

限定领域口语对话系统中超出领域话语的协处理方法*

王俊东, 黄沛杰, 林仙茂, 徐禹洪, 李凯茵

(华南农业大学数学与信息学院, 广东 广州 510642)

摘要: 领域外话语的开放性、口语化以及表达多样性, 使得现有的限定领域口语对话系统不能很好地处理超出领域话语。本文提出了一种限定领域口语对话系统协处理方案, 基于人工智能标记语言 AIML, 设计一套理解开放语义用户话语的理解模板, 并对未匹配话语基于话语相似度进行理解模板分类, 进而采用扩展有限状态自动机处理模式, 结合对话流程上下文的状态及信息, 实现理解模板到应答模板的转换, 改变了单纯模板匹配方法在对话流程控制方面的相对缺失。中文手机导购领域的测试表明, 本文所提出的协处理方法能有效地辅助口语对话系统完成限定领域完整对话流程, 得到更好的用户满意度。

关键词: 超出领域话语; 协处理; AIML; 有限状态自动机; 口语对话系统

中图分类号: TP391

文献标识码: A

A Coprocessor for Out-of-Domain Utterances in Restricted Domain Spoken Dialogue System

WANG Jundong, HUANG Peijie, LIN Xianmao, XU Yuhong, LI Kaiyin

(College of Mathematic and Informatics, South China Agricultural University, Guangzhou 510642, China)

Abstract: The openness, colloquialism and diversity of out-of-domain (OOD) utterances make the restricted domain spoken dialogue system cannot handle well with the OOD utterances. This paper tackles this problem by proposing a coprocessor for restricted domain dialogue system. Based on the artificial intelligence markup language, AIML, open semantic understanding templates are given, and understanding template classification is used to address the unmatched OOD utterances. And then the extended finite state machine (EFSM) is adopted which transforms the understanding template into answering template and realizes the control of the state and information of the dialogue process. The application in Chinese mobile phone shopping guide domain shows that the proposed coprocessor can effectively help the dialogue system to finish the restricted domain dialogue process and get better user experience.

Key words: out-of-domain utterance; coprocessor; AIML; FSM; spoken dialogue system

1 引言

口语对话系统(spoken dialogue system)指的是通过自然语言和人交流的计算机系统, 主要研究如何能让计算机理解并生成人们日常所使用的语言, 对人给计算机提出的问题, 通过对话的方式, 用自然语言进行回答。研究对话系统的目的是让人同计算机的交流更方便, 让计算机具有类似人类的思维从而帮助人们完成更多的工作, 从这个角度看, 能胜任某一专业领域工作的面向任务(task-oriented)的限定领域(restricted domain)对话系统^[1-8], 比开放领域(open domain), 如面向聊天(chat-oriented)的对话系统^[9-10]更有研究和应用价值^[11]。然而, 当使用自然语言对话时, 即使用户了解某对话系统的限定领域, 例如, 医疗信息咨询、导航或者导购, 用户在对话流程中仍然不可避免会使用一些超出领域(out-of-domain, OOD)话语(utterance), 如, 问候、个人问题、表达肯定等。事实上, OOD话语的现象很常见, 比如AT&T的“*How may I help you*”系统^[2], 以及BTaxeCT和Lucent Bell合作开发的“OASIS

*收稿日期:

定稿日期:

基金项目: 国家自然科学基金(71472068); 广东省大学生创新训练计划项目(201410564290, 201510564281)

call-steering”系统^[5]，大约有20%的用户问题是OOD的。尽管这些限定领域对话系统从完成任务角度看只需要专注于自己预定义的业务功能，但是，如果能较为妥善地处理好OOD话语，而不仅仅是提示用户话语超出领域，将会有效地提高用户体验^[12]。

目前已有的限定领域对话系统，如导航系统^[4, 8]、导游系统^[7]和信息查询系统^[1-3, 5-6]等，基于语义和语法相结合的技术，并结合近年来逐渐成为热点的意图追踪^[13]和对话管理^[14-16]，能有效地理解和处理领域内语义的话语。然而，面对OOD话语的开放性、口语化以及表达多样性，现有的限定领域口语对话系统在处理OOD话语时仍然存在一定困难。不少研究人员开始进行关于限定领域对话系统OOD话语问题的研究，比如Lane等人^[17]开发的机器辅助对话系统和Tür等人^[18]开发的虚拟个人助理系统使用基于SVM(support vector machine)主题分类方法计算源话语在领域内各子领域的分类置信度，再使用置信度向量训练一个线性判断模型，并检测源话语是否超出领域，若超出领域，系统可以提示用户当前系统无法处理该任务并引导用户回到领域可处理范围。Celikytmaz等人^[19]也提出了使用LDA(latent dirichlet allocation)主题模型对多领域问题做主题聚类并检测OOD情况。Reichel等人研究的汽车导航领域口语对话系统^[8]，将在多领域内切换过程中出现错误且不能被任何子领域接受的话语归为OOD情况，并返回给用户做领域选择。但这些工作中主要集中在OOD话语的检测识别，并根据检测结果简单响应用户，而缺少对OOD话语的有效处理方法。

