

基于框架的对话管理模型的研究与实现

袁 琰, 田怀凤, 杜 波, 陆汝占

(上海交通大学计算机系, 上海 200030)

摘 要: 对话管理模块在口语对话系统中占有极为重要的地位, 它控制系统与用户之间的交互, 确定交互过程中的每一步系统将采取何种操作。文中介绍的上海市交通信息智能查询系统中所采用的基于框架的对话管理模型由查询框架和建立在查询框架基础上的对话管理策略构成, 该模型实现了对话过程的混合主导, 保证了对话过程的一致性和灵活性, 并具有较好的移植性。

关键词: 对话系统; 对话管理模型; 框架; 对话管理策略

Research and Implementation of the Frame-based Dialogue Management Model

YUAN Yan, TIAN Huaifeng, DU Bo, LU Ruzhan

(Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030)

【Abstract】 Dialogue management module is a very important component in the spoken dialogue system, it controls the transaction between the system and user and decides which operation should be executed by system in each step. The frame-based dialogue management model, which is applied in Shanghai Intelligent Traffic Information System, is made up of query frame and strategy of dialogue management based on query frame. This model makes transaction both consistent and flexible with mixed-initiative dialogue control, and can be easily transplanted.

【Key words】 Dialogue system; Dialogue management model; Frame; Dialogue management strategy

新一代的口语对话系统综合了语音识别、自然语言理解、语言生成和对话管理等诸多技术^[1]。为了最大限度满足用户的使用要求, 对话系统的对话过程应该自然而有效。对话管理模块负责控制对话过程, 它根据语言分析模块所得到的分析结果, 决定是否已经从用户那儿得到了足够的信息, 并负责与应用程序通信来获取需要的信息, 最后将信息反馈给用户。

对话管理根据其所提供的控制类型可以分为3类: 系统主导, 用户主导和混合主导。在一个系统主导的对话中, 系统向用户提出一系列问题, 获取所需的任务信息。反过来, 用户主导的对话是由用户向系统提问, 将信息传递给系统。在混合主导的对话中, 用户可以向系统提问, 而系统也可以在必要时向用户提问。对话管理的方法大致可以分为有限状态的方法和自组织的方法两大类^[2]。在有限状态模型中, 对话结构是以状态转移网络的形式表现的, 很多商业性的对话系统采用某种形式的有限状态对话模型。但是这种模型缺乏灵活性, 比较适合于相对简单的对话过程, 当对话过程变得复杂而庞大时, 状态转移网络中的状态节点会成倍增加, 变得难以控制, 而用户的语句偏离系统计划, 使得系统无法处理的可能性也大大增加。自组织的对话管理包括基于框架和面向对象的方法、基于规划的管理以及定理证明等方法, 它提供了用户较大的自由度, 有越来越多的对话系统采用了自组织的方法。

由于交通查询领域具有对话主题较多、用户表达方式灵活等特点, 因此作为上海科委重点项目的“上海市交通信息智能查询系统”(SHJTQ)采用了基于框架的、混合主导的对话管理模型, 该模型主要由查询框架和建立在查询框架

基础上的对话管理策略构成。

1 系统概述

SHJTQ系统主要提供上海市区两点间的交通路线查询, 即由用户提供目的地、出发地和交通方式等信息, 系统根据用户的要求上网或从本地的数据库中获取交通路线信息反馈给用户。

从SHJTQ的总体框架结构图(图1)中可以看出, 对话管理模块在整个系统中占有极为重要的地位, 它控制系统与用户之间的交互, 确定交互过程中的每一步系统将采取何种操作。具体有如下几项功能: 从语言分析模块中接收关键信息, 并将其填入到系统的查询框架中去; 评估关键信息的可信度, 如有必要将提示用户进行澄清; 发送信息给上网查询模块, 便于其上网查询相应的路径信息; 发送信息给自然语言生成模块, 便于语言生成模块生成相应回答。

