

文章编号: 1007-5321(2019)06-0010-10

DOI: 10.13190/j.jbupt.2019-169

智能对话系统架构及算法

黄 毅, 冯俊兰, 胡 珉, 吴晓婷, 杜晓宇

(中国移动研究院, 北京 100053)

摘要: 智能人机对话系统综合应用了人工智能领域多项核心技术, 随着深度学习、强化学习等基础算法的快速发展, 人机对话系统的总体架构、算法体系以及应用模式都有了重大的改变和提升。从人机对话系统的架构、人机对话系统的核心算法和该领域面临的挑战以及技术研究方面做了综述。

关键词: 对话系统; 深度学习; 人机交互

中图分类号: TP391.1 **文献标志码:** A

Architecture and Algorithms of Intelligent Dialogue System

HUANG Yi, FENG Jun-lan, HU Min, WU Xiao-ting, DU Xiao-yu

(China Mobile Research Institute, Beijing 100053, China)

Abstract: Intelligent human-machine dialogue system comprehensively utilizes a number of core technologies in the field of artificial intelligence. In recent years, with the development of basic algorithms such as deep learning and reinforcement learning, the overall structure, algorithm system and application mode of human-machine dialogue system have been changed and improved greatly. In order to sort out and summarize this technology, the architecture of human-machine dialogue system, the core algorithm of human-machine dialogue system, challenges and technical directions in this field are reviewed.

Key words: dialogue system; deep learning; human-machine interaction

现代计算机之父阿兰·图灵1950年在哲学刊物《思维》上发表了“计算机与智能”一文, 提出了经典的图灵测试(the Turing test)^[1], 在测试者与测试者(一个人和一台机器) 隔开情况下, 通过一些装置, 如键盘, 向被测试者随意提问。进行多次测试后, 如果有超过30%的测试者不能确定被测试者是人还是机器, 那么这台机器就通过了测试, 并被认为是具有人类智能。由此可见, 人机对话技术在人工智能领域的重要性。

经过几十年的研究和发展, 一个非人类但能同人类对话的交互系统不再是科幻片中的画面, 它以各种各样的形式走入了日常生活的方方面面, 如手

机上的个人助理、代替公司客户服务代表的客户服务机器人、电话里的营销机器人、可以跟小朋友聊天的智能音响等。随着技术的进步, 这些机器人的智能程度和实用性有了极大提升, 但同时也引起了人们的担心, 如个人隐私、智能系统所遵循的道德规范等。

按照对话系统的技术架构来分, 人机对话系统可以分为任务型对话系统和非任务型对话系统两大类, 在实际商业应用中这2种对话方式常常结合在一起使用。任务型对话通过交互的方式帮助用户完成一项或多项特定的任务, 系统能够完成的任务通常是一个有限的集合, 每个任务有明

收稿日期: 2019-07-31

作者简介: 黄 毅(1989—), 男, 研究员。

通信作者: 冯俊兰(1974—), 女, 国家特聘专家, E-mail: fengjunlan@chinamobile.com.

确的流程、输入和输出,如智能订票系统、账单查询系统等。非任务型对话系统没有明确的任务列表,可以用于同用户闲聊,或者回答用户在某一个或多个领域的问题。多篇综述性文章中大多强调了学者在该领域前沿技术上的探索^[2-4],但人机对话系统技术架构的演进以及对商用系统所面临的挑战和现状描述相对较少。笔者综合考虑两个方面,从对话系统架构和算法的角度,探讨了人机对话的现状和未来。

1 人机对话系统的架构

1.1 任务型对话系统架构演进

当前任务型对话系统的典型架构主要分为 2 类:第 1 种是管道式架构,如图 1 所示,包含语音识别(ASR,automatic speech recognition)、自然语言理解(NLU,natural language under standing)、对话管理器(DM,dialogue manager)、语言生成(NLG,natural language generation)、语音合成(TTS,text to speech)等组件^[5]。纯文本的对话不含语音识别和语音合成模块。近几年随着深度学习的发展,一部分对话系统将对话管理器分成对话状态跟踪(DST,dialogue state tracking)及对话策略优化 2 个部分,使得对话管理更加依赖统计模型的方法,更加鲁棒。管道式架构中各个模块独立优化,也是目前商用系统的典型架构。

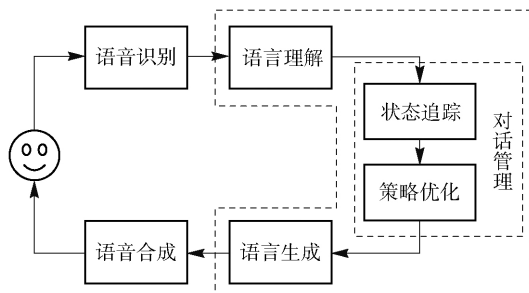


图 1 管道式架构

图 2 所示为目前常见的纯文本管道式架构,对话系统、模拟器和多个外部服务、资源组成多智能体的交互整体。对话系统第 2 种架构如图 3 所示,称之为端到端的人机对话系统架构,各个模块彼此依赖,统一优化。目前端到端的人机对话系统随着深度学习等一系列算法技术的进步而迅速发展,逐渐成为当前的研究热点。