面向聊天的开放领域对话系统，如国外的ELIZA^[20]、PARRY^[21]和ALICE^[9, 22]，国内的清华大学图书馆的“小图”^[23]、机器人小I^[24]等，由于设计初衷就是开放领域，在覆盖多领域的语料库的支持下，一定程度上满足了话语开放性、口语化和表达多样性的要求。这些系统可以与用户进行基于自然语言话语的交互，并已被广泛应用到机器助理比如反恐支持平台^[25]、计算机辅助英语学习^[26]和健康咨询^[27]等。然而，这些面向聊天的对话系统几乎都是基于完全精确匹配或者很弱的模糊匹配的模板来搜索回答^[28]，并没有真正地“理解”用户的问题，并且难于掌控对话流程的系统状态和信息。这类对话系统不能单独应用到复杂任务的限定领域对话。此外，语料库的构建是此类系统的瓶颈，尽管有一些关于语料库自动建设^[12, 29]、超出词汇处理^[10]等的研究，如Banchs^[29]和Ameixa等人^[12]分别采用电影剧本和电影字幕来构建语料库，但其话语覆盖程度仍然远远不足以应用于限定领域对话系统中OOD话语的处理。此外，单纯模板匹配方法由于对话上下文及对话状态方面的缺失，难以真正提高OOD话语回答的针对性。

本文提出一种限定领域口语对话系统中处理OOD话语的协处理方案，主要贡献包括：

(1) 根据基于模板匹配方法适合开放语义、口语化话语以及表达多样性的特点，基于人工智能标记语言AIML (artificial intelligence markup language)^[30]，设计一套用于理解OOD话语的理解模板，并采用基于话语相似度的分类方法，将未匹配话语分类到相应的理解模板，从某种程度上解决了语料库一次性完备构建的困难。

(2) 采用扩展有限状态自动机 (extended finite-state machine, EFSM) 处理模式，结合对话上下文信息及对话状态，实现理解模板到应答模板的转换，并给出不同应答模板相应的处理程序，改变了单纯模板匹配方法在对话流程控制方面的相对缺失。

本文的方案被应用到我们开发的手机导购对话系统^[31]，应用测试结果表明，能有效地辅助系统完成限定领域对话流程，得到更好的用户满意度。本文后续部分安排如下：下一节简要介绍协处理方案的系统框架及处理流程。第3节详细介绍了协处理方案。第4节给出了测试验证结果。最后，第5节总结了本文的工作并作了简要的展望。

2 系统框架

图1是我们开发的限定领域口语对话系统的系统框架，目前应用于手机导购领域，加粗

部分是本文提出的协处理方案。

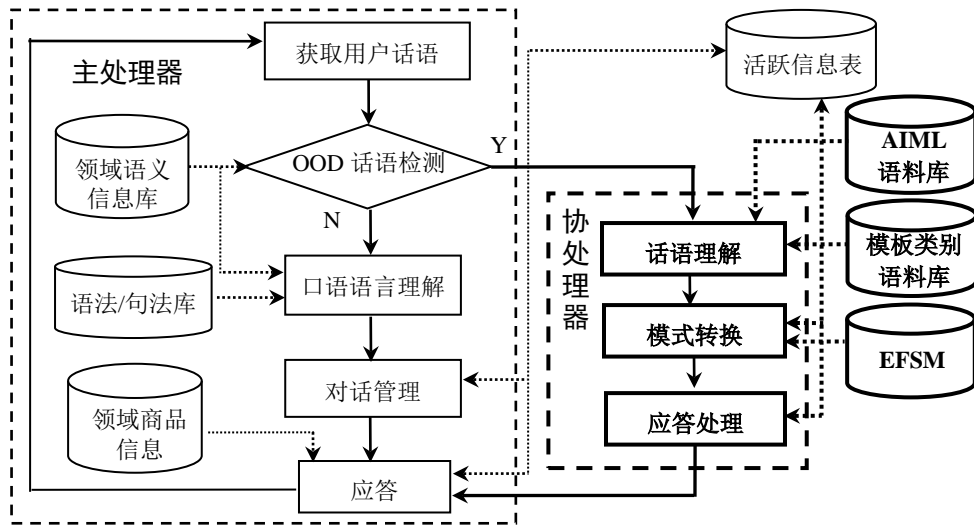


图1 限定领域口语对话系统的系统框架

在这个框架里，“主处理器”基于自然语言处理技术，完成携带领域语义信息的话语的处理。语义提取、语法分析等技术被应用于口语语言理解。活跃信息表象征着对话系统的“记忆”，目前保持了三种活跃信息，包括商品属性值、对话的上下文以及已推荐商品列表。其中，商品属性值也即是一般对话系统具有的对话语义框（dialogue frame）^[32-33]中的槽信息（slot）。

