

上海对外经贸大学

# 毕 业 论 文

论文名称：智能对话系统研究与智能对话场景设计

学 院：统计与信息学院

专 业：数据科学与大数据技术

学 号：19024075

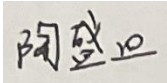
学生姓名：陶盛皿

指导教师：刘亮亮

20 23 年 5 月

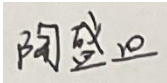
## 毕业论文独创性声明

本人所呈交的毕业论文是在指导教师指导下进行的工作及取得的成果。除文中已经注明引用的内容外，本论文不包含其他个人已经发表或撰写过的研究成果。对本文的研究做出重要贡献的个人和集体，均已在文中作了明确说明并表示谢意。

作者签名：  \_\_\_\_\_

## 毕业论文使用授权声明

本人完全了解上海对外经贸大学有关保留、使用毕业论文的规定，学校有权保留毕业论文并向国家主管部门或其指定机构送交论文的电子版和纸质版。有权将毕业论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅。有权将毕业论文的内容编入有关数据库进行检索。有权将毕业论文的标题和摘要汇编出版。

作者签名：  \_\_\_\_\_ 日期： 2023 年 5 月 2 日

# 目 录

摘要.....	3
ABSTRACT.....	4
1 绪论.....	5
1.1 选题初衷.....	5
1.2 智能对话系统研究背景与发展历史.....	5
1.3 研究意义.....	7
1.4 研究现状.....	7
1.5 研究思路与方法.....	8
2 智能对话系统设计.....	9
2.1 智能对话系统总体架构.....	9
2.2 自然语言理解模块.....	10
2.3 对话管理.....	10
2.4 联合建模自然语言理解和对话状态跟踪.....	15
2.5 自然语言生成模块.....	16
2.6 基于迁移学习方法的端到端智能对话系统.....	16
3 基于场景的对话和任务设计.....	17
3.1 配置简介.....	17
3.2 火车票查询.....	17
3.3 快递单号查询.....	19
3.4 中医症状查询.....	22
3.5 天气查询.....	23
4 工业界智能对话系统产品.....	25
4.1 基于端到端方法的智能对话系统：阿里小蜜.....	25
4.2 基于生成式模型和强化学习方法的智能对话系统：ChatGPT.....	25
5 总结与展望.....	25

5.1 本文工作总结.....	25
5.2 领域未来挑战.....	26
5.3 领域未来展望.....	27
5.4 本文不足总结与展望.....	28
参考文献.....	29
致谢.....	33

## 摘 要

本文针对智能对话系统展开体系的理论性研究，并追踪当下该领域最前沿应用：ChatGPT 人工智能聊天机器人。针对智能对话系统设计总体架构进行逻辑梳理，并对自然语言理解、对话管理、自然语言生成三大模块以及实现方式展开具体论述。其次，设计火车票查询、快递单号查询、中医症状查询、天气查询四个对话场景，通过数据处理与清洗后，三大模块使用深度学习算法联合建模的方式实现场景应用有效性，并进行功能展示。接着，概述工业界智能对话系统产品：阿里小蜜和 ChatGPT。接着，分析人工智能智能对话系统的未来发展趋势，得出：该领域在落地、人机交互、算法解释、算法合规、算力、网络安全和隐私保护方面可能面临挑战，该领域未来发展趋势是 5G 技术融合、多模态智能对话系统、情感理解。最后，总结本毕业设计的不足与展望。

关键词：人工智能，智能对话系统，对话场景设计

## **ABSTRACT**

This paper presents a systematic theoretical study of intelligent dialogue systems, and traces the current cutting-edge application in this field: ChatGPT Artificial Intelligence Chatbot. The overall architecture of the Intelligent Dialogue System design is logically sorted out, and the three modules of Natural Language Understanding, Dialogue Management, and Natural Language Generation and the implementation methods are discussed in detail. Secondly, four dialogue scenarios are designed: train ticket query, express single number query, Chinese medicine symptom query, and weather query. After data processing and cleaning, the three modules use Deep Learning algorithms to jointly model the effectiveness of the scenario application and demonstrate the functions. Then, an overview of the Intelligent Dialogue System products in industry: Ali Xiaomi and ChatGPT. next, analysis of the future development trend of AI Intelligent Dialogue System has been concluded that: the field may face challenges in landing, human-computer interaction, algorithm interpretation, algorithm compliance, arithmetic power, network security and privacy protection, the future development trend of the field is 5G technology integration, Multimodal Intelligent Dialogue System, emotional understanding. Finally, the shortcomings and my prospect of this graduation design are summarized.

Keywords: Artificial Intelligence, Intelligent Dialogue System, Dialogue Scene Design

# 智能对话系统研究与智能对话场景设计

## 1 绪论

### 1.1 选题初衷

本文选题任务型智能对话系统研究并进行对话场景设计，旨在为我过去的三到五年画上句号，算是完成高中时期工科学校的执念。作为一个高考数学没有及格并且转专业进大数据系的同学，在上海对外经贸大学本科四年期间是一件非常具有挑战性的事情。在这四年的时间里，学习了数理课程，如数学分析，概率论 A，数理统计，线性代数 A，应用时间序列分析，应用多元统计等，学习了计算机编程语言 C++，Python，R，GoLang，MySQL，机器学习，深度学习，人工智能，算法导论，分布式计算等前沿领域的课程。感觉自己好像对某一领域非常了解，也学到了不少新的东西。之所以选择智能对话系统领域，也是表达了我自己的局限性和内心对工科领域嵌入式系统等向往。

### 1.2 智能对话系统研究背景与发展历史

#### 1.2.1 概述

多轮聊天机器人(Chat Generative Pre-trained Transformer, ChatGPT)技术实现并大规模被应用，其属于智能对话系统领域的技术。智能对话系统是一种计算机程序，其优势可以总结为：（实时）即时反应、24 小时 365 天（全天候）可用性、解放人力并降低企业人工成本、最大化员工技能、个性化操作、更好地理解客户、增加业务收入等<sup>1</sup>。智能对话系统能够实现人机交互，将提供巨大的市场价值，正在成为趋势技术。

#### 1.2.2 发展历史

虽然“智能对话系统”一词对于大众而言，显得陌生而又遥远，但是纵观其发展史可知其并不是一个新兴的概念，本段落对智能对话系统发展历史梳理。

1950 年，图灵测试(Turing Test)的《计算机与智能》为智能对话系统奠定了基础。1966 年，麻省理工大学(Massachusetts Institute of Technology, MIT)研发的全球第一个对话系统 Eliza，是基于规则设定和自然语言处理的问答系统。20 世纪 90 年代微软公司(Microsoft)研发并大规模推向市场的对话系统为虚拟助手 Clippy 是问答系统。1995 年，Wallace 研发的基于模式匹配和启发式方法 Alice，是一个真正意义上的智能对话系统。2006 年，国际商业机器有限公司(The International Business Machines Corporation, IBM)发布了基于自然语言处理和机器学习的 Watson，是能够理解大量数据的智能对话系统。2010 年，苹果公司移动端语音助手 Siri 集成在移动端，智能对话系统开始融入在主流消费市场的消费者的生活中。2012 年，谷歌公司(Google)推出了能够通过自然语言在谷歌搜索引擎中与用户交互的 Now，是能在网页应用进行问答的智能对话系统。2015 年，亚马逊公司(Amazon)推出集成在智能家居中的智能对话系统 Alexa。以及后续微软公司推出的 Cortana、元宇宙公司(Metaverse)推出的聊天机器人，以及国内科技大厂推出阿里小蜜等都在智能对话系统的范畴之中<sup>2</sup>。

### 1.2.3 智能对话系统分类

智能对话系统可依据主要任务类型分成 3 类：社交闲聊型(Chit-Chat)、问答型(Frequently Asked Questions, FAQ)、任务导向型 (Task-based)<sup>3</sup>。闲聊型能够实时与用户展开多轮对话；问答型可回答用户的具体问题。该类型可以细分成：基于规则人工设定和融入有限状态机(Finite State Machine, FSM)思想的问答系统、基于检索关系型数据库或者非结构化文档的问答匹配系统<sup>4</sup>、基于知识图谱（图数据库）的信息检索系统<sup>5</sup>、基于神经网络(Neural Network, NN)或者端到端(End-to-End)深度学习模型的文本分类和自然语言理解的对话系统、基于信息检索(Information Research, IR)和基于序列到序列(Sequence-to-Sequence, Seq2Seq)模型架构的聊天机器人<sup>6</sup>、基于预训练模型(Pretrained Model, PM)和强化学习(Reinforce Learning, RL)的人工智能聊天机器人<sup>7</sup>。

