

doi:10.11835/j.issn.1000-582X.2022.07.008

# 基于自然语言处理的旅游景区智能讲解系统研究

刘文静<sup>1</sup>, 蔡章利<sup>2a,2b</sup>, 卢海<sup>2a,2b</sup>

(1. 重庆航天职业技术学院 人文社科系, 重庆 400000;

2. 重庆大学 a.信息物理社会可信服务计算教育部重点实验室; b.自动化学院, 重庆 400044)

**摘要:**智慧旅游建设的目的是要用新一代信息技术提升旅游服务水平和改善游客旅游体验。针对现有的旅游景区电子讲解系统缺乏自然语言互动交流能力的问题,研究一种基于自然语言处理和本体知识检索技术的旅游景区智能讲解系统,以便旅游景区能为游客提供自然语音对话式的个性化智能讲解服务。笔者设计了系统的物理架构和逻辑架构,然后分析实现文本向量化处理、命名实体识别、对话策略管理和讲解知识检索等关键技术,最后通过开发、测试系统原型,验证了系统的可行性。

**关键词:**自然语言处理;智慧旅游;智能讲解系统;本体;知识检索

中图分类号:TP391

文献标志码:A

文章编号:1000-582X(2022)07-070-09

## Intelligent tourism explanation system based on natural language processing

LIU Wenjing<sup>1</sup>, CAI Zhangli<sup>2a,2b</sup>, LU Hai<sup>2a,2b</sup>

(1. Humanity and Sociology Department, Chongqing Aerospace Polytechnic, Chongqing 400000, P. R. China;

2a. Key Laboratory of Dependable Service Computing in Cyber Physical Society, Ministry of Education;

2b. College of Automation, Chongqing University, Chongqing 400044, P. R. China)

**Abstract:** The goal of smart tourism construction is to improve tourism service and tourist experience by using the latest information technologies. To enable tourists to dialog with the electronic explanation system of scenic spots by natural voice, an intelligent tourism explanation system based on natural language processing and ontology knowledge retrieval was proposed in this study. Firstly, the physical view and logical view of the system architecture was designed. Secondly, the key technologies of text vector, named entity recognition, dialogue policy management and tourism knowledge retrieval were analyzed. Finally, a prototype was developed, verifying the feasibility of the system.

**Keywords:** natural language processing; smart tourism; intelligent explanation system; ontology; knowledge retrieval

收稿日期:2020-06-12 网络出版日期:2020-09-21

基金项目:国家自然科学基金资助项目(61771077);中央高校基本科研业务费专项基金(2019CDQYTX012)。

Supported by National Natural Science Foundation of China (61771077), and the Fundamental Research Funds for the Central Universities (2019CDQYTX012).

作者简介:刘文静(1981—),女,主要从事信息管理方向研究,(E-mail) 624372375@qq.com。

通信作者:蔡章利,男,副教授,主要从事智能信息处理方向研究,(E-mail) czl@cqu.edu.cn。

随着社会经济的快速发展,人们对生活质量的要求越来越高,外出旅游的人越来越多。为了更好地促进中国旅游业高质量发展,国家旅游局2015年1月发布了《关于促进智慧旅游发展的指导意见》,指出“智慧旅游是运用新一代信息网络技术和装备,充分准确及时感知和使用各类旅游信息,从而实现旅游服务、旅游管理、旅游营销、旅游体验的智能化,促进旅游业态向综合性和融合型转型提升”<sup>[1]</sup>。

游客到达旅游地后,景区导游的重要职责是讲好景点故事,引导游客进入观赏意境,提升旅游体验。好的讲解服务能带给游客好的旅游体验,增长旅游知识!由于景区导游自身知识水平、责任心和服务收费等原因,游客往往难以享受到高质量、低成本的个性化讲解服务,更多是一大群人跟着导游走马观花式地观赏,旅游沉浸感不强。为了提升讲解服务质量和降低服务成本,国内不少景区建设了电子讲解系统,研究人员提出了一些系统实现方案<sup>[2-4]</sup>,市场上也有美景听听、三毛游FM、旅游精灵等智能手机APP供游客下载使用。但这些系统目前都是通过红外感应/扫码/触屏启动的广播式解说,缺乏人与人之间的自然语音互动交流能力,游客参与感不强。

自然语言处理是人工智能的一个研究热点,目的是让人以自然语音或文本方式和智能设备进行交互,让机器理解人的意图并做出响应。随着语音识别技术的逐渐成熟,自然语言处理技术已开始应用于医疗健康、电商客服、智能家居、铁路售票、自动驾驶等领域。比如微软的小娜、苹果手机 Siri、百度搜索的小度、淘宝客服的小蜜等智能系统<sup>[5-6]</sup>。但在旅游景区讲解服务方面,笔者尚未看到自然语言处理应用于该领域。因此,笔者主要探讨如何综合应用自然语言处理、本体知识检索、移动互联网、云计算等技术实现一个支持游客用自然语言与其交互,能为游客提供个性化服务的旅游景区智能讲解系统。

## 1 系统架构设计

“五视图法”是信息系统架构设计常用的一种方法<sup>[7]</sup>。使用该方法设计了旅游景区智能讲解系统的物理架构和逻辑架构。

### 1.1 物理架构

随着云计算、大数据、人工智能、物联网、移动互联网等新一代信息技术的快速发展,支持“云大物移智”已成为信息系统设计遵循的一大准则。按照“云+端”模式设计的旅游景区智能讲解系统物理架构如图1所示。