而“协处理器”，顾名思义，就是主处理器的协处理机制，并不独立完成导购服务，仅在主处理器中的 OOD 话语检测模块识别不到有效领域语义信息时接收主处理器传来的用户输入话语，并协助完成该对话步骤。协处理方案包括一套有三个模块的协处理机制和三个信息库，其中，AIML 语料库支持对 OOD 话语到理解模板的映射，模板类别语料库支持未匹配话语的理解模板分类，EFSM 支持模式转换。基于 AIML 和 EFSM 的协处理流程是：首先加载 AIML 语料库到内存并初始化需要的数据，用户话语进入话语理解模块后匹配上 AIML 中的理解模板或者在未匹配时通过模板分类算法得到理解模板；接着通过模式转换模块，完成系统模式的转换和理解模板到应答模板的映射，转移条件中的上下文信息来自活跃信息表；最后是对应答模板进行分类处理，得到相应的应答并返回主处理。

3 AIML 和 EFSM 相结合的协处理方案

3.1 基于 AIML 的话语理解模板

AIML^[30] 是经典聊天机器人 ALICE^[9, 22]所采用的一种 XML 的人工智能标记语言，定义了一整套具有特定含义的标签。由包含 <pattern>...</pattern> 和 <template>...</template> 标签对的知识单元构成，前者表示用户输入需要进行匹配的模板，后者代表匹配成功后用于回答的模板。AIML 除了能够处理精确匹配，还能够处理模糊匹配以及优先匹配。

通过对 AIML 语料库的回答模板（<template>）进行合理的归类，有效降低语料库建设的难度。相比于每个话语都有专门的回答，虽然在一定程度上降低了系统对用户话语的理解的“准确”度，但并不会影响后续处理和应答。同时由于理解模板及后续介绍的模式转换抓住了多样化用户话语的共性，具有一定通用性，也易于扩展到其它限定领域。目前理解模板分为四大类：闲聊、表态、导购和其它，每一大类又分为若干小类，如表1所示。如“闲聊”大类的“问候”小类对应的理解模板为{#闲聊#问候#直接模板}，其中的“直接模板”是可选的附加内容，用于在后续组织应答处理中提供更有针对性回答。

表 1 理解模板的分类与话语示例

大类	小类	话语示例
闲聊	问候	“你好”、“早上好”等
闲聊	结束	“再见”、“我走啦”等
闲聊	时间	“现在几点啦”等
闲聊	天气	“今天天气好吗”、“好热”等
闲聊	身份信息	“你是谁”、“你的老板是谁”等
表态	肯定	“不错喔”、“挺好的”等
表态	否定	“不好”、“我不喜欢”等
表态	犹豫	“不知道哪个好”等
表态	随便	“随便”、“没有要求”等
导购	开始	“这里有什么手机”等
导购	更换	“换一款把”、“其它”等
导购	成交	“我要了”、“成交”等
其它	附和	“好好笑”、“我想静静”等
其它	捣乱	“混蛋”、“住口”等
其它	其它	其它理解模板未覆盖的话语

值得注意的是，对于模板未覆盖的话语，单纯的 AIML 机制只能将其归类为{#其它#其它#直接模板}，并采用最模糊的回答，也即是 AIML 机制中典型的单独通配符“*”^[30]，并选用一些通用的回答作为返回内容，比如“没听懂，请换一个说法”。

3.2 基于话语相关度的话语理解模板分类

对于 AIML 未匹配的用户话语，也即“其它”大类的“其它”小类，本文提出一种基于话语相关度的话语理解模板分类模型。该模型利用统计方法和 VSM (vector space model) 模型构造语料库，然后以类别语料库作为话语分类的知识库进行分类。该模型增强了系统对大部分未被单纯 AIML 机制良好处理的 OOD 用户话语的理解。

3.2.1 类别语料库的构建

类别语料库是主题、主题特征词及其权重的元组集合。因为构造涵盖各个领域的通用类别语料库工作量巨大，尤其在网络时代，新的词汇不断出现，所以这里我们只对我们感兴趣的领域类别建立类别语料库，类别语料库的类别与理解模板的小类一致，如“表态|犹豫”等。