### 1.2.4 智能对话系统组件

任务导向型智能对话系统擅长在特定领域中完成指定任务<sup>8</sup>，其包括：理解用户意图（即进行自然语言理解）、跟踪人机交互的多轮对话状态和对话策略计算下一步行动计算这两个子模块，对话状态跟踪使对话系统能专注于在搜索以前的



状态和跟踪信息以理解上下文语义、对话策略使其能交付实际操作实现与用户交互、自然语言生成能创建人类可读的语言并传递给用户。

### 1.3 研究意义

基于人工智能领域强化学习、深度学习和自然语言处理技术的智能对话系统在人机交互、智能语音助手、智能客服、问答咨询聊天机器人等领域广泛应用，并由于 ChatGPT 的问世将人们的实现聚焦在这个领域。这极大地促进了自然语言理解、自然语言生成、对话管理、端到端模型的理论研究与现实世界中的技术的发展，这些已是工业界关注焦点。

任务型智能对话系统因完成特定领域任务的原因能提供个性化服务，尤其在电商、教育和医学等方向具有商业价值。故本文将通过文献阅读和实验相结合的方式，学术上深入研究和实际感受，并跟进当下热点和趋势，并阐述该领域挑战和未来展望，以进行体系化的研究。

### 1.4 研究现状

#### 1.4.1 国外研究现状

基于智能对话系统的发展史，智能对话系统从图灵机问答发展到基于人工规则预定义的问答系统，从基于模式匹配和启发式算法的系统到智能语音系统，智能对话系统从初级离线系统一路发展至实时计算的真正意义上的智能系统，目前最前沿是融入强化学习实时训练和爬虫实时计算的 ChatGPT，故认为 GPT (Generative Pre-trained Transformer, GPT)系列生成式对话模型以及当前 ChatGPT 是基于国外最前沿的研究现状的应用。

针对国外智能对话系统头部企业 OpenAI 的 GPT 系列模型展开介绍，它们是基于单向 Transformer(Uni-Transformer)模型的人工智能自然语言处理领域的模型<sup>9</sup>。2018 年，GPT-1 被推出（早于 Google 为人熟知的 Bert 模型），采用预训练模型(pre-trained model)的迁移学习(Transfer Learning)进行参数微调(fine-tuning)处理下游任务，它能够解决动态语义的模型，其模型架构是基于监督学习和词嵌入(Word Embedding)的单向 Transformer。GPT-2 继承 GPT-1 的架构，使用基于无监督方法的深度学习预训练模型完成多任务学习<sup>10</sup>，刷新了著名的 7 大数据集基准

(General Language Understanding Evaluation benchmark, GLUE benchmark), 被认为是“最强通用自然语言模型”。2020 年, GPT-3 作为第三代生成式模型具有 1750 亿参数, 即情景学习模型, 用于语言预测任务<sup>11</sup>, 比如创造性写作 (包括: 诗歌、新闻等文字工作者产出物), 目前并未面向公众。ChatGPT 融入了强化学习进行训练并使用近端策略优化的奖励函数进行模型微调, 故其在目前是最前沿的 (State-of-Art, SOTA) 的人工智能聊天机器人程序<sup>12</sup>。

### 1.4.2 国内研究现状

对标国外智能对话系统产品市场渗透率, 本文将国内头部大厂阿里的产品, 即多语言会话的任务型对话机器人的“阿里小蜜”认为当前国内能够实现落地的最前沿研究现状 (产品)。“阿里小蜜”能与传统多轮任务型智能对话系统的显著区别开优势在于其应用生成对抗网络 (Generative Adversarial Network, GAN) 实现混合语言理解和翻译, 并基于知识蒸馏技术 (Knowledge Distill, KD) 进行多语言模型的自动迭代<sup>13</sup>。

针对当下智能对话系统可归纳为 3 类, 即基于自然语言理解算法和知识库方法的文本机器人、集成语音识别和语音合成的智能语音交互机器人、基于计算机视觉 (Computer Vision, CV) 算法和多模态 (Multi-Modal) 的多模态数字人, 技术水平已经非常超前。当下国内开始有多模态的趋势, 即在传统文字和语音的基础上融入了虚拟形象的技术, 以仿真人的方式进行人机交互<sup>14</sup>, 该领域依旧在起步阶段 (概念款设计), 距离真正能够在现实世界实现落地仍需年月。此外, 情感计算的概念正兴起, 以多层语义分析为主<sup>15</sup>, 未来将整合多模态信息, 实现上下文理解、预测客户的情绪, 真正迎来人机情感交互的时代还很遥远。

## 1.5 研究思路与方法

### 1.5.1 研究思路

人工智能是有数据驱动的, 模型底层逻辑来源于数学, 首先需要通过数学分析、线性代数、统计学学科的学习扎实算法基础。通过学习数据科学领域的统计学方法和机器学习算法进行数据分析学习。进而通过深度学习、自动化、物联网 (Internet of Things, IOT) 等领域的了解以及领域相关论文文献的阅读, 了解智能对

话系统领域的动态趋势和深入了解智能对话系统整体架构和功能模块的逻辑。

### 1.5.2 研究方法

本毕业论文主要是通过市场调研、网络信息检索，进行人工手动搜集智能对话系统领域情报和市场动态进行综合归纳。通过对智能对话系统产品进行实践和观察。文献研究法则主要涵盖：智能对话系统、人工智能、自然语言处理、深度学习、机器学习等，进行文献阅读并建立文献文档。本文通过对话场景设计创意设计 4 个任务型智能对话系统的功能：火车票查询、快递查询、中医症状查询和天气查询，通过现实场景落地使用实现总体架构 3 大模块深度学习联合建模，使用户直观感受智能对话场景的落地。

## 2 智能对话系统设计

### 2.1 智能对话系统总体架构

本智能对话系统总体架构包含输入层（即用户输入）、自然语言理解、对话管理、自然语言生成、输出层（即智能对话场景输出）这 5 大模块，其中对话管理包含对话状态追踪和动作生成（对话策略）2 个子模块。并且通过分布式部署知识图谱（图数据库）、关系型数据库来支撑输入和输出之间语义理解等模块的计算，通过这些模块的联合建模实现对话场景落地。

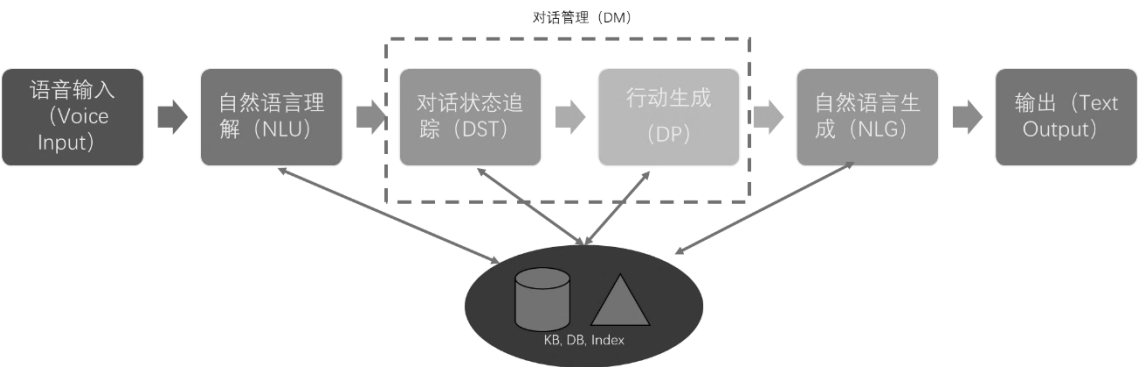


图 1 智能对话系统总体架构图

## 2.2 自然语言理解模块

### 2.2.1 自然语言理解模块总体架构

自然语言理解模块的核心思想：希望机器像人一样，具备正常人的语言理解能力，是在语音输入和语音识别模块后非常重要的模块。自然语言理解需要完成用户意图识别(Intent Detection)和对意图词槽填充(Slot Filling)<sup>16</sup>。

自然语言理解模块的目的是：利用语音识别技术将语音输入转化为文本，通过意图识别进行文本分类实现用户意图判断，再通过槽填充是序列标注或抽取句子中重要的语义成分，将其翻译成意图输入下一个模块。