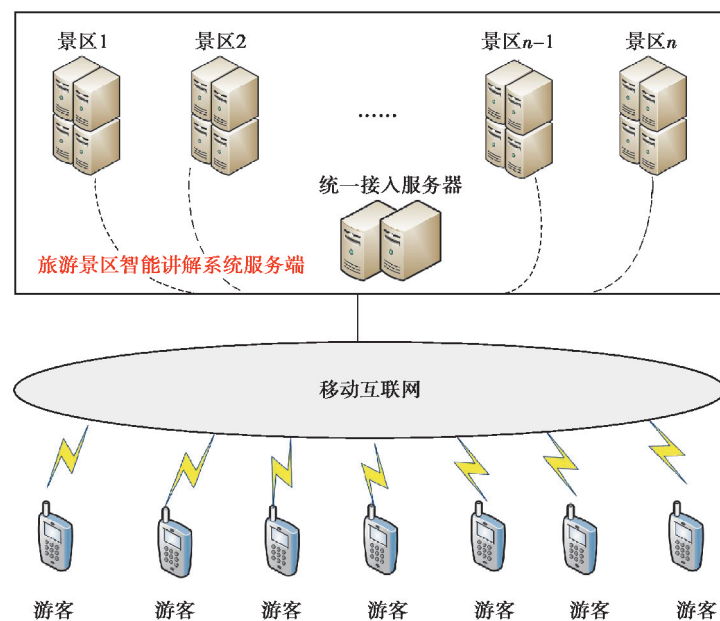


图1 系统的物理架构图

Fig. 1 Physical architecture of the system

在图 1 所示物理架构中,将整个系统分为游客移动端和讲解服务端 2 大部分,通过移动互联网连接起来。移动端指安装了本系统 APP 的游客智能手机;服务端由安装了本系统服务程序的云主机集群组成。为了让移动端 APP 具有通用性(即不局限于某个旅游景区才能使用),当 APP 启动时,系统首先与服务端的接入服务器通信,将定位数据发送给接入服务器查询出游客当前所处景区,然后返回该景区的讲解服务访问地址给移动端。移动端 APP 根据该地址直接与景区服务器通信完成后续讲解服务。为了兼顾各景区已有信息化投入和分流不同景区的游客并发访问压力,系统提供了 2 种方式供各景区选择。1)以独立访问域的方式接入平台,即各景区就近自建/租赁云主机并按本系统要求部署应用程序后将访问地址注册到接入服务器;2)以独立知识库的方式入驻平台,即各景区共用一套云主机和服务程序,但每个景区的讲解知识库相互独立。方法 1)的好处是不同景区的游客就近访问自己的服务器,可以减少网络拥堵和分流服务器访问压力,比较适合一些实力雄厚,游客量大的景区采用。方法 2)的好处是可以减少景区前期投入成本,在数据就是资产的当下拥有专属于自己的讲解知识库,可以自行决定是否免费提供讲解服务,比较适合一些规模较小的旅游景区采用。

## 1.2 逻辑架构

系统的逻辑架构如图 2 所示。从软件角度来看,系统分为 2 大部分。一是可安装于游客智能手机上的 APP 软件,二是部署在服务端云主机上的讲解服务程序。移动端 APP 是游客享受景点讲解服务的介质,提供了游客定位数据获取、游客问询语音输入、系统应答(包括讲解、问询、通知语音输出/图片展示/视频播放)等功能。服务端讲解服务程序主要完成语音识别、文本向量化处理、命名实体识别、对话状态跟踪、对话策略管理等自然语言处理功能,系统将根据游客意图到景区的讲解知识库中检索知识,最后生成讲解内容返回给移动端 APP。