设感兴趣的原始话语集为 $US(c)$ ，其中， c 为类别数，用传统的词袋表示法将话语 U 表示为 VSM 形式，由于分词是整个话语表示的基础，不同的分词结果有差异，我们采用效果比较好的中科院计算技术研究所的 ICTCLAS¹分词系统作为分词标准。经过分词处理和统计，话语 U 表示为 $U=\{(t_1, pos_1), (t_2, pos_2), \dots, (t_i, pos_i), \dots, (t_n, pos_n)\}$ ，其中， t_i 是分词后所得的词元（词或短语）， pos_i 是 t_i 的词性。类别语料库特征词提取算法的伪代码如图 2 所示。

¹ <http://ictclas.nlpir.org/>

算法: 类别语料库特征词提取算法
输入: 感兴趣的话语集 US(c)
输出: 提取的类别特征词集合 K(c)
过程:

1. init 类别词元集 U(c)和 K(c)
2. for US(c)中的所有话语 US
3. 使用 ICTCLAS 分词系统对 US 进行分词和词性标注, 并生成词元集合 U
4. 遍历 U 中所有词元并根据词性过滤掉标点符号、停用词、虚词(助词等)等
5. 将 U 添加到 U(c)中
6. end for
7. for U(c)中的所有类别的词元集 U
8. 计算 U 中的所有词元的权值 w_i 和频率 $freq_i$
9. 遍历生成四元组 $key=(t_i, pos_i, w_i, freq_i)$, 添加到特征词集合 K1, 按权值大小排序
10. 选取 K1 中的 80% 作为该类的特征关键词集合 K
11. 将 K 添加到 K(c)中
12. end for
13. 返回 K(c)

图 2 类别语料库特征词提取算法伪代码

算法第 8 步中的词元的权值采用期望交叉熵计算得到:

$$w_i = p(t_i) \sum_j [p(c_j | t_i) \lg(p(c_j | t_i) / p(c_j))] \quad (1)$$

由式(1)可知, 当词 t_i 在类别语料库中越分散(即出现的类别数越多), 权值 w_i 也会相应降低, 说明该词的辨识度不高; 如果词 t_i 越集中, 权值 w_i 也会相应升高, 表明该词的辨识度高。

3.2.2 话语理解模板分类

首先根据用户话语特征词的语料权值和频率计算话语的类别相似度矩阵, 进而根据相似度矩阵来找出最高相似度的理解模板类别。话语相关度计算算法伪代码如图 3 所示。

算法: 话语相关度计算算法
输入: 类别特征词集合 K(c)、待分类话语 U、参数 a
输出: 话语 U 的类别相似度矩阵 S
过程:

1. 使用 ICTCLAS 分词系统对 U 进行分词, 并生成 $U1=(t_1, t_2, \dots, t_n)$
2. init 相似度矩阵 S
3. for K(c)中的所有特征词集合 K
4. init 当前相似度分数 score = 0
5. for K 中的所有特征词 key
6. 计算 key 在 U1 中出现次数 n
7. 计算 $score=score+a*\log_2(n+1)*(1+key.w)*key.freq$
8. end for
9. 将 score 添加到 S 中
10. end for
11. 返回 S

图 3 话语相关度计算算法伪代码

经测试，本节提出的方法可以解决大部分的 OOD 话语分类问题。剩余的部分与各个类别的相似度都不高的话语，则继续归类为{#其它#其它#直接模板}。

3.3 基于 EFSM 的模式转换

在有效识别到相应理解模板的基础上，本文借鉴基于模式的对话模型^[16]在对话状态及信息控制的优点，结合对话上下文信息及对话状态，采用 EFSM 处理方法得到合适的应答策略，提高 OOD 话语的处理效果。

3.3.1 形式化模式转换

我们采用扩展的 FSM 来描述协处理模式及其转换，扩展包括状态和条件两方面：

(1) 状态的扩展。在对话系统中，系统状态通常伴随着系统动作，因此，在传统的 FSM 的状态的基础上，我们将系统状态和动作放到一起并称之为模式 P。

(2) 条件的扩展。本文综合了理解模板、系统事件和上下文信息作为模式转换的条件。其中系统事件主要是系统内部的一些机制，指系统掌控对话过程所做的必要监听事件，比如对用户相同输入的回合数计数事件等。系统会不断监控对话处于协处理的回合数，如果回合数过多，就需要及时引导回主处理程序。上下文包括了上一系统状态，以及历史交互中的所有语义信息和用户意图追踪的结果，还包括一些关键状态信息，比如上下文中会记录当前正谈论的是哪款商品、哪个参数等。

协处理的模式转换形式化定义如下：

- 模式 P
- 理解模板 UT
- 系统事件 E
- 上下文 C
- 应答模板 (AT)
- 输入集 $I \in (UT \cup E \cup C)$
- 输出集 $O \in (AT)$
- 转移函数 $T: P \times I \rightarrow P' \times O$

