mixed-initiative dialogue management in information-query systems. In: Proceedings of Eurospeech'99, v4. Budapest Hungary; European Speech Communication Association, 1999. 1519~1522
- 4 Jerry H Wright, Allen L Gorin, Alicia Abella. Spoken language understanding within DIALOGS using a graphic



- model of task structure. In: Proceedings of 5th International Conference on Spoken Language Processing, v5. Sydney Australia: Australian Speech Science and Technology Association (ASSTA), 1998. 2107~2110
- 5 Lin Bor-shen, Wang Hsin-min, Lee Lin-shan. Consistent dialogue across concurrent topics based on an expert system model. In: Proceedings of Eurospeech '99, v3. Budapest Hungary: European Speech Communication Association 1999. 1427~1430

**邬晓钧** 清华大学计算机科学与技术系博士研究生. 研究领域为人机口语对话系统中的对话模型、对话管理和对话生成.

(**WU Xiao-Jun** Ph. D. candidate of the Department of Computer Science and Technology, Tsinghua University. His research interests include dialogue modeling, dialogue management and response generation in human-machine spoken dialogue systems.)

**郑方** 清华大学副教授、智能技术与系统国家重点实验室语音技术中心主任. 研究领域包括语音识别与理解、说话人识别、数字信号处理等.

(**ZHENG Fang** Associate professor of Tsinghua University. He is the Director of the Center of Speech Technology, State Key Laboratory of Intelligent Technology and Systems. His research interests include speech recognition and understanding, speaker recognition and digital signal processing.)

**徐明星** 清华大学计算机科学与技术系讲师. 研究领域为语音控制、连续语音识别与理解、说话人识别、智能语音信息查询等.

(**XU Ming-Xing** Instructor of the Department of Computer Science and Technology, Tsinghua University. His research interests include speech command, continuous speech recognition, speaker recognition and intelligent speech information inquiry.)