### 2.2.2 自然语言理解模块架构详解

该模块主要完成文本分类，即多意图识别的任务，进阶版本则是在多轮对话中的多意图识别文本分类，此外还有词槽抽取和联合词槽意图的方法。

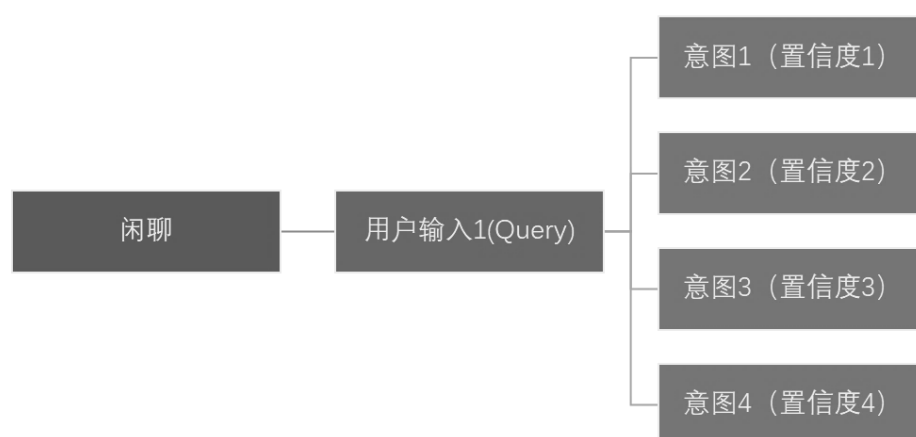


图2 多意图识别文本分类架构

自然语言理解模块内部架构是传统深度学习端到端的模型架构，一般部署 Bert 预训练(pre-trained Bert)模型，进行文本分类<sup>17</sup>。在文本分类的流程中，将用户语句输入自然语言理解模块，通过模型计算得到多意图和其置信度的输出，置信度最大的意图作为最终输出。

## 2.3 对话管理

对话管理(Dialogue Management, DM)模块接收自然语言理解输出的用户意图，

并通过其对话状态跟踪模块根据对话历史信息生成对话策略，输出系统行为，决定任务型智能对话系统下一步要做什么。最常见的应用任务性多轮智能对话系统，针对特定场景执行特定任务如本文设计的天气查询、火车票等。有时用户意图识别过程中可能存在很多限制条件，故需要多轮对话的方式，即用户和智能进行一问一答的方式让用户键入实体。用户在人机交互对话过程中可以通过多次输入修改输入的需求语句。此外，当用户需求不能被明确识别完成意图分类时，智能对话系统可以通过询问来确认用户需求。

### 2.3.1 对话状态跟踪

对话状态跟踪(Dialogue State Tracking, DST)模块输入当前用户意图和用户过去意图历史信息输出当前状态（即更新后的意图）或嵌入式张量<sup>18</sup>，即进行状态序列迭代运算。

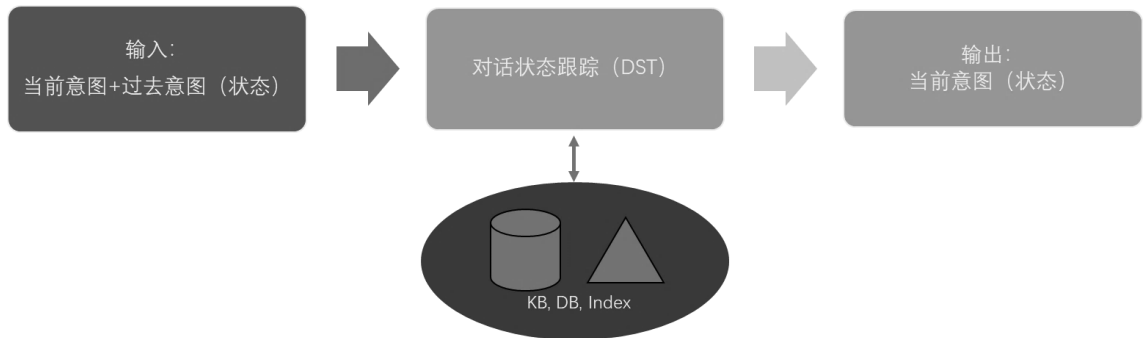


图3 对话状态跟踪模块

设  $T+1$  时刻的对话状态:  $S_{T+1}$ ，依赖于  $T$  时刻的对话状态  $S_T$ 、对应的系统行为  $a_T$ 、当前  $T+1$  时刻的用户行为  $a_{T+1}$ ，本质是对状态转移方程迭代运算。

$$S_{T+1} = S_T + a_T + a_{T+1} \quad (\text{公式 1})$$

### 2.3.2 基于深度学习预训练模型的有限状态跟踪

在任务型智能对话系统中，对话状态跟踪是从对话历史记录跟踪对话状态，当前研究现状是在该任务上使用基于 Bert 模型的有限状态跟踪(Bert-based Finite State Machine, Bert-based FSM)的方法。<sup>19</sup>其中 Bert 模型规模小且部署简便，使得

有限状态追踪的性能非常好。<sup>20</sup>该模块又称基于 Bert 预训练模型的意图词槽联合提取。

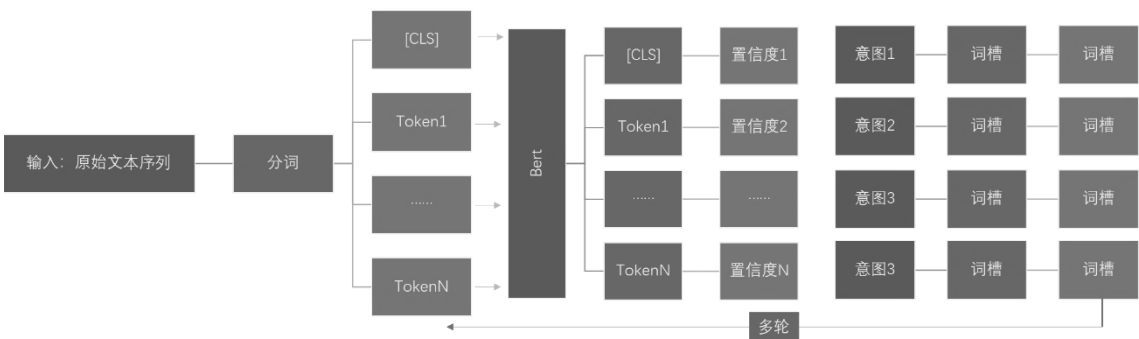


图 4 基于 Bert 预训练模型的多轮对话有限状态跟踪架构

本小节针对 Bert 模型运作方式展开具体解释，该预训练模型使用分词函数 (tokenizer)，直接调用 bert-baseduncased 的分词函数进行分词操作，直接调用使用编码函数(encoder)进行编码。该模型需要进行填充(Padding)和创建掩码注意力(masked attention)使用层训练的方式，即相当于把序列当做一个批量)。使用随机采样(RandomSampler)和序列采样(SequentialSampler)进行数据采样和数据增强，加强模型捕获语义信息的能力。<sup>21</sup>

2.3.3 对话策略

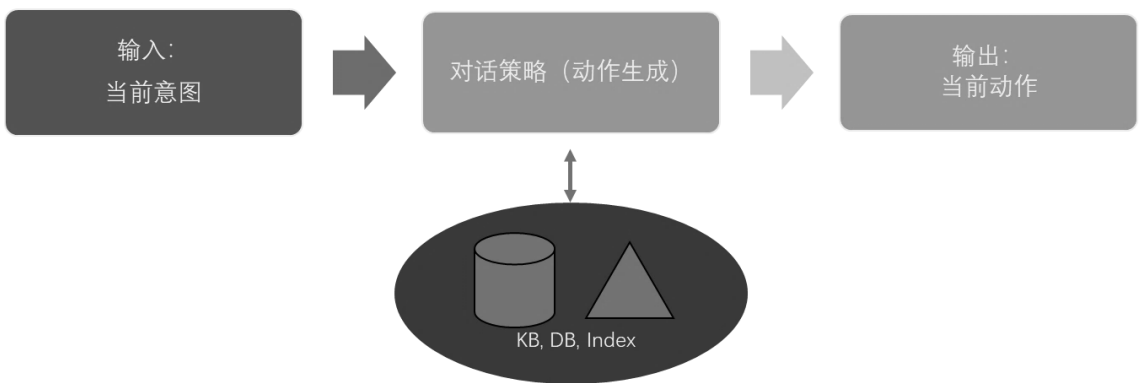


图 5 对话策略模块

对话策略(Dialogue Policy, DP)模块根据  $T+1$  时刻的对话状态  $S_{(T+1)}$ ，生成

T+1 时刻系统行为(Dialogue Action):  $a_{(T+1)}$ , 该模块通过当前状态决定下一步行动<sup>22</sup>。

该模块可以使用监督学习、强化学习和模仿学习得到，其中目前工业界产品比如：ChatGPT 已经应用强化学习的方式优化函数更新生成的动作。

2.3.4 基于规则的对话策略

基于规则的对话策略(Rule-based Dialogue Policy, Rule-based DP)借鉴有限状态机的思想，通过状态转移触发智能对话系统相应模块内的行为<sup>23</sup>，我们人为定义智能对话系统具有的意图域以及可采取的行动域，将动作函数(actions)和意图(intents)看作是包含有限数量元素的集合，其中的元素在实际的任务流中是行动接口(Application Programming Interface, API)。