总体来说,任务型对话系统的架构分类如图 4^[6]所示。

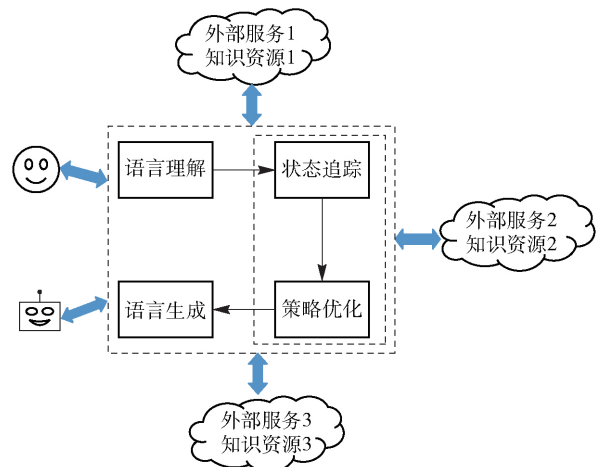


图 2 多智能体对话系统

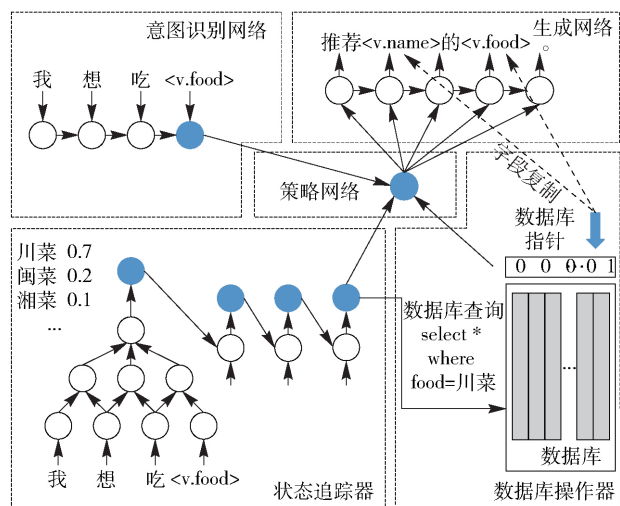


图 3 经典端到端任务型对话系统

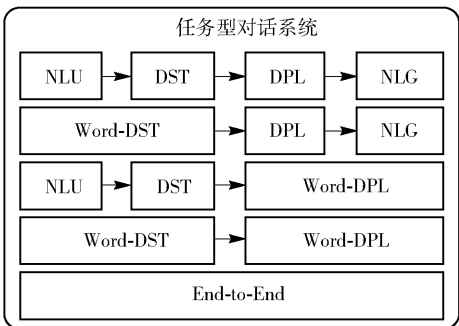


图 4 任务型对话系统的架构分类

1.1.1 管道式架构

1) 自动语音识别

将用户的语音信号转写为文字,在人机交互系统中语音识别系统有的是通用的,有的是面向具体任务场景专门训练的。ASR 的输出也是多样化的,有 1-Best 文本,即输出得分最高一段文字,也有的

ASR 输出 N -Best 文本,即得分高的多段文字,主要考虑是 ASR 的错误识别会给后续模块带来不可恢复的错误。ASR 也可以输出一个字混淆图,在此图中每一个词位都有一个 N -Best 的列表。早期依据关键词识别的对话系统从用户语音中抽取关键词。不同类型的对话系统根据业务需求选择 ASR 引擎。

2) 自然语言理解

完成的具体任务在不同对话系统中差异较大,比较共性的能力包括领域识别、意图识别、对话行为识别、槽位识别和槽位值抽取。领域识别是将用户文字输入或者 ASR 输出的文本分类到各个子领域,如查询天气、订机票等。用户意图指用户想要完成的具体任务,通常同具体任务紧密相关,如一个账单相关意图包括查询账单、账单投诉、账单比较等。对话行为也是用于描述用户意图,但不像用户意图那样从具体完成的任务流程角度来定义意图,而是从对话本身更通用的角度来识别用户当前对话的主要目的,如用户提问、用户确认一个问题、陈述一个事实等。在人机对话领域,对话行为识别是一个定义非常清晰的任务,国际上有多项比赛支持这些研究。槽位识别和槽位信息的提取是一个典型序列标注问题,对输入句子中的每一个字标注相应的槽位标识。下文将会从算法的角度进一步说明这些问题目前都是如何被解决的。

3) 对话管理器

对话管理器是一个人机对话系统的中心,它负责管理对话上下文的状态和控制对话流,根据当前 NLU 的输出及对话的上下文,来决定系统下一步的行为,包括说什么和做什么。对话管理器的主要功能包括对话状态管理、执行对话流和决策下一步行为。

① 对话状态管理。对话状态通常用槽位填充或者语义框架来表示。当子任务、槽位个数、槽位取值增多的时候,对话状态也相对复杂。早期的对话系统通常使用规则来选择最可能的槽位填充为系统的状态。但基于规则的系统往往导致系统鲁棒性不强,一旦发生错误,恢复困难。语音识别和自然语言理解均是基于概率统计意义上的优化,因此近年来对话状态跟踪也转变为更多基于统计的方法,不再是选择最佳的,而是输出一个针对对话状态的概率分布。国际上 DST Challenge 正是为这个任务而专门设计的,大量的研究工作推动了这个方向的技术进步,第2节将详细介绍该领域的主流算法。

② 对话流的执行。对话流可以表示为一个树的模型或者一个图的模型。传统的人机对话系统通常用 XML(extensible markup language) 等表示语言来定义一个对话树,近期许多工业界的人机对话系统将复杂的用户意图更灵活地嵌套和连接在一起形成一张图。对话管理器的主要功能之一即执行对话流。