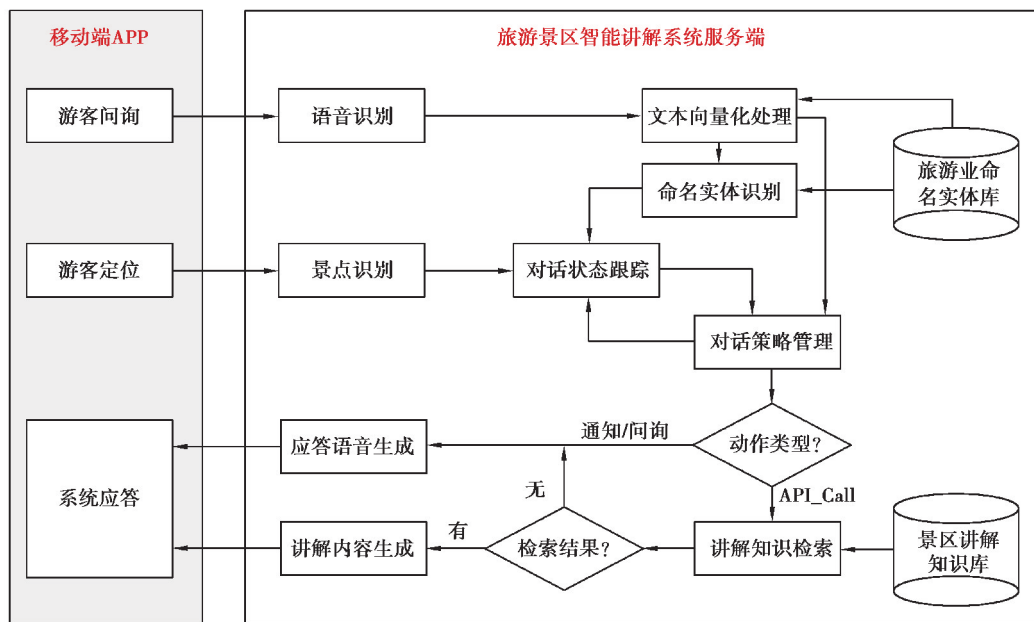


图 2 系统的逻辑架构图

Fig. 2 Logical architecture of the system

## 2 关键技术分析

如图 2 所示,运行于服务端的讲解服务程序是系统最重要部分,主要由语音识别、文本向量化处理、命名实体识别等功能模块组成。由于语音识别和生成技术目前已比较成熟,且有百度语音和科大讯飞等公司/机构提供的 SDK 包供使用,因此,研究主要探讨文本向量化处理、命名实体识别、对话策略管理、讲解知识检索 4 个核心功能模块的关键技术。

## 2.1 文本向量化处理

文本向量化处理的目的是要将语音识别得到的用户话语文本,经独热编码和词嵌入处理,表示为系统后续模块计算所需的实数域空间连续向量。词嵌入是指把一个独热编码得到的、维数为所有词数量的高维空间向量,映射为一个维数较低的实数域空间的连续向量,目前主要有 GloVe、Word2Vec、FastText、ELMo、OpenAI GPT、基于 Transformer 的双向编码表示(BERT, bidirectional encoder representations from transformers)等方法<sup>[8-9]</sup>。

BERT 是 Google 公司 2018 年提出的一种可用于多种语言处理的预训练模型,由多层 Transformer 编码器组成。由于 Google 发布的 BERT-base(Chinese)预训练模型是以字为粒度对中文进行切分的,没有考虑中文需要分词的特点,哈工大讯飞联合实验室于 2019 年 6 月发布了考虑中文分词的全词覆盖 BERT 中文预训练模型(简称 Chinese-BERT-wwm)<sup>[10]</sup>。文献<sup>[10]</sup>给出的实验结果表明:将 Chinese-BERT-wwm 用于机器阅读理解、自然语言推理、情感分类、句对匹配和文档分类等任务,性能均比 BERT-base(Chinese)提高了 1~2 个百分点。中文语义表示的最小单位是词,不是字。2 种模型相比,笔者认为后者更适用于研究系统,因此选用 Chinese-BERT-wwm 预训练模型来进行文本向量化处理。

## 2.2 命名实体识别

在本文系统中,命名实体识别(NER, named entity recognition)模块用于将用户话语中具有特定意义的实体(比如人名、地名、机构名、景点名、景物名、景观名、时间、数量、货币、比例数值等)辨识出来,填充到预定义好的语义槽中。该模块的输入是文本向量化模块输出的用户话语词向量,输出是辨识出的槽值对,其逻辑结构如图 3 所示<sup>[11]</sup>。

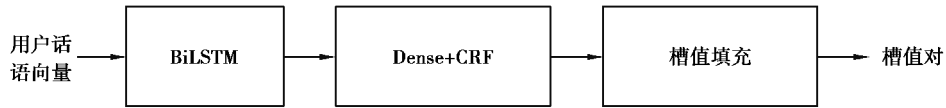


图 3 命名实体识别逻辑结构

Fig. 3 Logical structure of named entity recognition

在图 3 所示结构中,首先将用户话语词向量输入双向长短时记忆神经网络(BiLSTM, bi-directional long short-term memory)模型做编码处理,然后将处理结果输入条件随机场(CRF, conditional random field)模型计算,接着用 BIO 标签做序列标记,最后根据标记从序列中解析出对应不同语义槽的每个命名实体,将它们组合成槽值对输出。整个过程可抽象表示如下

$$\left. \begin{aligned}
 & \mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t, \dots, \mathbf{x}_n)^T, \\
 & in_t^{fw} = [h_{t-1}^{fw}, x_t]; in_t^{bw} = [h_{t-1}^{bw}, x_t], \\
 & \tilde{c}_t^{fw} = \tanh(in_t^{fw} * \mathbf{W}_c^{fw} + \mathbf{b}_c^{fw}), \\
 & \tilde{c}_t^{bw} = \tanh(in_t^{bw} * \mathbf{W}_c^{bw} + \mathbf{b}_c^{bw}), \\
 & c_t^{fw} = \text{sigmod}(in_t^{fw} * \mathbf{W}_i^{fw} + \mathbf{b}_i^{fw}) * c_{t-1}^{fw} + \\
 & \quad \text{sigmod}(in_t^{fw} * \mathbf{W}_f^{fw} + \mathbf{b}_f^{fw}) * \tilde{c}_t^{fw}, \\
 & c_t^{bw} = \text{sigmod}(in_t^{bw} * \mathbf{W}_i^{bw} + \mathbf{b}_i^{bw}) * c_{t-1}^{bw} + \\
 & \quad \text{sigmod}(in_t^{bw} * \mathbf{W}_f^{bw} + \mathbf{b}_f^{bw}) * \tilde{c}_t^{bw}, \\
 & h_t^{fw} = \text{sigmod}(in_t^{fw} * \mathbf{W}_o^{fw} + \mathbf{b}_o^{fw}) * \tanh(c_t^{fw}), \\
 & h_t^{bw} = \text{sigmod}(in_t^{bw} * \mathbf{W}_o^{bw} + \mathbf{b}_o^{bw}) * \tanh(c_t^{bw}), \\
 & \mathbf{H}_{fw} = (h_1^{fw}, h_2^{fw}, \dots, h_t^{fw}, \dots, h_n^{fw})^T, \\
 & \mathbf{H}_{bw} = (h_1^{bw}, h_2^{bw}, \dots, h_t^{bw}, \dots, h_n^{bw})^T, \\
 & \text{Score} = [\mathbf{H}_{fw}, \mathbf{H}_{bw}] * \mathbf{W}_{\text{dense}} + \mathbf{B}_{\text{dense}} \\
 & Y = \text{Viterbi}(\text{Score}, \text{Transitions}),
 \end{aligned} \right\} \quad (1)$$