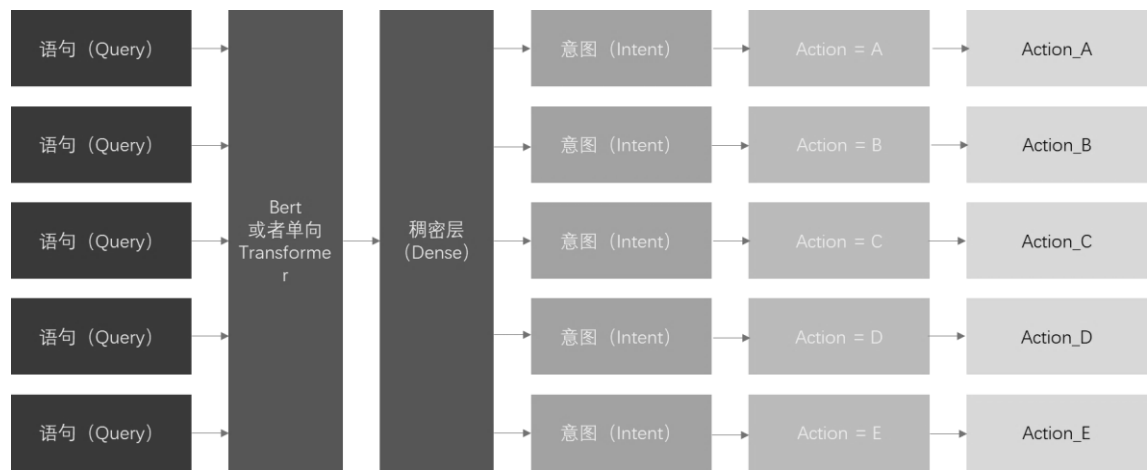


图 6 基于规则的对话策略

本智能对话系统的配置如下：

表 1 行动函数与对应动作目的

actions	动作含义或目的
utter_first	向用户进行初次见面问候
utter_donknow	向用户表达无法理解当前对话内容
action_first	执行第一步动作决策
action_donknow	返回结果在搜索域外的动作
action_echo	重复相同语句或反问
action_search_treat	表达玩笑
action_search_food	食品检索
action_search_symptom	症状检索

<b>action_search_neopathy</b>	病症检索
<b>action_search_drug</b>	药品检索
<b>action_search_prevention</b>	预防措施检索
<b>action_search_drug_func</b>	功效检索
<b>action_search_disease_treat_time</b>	治疗次数检索
<b>action_search_easy_get</b>	获取方式搜索
<b>action_search_disease_dept</b>	疾病科室检索
<b>utter_greet</b>	表达问候
<b>utter_howcanhelp</b>	向用户询问需要帮助的内容
<b>utter_goodbye</b>	表达再见

表 2 用户意图识别和意图含义

<b>Intents</b>	<b>意图含义</b>
<b>bye</b>	告别
<b>affirmative</b>	肯定
<b>search_treat</b>	用户想要了解治疗方式
<b>search_food</b>	用户想要知道食品名称
<b>search_symptom</b>	用户想要获取疾病相关症状
<b>search_neopathy</b>	用户想要了解病症详情信息
<b>search_drug</b>	用户想要获取治疗药品清单
<b>search_prevention</b>	用户想要了解预防方式
<b>search_drug_func</b>	用户想要了解药品全部功效
<b>search_disease_treat_time</b>	用户想要了解治疗周期
<b>search_easy_get</b>	用户想要知道药品获取方式
<b>search_disease_dept</b>	用户想要知道病系和科室信息

### 2.3.5 基于强化学习模型的对话策略

基于强化学习的模型本文无法实现，本文仅限传统自然语言理解、对话管理、自然语言生成的联合建模的方式进行对话场景实现，使用的模型均为预训练模型。据悉，目前最前沿的智能对话系统应用产品 ChatGPT 使用的基于强化学习的模型的对话策略，通过预训练模型和强化学习结合的深度强化学习方法在任务型智能对话场景有效<sup>24</sup>。前面章节提到的基于规则的方法需要人工自定义(Hand-crafted Approaches)，无论是意图集合还是动作集合都是有限域，然而人不可能穷尽有可能发生的场景。这种情况下，强化学习的优势明显。基于强化学习的对话管理(Reinforcement Learning Based Dialogue Management, RL-Based DM)能够对智能对话系统理解用户输入的不确定性进行建模学习<sup>25</sup>，利用奖励函数(Reward



Function)来模型学习到最优行为序列<sup>26</sup>。

相较于上述小节的对话策略架构，基于强化学习的方法在对话管理模块的对话状态跟踪和对话策略之间新增了奖励函数。



图7 基于强化学习的对话策略架构

## 2.4 联合建模自然语言理解和对话状态跟踪

基于本文上述小结自然语言理解模块详解和对话状态跟踪的详解，本章节其理论进行自然语言理解和对话状态跟踪的联合建模。智能对话系统自然语言理解模块首先处理输入的用户语句，按照自然语言理解模块内部的流程进行理解用户意图，以决定智能对话系统下一步采取的行为，并自定义行动程序(action.py)来出发行动接口，并且知识库交互获得对话系统的回答，即最终输出。

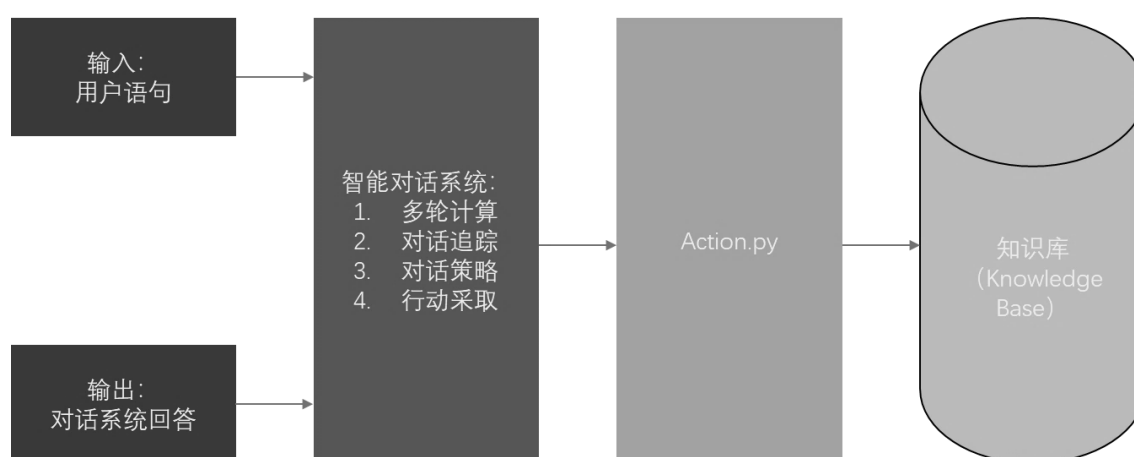


图8 联合建模逻辑图

注：上图是对于实验设计的逻辑业务展示，具体实现详见第三章的设计，模块功能对话场景的实现均需要使用联合建模方法，即规则(rules)路径下行动域(domain)和动作(action)两个模块的逻辑勾稽。

## 2.5 自然语言生成模块

### 2.5.1 自然语言生成模块详解

自然语言生成(Natural Language Generation, NLG)模块核心功能：将数据转换为人类可理解(human-friendly)的语言<sup>27</sup>，即从非语言输入中生成文本或语音的任务。

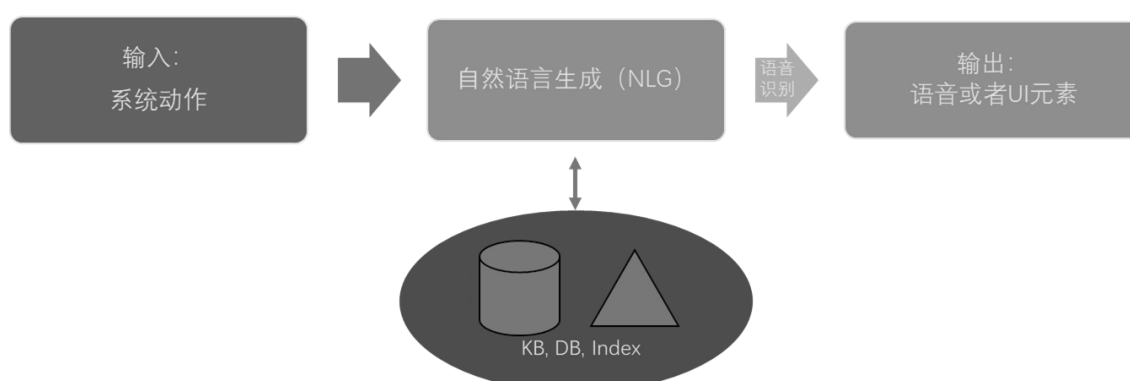


图9 自然语言生成模块

### 2.5.2 基于模板的自然语言生成

可以参照传统问答型对话系统的逻辑，在任务型智能对话系统当中，将对话管理模块输入的系统动作输入自然语言生成模块进行转化，通常有文本到语言(Text-to-Text)和数据到语言(Data-to-Text)两种方式，并通过语言识别或者转换的方式输出语音或者对应的交互界面(User Interface, UI)的元素。