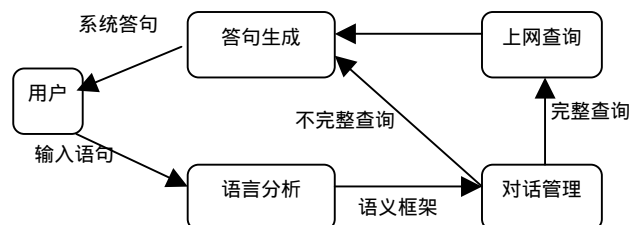


图1 SHJTQ总体框架结构

作者简介: 袁 琰(1978—), 男, 硕士生, 研究方向: 中文信息处理; 田怀凤、杜 波, 硕士生; 陆汝占, 教授、博导

定稿日期: 2004-06-01 **E-mail:** yuanyan@sjtu.edu.cn

2 查询框架

根据系统的实际使用需求,我们给出如下的定义:

定义1 查询项(Query Item)

查询项是一个包括<查询项名(ItemName) 查询项值(ItemValue)和概率值(ItemPossibility)>的三元组。此处的概率值不同于一般意义上的概率值,为了描述的方便,ItemPossibility的取值范围为[-1,1]。

本系统共定义了6种查询项,分别为Dest, Source, TransMethod, FrameType, RouteType, FrameID。其中:

Dest表示出发地,其缺省值为空;

Source表示目的地,其缺省值为空;

TransMethod表示交通工具,设T为交通工具的集合,则TransMethod的取值范围为T的幂集,其缺省值为空集;

FrameType表示查询框架的类型,本系统中有两种类型,分别为查询(Query) 叙述(State),其缺省值为Query;

RouteType是用户对所查询路线的要求,本系统中有4种类型,分别为无要求(NULL) 最快路线(Fastest) 最经济路线(Economical) 和换乘最少路线(Least),其缺省值为Null。

FrameID是查询框架的编号,取值范围为自然数,缺省值为1。FrameID的引入可以满足一次查询多条交通路线的需要。

定义2 查询框架(Query Frame)

查询框架(以下简称为QF)是查询项的集合 $QF=\{QItem|QItem是查询项\}$,其中QF满足以下条件:每一种查询项出现且只出现一次。

定义3 完整查询框架(Completed Query Frame)

查询框架QF是完整的查询框架当且仅当QF满足以下条件:

QF所含的Dest和Source查询项的值非缺省值;

设QF所含的TransMethod查询项的值为Ts,则 $|Ts|=1$;

QF所含的FrameType的值为Query。

显然,如果一个查询框架是完整的,则系统可以根据该框架上网查询。事实上,和用户交互的目的就是为了得到完整查询框架。下面是一个查询框架的例子,显然该查询框架不是完整的查询框架。

<	FrameType,	Query,	0.8	>
<	Source,	"交通大学",	0.8	>
<	Dest,	"外滩",	0.8	>
<	TransMethod,	{},	0	>
<	RouteType,	Null,	1.0	>
<	FrameID,	1,	1.0	>

图2 查询框架示例

查询框架对于用户是透明的,由对话管理模块根据句法分析模块的结果来填充框架中查询项的值,判断是否已经生成完整的查询框架,如果是,则将相关信息提交给上网查询模块;反之,则启动答句生成模块生成答句,通过与用户的交互以获取构成完整查询框架的信息。值得一提的是,在判断查询框架是否完整时,查询项RouteType对结果不产生影响,系统默认该项为可选项。

3 对话管理策略

在定义了查询框架的基础上,我们制定了若干对话管理策略,这些策略对于系统的运行起着重要作用^[3]。

3.1 查询框架的合并策略

查询框架合并是指将多个查询框架($\{QF1, QF2, \dots, QFn\}$)

合并生成一个新查询框架(NewQF)的过程,从用户意图的角度看,NewQF应该与 $\{QF1, QF2, \dots, QFn\}$ 等价。在以下两种情况下,需要合并查询框架,一种情况是用户的查询语句由几个小句组成,在这种情况下,如果能将每个小句对应的查询框架合并为一个,则合并为一个;另一种情况是将用户在一次对话过程中输入的所有语句所对应的查询框架合并,从而生成上网查询模块所需的输入参数。

设被合并的框架集合为 $QFs=\{QF1, QF2, \dots, QFn\}$,新生成的框架为NewQF。若 QFi 是一个框架,则 QFi 的任一查询项Fitem的值表示为Fitem(QFi),Fitem对应的概率为 $P(QFi.Fitem)$ 。

框架合并步骤如下:

第1步 根据准则1确定可合并的框架的集合QFs。

准则1 具有相同FrameID的查询框架才可以合并。即QFs中的框架可以被合并,当且仅当 $FrameID(QFi) = FrameID(QFj)$, QFi 和 QFj 是QFs的任意元素;