③ 做出决策。根据当前的系统状态来决定系统下一步的行为。有监督学习、强化学习等算法已经被广泛用于增强决策的智能度,详细算法介绍见下文。

4) 语言生成

语言生成模块将对话系统的行动转化为自然语言。传统的方法通常是基于模版,通过调用外部的资源库等方式来获得模版中需要更换的部分。现代 NLG 技术本身有了很大的进步,各种深度学习模型已经被用于语言的自动生成。

5) 语音合成

将生成的语言合成为人的声音,与用户沟通。最近两年,语音合成技术发展迅速,这里不展开描述。

1.1.2 端到端架构

近年来,随着深度神经网络模型的发展,研究人员提出了端到端的人机对话系统架构。由于内部独立模块的训练过程,管道式架构往往存在模块之间错误累加的问题。与管道式架构不同,端到端模型根据误差的反向传播共同调整和优化模型内部的网络结构和参数,直到模型收敛或达到预期的效果,中间所有的操作都包含在神经网络内部,不再分成多个独立模块分别处理。有的方法虽然是端到端的方法,但还是单独设计模型的部件,不同部件解决管道方法中某个或多个模块所承担的任务。

图3^[7]所示为一种基于神经网络的端到端可训练任务导向型对话系统,将人机对话的过程看作为从对话历史到系统回复的一种映射问题,通常应用编解码模型来进行端到端训练。端到端对话系统是在有监督的方式进行训练,需要大量人对话的训练数据,理论上有很多优点,但目前从结果上看仍需要在技术上做大的提升,以确保对话的逻辑性和鲁棒性。

1.2 非任务型对话系统架构演进

非任务型的对话系统也称为聊天机器人,在开放领域或某一个特定领域与人进行信息沟通,系统

没有特定完成任务的列表和槽位列表。

1.2.1 基于生成模型的架构

典型的非任务型对话系统的回复,一般通过生成方法生成或基于检索的方法抽取。序列到序列的对话生成模型将用户这轮对话的问题作为输入,借助神经网络将用户的问题编码为一个编码矢量,该编码矢量同对话上下文的隐含矢量结合,共同作为一个神经网络解码器的输入,逐词生成系统回复的句子。

1) 编码器和解码器的常用模型包括循环神经网络(RNN, recurrent neural network)、长短时记忆循环神经网络(LSTM, long short-term memory)、门控循环神经网络、卷积神经网络(CNN, convolutional neural networks)、注意力模型或者组合。

2) 对话上下文的编码。研究人员将对话上下文的语义信息有效地用矢量表示,提出了分层的架构,在编码和解码之间加了一层专门对上下文进行建模。

3) 对话的实义性和多样性。在当前以序列到序列为基础的对话系统中,一个具有挑战性的问题是,这些系统倾向于产生无关紧要的或不明确的、通用的、没有意义的回复,如“我不知道”、“哈哈”这样的无实际含义的回复。解决这类问题的方法之一是修改目标函数,在生成式对话中加上行列式点过程,使得生成每个词的时候既考虑质量又考虑多样性。另外一个变化是序列到序列模型同对抗神经网络的结合,以确保生成的答案同人工回复的相似性。

4) 主题一致性。为了确保对话过程中内在主题的一致性,对对话主题进行建模,作为额外的输入给对话生成的解码模型。

5) 系统人格的一致性。业界典型的做法是将用户的个性化信息用向量来表示,嵌入序列中进行学习,从而确保不同用户的人格和属性特征的一致性。

6) 情感。在传统的序列到序列模型的基础上,使用静态的情感向量嵌入表示,动态的情感状态记忆网络和情感词外部记忆的机制,使得新模型可以根据用户的输入以及指定情感分类输出相应情感的回复语句。

1.2.2 基于检索的架构

基于检索的方法从候选回复中选择回复。检索方法的关键是消息-回复匹配,匹配算法必须克服消息和回复之间的语义鸿沟。基于学习的搜索和排

序算法在基于检索的对话系统中起着至关重要的作用,基础算法在近几年取得了非常大的进步,基于检索的对话系统展现出非常客观的性能和准确率。在众多非任务型对话中,基于检索的方法输出的结构也常常作为生成对话模型的输入,或者进行平行组合使用。

检索式对话的基本思路就是将输入和候选输出投射到同一个语义空间中,判断两者是否相似。对于是否相似的计算而言,经典方式是基于独热编码或词袋模型,这是比较传统的表达方式。而当深度学习崛起之后,开始使用表示学习,也就是学习出向量嵌入表示,最后是基于抽象表示的匹配相似度计算。

近年来,有很多模型被提出,如微软的DSSM^[8]、CDSSM^[9-10]及相关系列模型;华为诺亚方舟实验室提出的ARC-I和ARC-II^[11];斯坦福的Tree-LSTM^[12];IBM的ABCNN^[13];Pang等^[14]提出的MatchPyramid;Qian等^[15]提出的ESIM等。通常,深度匹配模型可以从大类上分为表示型和交互型2种表示型模型,如DSSM等,侧重对表示层的学习和构建,它会在表示层将输入转换成整体表示向量,利用深度网络进行语义表示的加强。这种模型的问题是容易发生语义偏移,上下文信息的重要性难以显式衡量。交互型模型,如ARC-II等,这类模型在输入层就进行词语间的先匹配,并将匹配的结果作为灰度图进行后续的计算和建模,能够更精细地处理句子中的联系。它的优势是可以较好地把握语义焦点,对上下文重要性合理建模。近年来,基于检索的架构在工业界的实践和应用得到了快速的发展。