式中: $\mathbf{X} \in R^{n \times d}$  是 BERT 预训练模型编码输出的表示用户话语的矩阵向量; $n$  是用户话语长度; $d=768$  是向

量维度;  $\mathbf{W}_i^{fw}, \mathbf{W}_o^{fw}, \mathbf{W}_f^{fw}, \mathbf{W}_c^{fw} \in R^{(d+m) \times m}$  是 BiLSTM 模型训练后的前向权值矩阵,  $\mathbf{b}_i^{fw}, \mathbf{b}_o^{fw}, \mathbf{b}_f^{fw}, \mathbf{b}_c^{fw} \in R^m$  是训练后的前向偏置向量;  $\mathbf{W}_i^{bw}, \mathbf{W}_o^{bw}, \mathbf{W}_f^{bw}, \mathbf{W}_c^{bw} \in R^{(d+m) \times m}$  是 BiLSTM 模型训练得到的后向权值矩阵,  $\mathbf{b}_i^{bw}, \mathbf{b}_o^{bw}, \mathbf{b}_f^{bw}, \mathbf{b}_c^{bw} \in R^m$  是后向偏置向量;  $m = 128$  是 BiLSTM 的隐层大小;  $\mathbf{W}_{dense} \in R^{2m \times k}$  和  $\mathbf{b}_{dense} \in R^{n \times k}$  是训练得到的全连接层(Dense)权重矩阵和偏置矩阵,  $k$  是标记数量;  $\mathbf{Transitions} \in R^{k \times k}$  是 CRF 训练后得到的标记转移矩阵。Y 是使用维特比算法 Viterbi( $\cdot$ )求得的、对应输入 X 的标记序列, 根据预定义槽解析 Y 即可完成槽值填充。

### 2.3 对话策略管理

对话策略管理是系统实现自然语言处理的另一核心模块, 主要完成游客意图识别、系统动作选择、系统动作生成等功能, 其输入是用户话语词向量和对话特征向量, 输出是待执行的系统动作, 其逻辑结构如图 4 所示。

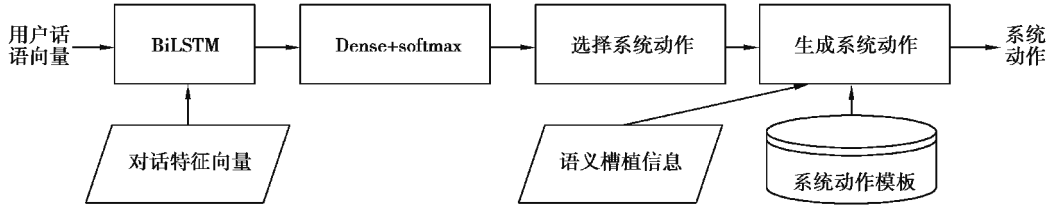


图 4 对话策略管理逻辑结构

Fig. 4 Logical structure of dialogue strategy management

首先将用户话语词向量与表示历史对话的对话特征向量拼接起来输入 BiLSTM 模型做编码处理, 然后对 BiLSTM 模型的输出做全连接(Dense)变换处理和 softmax 分类计算, 得到用户话语对应每个系统动作的概率, 其求取过程可抽象表示为式(2), 接着经过阈值比较筛选, 排序选出概率最大的系统动作, 最后根据系统动作模板和语义槽植信息生成完整的系统动作后输出<sup>[12]</sup>。

$$\left. \begin{aligned}
 \mathbf{X} &= (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t, \dots, \mathbf{x}_n)^T, \\
 in_t^{fw} &= [h_{t-1}^{fw}, \mathbf{x}_t]; in_t^{bw} = [h_{t-1}^{bw}, \mathbf{x}_t], \\
 \tilde{c}_t^{fw} &= \tanh(in_t^{fw} * \mathbf{W}_c^{fw} + \mathbf{b}_c^{fw}), \\
 \tilde{c}_t^{bw} &= \tanh(in_t^{bw} * \mathbf{W}_c^{bw} + \mathbf{b}_c^{bw}), \\
 c_t^{fw} &= \text{sigmod}(in_t^{fw} * \mathbf{W}_i^{fw} + \mathbf{b}_i^{fw}) * c_{t-1}^{fw} + \\
 &\quad \text{sigmod}(in_t^{fw} * \mathbf{W}_f^{fw} + \mathbf{b}_f^{fw}) * \tilde{c}_t^{fw}, \\
 c_t^{bw} &= \text{sigmod}(in_t^{bw} * \mathbf{W}_i^{bw} + \mathbf{b}_i^{bw}) * c_{t-1}^{bw} + \\
 &\quad \text{sigmod}(in_t^{bw} * \mathbf{W}_f^{bw} + \mathbf{b}_f^{bw}) * \tilde{c}_t^{bw}, \\
 h_t^{fw} &= \text{sigmod}(in_t^{fw} * \mathbf{W}_o^{fw} + \mathbf{b}_o^{fw}) * \tanh(c_t^{fw}), \\
 h_t^{bw} &= \text{sigmod}(in_t^{bw} * \mathbf{W}_o^{bw} + \mathbf{b}_o^{bw}) * \tanh(c_t^{bw}), \\
 \text{score} &= [h_n^{fw}, h_n^{bw}] * \mathbf{W}_{dense} + \mathbf{b}_{dense}, \\
 p &= e^{\text{score}} / \sum e^{\text{score}},
 \end{aligned} \right\} \quad (2)$$