### 2.5.3 基于模型的自然语言生成

基于模型的自然原因生成往往使用深度学习模型执行类似文本分类的功能，而当下基于最前沿的产品 ChatGPT 中在自然语言生成模块使用了条件循环神经网络语言模型(Conditioned Recurrent Neural Network Language Model, CRNNLM)捕获上下文信息，生成符合所需要的语言风格且内容连贯<sup>28</sup>的句子。

## 2.6 基于迁移学习方法的端到端智能对话系统

GPT 系列生成式对话模型隶属于 OpenAI 公司。2022 年 5 月该系列 GPT-3 被

推出，GPT-3 比 GPT2 更加通用对于任何输入文本都能生成文本来响应，微软公开主题为“两个人工智能通过哲学讨论人类”的视频。由于 GPT-3 未开源无法大规模使用在 2022 年 9 月之前（即 ChatGPT 的问世前）使用 GPT-2 模型的端到端智能对话系统一直是最前沿的应用产品。GPT-2 核心思想是使用无监督的预训练模型去做有监督学习任务，是相对于 Bert 模型的巨大进步，通过多任务学习和零次学习的方式来展开下游任务<sup>29</sup>。

## 3 基于场景的对话和任务设计

### 3.1 配置简介

在本自定义智能对话系统中自然语言理解模型训练样本数据使用标记语言(Markdown, MD)文件格式存储，在这个文件可以使用基于规则的方法人为手动设置智能对话系统需要识别意图和词槽等。文件 story.md 是核心(core)模型训练样本，即对话模型。文件 Domain.yml 存储了智能对话系统的所有信息。文件 Config.yml 是训练自然语言理解和核心模型的配置文件。

### 3.2 火车票查询

#### 3.2.1 火车票查询意图配置

在 nlu.md 中，人工设置了查询火车票意图属性(intent: request\_ticket)，该意图包含的通用话术有：“帮我查下火车票”、“查火车票”、“我要查火车票”等。带有槽位的时间信息的话术，举例：“我想知道[明天](date-time)的火车票”。带有槽位的地点信息的话术，举例：“去[北京](address)的火车票有哪些”。

#### 3.2.2 火车票查询情景自定义

在文件 story.md 中，可以设置通过自己的喜好命名情景(Scenario)，使用 \*intent{ “entity1” : “value” } 的字典形式存储发送的信息，单个用户表示单个实体(entity)，信息内容是取值(value)。通过“-action”表示动作生成，单个动作执行之后返回时间，则返回 slot(“slot\_name” : “value”)，每个字段对应一个槽位(slot)，词槽信息(value)存储在槽位当中。

以火车路径查询的情景为例，伪代码及对应解释如下：

表 3 快乐车票情景代码执行及其含义

代码顺序	代码内容	代码含义
1	happy ticket scenario path	快乐车票获取路径
2	request_ticket ticket_form	火车票表单检索
3	form{"name": "ticket_form"}	返回检索的表单信息
4	form{"name": null}	返回当前表单检索为空的信息
5	goodbye	表达告别的意图
6	utter_goodbye	表达告别的动作函数执行

假设识别槽位得到的时间是明天、地点是北京，代码实例如下：

表 4 明天北京车票检索示例

代码顺序	代码内容	代码含义
1	ticket scenario path2	实例车票获取路径 2
2	request_ticket{"date_time": "明天", "end": "北京"}	明天北京的火车票表单检索
3	create ticket_form	创建明天北京的张量
4	form{"name": "ticket_form"}	返回这条张量检索的表单信息
5	form{"name": null}	返回如果检索是否为空的信息
6	goodbye	表达告别的意图
7	utter_goodbye	表达告别的动作函数执行

### 3.2.3 火车票查询对话管理自定义

在文件 domain.yml 中手动设置订票回复模块的话术，代码如下：

表 5 参数设置对照表

参数名称	名称释义	参数数值	数值释义
Language	场景语言	zh	中文
Pipeline	管道	无	无
Model_Name	自然语言处理模型	MitieNLP	中文文本特征提取
Model_Path	模型路径	total_word_feature_extractor_zh.dat	中文文本特征提取路径
Tokenizer	分词器名称	JiebaTokenizer	结巴分词器
Dict_Path	词典路径	data/dict	词汇存储的路径
Entity_Extract	实体抽取器	MitieEntityExtractor	中文文本实体

			抽取器
<b>Entity_Syn</b>	同义词获取器	EntitySynonymMapper	中文实体同义词映射
<b>Featurizer</b>	文本特征编码器	RegexFeaturizer, MitieFeaturizer	中文文本特征编码器
<b>Intent_Class</b>	意图识别分类器	SklearnIntentClassifier	基于 Sklearn 意图分类器
<b>Response_Select</b>	话术选择器	ResponseSelector	选择回复话术
<b>Response_Epc</b>	回复话术训练轮数	100	默认设置参数 100
<b>Policy</b>	策略名称	FormPolicy	表单检索策略函数
<b>Memory</b>	机器人记忆	MemoizationPolicy	历史记忆策略函数
<b>Transformer Embedding Dialogue Policy</b>	Transformer 模型 嵌入对话策略	TEDPolicy	基于 Transformer 模型嵌入对话策略
<b>Max_History</b>	最大对话轮数	5	默认设置参数 5
<b>Model_Epochs</b>	模型训练迭代次数	100	默认设置参数 100
<b>Mapping</b>	动作映射器	MappingPolicy	动作映射函数

### 3.2.4 火车票查询功能展示

由于不知道这个功能模块实现是否数据安全合规，综合考虑之下选择不进行展示。

## 3.3 快递单号查询

### 3.3.1 快递单号查询情景

在文件 story.md 中定义查询的情景：快递搜索和快递查询，并标注场景对应的意图和动作，具体实现，如下：

表 6 快递信息检索策略执行顺序

代码顺序	代码内容	代码含义
1	story: express_search	快递信息检索情景
2	intent: search_express	快递信息检索意图
3	action: action_search_express_form	检索快递表单动作策略
4	active_loop: action_search_express_form	循环表单检索动作策略

### 3.3.2 快递单号查询意图自定义

在文件 nlu.md 中，设置了查询快递单号意图(intent: search\_express)，包含的

话术有：“帮我查个快递”、“我要查快递”、“快递”、“查询快递”等。本模块核心功能函数是搜索快递(search\_express)，查询快递使用到的方法包含：基于有限状态机的规则设定的方法，基于意图识别与词槽联合的方法，自定义3家公司的快递单号查询，顺丰、中通和圆通。

带有槽位的快递公司或者快递单号信息的话术，举例：“帮我查下[中通](express)快递”、“我想查询[圆通](express)快递，单号[YT1234567890112](number)”、“[中通](express)单号为[3456689988662](number)的信息”、“我想知道我的[顺丰](express)快递到哪里了，单号[SF10ASDFGH31731](number)”、“单号为[75ZXCVB420290](number)的[中通](express)信息”。通过意图识别获得对应的词槽信息，使用交叉表(lookup table)触发动作接口，查询（抓取）快递单号详细信息。（注：由于快递单号涉及他人隐私，故使用随机编码编辑入本报告。）

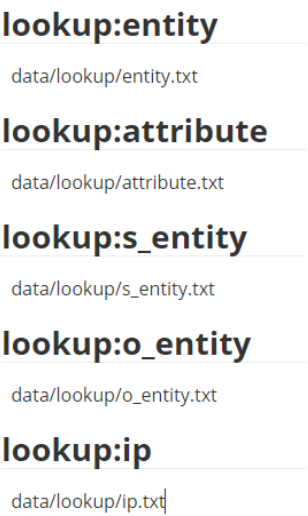


图 10 查询快递单号表单联查建模

当客户输入了快递公司信息和快递单号就可以实现快递信息的查询。在文件 domain.yml 中设置表单(forms)，当表单填写完成之后就会调用行动函数(actions)来进行快递信息抓取。因此，需要认为定义规则使用交叉表单触发填表操作，不展开赘述。