2 人机对话系统的核心算法

随着人工智能技术的快速发展,人机对话系统中的算法和模型较之前有了很大的变化,深度学习、强化学习、迁移学习等一大批技术和算法被应用到人机对话系统中。

深度学习已成为对话系统的一项基本技术。研究人员将神经网络应用于传统任务导向型对话系统的不同组成部分,包括自然语言理解、对话状态追踪、对话策略优化等。近年来,端到端的框架不仅在非任务型的聊天对话系统中流行,而且在面向任务的对话系统中逐步发展起来。

深度学习能够利用大量的数据,模糊了任务导

向型对话系统和非任务导向型对话系统之间的界限。但是,目前基于端到端模型的对话系统仍然远非完美。

下面从几个方面来描述相关核心算法的进展。

人机对话技术中的核心问题及相关算法比较见表 1,其中的任务也是近几年研究比较多的几个子任务。相关经典数据和评测指标与对应任务的对比见表 2。

表 1 对话系统中经典任务和相关算法比较

任务	算法	优势	劣势
自然语言理解	Classical Classifier	实现简单,经典分类算法常用于意图分类、领域分类、对话动作分类等	需要特征工程
	CRF	考虑结合标签的转移概率,适合槽位识别任务	只使用了固定窗口尺寸,难以扩展
	RNN	能够对语言内容长距离建模,适合于上下文相关的多轮意图分类	梯度消失或者梯度爆炸问题
	LSTM	结合“门”机制,能够选择记忆和忘记相关信息	需要更多的训练数据
对话状态追踪	Hidden Information State Model	理论上能够对任意槽值的依赖进行建模	只能保持 top N 个状态
	Bayesian Update of Dialogue States	能够对状态的概率进行建模	只能对简单的依赖进行表征
	Static Classifier	不需要依赖之前的状态,可对任意的分类器工作	没有考虑到状态的转移,使用了人工定义的特征
	Sequential Classifier	考虑到状态的转移概率	发生在最后一个状态的错误可能会影响当前状态的估计
	Transfer Learning	能够基于现有槽值信息对新任务或领域进行迁移	需要大量的特征设计
对话策略优化	Q-Learning	经典算法,易于实现	不能处理连续的动作空间问题
	Policy Iteration	能够处理连续动作空间问题	优化过程不稳定
	Actor Critic	能够处理连续动作空间问题,稳定地优化	需要严格选择函数的逼近方法
	Transfer Learning	对于原领域和目标领域的槽值适应性	计算复杂度高
端到端	End-to-End LSTM Policy Network with Answer Selection	能够进行对话状态的自动学习	仍然有组件需要手工定制
	End-to-End Training of Modular Dialogue System	只需要少量的人工参与	槽的数量和槽值信息需要提前定义,DST 是分开训练的,需要更多的数据
	Memory Network based End-to-End Dialogue System	对确定槽和槽值信息的依赖性强	难以与人类知识集成

表 2 经典任务和数据与评测指标对比

任务	数据集	评价指标
NLU	ATIS	Accuracy, F1
NLU	Snips	Accuracy, F1
NLU, NLG	TouristInfo	Accuracy, F1, BLEU
NLU	DARPA Communicator	Accuracy, F1
DST	DSTC1 ~ DSTC3	Joint Goal Accuracy
DST	WOZ2, mutli WOZ	Joint Goal Accuracy
DST Policy	TownInfo	Accuracy, Reward, Success Rate
DST Policy	Cambridge Restaurant	Reward, Success Rate, BLEU

2.1 NLU

NLU 是对话系统中的经典任务之一。传统模型有最大熵或支持向量机(或其不同的变种)。基于用户输入语句的词法、词性、句法等语言学特征对问题进行分类,分类的标签集合需要提前确定。

Tur 等^[16]提出了基于深度凸网络(DCN, deep convex network)的可扩展分类体系结构,把集成算法用到意图分类中,用语言模型 N -grams 对用户的输入语句做特征选择,然后把简单的分类器做集成 Stacking: 首先训练多个不同的模型,然后把训练的各个模型的输出作为输入来训练一个模型,以得到最终输出。

针对对话中的行为分类问题, Lee等^[17]提出了基于RNN和CNN并融合短文本信息的混合模型, 从序列化的角度来建模, 提高了分类效果。

随着注意力机制的发展, Liu等^[18]提出带注意力的RNN模型, 把意图识别与槽填充联合学习。模型建立在编码-解码框架的基础上, 融合翻译模型中的对齐概念。解码器的隐层输出加权得到最后的意图分类。双向RNN得到特征并与文本向量拼接后作为解码器的输入, 然后识别槽位。

Goo等^[19]把意图识别和槽填充联合学习。相较于通过损失函数将两者隐式地关联, 提出了槽位门显式结构, 其关注于学习意图和槽注意力向量之间的关系, 通过全局优化获得最终的输出结果。

Jaech等^[20]提出将迁移学习用到领域适配的问题中。领域适配解决这类任务的主要思路是: 利用源领域与目标领域分布之间的KL散度对目标领域模型进行正则化。

经过多年的发展, 基于深度学习可以商业化的NLU技术研究已经取得了一定的进展。目前, 用户意图识别之后所提供的“后续服务”开始成为对话机器人差异化的核心之一。

对于对话系统而言, “后续服务”是指对话状态追踪以及对话策略优化, 或者统一称其为对话管理: 当接收到NLU模块的输出、对话历史及用户特征信息之后, 判断系统应该跳转到什么状态以及执行什么样的动作。