式中:  $\mathbf{X} \in R^{n \times d}$  是 BERT 预训练模型编码输出的用户话语词向量  $\mathbf{x}_n$  与最近  $n-1$  轮对话中表示每轮对话状态的特征向量  $\mathbf{x}_i (1 \leq i \leq n-1)$  拼接所得的矩阵向量;  $d = 768$  是向量维度;  $\mathbf{W}_i^{fw}, \mathbf{W}_o^{fw}, \mathbf{W}_f^{fw}, \mathbf{W}_c^{fw} \in R^{(d+m) \times m}$  是 BiLSTM 模型训练后的前向权值矩阵,  $\mathbf{b}_i^{fw}, \mathbf{b}_o^{fw}, \mathbf{b}_f^{fw}, \mathbf{b}_c^{fw} \in R^m$  是训练后的前向偏置向量;  $\mathbf{W}_i^{bw}, \mathbf{W}_o^{bw}, \mathbf{W}_f^{bw}, \mathbf{W}_c^{bw} \in R^{(d+m) \times m}$  是 BiLSTM 模型训练得到的后向权值矩阵,  $\mathbf{b}_i^{bw}, \mathbf{b}_o^{bw}, \mathbf{b}_f^{bw}, \mathbf{b}_c^{bw} \in R^m$  是后向偏置向量;  $m = 128$  是 BiLSTM 的隐层大小;  $\mathbf{W}_{dense} \in R^{2m \times k}$  和  $\mathbf{b}_{dense} \in R^k$  是训练得到的全连接层(Dense)权重矩阵和偏置向量,  $k$  是预先定义系统动作模板数;  $p$  是使用 softmax 分类函数求得的当前用户话语对应每个预定义系统动作模板的概率向量。

如图 4 所示, 对话策略管理模块除了需要对话状态跟踪模块提供对话特征向量和语义槽植信息外, 还需开发人员预先定义系统动作模板。系统动作是用户意图的反映, 一个系统动作模板可视为用户话语要表达

的一种意图。将系统动作分为通知(Inform)、问询(Quest)、API\_Call 3 种类型。通知指系统在没有游客问询的情况下主动发出消息,比如致词欢迎游客到达景区等。问询指系统无法识别出游客意图或游客意图明确但信息缺乏或不模糊时,有针对性地引导游客提供信息。API\_Call 指系统准确识别出游客意图和所需语义槽值后向讲解知识管理模块发出的知识检索请求。用本体方法来构建旅游景区讲解知识库,本体的每种属性都可能成为游客问询的问题,因此 API\_Call 类型的系统动作模版可根据本体模型定义的属性来生成。

### 2.4 讲解知识检索

当系统识别出游客意图后,讲解知识检索模块将从景点讲解知识库中找出系统所需的讲解知识,将其输出给讲解内容生成模块,由该模块根据预先定义的讲解模板和讲解规则生成回复游客的讲解词。要检索知识需要先做好知识建模。在知识管理研究领域,知识建模目前最常用的是本体(Ontology)方法。本体是一个源于哲学的概念,哲学用其来描述世界上客观存在的所有人、事、物。在人工智能领域,最流行的本体定义是 Gruber 于 1993 年给出的“本体是共享概念的规范说明”<sup>[13]</sup>。按此定义,本体模型可抽象表示为三元组(概念、关系、属性);其中,概念是对应用领域中某类人、事、物的抽象表示,等同于面向对象软件开发方法中的类;关系用于反映概念间的继承、泛化等层次关系和组合、聚合、依赖、扩展、控制等关联关系;属性用于描述概念的静态特征。

在旅游信息化领域,研究人员设计了 Mondeca-Tourism-Ontology、Harmonis-Ontology、OTA-Specification、WDTourism 等本体模型来满足个性化旅游线路规划、旅游景点推荐、旅游活动推荐等旅游服务需求<sup>[14]</sup>。为了满足游客的个性化讲解服务需求,选用本体方法来建模旅游景区讲解知识。通过分析旅游讲解服务特点本文使用 Protégé 工具定义了景区、景点、景物、景观、人物、事件、民族、地方特产、文化作品、节庆活动等 45 个类(Class),定义了拥有、隶属、写作、品尝、到访、居住、拍摄等 14 种反映类间关联关系的对象属性(Object Property);定义了名称、经度、纬度、创建时间、面积大小、高度、长度、价格等 56 种数据属性(Data Property)来反映类的静态特征。研究所构建的旅游本体模型如图 5 所示。

