

状态转移控制的多领域任务型对话系统

刘伟杰 劳雅迪 高升

北京邮电大学信通院

摘 要:

本文将介绍任务二中基于规则搭建的面向任务型的多轮问答系统。领域划分为机票、动车票以及酒店。本系统主要分为自然语言处理模块(NLU)、对话状态管理(DM)以及自然语言生成模块(NLG)。

1 简介

人机对话系统是一种类似于人与人对话的、持续的、前后相关的、双向的信息交流。而在近十年里,任务驱动的对话系统是最主要的研究领域。为了完成面向任务的对话,一般将对话系统分为三个模块,即自然语言处理模块(NLU)、对话状态管理(DM)以及自然语言生成模块(NLG)。其中 NLU 是通过解析用户问句来识别对话领域(如订机票、订旅馆等)以及意图分类和对话槽位填充(slot-filling)。领域和意图分类都可以建模成短文本分类问题,如传统机器学习中的 SVM,深度学习中的循环神经网络(RNN)、卷积神经网络(CNN)等深度模型都在文本分类问题上取得了不错的效果。槽位填充则可以当做序列标注问题,即使用 IOB(in-out-begin)的格式来表示槽位标签。常用的模型有隐马尔科夫模型(HMM)、条件随机场(CRF),以及和 RNN 结合使用的深度模型等。自然语言生成模块则一般使用基于模板的规则生成和基于 seq2seq 的端到端生成器。

2 系统设计

本文介绍的针对特定域的任务型人机对话系统主要由三大模块组成:1)语言理解模块,包括用户意图分类以及捕获用户输入的有效信息,从而调整对话策略与输出应答;2)对话管理模块,包括状态追踪以及决策管理等功能,控制整个多轮对话的进行;3)语言生成模块,根据当前的完成进度与状态,生成回答返回给用户。

针对特定域的任务型人机对话系统的结构组成如图 1 所示。用户的对话意图 A_u 可以定义为一个用语表达某种意向的具有槽/值对内部结构的信息框架形式【5】,表示为:

$$A_u = \{I_{u_1}: [Slot_1 = I_{t_1}, Slot_2 = I_{t_2}, \dots, Slot_n = I_{t_n}], I_{u_2}: [Slot_n = I_{t_n}, \dots], \dots, I_{u_n}\}$$

其中 I_{u_i} 代表用户的对话意图,本任务型人机对话系统共涉及包括 3 种意图:机票 $flight$,火车票 $train$ 与住宿 $hotel$ 的购买与预订,一个对话中用户可能有多个意图或混合意图。