### 3.3.3 快递单号对话管理自定义

在文件 domain.yml 中设置了该场景可能使用到的话术，代码如下：

表 7 快递查询话术对照表

responses	回复含义	回复内容(text)
utter_ask_expressinfo	询问快递名称	“请输入需要查询的快递公司，目前只支持顺丰，中通，圆通。”
utter_ask_numberinfo	询问快递单号	“请输入要查询的快递单号”
utter_search_stop_numberinfo	查询任务执行结束	“关于{expressinfo}，快递单号{numberinfo}查找结束。”

### 3.3.4 快递单号场景测试

运行 Shell 命令，输入快递用户输入检索语句，可以看到该模块对话术模板的实体识别正确，置信度在基本在 99%以上，效果理想。（快递单号已经使用马赛克标注并不涉及泄露个人隐私，本人操作合规仅限测试单个样例）。

```
Next message:
我想查询中通快递，单号为[REDACTED]
{
  "text": "我想查询中通快递，单号为[REDACTED]。",
  "intent": {
    "id": 1403204364361344680,
    "name": "search_express",
    "confidence": 0.9999998307907104
  },
  "entities": [
    {
      "entity": "express",
      "start": 4,
      "end": 6,
      "confidence_entity": 0.998625636100769,
      "value": "中通",
      "extractor": "DIEITClassifier"
    },
    {
      "entity": "number",
      "start": 12,
      "end": 26,
      "confidence_entity": 0.9984984397888184,
      "value": "[REDACTED]",
      "extractor": "DIEITClassifier"
    }
  ],
  "intent_ranking": [
    {
      "id": 1403204364361344680,
      "name": "search_express",
      "confidence": 0.9999998307907104
    }
  ]
}
```

图 11 快递单号查询场景测试

## 3.4 中医症状查询

### 3.4.1 中医症状查询情景

在文件 stories.yml 中自定义了中医症状查询的情景。

表 8 中医症状查询对照表

代码顺序	代码内容	代码含义
1	story: syndrome_search	中医症状检索情景
2	intent: search_syndrome	中医症状检索意图
3	action: action_search_syndrome	检索中医症状数据库
4	active_loop: action_search_syndrome_db	循环中医症状数据库

### 3.4.2 中医症状意图自定义

在 nlu.yml 文件中基于有限状态的方法，自定义了带有词槽信息的意图，比如“我想知道[胃热证](syndrome)有哪些？”、“[胃气虚寒证](syndrome)”、“[脾胃气虚，湿郁生热证](syndrome)”等。通过意图识别获得对应的词槽信息，使用动作触发图数据库查询的接口，获取该中医症状的信息。通过人工手动标注查询话术模板总计 300 个，具体标注格式展示如下图所示：

```
examples: |
- [肝郁脾虚证](syndrome)
- [阳明病证](syndrome)
- [阳明经证](syndrome)
- [阳明腑证](syndrome)
- [太阴病证](syndrome)
- 我想知道[气分证](syndrome)有哪些?
- [中焦病证](syndrome)
- [脾肾阳虚证](syndrome)
- [寒湿中阻证](syndrome)
- [湿热内蕴证](syndrome)
- [脾虚湿困证](syndrome)
- [肺胃津伤证](syndrome)
- [肺脾气虚证](syndrome)
- 我想知道[胃热证](syndrome)有哪些?
- 我想知道[胃寒证](syndrome)有哪些?
```

图 6 中医症状实体标注

注：知识图谱数据来源为公开可见数据：中医药知识服务平台，数据库建立和数据结构自动转换使用产品工具，数据查询使用多表联查的方式，针对难以识



别的中医数据对文本进行手工标注。数据处理与场景设计无关，故不展开论述。

### 3.4.3 中医症状话术自定义

在 domain.yml 中针对该场景设置话术，举例如下：

表 9 中医症状查询话术对照表

responses	回复含义	回复内容(text)
utter_first	表达自我介绍和问询	“您好，我是中医证候知识助手， 请问有什么可以帮您？”
utter_greet	表达问候话术	“您好，我是中医证候知识助手， 您可以发送“寒证有哪些？”
utter_goodbye	表达再见问候	“再见，祝身体健康～”
utter_donknow	表达不理解和不知道	“不好意思，我没有理解您说的话。”

## 3.5 天气查询

### 3.5.1 天气数据准备

通过定义 fetch\_weather 和 get\_weather\_by\_day 函数爬取天气数据并转化成树形结构格式，默认参数 location（地点）设置为：上海，默认获取 15 天天气数据。

### 3.5.2 天气意图设计

查询天气需要两个维度的实体抽取，即地点和日期，故通过手工标注北上广深的意图作为样例。比如：我想知道[上海](address)[明天](date\_time)的天气、请问[北京](address)[今天](date\_time)的天气如何？、告诉我在[广州](、address)的怎么样、[深圳](address)昨天雨了，[明天](date\_time)还会下吗？等。通过设置 get\_slot 函数进行实体抽取。

### 3.5.3 天气查询话术自定义

话术适用的场景逻辑为：天气 + 时间 + 地点 + 时间、天气/时间/地点 + 地点、天气/时间/地点 + 闲聊 + 询问天气。使用多维数组或者矩阵结构存储检索返回数据。对应数组名或者字段名以及参数对照表，如下：

表 10 数组名称及取值对照表

字段名称	字段含义	取值
location	地点	北京，上海，广州，深圳

<b>date_time</b>	时间	昨天, 今天, 明天, 后天
<b>text_day</b>	时段	白天
<b>text_night</b>	时段	夜晚
<b>temperature</b>	温度	最高, 最低

在 domain.yml 中针对该场景设置话术, 举例如下:

表 11 天气查询话术对照表

responses	回复含义	回复内容(text)
<b>utter_sorry</b>	表达抱歉	"暂不支持查询 {} 的天气"
<b>utter_answer</b>	表达天气查询结果	"{} {} ({} ) 的天气情况为: 白天: {}, 夜晚: {}, 气温: {}-{} ° C"
<b>utter_error</b>	表达查询有误	"查询失败。"
<b>utter_wait</b>	安抚用户等待心情	"正在查询, 请稍后..."

### 3.5.4 天气查询功能使用

天气查询功能展示如图所示, 通过地点(Location)和日期(date\_time)的实体输入, 自动获取具体格式为 YYYY-MM-DD 的日期。数据来源详见本模块第一小节内容, 仅限个人使用, 不作为任何商业用途, 详见下图:

```
Text message:
你好, 请说出需要提供天气预测服务的地点和时间
Your input -> 
什么时候?
Your input -> 后天
正在查询中, 请稍后 ...
后天 的天气情况为: 白天: 多云; 夜晚: 多云; 气温: 23-15 ° C
Your input -> 明天
正在查询中, 请稍后 ...
明天 的天气情况为: 白天: 多云; 夜晚: 多云; 气温: 23-15 ° C
Your input -> 今天
正在查询中, 请稍后 ...
今天 的天气情况为: 白天: 晴; 夜晚: 多云; 气温: 25-15 ° C
```

图 7 天气查询功能展示

## 4 工业界智能对话系统产品

### 4.1 基于端到端方法的智能对话系统：阿里小蜜

阿里小蜜势是：在自然语言理解模块使用了基于零次学习和感应网络模型的小样本方法和基于迁移学习、知识蒸馏、多任务学习的多样本方法<sup>30</sup>。通过小样本归纳和基于类信息<sup>31</sup>。在对话管理模型使用了业务建模方法开发业务状态机，衔接自然语言模块提供样例描述，并衔接下轮触发以实现多轮对话，并融入了强化学习模型具备鲁棒性能够支持持续在线学习。最突出的优势是它具有用户模拟模块，通过对话诊断来进行人机协同<sup>32</sup>。

### 4.2 基于生成式模型和强化学习方法的智能对话系统：ChatGPT

ChatGPT 是由 OpenAI 开发的人工智能聊天机器人程序，基于 GPT3.5 的大型生成式语言模型并融入了强化学习进行训练，优势是文本自动生成相似于真人写的文章<sup>33</sup>，但目前的局限是仍只能使用文字方式进行交互和接口调用，与传统工业界只能对话系统相比，它无法进行语音交互。

## 5 总结与展望

### 5.1 本文工作总结

本文针对智能对话系统的发展史进行体系理论性研究并追踪当下智能对话系统领域最前沿的趋势：人工智能聊天机器人。由于实验环境局限本文仅能完成基于深度学习方法的智能对话系统（一般智能对话系统领域龙头企业才有财力实现），故本文关注任务型智能对话系统即对话场景的设计和现实世界落地。任务型智能对话系统在各行各业由于能完成特定领域的任务，在当下具有明显的商业价值。（注：本文对话场景设计仅限个人业余研究，非面向商业用途，数据均来自于公开可见数据，所有数据的使用均合法合规，针对设计个人隐私和数据安全的数据均使用马赛克或者随机编码的方式隐去。）并设计 4 个任务型智能对话场景：火车票查询、快递单号查询、中医候症检索、天气查询。