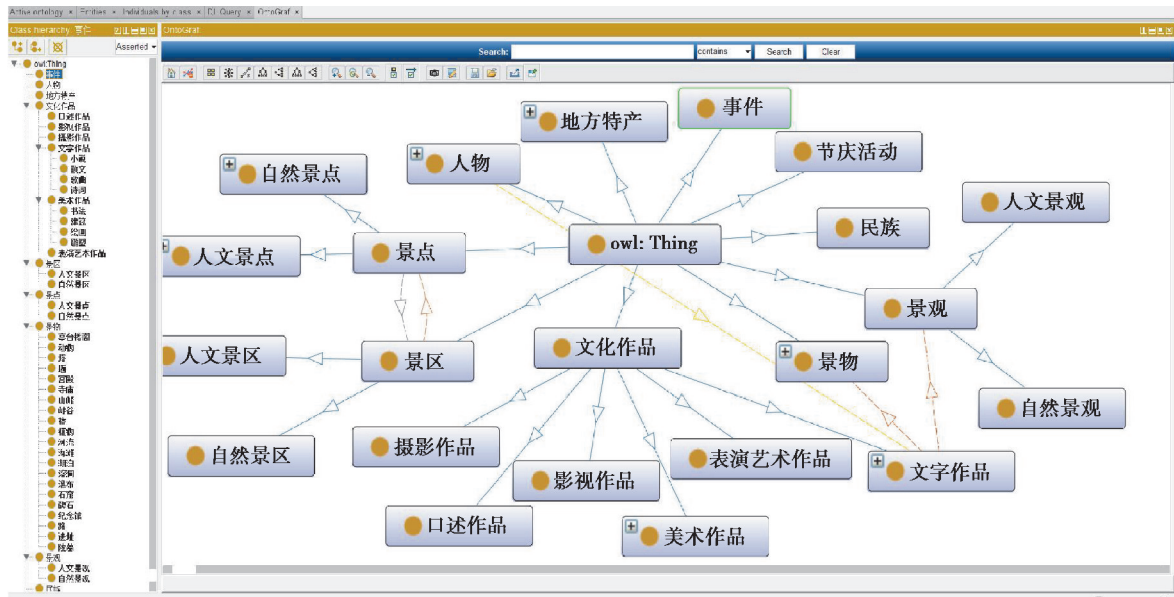


图 5 旅游本体模型

Fig. 5 Tourism ontology model

知识检索指根据用户话语从知识库中查询或推理出满足条件的知识。系统中,用户话语经过命名实体识别模块、对话策略管理模块的语义分析,已形式化为计算机可理解的系统动作和语义槽值对,知识检索将通过执行 API\_Call 类型的系统动作来完成。如何检索知识与知识的表示方法和存储方式紧密相关。知识表示主要有资源描述框架(RDF, resource description framework)、资源描述框架模式(RDFS, resource description framework schema)和 Web 本体语言(OWL, web ontology language)等方法<sup>[15]</sup>。RDF 使用类

似于主谓宾的 SPO 三元组来表示知识,一个三元组就是一条知识。RDFS 预定义了 `rdfs:Class` 等专用词汇,解决了 RDF 无法区分类和对象,无法描述类的关系和属性等不足。OWL 是对 RDFS 的扩展,增加了“`owl:ObjectProperty`”等词汇来支持本体建模和知识推理。因此,选用 OWL 方法来表示本体知识,使用 RDF/XML 语法来序列化保存本体知识,使用知识检索语言 SPARQL 来查询本体知识。

### 3 可行性验证

为了检验前述系统架构和方法是否可行,作者通过实际开发,验证了所研究系统是可行的,将从实验环境搭建、知识库构建、软件功能实现、实验效果分析 4 个方面做简要阐述。

#### 3.1 实验环境搭建

为了模拟系统的运行环境,使用 3 台 PC 电脑(CPU 2.4GHz/内存 32G)和 2 部安卓智能机构建了实验环境。3 台电脑均安装 CentOS6.5 服务器操作系统。其中,1 台电脑用作服务端的统一接入服务器,1 台电脑用于旅游本体知识管理,1 台电脑用于自然语言处理。2 部安卓智能手机安装本系统 APP 软件,用来模拟游客使用。

#### 3.2 讲解知识库构建

基于设计的旅游本体模型,以重庆瓷器口古镇为例,通过人工整理网上收集的有关磁器口景区的介绍资料,在本体开发工具 Protégé 中构建了磁器口的本体知识库,并选用 RDF/XML 语法格式,以 OWL 文件方式保存在知识管理服务器上。如图 6 所示。