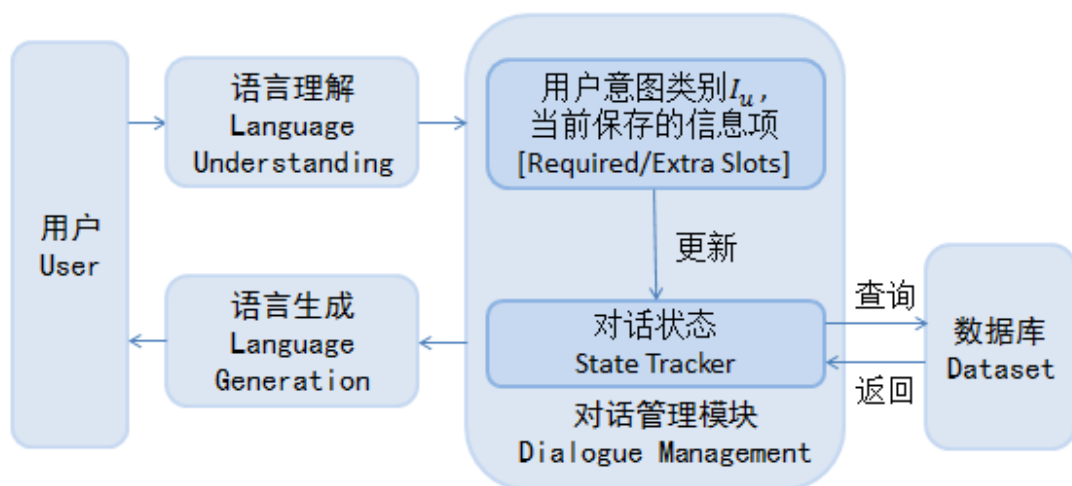


图 1 针对特定域的任务型人机对话系统的结构组成图

$Slot_i-I_t$ 为槽/值对，包含两类槽位：1）必需信息项，如用户需要购买机票时，起始城市等信息不可缺少，故为必须信息；2）可缺失信息项，如购买机票时，缺失航空公司信息亦能给用户推荐航班，但如果用户提出航空公司的要求，则在返回结果时将考虑该信息。此类槽位在对话中途作保存，但不是必须询问项。

2.1 语言理解模块

该人机对话系统用于解决机票，车票与住宿预订等三个特定领域的任务型人机对话，故语言理解模块的主要任务是领域分类，意图理解和槽位填充。即填充一组槽来形成语义框架，并获取到用户提及的信息项 I_t 。如下图：

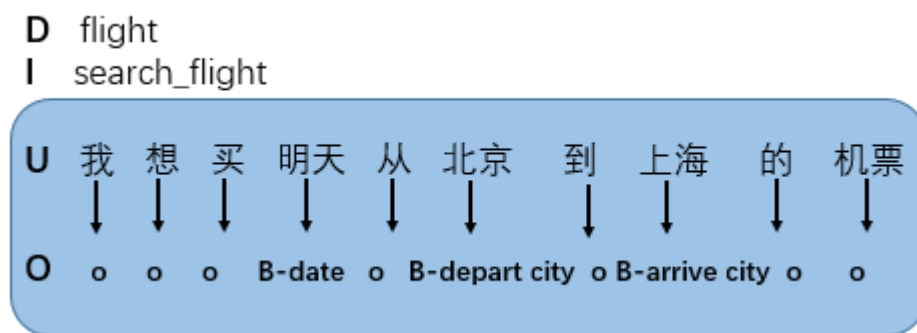


图 2. NLU 模块示例句子

首先系统处于无明确意图状态，获取到用户输入语句 S_u 后，将尝试识别用户意图 I_u 。如 S_u = “我想买一张明天去北京的机票”，则用户意图 $I_u = flight$ ，之后的对话将在此领域内完成槽值填充，直到通过多轮对话将该意图 I_u 的各个槽位 $Slot_t$ 的信息项 I_t 获取完全。该意图完成后或者在对话中途，用户随时可以更改意图，此时切换意图状态，开始对另一列槽位进行填充。如 S_u = “我还想预订到达之后的入住酒店”，此时用户意图 $I_u = hotel$ ，接下来对话将围绕住宿预订这一任务进行。

用户意图确认之后，需要获取到句中的有效信息。 S_u 中包含的信息项 I_t 有： $I_{t_{DepartTime}} =$ 明天， $I_{t_{ArriveCity}} =$ 北京。

2.2 对话管理模块

任务驱动的多轮对话是一个决策过程，需要在对话过程中不断根据用户的意图与当前状态来制定对话策略并不断更新，从而帮助用户高效自然地完成对话的意图或目的。

对话策略分为 3 类^[2]：系统主导、用户主导和混合主导。系统主导的对话时指由系统向用户提出一系列的问题，根据用户的回答来提供信息。用户主导方式是指在对话过程中，用户是对话的主导者，可以非常自由地按照自己的意愿来提问。本系统采用的是混合主导方式，系统针对需要获知的信息与当前策略向用户提问，用户也可以自由地提出自己的要求^[3]。

在获取到用户的具体意图类型后，开始填充槽位。整个对话流程共有 5 个状态：

- 1) 初始状态。此时用户意图 I_u 未知，对话策略为询问用户意图。一般此类情况出现在对话刚开始，用户的问候用于如“你好”等。

对话 1:

User: 明天我想去上海。(I_u 未知，获取并记录 $I_{t_{ArriveCity}} =$ 上海， $I_{t_{DepartDay}} =$ 明天)

Chatbot: 您好，需要订去上海的机票/车票，还是需要上海的住宿预订？

User: 机票 (获取到用户意图 $I_u = flight$ ，状态转为槽位填充状态)

- 2) 槽位填充状态。根据当前的用户意图 I_u ，确立当前系统任务 T_s 。系统任务 T_s 的完成前提是获取到所有的必需信息项，此状态下系统应不断提示询问，得到用户的补充和修正，多回合后构成一个完整的信息查询。

对话 2:

Chatbot: 您需要从哪个城市飞往上海？

User: 从北京，中午左右出发，要南航的。($I_u = flight$ ， $I_{t_{DepartCity}} =$ 上海， $I_{t_{DepartTime}} = 12:00$ ， $I_{t_{AirLine}} =$ 南方航空。此处 $I_{t_{AirLine}}$ 为额外信息项，假若用户尚未提及，系统默认对此无要求。后续如果不满意可以提出对此信息的补充，系统会更改推荐结果)

Chatbot: 请问您想定什么舱位呢？

User: 公务舱吧 ($I_{t_{CabinInfo}} =$ 公务舱，此时所有的必需信息项获取完全，转为结果推荐状态)

- 3) 结果推荐状态。上述过程中将一直记录当所有必需信息项 I_t ，当用户提供了 I_u 所需要的所有必需信息项 I_t 后，将开始从数据库检索查询，并展示给用户符合要求的信息。

对话 3.1:

Chatbot: 为您查询到的从北京飞往上海的航班信息如下，请问您需要预定第几个航班？

User: 第三个吧/CZ6412 (用户根据推荐结果与自身需求，选择具体航班号等)。

对话 3.2:

User:我想要到虹桥机场的 (I_{taPort} =虹桥国际机场是额外信息项, 一般不作询问, 此时获取到则更新检索条件, 重新推荐)

4) 用户确认

对话 4.1:

Chatbot:即将帮您预定 2017 年 4 月 19 日从北京到上海的机票, CZ6412, 12:10:00 出发, 7.3 折, 910 元。确认要预定吗? (转为用户确认状态)

User:确定 (任务完成, 转为结束状态)

对话 4.2:

User:有点太贵了, 帮我看一下火车票吧/我还要预定一下到达上海后入住的旅馆。(用户意图转换 $I_u = train/hotel$)

Chatbot:您需要订什么车型? /您想要住在上海的什么位置附近呢? (转为槽位填充状态)

5) 结束状态, 初始化所有信息。

2.3 语言生成模块

将对话管理模块维护的信息转换成 SQL 语句进行查询。对于酒店类, 则调用了百度经纬度 API, 对该城市所有的酒店计算距离, 取前五个最近的酒店进行推荐。同时, 我们预设了对话模板。当用户的对话动作可以在预定义的答案模板中找到时, 则调用基于模板生成的 NLG 模型, 把对应从记忆模块或者知识库中索引到的槽位及槽位值去替换占位符。

参考文献:

- [1] Young S. Talking to machines (statistically speaking). In Proc. ICSLP 02, Denver, USA, 2002:9-16
- [2] 口语对话管理综述. 王菁华, 钟义信, 王枞, 等. 计算机应用研究, 2005, 22(10):5-8.
- [3] 基于实例推理的人机对话系统的设计与实现. 姚琳, 梁春霞, 张德干. 计算机应用, 2007, 27(3):765-768.
- [4] Xiujun Li, Yun-Nung Chen, Lihong Li, Jianfeng Gao. 2017. End-to-End Task-Completion Neural Dialogue Systems. *arXiv: 1703.01008v2*.
- [5] 面向任务口语对话系统中期待模型的实现算法. 刘蓓, 杜利民, 于水源. 电子与信息学报, 2004, 26(11):1721-1727.