智能对话系统领域聚焦落地使用，非完全面向学术的研究，注重技术使用于创新，该领域的发展史史实则可追溯至 18 世纪前中期，是从通过图灵测试的机器到问答系统到信息检索系统<sup>34</sup>，再从智能语音助理到智能对话系统，一路

迭代过来。其本质是：能够实现人机语言交互并解放人力的人工智能计算机程序。它的迭代进化历程为：基于有限状态规则设定的问答系统、基于关系型数据库的问答匹配系统、基于知识图谱信息检索系统、基于深度学习预训练模型的对话系统、基于迁移学习的端到端的聊天机器人、基于强化学习的人工智能聊天机器人。人气高涨的 ChatGPT 商业估值近千亿，是截止目前在模型和训练方式上是当下最前沿智能对话系统产品（即人工智能聊天机器人），再次引爆人们对智能对话系统领域的关注。

## 5.2 领域未来挑战

### 5.2.1 落地层面

深度学习需要大量数据来训练模型，目前工业界缺少标注的优质语料数据，使得大部分智能对话系统难落地。大部分小型智能对话系统是基于有限状态机的方法，即认为设定动作集合，从而实现特定场景中特定任务上落地。随着国内更多的企业开始从信息化走向数字化转型，有望在数智化场景应用到任务型智能对话系统完成，例如：办公自动化(Office Automation, OA)、机器人流程自动化(Robotic Process Automation, RPA)、流程审批等场景落地。此外，多轮对话在对话轮数比较大的场景难以控制，需要算法调优和性能测试，需要在准确性和计算资源性能之间做平衡。

### 5.2.2 人机交互层面

目前工业界落地并投入使用的智能对话系统，比如：阿里小蜜、美团客服等，能够完成语音识别方面层面的人机交互，并使用了强化学习的方法能够实现其可持续性<sup>35</sup>学习。并且通过业务层面和产业运营层面开发面型不同人群的智能度化功能进行落地。因此，这也导致智能对话系统面向人机交互的更新迭代的挑战，人机交互设计将从人工智能视角得到新的发展和推动。

### 5.2.3 算法可解释性

目前，该领域工业界产品普遍使用基于迁移学习和微调的方法部署预训练 Bert 模型和预训练单向 Transformer 完成自然语言理解模块，并在多轮对话管理当中使用深度学习的方法进行对话跟踪，完成自然语言生成，使用特性场景的应用。工业界和学术界公认深度学习算法模型类似黑盒，内部复杂不可解释。

算法可解释性始终是智能对话系统难以逾越的挑战之一。

#### 5.2.4 人工智能算法合规

国家先后出台了《国家新一代人工智能标准体系建设指南》、《人工智能法》<sup>36</sup>等法规，其中明确提出了几个要点：需要保证算法透明度，即应明确告知用户所部署的人工智能技术和潜在风险点；针对可解释性不强的痛点，官方应向所有相关方解释部署人工智能技术的应用程序逻辑；需要确保输入人工智能模型进行训练的数据具有良好的质量和相关性以满足数据保护的相关要求；需要确保人工智能算法产品，其决策不会产生种族歧视或对任何产生群体偏见；产生人工智能意外结果时需落实应急措施进行人工干预。

#### 5.2.5 网络安全与隐私保护层面

就网络安全层面需实施有效的网络安全措施，以保证人工智能模型和程序的可靠性、弹性和稳健性，能够应对黑客网络攻击<sup>37</sup>。就隐私保护层面，用于训练人工智能的数据都是经过精细标注，针对敏感性数据存在隐私隐患。结合网络安全要求，涉及个人数据可能会在没有足够网络安全保护下存在风险。

#### 5.2.6 算力层面

算法、算力、大数据作为人工智能“三驾马车”，即核心驱动力和生产力。就当下 CHATGPT 而言，其模型训练需要通过网页爬虫等方式获取超大规模的训练预料数据，并设置爆炸级的模型参数，整个模型的训练需要超强的计算资源<sup>38</sup>。对底层算力要求会愈发提高，而算力是人工智能技术落地的保障。

### 5.3 领域未来展望

#### 5.3.1 5G 时代的智能对话系统

5G 是无线通信产业一次重大的技术变革，是国民经济转型升级的重要推动力，是迈向智能化世界的必经之路。应用 5G 网络的智能对话系统能让人工智能应用更通用<sup>39</sup>，与此同时人工智能将优化 5G 网络<sup>40</sup>。

#### 5.3.2 多模态的人工智能聊天机器人

人工智能专家刘菁菁提出，多模态人工智能的目标是模拟人脑理解多模态数据信号并完成推断、预测、决策等任务<sup>41</sup>。多模态信号包括语音、文本、图片、视频等，多模态人工智能聊天机器人是工业界现有智能会话产品未来迭代

的方向。

### 5.3.3 融入情感计算的人工智能聊天机器人

情感计算(Emotion Computing)是个跨学科领域，包括：计算机科学、心理学、认知科学等，可能涉及人体姿态估计、面部情感检测、语音识别等模块，旨在使机器能够理解人类情绪并做出相适应的行为，实现人机交互<sup>42</sup>。本文认为情感计算是人工智能聊天机器人发展的蓝海方向之一。

## 5.4 本文不足总结与展望

本文不足之处在于无法穷尽智能对话系统领域，因为该领域是多个领域的知识实践的交集，并且个人认为智能对话系统领域实践更面向场景落地，但是写实现的功能模块又需要对各个领域的背景知识了解进行展开。因此，本文通过各方面的知识拼接对我自己设计的智能对话系统总体架构展开介绍，并对 4 个对话场景设计进行展示，会先的没有重点和无法体现工作量。但是本文章旨在通过完整故事性的方式，论述本智能对话系统整体架构和业务逻辑展示。此外，由于涉及一些操作规范性、数据安全合规性，并且本文 4 个对话场景设计均面向个人而不具有任何商业性用途，旨在体现个人的技术创新的实践。

针对本文的展望即对于个人的期待是希望有机会真正进入这个领域，而非通过各领域知识交集来辅助对这个领域的研究，但愿有机会能对这个领域深入研究，通过大型场景的技术赋能来体会场景落地的魅力，成为大规模技术落地和智能场景人类美好愿景实现的一份子。

## 参考文献

---

- [1] Chen, H., Liu, X., Yin, D. and Tang, J., 2017. A Survey on Dialogue Systems: Recent Advances and New Frontiers. *Acm Sigkdd Explorations Newsletter*, (Volume: 19), pp.25-35.
- [2] Schuetzler, R.M., Grimes, G.M., Giboney, J.S. and Rosser, H.K., 2021. Deciding Whether and How to Deploy Chatbots. *MIS Quarterly Executive*, (Volume: 20).
- [3] Adiwardana, D., Luong, M.T., So, D.R., Hall, J., Fiedel, N., Thoppilan, R., Yang, Z., Kulshreshtha, A., Nemade, G., Lu, Y. and Le, Q.V., 2020. Towards a Human-like Open-domain Chatbot. *arXiv preprint arXiv:2001.09977*.
- [4] Y Yan, Z., Duan, N., Bao, J., Chen, P., Zhou, M., Li, Z. and Zhou, J., 2016, August. Docchat: An Information Retrieval Approach for Chatbot Engines Using Unstructured Documents. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, (Volume 1), pp.516-525.
- [5] Shen, Y., Ding, N., Zheng, H.T., Li, Y. and Yang, M., 2020. Modeling Relation Paths for Knowledge Graph Completion. *IEEE Transactions on Knowledge and Data Engineering*, (Volume 33), pp.3607-3617.
- [6] Qiu, M., Li, F.L., Wang, S., Gao, X., Chen, Y., Zhao, W., Chen, H., Huang, J. and Chu, W., 2017. Alime Chat: A Sequence to Sequence and Rerank Based Chatbot Engine. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, (Volume 2), pp. 498-503.
- [7] Roose, K., 2022. The Brilliance and Weirdness of ChatGPT. *The New York Times*.
- [8] Chen, H., Liu, X., Yin, D. and Tang, J., 2017. A Survey on Dialogue Systems: Recent Advances and New Frontiers. *Acm Sigkdd Explorations Newsletter*, (Volume: 19), pp.25-35.
- [9] Huang, Y., Fei, T., Kwan, M.P., Kang, Y., Li, J., Li, Y., Li, X. and Bian, M., 2020. GIS-based Emotional Computing: A Review of Quantitative Approaches to Measure the Emotion Layer of Human–environment Relationships. *ISPRS International Journal of Geo-Information*, (Volume: 9), pp.551.
- [10] Floridi, L. and Chiriatti, M., 2020. GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds and Machines*, (Volume: 30), pp.681-694.
- [11] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D. and Sutskever, I., 2019. Language Models are Unsupervised Multitask Learners. *OpenAI Blog*, (Volume: 1),

---

pp.9.