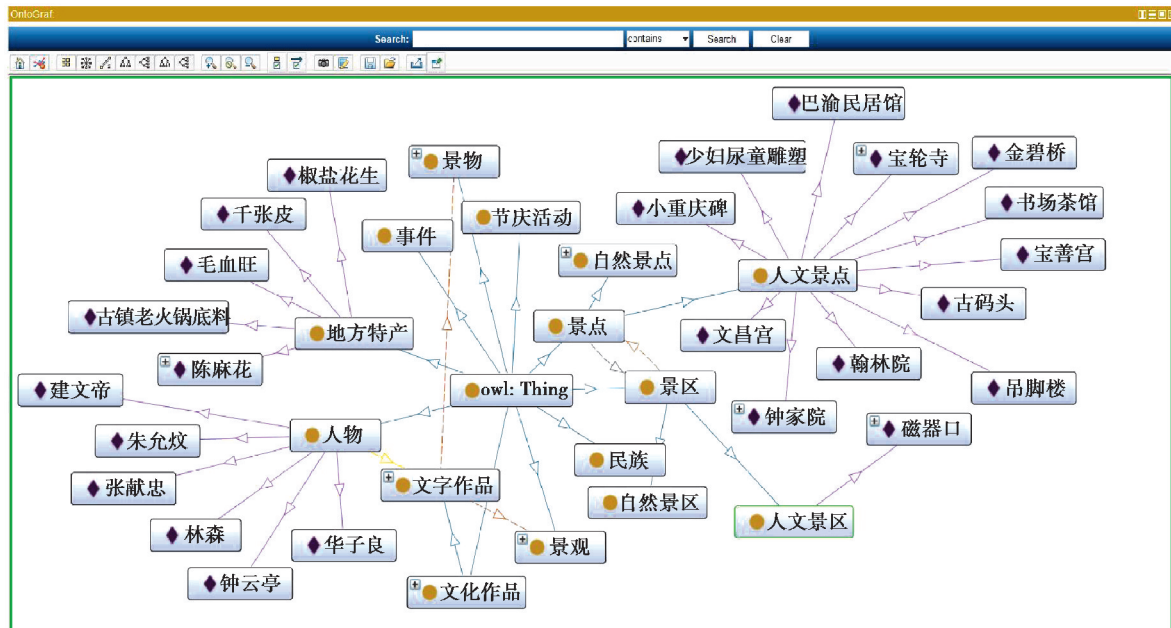


图 6 重庆磁器口本体知识库

Fig. 6 The ontology knowledge base of Chongqing Ciqikou

#### 3.3 软件功能实现

##### 3.3.1 统一接入服务程序

统一接入服务程序主要完成游客账号验证、所在景区定位等功能。使用 Java 编程语言和 Spring Boot 框架按 Web 服务方式实现,对外提供 Web API 接口供移动端 APP 启动时访问。

##### 3.3.2 自然语言处理程序

自然语言处理程序主要完成语音识别、文本向量化处理、命名实体识别、对话策略管理、对话状态跟踪、应答语音生成、讲解内容生成等功能,是整个系统最核心的部分。首先基于 TensorFlow1.15.2 用 Python 语言编程实现文本向量化处理、命名实体识别、对话状态跟踪、对话策略管理功能,然后用 Java 编程语言和

Spring Boot 框架将这些功能模块封装成 Web 服务,对外提供 Web API 接口供移动端 APP 访问。

### 3.3.3 讲解知识检索程序

讲解知识检索程序主要完成旅游本体知识检索等功能。使用 Java 编程语言和 Spring Boot 框架及 Jena 框架按 Web 服务方式实现,提供 Web API 接口供自然语言处理程序内部调用(即执行 API\_Call 类型的系统动作)。当 Web API 接口收到自然语言处理程序发来的知识检索请求后,系统会首先将检索请求解析为 SPARQL 语句,然后调用 Jena 框架提供的 QueryFactory 类 create() 方法以及 QueryExecution 类 execSelect() 方法查询知识。

### 3.3.4 移动端 APP

移动端 APP 是游客与系统交互的介质,主要完成游客定位数据获取、游客问询语音输入、系统应答等功能。研究使用 Android Studio 集成开发工具和 Java 编程语言实现了安卓版 APP,其运行效果如图 7 所示。当游客在智能手机上启动该 APP 时,系统会首先获取游客当前定位数据,然后传回服务端接入服务器识别出所在景区后,自动播放景区预设的欢迎语音/图片/视频,游客可以按下页面底端的圆形话筒按键开启语音询问功能,询问语音传回服务端后将由自然语言处理程序先请求百度语音识别接口将其转换为文本,然后再继续后续处理。为了便于查看效果,图 7 所示 APP 界面在播放对话语音的同时显示了对话文本,但游客实际使用时可关闭文本显示功能,直接收听或查看 APP 播放的图片和视频。

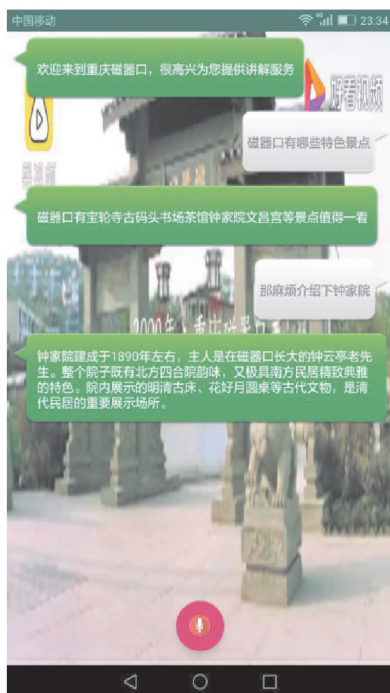


图 7 游客讲解服务 APP 运行效果

Fig. 7 Operation effect of tourist explanation service app

## 3.4 实验效果分析

为了检验系统的讲解服务效果,首先从讲解知识库中随机抽取部分实例和属性,设计了 80 个游客可能提出的问题,比如“磁器口最有名的小吃是什么?磁器口有多大?建文帝是谁?华子良在哪儿脱险的?”等;然后人工标定出每个问题的正确系统响应并用做测试数据;然后模拟游客向系统提问并记录下每次系统响应结果;最后计算系统的每轮对话响应精度(PRA, per-response accuracy)值。PRA 表示每轮对话系统回应用户提问的正确率<sup>[16]</sup>。经测算,系统的 PRA 分别为 56%,达到了预期实验效果,能初步满足旅游讲解服务需求。

