

基于深度学习的交互式问答研究综述

黄欣¹ 雷刚¹ 曹远龙¹ 陆明名²

1 江西师范大学软件学院 南昌 330022

2 同济大学电子与信息工程学院 上海 200092

(xinhuang@jxnu.edu.cn)

摘要 相比传统的一问一答,交互式问答增加了对话上下文和背景等信息,这为理解用户输入和推理答案带来了新的挑战。首先,用户输入不再局限于问题,还可以是告知问题细节、反馈答案可行与否等带有意图的语句,因此需要理解对话中每个语句的意图。其次,交互式问答允许多个角色同时参与一个问题的讨论,为每个角色生成个性化的答案,因此需要理解对话中存在的角色。再次,当交互式问答围绕一段背景文本展开时,需要理解这段背景文本,并从中抽取出现问题的答案。文章对交互式问答的发展及前沿动态进行了调研,分别对无背景交互式问答、有背景交互式问答以及迁移学习在交互式问答的应用3个子方向的新方法和新发现进行了介绍,并对交互式问答的研究前景进行了分析和展望。

关键词 问答系统;交互式问答;背景信息;预训练模型

中图分类号 TP391

Review on Interactive Question Answering Techniques Based on Deep Learning

HUANG Xin¹, LEI Gang¹, CAO Yuan-long¹ and LU Ming-ming²

1 School of Software, Jiangxi Normal University, Nanchang 330022, China

2 School of Electronics and Information Engineering, Tongji University, Shanghai 200092, China

Abstract Compared to the traditional question answering(QA), interactive question answering(IQA) considers dialogue context and background information, which brings new challenges to understand user input and reason answers. First of all, user input is not only limited to questions, but can also be utterances that inform the details of the question and give feedback on whether the answer is feasible or not. Therefore, it is necessary to understand the intent of each utterance in the dialogue. Secondly, IQA allows multiple characters to discuss a question at the same time, generating personalized answers. So, it is necessary to understand different characters and identify them from each other. Thirdly, when IQA revolves around a background document, it is necessary to understand this document and extract answers from it. This paper reviews recent development in three subareas: IQA without background, IQA with background, and the application of transfer learning in IQA, and finally discusses the future perspective of interactive question answering.

Keywords Dialogue system, Interactive question answering, Background, Pre-trained models

智能互联时代的到来让智能语音助手、智能客服机器人、智能音箱等系统蓬勃发展,成为新的交互方式和流量入口。以 AppleSiri¹⁾、百度 DuerOS²⁾ 为代表的智能语音助手进入了人们的日常生活,这些系统广泛应用了交互式问答(Interactive Question Answering)技术,通过自然连续的对话,为用户解答各类问题。相比传统的一问一答,交互式问答增加了对话上下文和背景等信息,这为理解用户输入和推理答案带来了新的挑战。

区别于之前各类问答系统的一问一答,交互式问答系统采用多轮对话机制,更符合人类交流习惯,也能为用户提供更加多元的服务:主动预测用户需求,提醒即将发生的事件或推

荐有价值的内容^[1]。交互式问答可以看作对话与问答的一个结合,因此也有学者称其为对话问答^[2](Conversational Question Answering)。从对话的角度来看,交互式问答与任务型(task-oriented)对话^[3-4]和闲聊(chitchat)对话^[5-6]的区别在于,交互式问答要求返回答案给用户。而任务型对话的目标是帮助用户完成某项事务,如安排会议、规划行程;闲聊对话则关注与用户之间的交流。从问答的角度来看,交互式问答不再是“用户提出一个问题,系统返回一个答案,问题之间彼此独立”这种单次、无交互的形式^[7],而是在问题与答案之间存在多轮交互,问题与问题之间、问题与答案之间互相联系,构成紧密的上下文关系^[8]。正因如此,交互式问答表现出了

¹⁾ <https://www.apple.com/ios/siri/>

²⁾ <https://dueros.baidu.com/>

到稿日期:2021-01-27 返修日期:2021-04-24

基金项目:江西省教育厅科技研究项目(GJJ200318);国家自然科学基金(61962026)

This work was supported by the Science and Technology Research Project of Jiangxi Provincial Department of Education(GJJ200318) and National Natural Science Foundation of China(61962026).

通信作者:雷刚(leigang@jxnu.edu.cn)

足够的多样性和复杂性。

本文对交互式问答的发展及前沿动态进行了调研,分别从无背景交互式问答、有背景交互式问答以及背景迁移的交互式问答3个子方向梳理和介绍了最近发表的新方法和新发现。最后,本文对交互式问答的研究前景进行了分析和展望。

1 任务的定义

不同角度的研究对问答分类不一,本文认为交互式问答的多样性主要体现在形式上。本文依据问答是否围绕某一背景进行,将交互式问答分为无背景交互式问答和有背景交互式问答两种,并分别对其进行阐述。借鉴 Mitchell 等^[9]将阅读区分为微阅读(micro-reading)和宏观阅读(macro-reading)的思想,本文将问答分为无背景交互式问答和有背景交互式问答。无背景交互式问答可以依赖任何资源,包括结构化的知识库、半结构的表、非结构化的文本等,其交互场景如图1所示。有背景交互式问答强调对背景的理解,回答所需的信息都来自背景本身,不需要其他信息作支撑,其交互场景如图2所示。