图 4 是一个模式转换图例，提供执行模型和可理解的图表达。目前只考虑“一跳”的模式转移。图中阴影模式表示任意模式，也即是任意模式可以进入协处理。

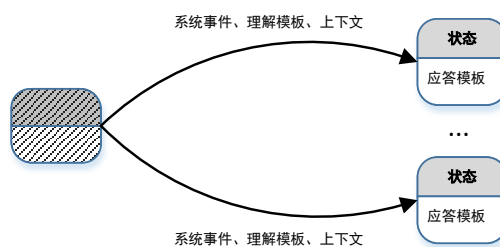


图 4 模式转换图例

3.3.2 模式

传统状态机的状态可用于描述系统当前所处阶段的相关信息，本文提出的协处理模式提供了更完善的动作信息，下面给出其定义：

- 模式状态 (PS)：系统接到用户动作后所处的状态。
 - 模式动作 (PA)：系统状态发生改变时的动作，本文将其定义为应答模板的形式。
- 表 2 是分析不同理解模板在不同条件下的可能状态及动作得到的协处理模式。

表 2 协处理模式

模式状态 (PS)	模式动作 (PA)
接受问候	{#直接返回#返回内容}
结束问候	{#主动引导#应答句式}
接受结束	{#返回并结束#返回内容}
挽留	{#查询其它特性#应答句式}
挽留	{#选择推荐手机#应答句式}
接受闲聊	{#直接返回#返回内容}
结束闲聊	{#主动引导#应答句式}
接受肯定	{#直接返回#返回内容}
接受肯定	{#主动引导#应答句式}
接受肯定	{#购买确认#应答句式}
接受肯定	{#返回并结束#返回内容}
接受否定	{#查询其它特性#应答句式}
接受否定	{#选择其它手机#应答句式}
接受否定	{#返回并结束#应答句式}
反驳否定	{#选择若干优势#应答句式}
犹豫推荐	{#选择推荐手机#应答句式}
犹豫诱导	{#选择若干优势#应答句式}
犹豫分析	{#选择差异参数#应答句式}
接受随便表态	{#主动引导#应答句式}
拒绝随便表态	{#属性值确认#应答句式}
接受开始导购	{#选择候选属性#应答句式}
接受更换	{#选择候选手机#应答句式}
接受成交	{#购买确认#应答句式}
接受捣乱	{#直接返回#返回内容}
拒绝捣乱	{#主动引导#应答句式}
接受附和	{#直接返回#返回内容}
拒绝附和	{#主动引导#应答句式}
其它	{#直接返回#返回内容}

3.3.3 模式转换例子

图 5 给出一个在对话初始阶段用户问候进入协处理的例子, 模式转换条件中的上下文信息依据的是交互历史中系统是否已有问候语义; 图 6 是用户犹豫表态后进入协处理的例子, 模式转换条件中的上下文信息依据的是关键语义信息中是否有谈论手机以及谈论的商品的数量。系统在任何状态下接收到图中所示的条件后, 随即转换到相应的协处理模式, 在转换到对应的状态的同时, 执行相应的动作, 下一次的模式转换是未知的, 可能重新进入到协处理模式, 也可能回到主处理中的模式, 这需要看用户的下一个输入话语。

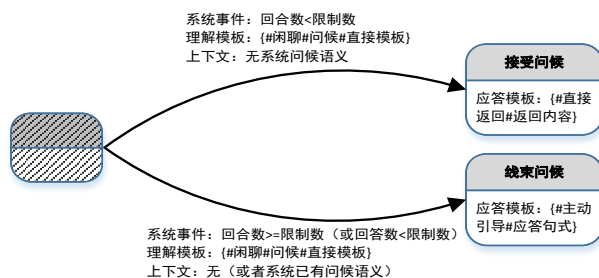


图 5 对话开始阶段进入协处理的例子

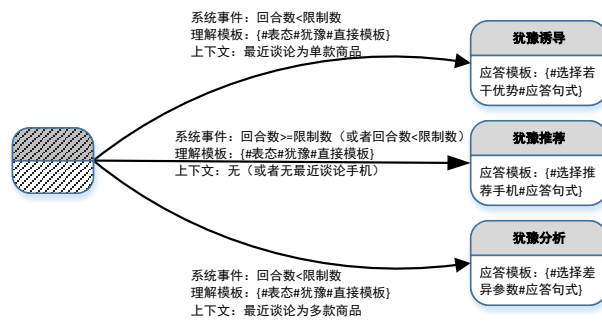


图 6 用户犹豫表态后进入协处理的例子

3.4 应答模板处理程序

本节以两个代表例子简要介绍应答模板的设计思想。

(1) “#主动引导#”的处理

当系统识别到应答模板为{#主动引导#应答句式}时,先在活跃信息表中根据优先级查找先前未提及的参数作为引导提问参数,然后将选出来的参数名称记录到活跃信息表中的“当前关联参数”,同时将该参数填到应答模板中应答句式的相应位置。

