

特定域任务型人机对话测评报告

邱小虎 陈诗芸 张辉敏*
(上海葡萄纬度科技有限公司 上海 201100)

摘要

本文论述多轮对话在特定域中的应用，设计了一个独立于领域的任务型对话系统，可以移植于不同的域中。本文以任务二机票、火车票，酒店查询预订为例，阐述该系统的架构以及各个组成部分。

关键词 framed-based DM 意图识别 对话管理 基于上下文信息抽取 多域中控决策

1 引言

近年来，人机对话技术越来越受到学术界以及工业界的广泛关注。主要可以分为开放域聊天、任务驱动的多轮对话、问答和推荐。目前任务驱动的人机多轮对话已经在工业界中得到广泛应用。

在任务驱动的多轮对话中，用户通常具有明确的目的倾向，希望得到特定的信息或者服务，例如：订餐，订票，寻找音乐、电影或某种商品等等。该类任务具备一定的规律性，比较复杂，难以在单轮对话中得到有效地解决，可能需要分多轮进行陈述，用户也可能在对话过程中不断修改或完善自己的需求。此外当用户的陈述的需求不够具体或明确的时候，机器也可以通过询问、澄清或确认来帮助用户找到满意的结果。

在多域的情况下，还涉及域之间的切换等复杂问题。衡量一个任务型的对话系统的好坏，通常有任务完成率（越高越好）、用户满意度、回复语言的自然度、对话轮数（越少越好）和资源未覆盖情况的引导能力等指标。

2 任务型对话系统设计

针对人机对话任务二-任务型人机对话本文设计了如图 2 所示的架构。针对跨域（多域切换）问题，增加了中控决策模块。下文会对各个模块进行介绍。

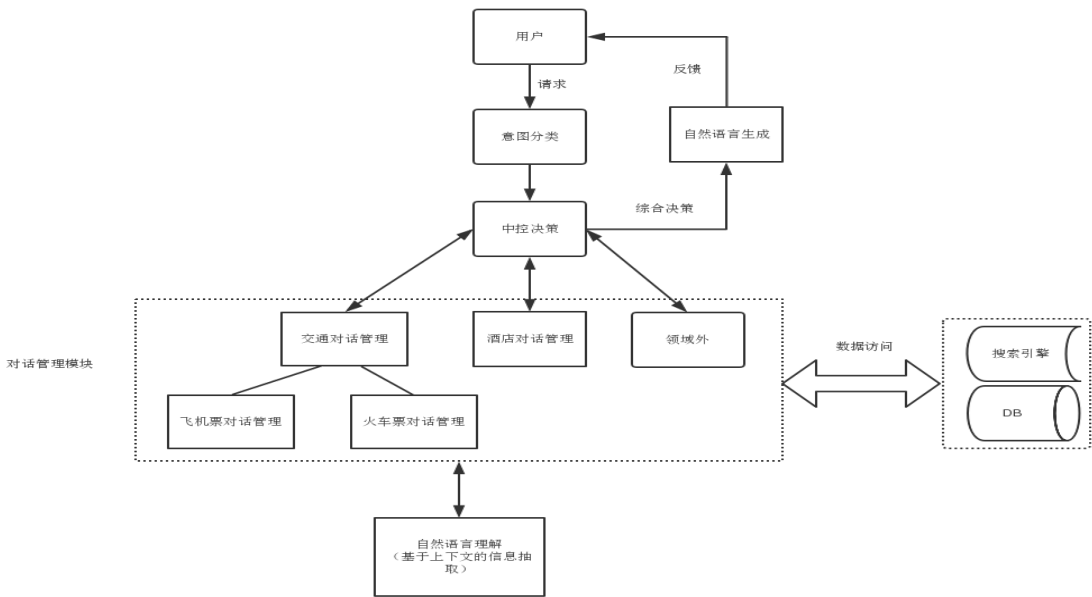


图 2. 任务二任务型人机对话架构

通讯作者：张辉敏 zhanghm0310@163.com

2.1 自然语言理解模块

用户的输入首先会到自然语言处理(NLP)模块中进行分析(例如:分词、词性标注,句法分析等),随后自然语言理解模块进行对话意图识别(Dialog Act Recognition)、命名实体识别(Named Entity Recognition),语义角色识别(Semantic role Recognition)等把用户的输入转为语义表示。

任务二中自然语言理解,需要确定用户的任务意图(订机票,火车票或者酒店),并根据上下文特征抽取相应的语义角色标记。

2.1.1 意图识别模块

识别用户的任务行为意图。任务二包含飞机票、火车票、酒店意图。针对该应用,我们把飞机票和火车票类目规约成交通类。我们定义了交通类别(包括飞机票,火车票)、酒店类别,域外3个大类。对于交通类别,进一步细分为飞机票、火车票或交通工具未定(需进一步询问)。