2.2 对话状态追踪

DST基于多轮对话过程中的用户话语信息, 更新对话状态, 为对话决策提供支持。这个领域的研究和技术经过多年的发展, 形成的状态追踪方法主要包括: 基于规则的方法^[21]、生成模型^[22]、判别模型^[23-26]。其中, 判别模型可以反映不同类别数据之间的差异, 准确率更高, 因此判别模型是状态追踪任务中主要使用的模型。

Nikola等^[23]提出NBT(neural belief tracking)框架, 该模型将DST和NLU联合建模, 对每个槽-槽值对二分类, 得到对应的概率值, 不需要大规模的标注语料训练NLU模块和手工构建词汇字典, 在状态追踪的任务中取得很好的结果。该研究成果成为后续很多对话状态追踪任务的基线比较系统。NBT模型基于CNN得到上下文向量、用户向量和槽-槽值对向量3种类型的向量。通过相似度计算和非线性变换后最终进行二分类, 得到当前槽-槽值对的取

值概率, 每个槽选择最高概率对应的槽值作为当前状态值。

Xu等^[24]提出指针网络(PointerNet, pointer network), 旨在解决新槽值问题。相较于之前的状态追踪模型需要提前知道槽和槽值信息, 针对状态追踪过程中的新槽值问题, 提出了一种基于分类模型和编解码模型的混合架构。在编码端, 使用LSTM模型对系统和用户的对话编码, 得到当前的对话向量。在解码端, 模型使用槽对应的向量作为解码端的输入, 使用带注意力机制的LSTM模型得到槽值的开始位置和结束位置, 最后得到槽对应的槽值。因为有些情况下槽值并没有出现在系统与用户的对话文本中, 如none, 所以当模型无法获取槽值的位置时, 使用分类模型得到槽对应的槽值。

Zhong等^[25]使用GLAD(global-locally self attentive dialogue state tracker)模型进行状态追踪任务, 最终返回槽-槽值对在对话中的概率值。首先使用局部和全局注意力机制的LSTM模型分别对用户动作、用户语句和槽-槽值对编码, 得到对应的隐层向量, 然后分别计算用户输入语句与槽-槽值和系统输出动作之间的相似度, 最终通过非线性变换得到槽-槽值的取值概率。将不同槽之间的模型参数共享, 并且使用全局注意力机制捕捉整体信息, 使用局部注意力机制捕捉不同槽的特有信息, 提高了状态追踪的准确率。

Ren等^[26]提出简洁的框架状态网络StateNet, 通过层次模型计算槽和槽值的概率值。在下层模型中, 根据权重将N-best对应位置的词向量相乘、相加后得到输入词向量, 使用N-gram的方法将词向量做线性变换并拼接得到句子向量, 使用词袋模型得到系统动作向量, 将每一个槽对应的词向量作为槽向量。然后将用户句子向量与系统动作向量拼接后与槽向量相乘得到的向量值作为上层LSTM的输入, 上层的输出向量做线性变换后得到当前槽的对话向量。最后计算对话向量与槽值向量的相似度得到槽-槽值对的概率值, 并确定最终的结果。StateNet相比于GLAD, 去掉了局部注意力的结构, 只保留全局的结构。相比于NBT, 没有采用CNN做特征提取, 也没有显式地进行状态更新, 直接通过LSTM进行当前轮信息和历史信息的融合。

为了解决领域数据不足的问题, DST还有很多迁移学习方面的工作, 如基于特征的迁移学习、基于模型的迁移学习等^[27-28]。

2.3 对话策略优化

在对话系统中,对话策略致力于建立状态空间到动作空间的映射,控制着系统和用户之间的交互,对于用户的整体体验至关重要。最直接的对话策略是人工制定对话流程,结合专家知识形成对话流的树状结构,这种方式能够严格控制每一轮交互并保证系统回复的质量,但需要投入大量人工设计和构建对话流程,并且当知识发生变化或动态增加时,对话流程也需要进行相应修改,缺乏灵活性。基于统计学的方法由模型做出决策,能减少人工成本且更具有可扩展性。将对话策略看作分类任务^[29],根据统计的“状态-动作”数据集训练模型,然后训练好的模型将新一轮的对话状态分类到某一动作上。但在实际情况中,训练数据集不可能覆盖所有可能的情况,并且系统仅根据历史数据集中的信息不能完全适应新的交互。强化学习是一种边交互边学习的方式,基于强化学习的对话系统能够在与用户交互的过程中根据反馈持续优化对话策略,克服缺乏训练数据和模型不能完全适应新交互的缺点,并且随着深度学习的发展,深度强化学习^[30-31]在对话场景中的表现越来越好。

强化学习在对话系统中的主要应用是学习对话策略,如图 5 所示。对话系统的策略器选择合适的动作,环境接受动作后状态发生变化,并将奖励值反馈给策略器,然后策略器再继续选择下一个动作。