- [12] Dale, R., 2021. GPT-3: What's It Good For?. *Natural Language Engineering*, (Volume: 27), pp.113-118.
- [13] Goel, V., Kuo, H.K., Deligne, S. and Wu, C., 2005. Language Model Estimation for Optimizing End-to-end Performance of a Natural Language Call Routing System. *IEEE International Conference on Acoustics, Speech, and Signal Processing*. (Volume: 1), pp. 1-565.
- [14] Wu, X., 2019. *Multimodal Ultrasound Photoacoustic Imaging for Monitoring Focused Ultrasound Ablation Therapy: Machine-Learning-Based Tissue Classification and System Development*. North Carolina State University.
- [15] Licht, C. and Chabot, D., 2006. The Chabot Emotional Differentiation Scale: A theoretically and psychometrically sound instrument for measuring Bowen's intrapsychic aspect of differentiation. *Journal of Marital and Family Therapy*, (Volume: 32), pp.167-180.
- [16] Huang, Y., Fei, T., Kwan, M.P., Kang, Y., Li, J., Li, Y., Li, X. and Bian, M., 2020. GIS-based Emotional Computing: A Review of Quantitative Approaches to Measure the Emotion Layer of Human–environment Relationships. *International Journal of Geo-Information*, (Volume: 9), pp.551.
- [17] Goo, C.W., Gao, G., Hsu, Y.K., Huo, C.L., Chen, T.C., Hsu, K.W. and Chen, Y.N., 2018. Slot-gated Modeling for Joint Slot Filling and Intent Prediction. *Computational Linguistics: Human Language Technologies*, (Volume: 2), pp. 753-757.
- [18] Trivedi, K., 2019. *Multi-label Text Classification Using Bert—the Mighty Transformer*. Hugging Face.
- [19] Kim, S., Yang, S., Kim, G. and Lee, S.W., 2019. Efficient Dialogue State Tracking by Selectively Overwriting Memory. *arXiv preprint arXiv:1911.03906*.
- [20] Acheampong, F.A., Nunoo-Mensah, H. and Chen, W., 2021. Transformer Models for Text-based Emotion Detection: A Review of BERT-based Approaches. *Artificial Intelligence Review*, pp.1-41.
- [21] Lai T M, Tran Q H, Bui T, 2020. A Simple But Effective Bert Model for Dialog State Tracking on Resource-Limited System. *Speech and Signal Processing, IEEE*, pp.8034-8038.
- [22] G González-Carvajal, S. and Garrido-Merchán, E.C., 2020. Comparing BERT Against Traditional Machine Learning Text Classification. *arXiv preprint*



---

arXiv:2005.13012.

- [23] Peng B, Li X, Gao J, 2018. Deep Dyna-q: Integrating Planning for Task-completion Dialogue Policy Learning. arXiv preprint arXiv:1801.06176.
- [24] Van Gurp, J. and Bosch, J., 2000. On the Implementation of Finite State Machines. Variability in Software Systems the Key to Software Reuse, pp.45.
- [25] Su P H, Budzianowski P, Ultes S, et al, 2017. Sample-efficient Actor-critic Reinforcement Learning with Supervised Data for Dialogue Management. arXiv preprint arXiv:1707.00130.
- [26] Saha, T., Gupta, D., Saha, S. and Bhattacharyya, P., 2018. Reinforcement Learning Based Dialogue Management Strategy. Springer International Publishing, pp. 359-372.
- [27] Yan, Z., Duan, N., Chen, P., Zhou, M., Zhou, J. and Li, Z., 2017. Building Task-oriented Dialogue Systems for Online Shopping. In Proceedings of Conference on Artificial Intelligence (Volume: 31).
- [28] Gatt, A. and Krahmer, E., 2018. Survey of the State of the Art in Natural Language Generation: Core Tasks, Applications and Evaluation. Journal of Artificial Intelligence Research, pp.65-170.
- [29] Oraby, S.M., 2019. Stylistic Control for Neural Natural Language Generation. University of California Santa Cruz.
- [30] Radford, Alec and Karthik Narasimhan, 2018. Improving Language Understanding by Generative Pre-Training. Semantic Scholar.
- [31] Qiu M, Li F L, Wang S, et al, 2017. Alime Chat: A Sequence to Sequence and Rerank Based Chatbot Engine. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, (Volume 2: Short Papers), pp. 498-503.
- [32] Geng R, Li B, Li Y, et al, 2019. Induction Networks for Few-shot Text Classification. arXiv preprint arXiv:1902.10482.
- [33] Wu, X., 2019. Multimodal Ultrasound Photoacoustic Imaging for Monitoring Focused Ultrasound Ablation Therapy: Machine-Learning-Based Tissue Classification and System Development. North Carolina State University.
- [34] Dwivedi, Y.K., Kshetri, N., Hughes, L., Slade, E.L., et al, 2023. So What If ChatGPT Wrote It: Multidisciplinary Perspectives on Opportunities, Challenges and Implications of Generative Conversational AI for Research. International Journal of Information Management, pp. 102642.
- [35] Ji Z, Lu Z, Li H, 2014. An Information Retrieval Approach to Short Text

---

Conversation. arXiv preprint arXiv:1408.6988.

- [36] Buhalis, D. and Moldavska, I., 2022. Voice Assistants in Hospitality: Using Artificial Intelligence for Customer Service. *Journal of Hospitality and Tourism Technology*, (Volume 13), pp.386-403.
- [37] National Standardization Administration Committee, Central Internet Information Office, National Development and Reform Commission, Ministry of Science and Technology, Ministry of Industry and Information Technology, 2020. *National New Generation of Artificial Intelligence Standard System Construction Guide*.
- [38] Bécue, A., Praça, I. and Gama, J., 2021. Artificial Intelligence, Cyber-threats and Industry 4.0: Challenges and Opportunities. *Artificial Intelligence Review*, (Volume: 54), pp.3849-3886.
- [39] Zhang, M. and Li, J., 2021. A Commentary of GPT-3 in MIT Technology Review. *Fundamental Research*, (Volume: 1), pp.831-833.
- [40] Friis, K. and Lysne, O., 2021. Huawei, 5G and Security: Technological Limitations and Political Responses. *Development and Change*, (Volume: 52), pp.1174-1195.
- [41] Liao, L. and Gu, F., 2022. 5G and Artificial Intelligence Interactive Technology Applied in Preschool Education Courses. *Wireless Communications and Mobile Computing*.
- [42] Cukurova, M., Kent, C. and Luckin, R., 2019. Artificial Intelligence and Multimodal Data in the Service of Human Decision-making: A Case Study in Debate Tutoring. *British Journal of Educational Technology*, (Volume: 50), pp.3032-3046.
- [43] Huang, Y., Fei, T., Kwan, M.P., et al, 2020. GIS-based Emotional Computing: A Review of Quantitative Approaches to Measure the Emotion Layer of Human–environment Relationships. *International Journal of Geo-Information*, (Volume: 9), pp.551.

---

## 致谢

首先，感谢刘亮亮老师在本毕业论文中的耐心而专业的指导，并且刘亮亮老师的自然语言处理课程让我受益匪浅，让我有了解智能对话系统领域的机会。经过几个月的努力，从开始论文选题到开题报告，从文献检索到文献综述，从理论性研究到实验，每一个部分都是对自我的考验，历经千辛万苦终于完成论文撰写工作。其实，在真正研究与实验过程中遇到了很多困难和认知障碍，每一次都是非常自耗的过程。

很庆幸，自己大一时努力学习获得了通过了系里转专业审批，本专业的几乎每一门课对我针对这个领域的理解与研究均有帮助，感谢每一门任课老师精心备课和耐心指导。在校期间遇到的很多良师益友以及努力积极共同进步的伙伴，无论是学习、工作还是生活，都给予了我很多帮助。

此外，我要特别感谢一个人，就是我自己，希望能够在持续遇到志同道合的人一起持续进步，我始终坚信技术是为创新赋能重要实现途径，祝愿我自己能在技术与创新路上越走越远。

最后，由于我的数学水平、科研水平、实验环境资源有限，所写论文难免比较粗浅，恳请各位老师和同学谅解、提出批评和指正。