(2) “#选择差异参数#”的处理

差异参数是指多款手机的参数值存在一定程度的差异的参数信息,比如两款手机,一款屏幕尺寸 4.0 寸,另一款 4.5 寸,因此屏幕尺寸就是差异信息。而对于两款手机,如果某个属性只存在细微差异,如价格分别为 1990 元和 2000 元,则不认为价格是差异信息。

4 中文手机导购系统中的应用

4.1 应用例子

本文的研究已应用到我们开发的中文手机导购系统。图 7 是一个开发人员测试的例子。可以看到本文提出的协处理方案能有效地辅助对话系统完成限定领域完整对话流程,包括对第(2)、(10)、(14)和(18)句用户话语的处理。其它用户话语由于都携带领域语义信息,由主处理器处理。

- (1) 系统: 您好!
- (2) 用户: **你好!**
- (3) 系统: 您想买什么品牌的手机呢?
- (4) 用户: 我要三星的。
- (5) 系统: 您打算买什么价位的?
- (6) 用户: 2000 块左右的吧。
- (7) 系统: 您想要什么输入方式的?
- (8) 用户: 触屏手机有什么好介绍?
- (9) 系统: 有 E120L 和 I9050。
- (10) 用户: **我也不知道哪个好。**
- (11) 系统: 您可以考虑屏幕大小。
- (12) 用户: 哪个大一点?
- (13) 系统: E120L。
- (14) 用户: **嗯, 不过我不喜欢。**
- (15) 系统: 那要不这款 I9050。
- (16) 用户: 这款摄像头像素高吗?
- (17) 系统: 还可以, 500 万像素。
- (18) 用户: **谢谢你的介绍。**
- (19) 系统: 欢迎再次观临!

图 7 中文手机导购系统的一个运行例子

其中，第(2)句用户话语对应的理解模板为{#闲聊#问候#直接模板}，结合图 5 的模式状态转移，转到应答模板{#主动引导#应答句式}，进而协处理器根据主动引导处理算法，从活跃信息表中查询当时用户未提及且优先级最高的参数进行询问引导，在这个运行实例中，该参数是“手机品牌”。对第(10)句用户话语的理解模板为{#表态#犹豫#直接模板}，由于上下文记录系统正在谈论的是多款手机，根据图 6 的模式转换，转到应答模板{#选择差异参数#应答句式}，选择出差异参数“屏幕尺寸”进行引导。对第(14)句的处理分别对应的是理解模板{#表态#否定#直接模板}和应答模板{#选择其它手机#应答句式}。对第(18)句的处理分别对应的是理解模板{#闲聊#结束#直接模板}和应答模板{#返回并结束#返回内容}。

4.2 测试报告

我们在实现的中文手机导购对话系统中进行了测试。系统的测试人员是 15 名学生志愿者，每位测试者测试 10 段，共 150 段对话。这些测试人员只知道系统的功能是手机导购，对系统的实现细节并不知情，这样可确保测试的真实性和自然性。为了提高测试的多样性，在测试时规定了每人 3 段自由发挥，7 段在给定的 10 个场景中挑选，比如“老妈开始眼花了，想给她买个大屏的手机，5 寸以上的”。

4.2.1 测试总体情况

测试的总体情况如表 3 所示。

表 3 测试总体情况

总对话段数	总回合数	每段对话回合数（平均/最大/最小）	协处理总回合数	每段对话协处理回合数（平均/最大/最小）
150	1840	12.27/30/5	412	2.75/11/0

150段对话的平均回合数为12.27，进入协处理的平均次数为2.75，占了22.4%，表明了协处理器存在的价值。从最大和最小的协处理次数分别是11和0也可以看出测试者的差异性很大，尽管都知道系统的功能是手机导购，有的测试者还是会说很多OOD话语，而有的则全部都是领域内话语。经过分析发现四大类中，闲聊、表态、导购、其它分别占了40.2%、30.1%、24.3%和5.4%。

4.2.2 评价标准和方案对比

超出领域话语处理不好通常带来的是用户体验的下降，因此本文以用户满意率为度量标准。用户不满意的回答主要包括：答非所问和回答没有针对性。主要引起不满意的原因包括 OOD 话语检测出错、理解模板分类出错、没有考虑上下文对话环境、以及系统其它模块出错导致的上下文不正确或缺失等。

为了验证本文提出的协处理方案，也即AIML、理解模板分类、EFSM三者相结合的方案。我们对比了另外两个方案（baseline methods）：

- (1) 方案 1：单纯 AIML，采用该方案的对话系统只使用纯 AIML 机制应对 OOD 话语；
- (2) 方案 2：采用 AIML 机制，并对未匹配话语，采用理解模板分类算法。