该模块复用了任务一的相关意图识别语料,并再人工标注了一部分训练数据,通过svm(支持向量机)模型来做多分类,使用了词、Bigram(二元语法模型)和字典等作为分类特征。

意图识别示例:

例1:我想从北京到上海去

系统判别为交通意图

例2:预订上海东方明珠附近的酒店

系统判别为酒店意图

2.1.2 语义角色抽取模块

该模块根据用户输入及上轮系统的决策输出,使用**序列标注模型**抽取我们需要的语义角色。如飞机票中的出发地、目的地和起飞时间等。并根据不同的状态,给出抽取的置信度得分(0-1之间的连续值),值越高代表抽取的置信度越高。语义角色抽取模块流程图如图3所示。

在该序列标注模型中我们采用了词、词性、窗口词,词典和上一轮的系统请求等作为特征。

该模块特点:

- 考虑上下文等多维度特征
- 考虑实体检索打分
- 支持模糊匹配特征

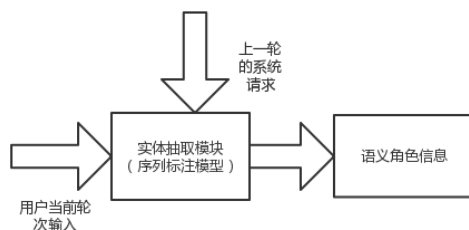


图3 语义角色抽取模块

语义角色抽取示例:

系统:亲爱的你要到哪去

用户:上海

抽取出的语义角色: 上海{语义角色: 目的城市, 置信度得分:1.0}

系统:亲爱的你要订哪个城市的酒店

用户:上海东方明珠附近的酒店,明天的

抽取出的语义角色: 上海{语义角色: 入住城市, 置信度得分:1.0}, 东方明珠{语义角色: 地标, 置信度得分:1.0}, 明天{语义角色: 入住时间, 置信度得分: 0.5}

2.2 对话管理模块

开发人机交互系统的一个核心问题是对话管理策略, Traum and Larson(2003)^[3] 阐述了对话管理系统主要包含4个任务。

1. 更新对话上下文

2. 对上下文进行解释
3. 与其它模块进行协作
4. 决定传递什么信息给用户

相对于 Finite state-based dialog managers (有限状态对话管理：只能问什么答什么，并按照固定的次序询问) 的对话策略，本系统采用了 Framed-based dialog managers (每轮能任意次序填充字段) 的灵活的对话策略。

系统设定了对话注册信息 (Dialog register) 数据结构，后文简称 DR，存储着各域在信息抽取模块中抽取出的历史语义角色词。针对错误处理问题我们还使用了置信值 (0.0-1.0 之间的连续值) 来记录字段的填充情况，值越高代表填充置信度越高。本系统设定的对话管理框架如图 3 所示。