## 4 结束语

用人工智能、云计算、物联网、大数据、移动互联网等新一代信息技术提升旅游服务、旅游管理、旅游营销、旅游体验是中国开展智慧旅游建设的目的。自然语言处理是人工智能研究领域的一个热点,相关技术可广泛用于人机对话、知识获取、舆情监测等应用。主要从技术应用的角度详细探讨了如何综合运用自然语言处理、本体知识管理、移动互联网、云计算等技术来研发旅游景区智能讲解系统,目的是让游客能通过自然语音对话方式享受低成本、高质量的个性化旅游景区电子讲解服务。通过实际开发,验证了系统方案的可行性。但受限于自然语言处理目前在深层次推理、本体知识自动获取等方面的不足,系统在一些需要推理才能准确识别游客意图的复杂对话场景,仍会出现错误应答和无法回答等问题。对此,笔者将继续对自然语言处理的命名实体识别方法、意图识别方法,以及本体知识的自动获取方法等做进一步研究。

### 参考文献:

[1] 中华人民共和国国家旅游局. 关于促进智慧旅游发展的指导意见[R]. Beijing: [s.n], 2015-01-12.

China National Tourism Administration. Guidance on promoting the development of smart Tourism[R]. Beijing: [s.n],



- 2015-01-12.(in Chinese)
- [2] 李入元, 张延军, 王志安, 等. 旅游云讲解管理系统的实现[J]. 电子商务, 2012(1): 36-37.  
Li R Y, Zhang Y J, Wang Z A, et al. Implementation of tourism cloud explanation management system[J]. E-Business Journal, 2012(1): 36-37. (in Chinese)
- [3] 刘应成. 基于智能手机的导游系统设计[J]. 计算机应用与软件, 2014, 31(11): 248-250.  
Liu Y C. The design of smart phone-based tour guide system[J]. Computer Applications and Software, 2014, 31(11): 248-250. (in Chinese)
- [4] 马磊. 基于智能解决方案的自助旅游系统[J]. 计算机系统应用, 2017, 26(3): 57-62.  
Ma L. Independent travel system based on intelligent solution[J]. Computer Systems & Applications, 2017, 26(3): 57-62. (in Chinese)
- [5] 赵一鸣, 朱奕蓉, 吴林容. 智能语音助手的知识服务能力评价研究[J]. 图书与情报, 2019(4): 132-140.  
Zhao Y M, Zhu Y R, Wu L R. Evaluating the knowledge service capability of intelligent voice assistants[J]. Library & Information, 2019(4): 132-140. (in Chinese)
- [6] Li F L, Qiu M, Chen H, et al. AliMe Assist: An Intelligent Assistant for Creating an Innovative E-commerce Experience[J]. 2018.
- [7] 温昱. 软件架构设计: 程序员向架构师转型必备[M]. 北京: 电子工业出版社, 2012.  
Wen Y. The transition from programmer to architect is essential[M]. Beijing: Publishing House of Electronics industry, 2012. (in Chinese)
- [8] Devlin J, Chang M W, Lee K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[EB/OL]. arXiv: 1810.04805v1[CS.CL],<http://arxiv.org/abs/1810.0480511>.
- [9] 刘胜杰, 许亮. 基于词嵌入技术的文本表示研究现状综述[J]. 现代计算机, 2020(1): 40-43.  
Liu S J, Xu L. Review of research status of text representation based on word embedding technology[J]. Modern Computer, 2020(1): 40-43. (in Chinese)
- [10] Cui Y M, Che W X, Liu T, et al. Pre-training with whole word masking for Chinese BERT[J]. ACM Transactions on Audio, Speech, and Language Processing, 2021, 29: 3504-3514.
- [11] Lê T A, Burtsev M S. A deep neural network model for the task of Named Entity Recognition[J]. International Journal of Machine Learning and Computing, 2019, 9(1): 8-13.
- [12] Williams J D, Asadi K, Zweig G. Hybrid Code Networks: practical and efficient end-to-end dialog control with supervised and reinforcement learning[EB/OL]. 2017; arXiv: 1702.03274[cs.AI],<https://arxiv.org/abs/1702.03274>.
- [13] Gruber T R. Toward principles for the design of ontologies used for knowledge sharing? [J]. International Journal of Human-Computer Studies, 1995, 43(5/6): 907-928.
- [14] 戴代新, 姬鹏飞, 戴开宇. 基于语义 web 的旅游景区本体构建及其应用研究[C]// 中国第二届数字景观国际论坛论文集. 南京: 万方数据, 2015:102-111.  
Dai D X, Ji P F, Dai K Y. Research on ontology construction and application of scenic spots based on Semantic Web[C]// Process of The 2nd China Digital Landscape International Forum. Nanjing: Wanfang Data, 2015:102-111.(in Chinese)
- [15] Pulido J R G, Ruiz M A G, Herrera R, et al. Ontology languages for the semantic web: a never completely updated review[J]. Knowledge-Based Systems, 2006, 19(7): 489-497.
- [16] Bordes A, Boureau Y L, Weston J. Learning end-to-end goal-oriented dialog[EB/OL]. 2016; arXiv: 1605.07683[cs.CL].  
<https://arxiv.org/abs/1605.07683>.