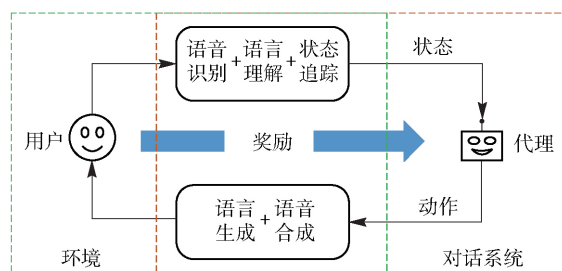


图 5 强化学习在对话系统中的应用

Heriberto 等^[32]提出 SimpleDS (simple dialogue system) 框架,用户模拟器输出原始文本,状态先用独热编码向量表示,再用 ASR 置信度将向量中的非零值替换,代理为全连接神经网络,动作空间包含 35 个订餐领域的动作,奖励函数是槽位填充比例与动作条件概率的加权求和,该方法将原始文本直接输入到系统中,简化了状态表示及状态转移过程。随后 Heribert 等^[33]又引入领域转移方程,将单领域任务扩展到多领域。

为研究基于强化学习的对话策略,剑桥大学团队^[34]开源了一种多领域任务型对话系统实验环境 PyDial,其中包含 10 个领域的任务型对话数据,对话系统的 NLU、DST、NLG 模块基于规则实现,策略模块包含了手工配置和高斯过程强化学习策略 2 种方式。基于 PyDial 环境,Casanueva 等^[35]、Weisz 等^[36]研究了不同对话策略学习方式,包括 DQN (deep Q-network)、A2C (advantage actor-critic)、eNAC (episodic natural actor-critic)、ACER (actor critic with experience replay) 等,并且采用 action mask 和 action summary 机制,避免了动作集过大导致的学习效率低下。

在管道方法中,NLU 和 DST 模块的输出是强化学习代理的输入,因此 NLU 和 DST 的结果会影响下游代理的学习效果,深度强化学习能够实现自动特征表示以及端到端的学习,避免模块与模块的误差累积。Li 等^[37]提出了一种多领域端到端的方法,将 DST 和对话策略优化作为一个对话管理模块进行训练,NLU 模块使用 LSTM 进行意图识别和槽位填充,DM 模块首先根据 NLU 模块的结构化输出查询数据库,并更新槽位,然后采用 DQN 网络学习策略,输出代理动作。

除上述方法中将代理执行的动作视为动作空间外,也可将生成模型的词汇表作为动作空间,并与 Policy Gradient 强化学习方法结合^[38]。Budzianowski 等^[39]在跨领域任务型对话数据集 (MultiWOZ) 的基础上开源了一种基于编码-解码的端到端对话系统,这种基于生成模型的方法以词汇表为动作空间,其对话策略也被称为词级别的对话策略。由于动作空间过大,词级别的对话策略常收敛缓慢且不能得到最优结果。因此 Zhao 等^[40]提出潜在变量模型,该模型在编码-策略-解码基础上加入潜在动作的嵌入表示,每个潜在动作对应一个句子,环境反馈的奖励作用于潜在动作而非单个词汇,这使得对话系统的决策过程与 NLG 分离,并且可降低策略优化的难度,该方法目前在 MultiWOZ 数据集上策略优化可取得最优效果。

基于深度强化学习的任务型对话系统的策略优化常以任务的成功率作为奖励^[36,40],非任务型对话系统更多的是以 BLEU (bilingual evaluation understudy) 等评估指标作为奖励^[41]。在强化学习与生成模型结合的方法中,除基于策略的方法外,Keshneshloo 等^[42]提出了 actor-critic、Q-learning 等多种强

化学习与生成模型结合的方法,目前主要被应用于自动摘要生成任务上。在对话生成任务上的应用等待更多的研究者去探索。

3 问题和挑战

无论从学术界还是工业界的角度看,当前的人机对话系统仍然存在很多问题和挑战。

1) 语义表征。人类的语言代表了文明。在人机对话系统中的自然语言表达具有多样性、歧义、口语化等多种特征,如何更加合理地将对话中的自然语言、上下文历史等相关信息用机器表征,从而让系统能更好地理解是有待研究的问题。基于深度学习的神经网络模型与传统的语言学做融合,基于符号的表示、基于嵌入的表示在特征表示、语义表示的角度上还有许多问题可以进行深入探索,基于大规模语料的预训练技术(ELMo^[43]、GPT^[44]、BERT^[45]系列)成为近来新的研究方向。

2) 逻辑性。对话过程中的逻辑性是人与人交谈过程中的显著特征。如何让机器在开放域能够具备这个能力,目前仍然存在比较多的困难和问题。通过借助知识图谱和知识嵌入的方式,让机器在一定程度上能够进行引用推理和逻辑判断,但是跨领域、跨行业模型的迁移和泛化仍然是目前的一个研究热点。

3) 一致性。用户在和机器的交互过程中,往往具有鲜明的个性化特点,不同用户的表达方式、行为习惯、个人属性等方面都存在差异。如何让机器首先具备个性化特点,并且在人机交互的过程中,具备表达内容的前后一致性,未来可以通过对抗学习、元学习等模型和技术进行优化。

4) 交互性。当前的对话系统还是以被动回答为主。在真实的对话场景中,良好的对话情形往往是由双方共同把控。主动对话和话题引导、话题的开启和结束等能力,需要机器能够更好地和外部资源、服务等连通,多智能体之间需要进行多任务训练和学习。更好地交互体验需要多模态(语音、语言、视觉、情感等)的共同作用,这也是未来的趋势之一。

更加友好的产品用户体验、更加标准化的对话系统组件、更加高效的模型训练与部署、更加合理的系统评价方式和更加自主的学习更新能力是对话系统在工业界实用产品化面临的几大难点。