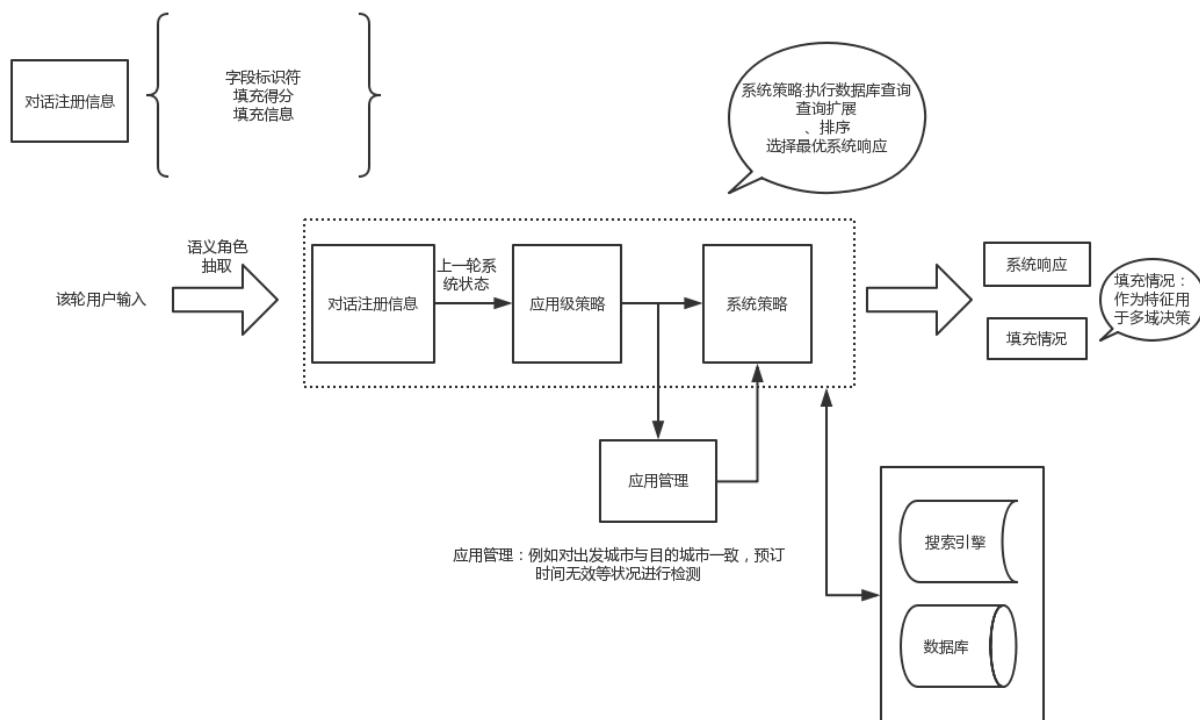


图 4. 对话管理架构图

2.2.1 对话策略模块

对话策略模块是对话管理中核心模块，本系统把对话策略分为如下两部分。

应用级策略：进行约束判断（如判断飞机票出发时间是否已过，出发地目的地是否一致等约束）、决策是否进行数据库查询。

系统策略：通过应用管理的输出、DR 填充情况和数据库查询结果等因素，来决定系统对用户的最终响应，如输出询问字段请求。本文把这个问题当成一个多标签分类 (Multi-label classification) 问题，以减少对话轮数做为主要目标。系统响应主要分为以下几类：

- ask (询问) 语句。对某些字段进行询问。例：ask_fromCity：询问用户出发地
- confirm (确认) 语句。对某些填充置信度不高的字段或者预订情况进行确认。例：confirm_fromCity (上海)：确认出发地是否上海。其中 confirm 又包含显式或隐式的 confirm。
- option (选项) 语句。给出候选集让用户进行选择。例：option_cabin (头等舱，经济舱，商务舱)：让用户在头等舱、经济舱，商务舱中做出选择。能有效的对用户进行引导。

交互模块是与对话管理与其它模块进行衔接的重要模块，在对话中需要与搜索引擎和数据库进行查询交互。

- 搜索引擎：使用 Lucene 建立酒店搜索引擎。搜索引擎提供酒店地标检索、根据经纬度进行空间检索等功能。并可以根据检索结果做搜索排序（考虑编辑距离、公共子串，地标距离等多维度综合打分）选出最相关的酒店信息给最终用户。
- Mysql 数据库：提供机票，火车票等查询服务。
- 本系统在进行数据库查询时，对无结果或数据库未覆盖用户需求时，采用了**查询扩展**的策略，能有效对用户进行引导。

2.3 中控决策模块

在完成特定任务时，用户有可能意图发生变化。如订飞机票时，觉得价格不合适，转订火车票，这时系统需要判断用户的真实意图，并给以相应的反馈。本系统使用了**判别式多分类模型**，使用单轮意图识别结果、抽取任务完成率等作为特征，综合做出决策。

综合决策示例：

用户：订北京到上海的机票

系统：好的，您要订北京到上海的机票，亲爱的，你要订哪天的

用户：机票太贵了，看看火车票吧

系统：亲爱的，你是要订北京到上海的火车票吗 （综合决策）

2.4 自然语言生成

Natural language generation (NLG) 的任务是将对话策略输出的语义表示转化成自然语言的句子，反馈给用户。

为准确生成流畅的对话，本系统使用模板填槽模式，并根据句子的类型制定相应的模板，可方便的进行增删以及回复语句修改。

以下为询问订机票中用户出发地与目的地的自然语言生成模板示例

机票请求:ask_fromCity, ask_toCity 对应的模板为“亲爱的请问您要订哪到哪的机票”

参考文献

- [1] David Griol, Zoraida Callejas, Ramón López-Cózar, and Giuseppe Riccardi. A domain-independent statistical methodology for dialog management in spoken dialog systems. 2014.
- [2] Cheongjae Lee, Sangkeun Jung, Kyungduk Kim, Donghyeon Lee, and Gary Geunbae Lee. Recent Approaches to Dialog Management for Spoken Dialog System. Journal of Computing Science and Engineering, Vol. 4, No. 1, March 2010, Pages 1-22.
- [3] Traum DR, Larsson S (2003) The information state approach to dialog management. In: Smith R, Kuppevelt J (eds) Current and new directions in discourse and dialog. Kluwer Academic Publishers, Dordrecht: 325 - 353. doi:10.1007/978-94-010-0019-2_15
- [4] Sarawagi S. Information Extraction[J]. Databases, 2007, 1(3): 261-377.
- [5] Lafferty, J., A. McCallum, and F. Pereira. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. ICML 2001.
- [6] Ye-Yi Wang and Alex Acero. Discriminative Models for Spoken Language Understanding. International Conference on Spoken Language Processing, pp. 1766-1769, ISCA, Pittsburgh, PA, USA, 2006