第2步 根据准则2~6调整 QFi 的各查询项的概率值, QFi 为QFs中的任一元素。

准则2 对于查询项Dest、Source、TransMethod,当其取缺省值时,其对应的概率为0;

准则3 对于查询项FrameID、FrameType,取缺省值时,其对应的概率为1;

准则4 否定词极大影响查询项的概率值,具体处理见否定的处理策略;

准则5 对于同类型的查询项,同等条件下,出现于后面框架的,其概率值大于出现于前面框架的。

准则6 如果Dest、Source、TransMethod的值不为缺省值,则其概率值从句法分析的结果中获得。

第3步 利用准则7确定NewQF中Dest查询项的值。

准则7 如果 $\{QF1, QF2, \dots, QFn\}$ 对同一个特征项Fitem取值不同,则新查询框架Fitem(NewQF) = Fitem(QFi), QFi 满足条件: $P(QFi.Fitem) > P(QFj.Fitem)$, QFj 为QFs中的任一元素。该准则所得结果不能与准则9相矛盾。

第4步 根据准则8重新调整 QFi 的各查询项的概率值, QFi 为QFs中的任一元素。

准则8 设 $Dest(NewQF)=Dest(QFi)$,则如果Source(QFi)、TransMethod(QFi)不为缺省值,则增大 $P(QFi.Source)$ 和 $P(QFi.TransMethod)$ 的值,然后再计算Source(newQF)和TransMethod(NewQF);

第5步 按准则7确定TransMethod(newQF)和Source(newQF)。所得结果不能与准则9相矛盾,如果结果矛盾,则根据准则7选择概率次大的;

准则9 $Dest(NewQF)$ 不能等于 $Source(NewQF)$ 。

3.2 否定的处理策略

对于含否定词的句子,同样建立专门的查询框架,否定词否定的内容对应到查询框架的查询项,通过将对应查询项的概率值置为-1表示否定。在框架合并时应满足如下准则:

准则10 设 $P(QFi.Fitem)=-1$, $Fitem(QFj)=Fitem(QFi)$,且 $j<I$,则 $p(QFj.Fitem)=-1$ 。其中 $\{QF1, QF2, \dots, QFn\}$ 表示被合并的框架集合。Fitem(QFi)表示框架 QFi 中查询项Fitem的值, $P(QFi.Fitem)$ 表示Fitem对应的概率。

以图3的对话片断为例,共生成4个查询框架见图4。

S: 您要从上海图书馆出发?

U: 不,我要从上海体育馆而不是上海图书馆出发去人民广场。

图3 一个带否定的对话片断(S代表系统,U代表用户)

<	FrameType,	Query,	0.8	>
<	Source,	"上海图书馆"	0.8	>
<	Dest,	"",	0	>
<	TransMethod,	{},	0	>
<	RouteType,	Null,	1.0	>
<	frameID	1,	1.0	>

(a) 查询框架 1

<	FrameType,	Query,	0.8	>
<	Source,	上海图书馆	-1.0	>
<	Dest,	"",	0	>
<	TransMethod,	{},	0	>
<	RouteType,	Null,	1.0	>
<	frameID	1,	1.0	>

(b) 查询框架 2

<	FrameType,	Query,	0.8	>
<	Source,	上海体育馆,	0.8	>
<	Dest,	人民广场,	0.8	>
<	TransMethod,	{},	0	>
<	RouteType,	Null,	1.0	>
<	frameID	1,	1.0	>

(c) 查询框架 3

<	FrameType,	Query,	0.8	>
<	Source,	"上海图书馆",	-1.0	>
<	Dest,	"",	0	>
<	TransMethod,	{},	0	>
<	RouteType,	Null,	1.0	>
<	frameID	1,	1.0	>

(d) 查询框架 4

图 4 对话片断生成的查询框架 1~4

系统问句对应查询框架 1, 用户输入的第一个小句“不,” 对应查询框架 2, 用户输入的第 2 个小句“我要从上海体育馆而不是上海图书馆出发去人民广场。”对应查询框架 3 和查询框架 4, 将上述 4 个框架合并后就得到该对话片断的查询框架:

<	FrameType,	Query	0.8	>
<	Source,	"上海体育馆"	0.8	>
<	Dest,	"人民广场",	0.8	>
<	TransMethod,	{},	0	>
<	RouteType,	Null,	1.0	>
<	frameID	1,	1.0	>

图 5 合并后的查询框架

3.3 用户确认策略

用户输入信息和语音识别的过程中, 难免会引入一定数量的错误, 这些错误有些比较严重, 导致系统的拒识别, 但统计结果表明, 这种类型的错误只占所有错误的一小部分; 绝大部分的错误只是导致关键信息的模糊, 针对这种情况, 我们制定了用户确认策略。当查询框架中的某个查询项 Fitem(QFi)所对应的概率 $0 < P(QFi.Fitem) < (0 < \leq 1)$, 为系统预定义的常数, 称为阈值, 也可根据情况动态调整), 则激活答句生成模块, 针对该查询项的信息要求用户确认, 若查询项是地名信息, 则还提供与之最相似的几个地名供用户选择。

除了上述这些策略, 系统中还有如缺省的处理等相对次

要的策略。从下一节的实例中就可以进一步看出这些策略的运用和作用。

4 对话实例

下面为一个实际对话的例子, 在这段对话中, 用户首先主动提出了目的地, 在系统作出回应后, 又否定了原来的目的地, 系统这时应用了否定的处理策略, 通过框架的合并, 得到正确的查询框架; 随后在未生成完整的查询框架时, 用户又提出了新的查询要求, 系统则为新的查询建立新的查询框架, 其 FrameID 为 2; 在所有的查询框架都成为完整的查询框架后, 系统才开始进行查询。

S: 您好! 我能为您做什么? (系统问候信息)

U: 我想去广州路。(用户主动)

S: 您要从什么地方出发去广州路? (系统主动, 提问出发地)

U: 外滩。(用户回答)

S: 请问您选择什么交通工具从外滩去广州路? (系统主动, 提问交通工具)

U: 不, 我要从外滩去南京路而不是广州路。(用户回答, 否定目的地)

S: 请问您选择什么交通工具从外滩去南京路? (系统应用否定处理策略, 主动提问交通工具)

U: 公交吧。噢, 随便问一下从那儿再去人民广场该怎么走?

(用户回答, 提出新的查询)

S: 从南京路去人民广场, 您选择什么交通工具? (系统建立新的查询框架, 主动提问交通工具)

U: 还是公交吧。(用户回答)

S: 系统正在查询, 请稍候…… (系统判断已生成完整的查询框架, 进行查询)

5 分析和总结

本文中提出的基于框架的对话管理模型支持混合主导的对话过程, 能引导用户进行高效的交流, 从而也防止了用户的对话超出系统的计划之外。同时这一模型还具有以下两个优点: (1) 语种的独立性, 应用该模型的 SHJTQ 系统正是一个中英双语的对话系统, 如果要开发新的语种的同类系统, 也可以很容易进行移植; (2) 领域的通用性, 该模型完全可应用于其他领域的同类系统, 如航班查询、图书查询等。

这种基于框架的对话管理模型能较好适用于特定领域的口语对话系统。但还可以在以下两方面作一些改进: (1) 将模型中与领域相关的和与领域无关的部分进一步加以分离, 增强它的可移植性; (2) 细化查询框架中查询项和每个查询项中信息项, 使其能够适应更为复杂的查询要求和任务。

参考文献

- Goddeau D, Meng H, Polifroni J. A Form-based Dialogue Manager for Spoken Language Applications. International Conference on Spoken Language Processing, Philadelphia, PA, 1996-10: 701-704
- 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)
- 郭荣. 汉语口语对话系统的语言分析[学位论文]. 上海: 上海交通大学, 2003-01
- Bell E. 用.NET 和 XML 构建 Web 应用程序. 北京: 清华大学出版社, 2003
- Short S. 构建 XML Web 服务——基于 Microsoft.Net 平台. 北京: 清华大学出版社, 2002
- 柴晓路, 梁宇奇. Web service 技术、架构和应用[M]. 北京: 电子工业出版社, 2003

(上接第 191 页)

参考文献

- 法克特. 应用服务供应商 (ASP) 解决方案[M]. 北京: 电子工业出版社, 2003
- 钟华. ASP 的发展和优势. 开目的电子商务观, 2001, (1): 15-16, 19-20, 23-24, 31