人工智能技术飞速发展的今天,它给人类带来

生活便利的同时,一些安全和隐私方面的问题也随之产生。如何让人工智能技术充满温度,更加人性化地服务于行业和社会,符合纲常伦理、公序良俗的创造将是所面临的更大挑战。

参考文献:

- [1] Turing Alan M. Computing machinery and intelligence [J]. Mind, 1950, LIX(236): 433-460.
- [2] 陈晨,朱晴晴,严睿,等. 基于深度学习的开放领域对话系统研究综述[J]. 计算机学报, 2019, 42(7): 1439-1466.
Chen Chen, Zhu Qingqing, Yan Rui, et al. Survey on deep learning based open domain dialogue system [J]. Chinese Journal of Computers, 2019, 42(7): 1439-1466.
- [3] López-Cózar Ramón, Zoraida Callejas, Griol David, et al. Review of spoken dialogue systems [J]. Loquens, 2015, 1(2): 2386-2637.
- [4] Chen Hongshen, Liu Xiaorui, Yin Dawei, et al. A survey on dialogue systems: recent advances and new frontiers [J]. SIGKDD Explorations, 2017, 19(2): 25-35.
- [5] Young Steve, Gašić Milica, Thomson Blaise, et al. POMDP-based statistical spoken dialogue systems: a review [J]. Proceedings of the IEEE, 2013, 101(5): 1160-1179.
- [6] Lee Sungjin, Zhu Qi, Takanobu Ryuichi, et al. ConvLab: multi-domain end-to-end dialog system platform [C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL). Florence [s. n.], 2019: 64-69.
- [7] Wen Tsung-Hsien, Vandyke David, Mrkšić Nikola, et al. A network-based end-to-end trainable task-oriented dialogue system [C]// European Chapter of the Association for Computational Linguistics (EACL). Valencia [s. n.], 2017: 438-449.
- [8] Huang Po-Sen, He Xiaodong, Gao Jianfeng, et al. Learning deep structured semantic models for web search using click through data [C]// Proceedings of the 22nd ACM International Conference on Information and Knowledge Management. Amazon [s. n.], 2013: 2333-2338.
- [9] Shen Yelong, He Xiaodong, Gao Jianfeng, et al. A latent semantic model with convolutional-pooling structure for information retrieval [C]// Proceedings of the 23rd ACM International Conference on Information and Knowledge Management (CIKM). New York: ACM Press, 2014: 101-110.
- [10] Shen Yelong, He Xiaodong, Gao Jianfeng, et al. Learn-

- ing semantic representations using convolutional neural networks for web search [C] // Proceedings of the 23rd International Conference on WWW. Seoul [s. n.], 2014: 373-374.
- [11] Hu Baotian, Lu Zhengdong, Li Hang, et al. Convolutional neural network architectures for matching natural language sentences [C] // Proceedings of the Advances in Neural Information Processing System. Montreal [s. n.], 2014: 2042-2050.
- [12] Tai Kai Sheng, Socher Richard, Manning Christopher D. Improved semantic representations from tree-structured long short-term memory networks [C] // Proceedings of the 53rd Annual Meeting of the Association for Computational (ACL). Beijing [s. n.], 2015: 15-26.
- [13] Yin Wenpeng, Schutze Hinrich, Xiang Bing, et al. AB-CNN: attention-based convolutional neural network for modeling sentence [J]. arXiv preprint arXiv: 1512.05193.
- [14] Pang Liang, Lan Yanyan, Guo Jiafeng, et al. Text matching as image recognition [C] // Proceedings of the 30th AAAI Conference on Artificial Intelligence. Phoenix [s. n.], 2016: 2793-2799.
- [15] Chen Qian, Zhu Xiaodan, Ling Zhenhua, et al. Enhanced LSTM for natural language inference [C] // Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL). Vancouver [s. n.], 2017: 17-29.
- [16] Tur Gokhan, Deng Li, Hakkani-Tur Dilek, et al. Towards deeper understanding: deep convex networks for semantic utterance classification [C] // 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Kyoto [s. n.], 2012: 5045-5048.
- [17] Lee Ji Young, Démoncourt Franck. Sequential short-text classification with recurrent and convolutional neural networks [C] // Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. San Diego California [s. n.], 2016: 515-520.
- [18] Liu Bing, Ian Lane. Attention-based recurrent neural network models for joint intent detection and slot filling [C] // Interspeech 2016, 17th Annual Conference of the International Speech Communication Association. San Francisco [s. n.], 2016: 685-689.
- [19] Goo Chih-Wen, Gao Guang, Hsu Yun-Kai, et al. Slot-gated modeling for joint slot filling and intent prediction [C] // Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. New Orleans [s. n.], 2018: 753-757.
- [20] Aaron Jaech, Heck Larry, Ostendorf Mari. Domain adaptation of recurrent neural networks for natural language understanding [C] // Interspeech 2016, 17th Annual Conference of the International Speech Communication Association. San Francisco [s. n.], 2016: 690-694.
- [21] Franck Démoncourt, Lee Ji Young, Bui Trung H. Robust dialog state tracking for large ontologies [M]. Singapore: Springer, 2016: 475-485.
- [22] Williams Jason, Young Steve. Partially observable Markov decision processes for spoken dialog systems [J]. Computer Speech & Language, 2007, 21 (2): 393-422.
- [23] Mrkšić Nikola, Séaghdha Diarmuid Ó, Wen Tsung Hsien, et al. Neural belief tracker: data-driven dialogue state tracking [C] // Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. Vancouver [s. n.], 2017: 1777-1788.
- [24] Xu Puyang, Hu Qi. An end-to-end approach for handling unknown slot values in dialogue state tracking [C] // Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne, Australia [s. n.], 2018: 1448-1457.
- [25] Zhong Victor, Xiong Caiming, Socher Richard. Globally self-attentive dialogue state tracker [EB/OL]. 2018 (2018-09-01) [2019-12-15]. <http://arxiv.org/abs/1805.09655>.
- [26] Ren Liliang, Xie Kaige, Chen Lu, et al. Towards universal dialogue state tracking [C] // Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels, Belgium [s. n.], 2018: 2780-2786.
- [27] Williams Jason. Multi-domain learning and generalization in dialog state tracking [C] // Proceedings of the SIGDIAL 2013 Conference, the 14th Annual Meeting of the Special Interest Group on Discourse and Dialogue. Metz, France [s. n.], 2013: 433-441.
- [28] Mrkšić Nikola, Séaghdha Diarmuid Ó, Thomson Blaise, et al. Multi-domain dialog state tracking using recurrent neural networks [C] // Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing. Beijing, China [s. n.], 2015: 794-799.