4.2.3 测试结果

本文提出的协处理方案以及两个对比方案的处理满意率如表4所示。

表 4 各方案的处理满意率对比

方案	处理满意的OOD话语	满意率
AIML	170	41.3%
AIML+分类	227	55.1%
AIML+分类+EFSM	292	70.9%

从表 4 可以看到，本文提出的协处理方案比只有理解模板（AIML+模板分类）的方案，以及面向聊天的对话系统常用的纯 AIML 方案分别提高 15.8%和 29.6%的 OOD 话语处理满意率。如果去除由于主处理传递过来的 64 处错误（主要包括 OOD 话语检测错误、对话上下文缺失和候选属性缺失），协处理的处理满意度达到 84%，表明了协处理器能有效地辅助对话系统完成限定领域完整对话流程。

4.2.4 不满意 OOD 话语处理分析

用户体验不满意的120个OOD话语应答的原因分类如表5所示。

表 5 OOD 话语处理不满意分类

不满意原因	错误数量	占错误比例
OOD话语检测错误	19	15.8%
上下文缺失错误	32	26.7%
候选属性缺失	13	10.8%
模板分类错误	56	46.7%

从表5可以看到，目前用户不满意的OOD话语处理有接近一半是由于理解模板分类错误造成的，进一步研究和完善未匹配模板分类的模板类别语料库，可望提高分类算法的分类正确率，降低这类的不满意处理比例。另外的一半属于主处理带来的错误，包括OOD话语检测错误、对话上下文缺失和候选属性缺失三类。其中OOD话语检测错误可望在进一步完善主处理的OOD话语检测（如进一步完善领域语义信息库）后得到有效降低；候选属性缺失经核查主要是由于部分手机的属性值缺失造成，解决方法是进一步完善手机商品信息抓取模块；而对话上下文缺失的错误原因相对比较复杂，需要深入分析主处理各模块导致某一（类）特定上下文的错误逐步改善。

5 结束语

本文基于 AIML 和 EFSM，提出了一种限定领域口语对话系统的协处理方法。在中文手机导购领域的测试表明，本文的方案能有效地辅助对话系统完成限定领域完整对话流程。方案中采用理解模板的方式，实现了多种同类的用户输入归入相同的理解模板，解决开放语义、口语化话语以及表达多样性的问题；而对未匹配话语进行理解模板分类，则在一定程度上解决了理解模板覆盖完整性的困难，并且随着语料库的增量训练，分类性能还可以持续提高。另一方面，通过综合理解模板、系统事件和上下文信息为条件的模式转换，实现理解模板到应答模板的映射以及状态的转换，改变了单纯依靠模板匹配方法在对话上下文控制方面的相对缺失。

未来工作一方面通过分析协处理测试出来的主处理错误，反馈至主处理器进行有针对性的系统完善，减少主处理带给协处理的错误；另一方面是进行更大范围的测试，根据测试结果进一步完善语料库，包括直接匹配的 AIML 语料库和模板类别语料库，提高协处理器的适应性。

参考文献

- [1] Price P J. Evaluation of spoken language systems: the ATIS domain[C]//Proceedings of DARPA Workshop on Speech and Natural Language, Hidden Valley, PA, 1990.
- [2] Gorin A, Riccardi G, Wright J. How may I help you?[J], Speech Communication,1997, 23(1-2): 113-127.
- [3] Zue V, Seneff S, Glass J, et al. JUPITER: a telephone-based conversational interface for weather information[J]. IEEE Transactions on Speech and Audio Processing, 2000, 8(1): 85-96.