- [29] Zhou Keyan , Zong Chengqing , Wu Hua , et al. Predicting and tagging dialog-act using MDP and SVM [C]// 2008 6th International Symposium on Chinese Spoken Language Processing. Kunming: IEEE , 2008: 293-296.
- [30] Li Yuxi. Deep reinforcement learning: an overview [EB/OL]. 2017 (2017-01-01) [2019-12-10]. <http://arxiv.org/abs/1701.07274>.
- [31] Mnih Volodymyr , Kavukcuoglu Koray , Silver David , et al. Human-level control through deep reinforcement learning [J]. Nature , 2015 , 518(7540) : 529-533.
- [32] Cuayáhuil Heriberto. SimpleDS: a simple deep reinforcement learning dialogue system [J]. Springer , 2017 (427) : 109-118.
- [33] Cuayáhuil Heriberto , Yu Seunghak , Williamson Ashley , et al. Scaling up deep reinforcement learning for multi-domain dialogue systems [C]// 2017 International Joint Conference on Neural Networks (IJCNN) . Anchorage , AK , USA: [s. n.] , 2017: 3339-3346.
- [34] Ultes Stefan , Rojas-Barahona Lina , Su Pei-Hao , et al. Pydial: a multi-domain statistical dialogue system toolkit [C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics-System Demonstrations. Canada: Vancouver , 2017: 73-78.
- [35] Casanueva Inigo , Budzianowski Paweł , Su Pei-Hao , et al. A benchmarking environment for reinforcement learning based task oriented dialogue management [EB/OL]. 2018 (2018-04-03) [2019-12-09]. <http://arxiv.org/abs/1711.11023>.
- [36] Weisz Gellert , Budzianowski Paweł , Su Pei-Hao , et al. Sample efficient deep reinforcement learning for dialogue systems with large action spaces [J]. IEEE/ACM Transactions on Audio Speech and Language Processing , 2018 , 26(11) : 2083-2097.
- [37] Li Xiujun , Chen Yun-Nung , Li Lihong , et al. End-to-end task-completion neural dialogue systems [C]// Proceedings of the Eighth International Joint Conference on Natural Language Processing (IJCNLP) . Taipei : [s. n.] , 2017: 733-743.
- [38] Paulus Romain , Xiong Caiming , Socher Richard. A deep reinforced model for abstractive summarization [C]// 6th International Conference on Learning Representations (ICLR) . Vancouver , BC , Canada: [s. n.] , 2018.
- [39] Budzianowski Paweł , Wen Tsung-hsien , Tseng Bo-hsiang , et al. MultiWOZ-A large-scale multi-domain wizard-of-oz dataset for task-oriented dialogue modelling [C]// Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels [s. n.] , 2018: 5016-5026.
- [40] Zhao Tiancheng , Xie Kaige , Eskenazi M. Rethinking action spaces for reinforcement learning in end-to-end dialog agents with latent variable models [C]// Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies(NAACL-HLT) . Minneapolis , MN , USA: [s. n.] , 2019: 1208-1218.
- [41] Galley Michel , Brockett Chris , Gao Xiang , et al. End-to-end conversation modeling: Dstc7 task 2 description [C]// Dialog System Technology Challenges. Honolulu: [s. n.] , 2018.
- [42] Keneshloo Yaser , Shi Tian , Ramakrishnan Naren , et al. Deep reinforcement learning for sequence-to-sequence models [EB/OL]. 2019 (2019-04-01) [2019-12-15]. <http://arxiv.org/abs/1805.09461>.
- [43] Peters Matthew E , Neumann Mark , Iyyer Mohit , et al. Deep contextualized word representations [C]// Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT) . New Orleans , Louisiana , USA: [s. n.] , 2018: 2227-2237.
- [44] Radford Alec , Narasimhan Karthik , Salimans Tim , et al. Improving language understanding by generative pre-training [EB/OL]. 2018 (2018-05-08) [2019-09-23]. https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf.
- [45] Devlin Jacob , Chang Ming-Wei , Lee Kenton , et al. BERT: pre-training of deep bidirectional transformers for language understanding [C]// Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies(NAACL-HLT) . Minneapolis , MN , USA: [s. n.] , 2019: 4171-4186.