- [4] 黄寅飞, 郑方, 燕鹏举, 徐明星, 吴文虎. 校园导航系统EasyNav的设计与实现[J]. 中文信息学报, 2001, 15(4): 35-40.
- [5] Durston P, Farrell M, Attwater D, et al. OASIS natural language call steering trial[C]//Proceedings of 7th European Conference on Speech Communication and Technology (Eurospeech 2011), 2001: 1323-1326.
- [6] 张琳, 高峰, 郭荣, 等. 汉语股票实时行情查询对话系统[J]. 计算机应用, 2004, 24(7): 61-63.
- [7] Pappu A, Rudnicky A. The structure and generality of spoken route instructions[C]//Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL 2012), 2012: 99-107.
- [8] Reichel C S, Sohn J, Ehrlich U, et al. Out-of-domain spoken dialogs in the car: a WoZ study[C]//Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL 2014), 2014: 12-21.
- [9] Wallace R S. A.L.I.C.E. artificial intelligence foundation[EB/OL]. [2015-08-10]. <http://www.alicebot.org>.
- [10] Banchs R E, Li H. IRIS: a chat-oriented dialogue system based on the vector space model[C]//Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (ACL 2012), demo session.
- [11] Mollá D, González J L V. Question answering in restricted domains: an overview[J]. Computational Linguistics, 2007, 33(1): 41-61
- [12] Ameixa D, Coheur L, Fialho P, et al. Luke, I am your father: dealing with out-of-domain requests by using movies subtitles [J]. In: T. Bickmore et al. (Eds.) IVA 2014. LNCS (LNAI), vol. 8637, pp. 13-21. Springer, Heidelberg (2014)
- [13] Metallinou A, Bohus D, Williams J D. Discriminative state tracking for spoken dialog systems[C]//Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013), 2013: 466-475.
- [14] Xu W Q, Xu B, Huang T Y, et al. Bridging the gap between dialogue management and dialogue models[C]//Proceedings of the Third SIGdial Workshop on Discourse and Dialogue (SIGDIAL 2002), 2002: 201-210.
- [15] 邬晓钧, 郑方, 徐明星. 基于主题森林结构的对话管理模型[J]. 自动化学报, 2003. 29(2):275-283.
- [16] Peltason J, Wrede B. Pamini: a framework for assembling mixed-initiative human-robot interaction from generic interaction patterns[C]//Proceedings of the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL 2010), 2010: 229-232.
- [17] Lane I R, Kawahara T, Matsui T, et al. Out-of-domain utterance detection using classification confidences of multiple topics[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2007, 15(1):150-161.
- [18] Tür G, Deoras A, Hakkani-Tür D. Detecting out-of-domain utterances addressed to a virtual personal assistant[C]//Proceedings of the 15th Annual Conference of the International Speech Communication Association (INTERSPEECH 2014), 2014: 283-287.
- [19] Celikyitmaz A, Hakkani-Tür D, Tür G. Approximate inference for domain detection in spoken language understanding[C]//Proceedings of the 12th Annual Conference of the International Speech Communication Association (INTERSPEECH 2011), 2011: 1293-1296.
- [20] Weizenbaum J. ELIZA—a computer program for the study of natural language communication between man and machine[J]. Communications of the ACM, 1966, 9(1):36-45.
- [21] Colby K M, Weber S, Hilf F D. Artificial paranoia[J]. Artificial Intelligence, 1971, 2(1): 1-25.
- [22] Schumaker R P, Chen H. Interaction analysis of the ALICE chatterbot: a two-study investigation of dialog and domain questioning[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part A, 2010, 40(1): 40-51.
- [23] 清华大学图书馆智能机器人小图[EB/OL]. [2015-08-10]. <http://166.111.120.164:8081/programd/>.
- [24] 小I机器人[EB/OL]. [2015-08-10]. <http://www.xiaoi.com/index.html>.

- [25] Schumaker R P, Chen H. Leveraging question answer technology to address terrorism inquiry[J]. Decision Support Systems, 2007, 43(4): 1419-1430.
- [26] Jia J Y. CSIEC: a computer assisted English learning chatbot based on textual knowledge and reasoning[J]. Knowledge-Based Systems, 2009, 22 (4): 249-255.
- [27] Crutzen R, Peters G Y, Portugal S D, et al. An artificially intelligent chat agent that answers adolescents' questions related to sex, drugs, and alcohol: an exploratory study[J]. Journal of Adolescent Health, 2011, 48(5):514–519.
- [28] Huang J Z, Zhou M, Yang D. Extracting chatbot knowledge from online discussion forums[C]//Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007), 2007: 423-428.
- [29] Banchs R E. Movie-DiC: a movie dialogue corpus for research and development[C]//Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (ACL 2012), 2012: 203–207.
- [30] Wallace R S. The anatomy of A.L.I.C.E. [EB/OL]. [2015-08-10]. <http://www.alicebot.org/anatomy.html>.
- [31] Huang P J, Lin X M, Lian Z Q, et al. Ch2R: a Chinese chatter robot for online shopping guide[C]//Proceedings of the 3rd CIPS-SIGHAN Joint Conference on Chinese Language Processing (CLP-2014), 2014: 26-34.
- [32] 黄民烈, 朱小燕. 对话管理中基于槽特征有限状态自动机的方法研究[J]. 计算机学报, 2004, 27 (08): 1092-1101.
- [33] Chen Y N, Wang W Y, Rudnicky A. Unsupervised induction and filling of semantic slots for spoken dialogue systems using frame-semantic parsing[C]//Proceedings of the 2013 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), 2013.

作者简介:



王俊东（1992—），硕士研究生，主要研究领域为自然语言处理。

Email: jdwang@stu.scau.edu.cn



黄沛杰（1980—），通讯作者，博士，副教授，主要研究领域为人工智能、自然语言处理、口语对话系统。

Email: pjhuang@scau.edu.cn



林仙茂（1990—），本科，主要研究领域为口语对话系统。

Email: xianmaulin@gmail.com