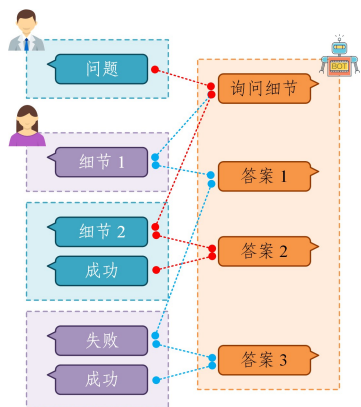


图1 无背景交互式问答示意图

Fig. 1 Interactive Q&A without background

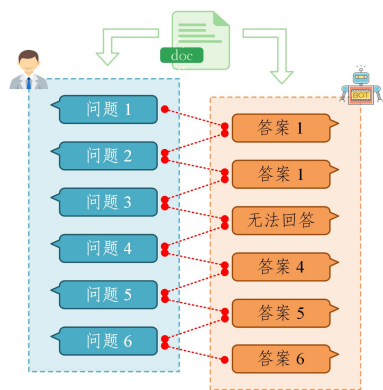


图2 有背景交互式问答示意图

Fig. 2 Interactive Q&A with background

无背景交互式问答允许一名或多名用户参与,如图1所示,其中相同颜色的对话框表示同一人所述,蓝色线段表示男士与对话系统的对话,红色线段表示女士与对话系统的对话。通常,先由一名用户发起提问,短时间内若其他用户碰到相同问题,其他用户不重新提问,而是直接参与已有问答。遇到用

户提问不明确、系统无法回答时,系统会要求用户补充问题的细节,直至找到一个可能的答案返回给用户,如图1中“询问细节-细节1-答案1”和“询问细节-细节2-答案2”过程所示。具体地,以“电脑蓝屏了怎么办?”这个问题为例,系统会首先询问操作系统版本、错误代码等信息,再给出可能的解决方案。补充细节的过程中,系统变成了提问者,用户变成了回答者,因此无背景交互式问答模糊了问与答的界限。对于系统给出的答案,用户还会进行验证,并反馈成功与否,方便系统构建知识库以及其他用户参考。如果答案可行,问答结束(“答案2-成功”);如果问题未解决,系统会根据用户反馈信息生成新的答案,直至问题解决或用户主动结束问答(“答案1-失败-答案3-成功”)。上述过程在技术问答社区(如 Microsoft Community¹⁾, StackOverflow²⁾)十分普遍,许多无背景交互式问答语料^[9-12]也采集自这些社区。

有背景交互式问答(见图2)因为围绕一段背景文本展开,问答过程比无背景交互式问答简单许多,只是重复多轮的问与答,参与用户也只有一名。此外,与无背景交互式问答系统尽可能给出答案不同,有背景交互式问答系统在无法回答问题时会直接告知用户答案不存在。虽然问题有多个,但除第一个问题外,其余问题不能独立看待,因为它们使用代词指代之前的问题或答案中已出现内容或省略问题中相同部分,使语言变得简洁。比如问题1是“姚明多高?”,接着问题2在对科比身高提问时可以省略身高而简单地表达为“科比呢?”。如果还想询问科比的出生地,可以使用“他”指代科比,将问题3表达为“他出生在哪?”。有背景交互式问答的典型应用场景是课堂上的师生问答。

交互式问答的复杂性主要体现在技术上。无背景交互式问答除了用户提问和系统回答之外,还包括询问与告知细节、反馈答案可行与否等功能,准确识别这些功能是系统作答的基础。此外,多用户参与还增加了辨析用户的工作,通过理解不同用户的发言内容,生成有针对性、个性化的答案。

表1列出了交互式问答中的一些基本概念及其定义。

表1 交互式问答中的一些基本概念
Table 1 Basic concepts of interactive Q&A

概念	定义
发言者	每个交互式问答最少有两名发言者:提问者和回答者
语句	切换其他发言者之前,当前发言者所说的全部内容
轮次	无背景交互式问答中,一轮即一个语句 有背景交互式问答中,一轮指连续的问题和答案对
对话	多个轮次构成一个对话。一个对话等价于一次完整的交互式问答过程
上下文	也称对话历史。指到当前为止,已出现的所有语句
意图	交互式问答的一种分类标准,表示语句的功能
角色	包括对话的发言者以及对话中提及的所有人名
背景	发言者围绕某件事物展开交互式问答,可以是一段文本或一张图片

其中,大多数概念都借鉴自对话系统,并就交互式问答做了一些新的定义。后续行文将使用这些概念统一表述,避免前后不一致、含义模糊等问题。

2 无背景交互式问答

因为无背景交互式问答系统的回复除答案外,还可能是

¹⁾ <https://answers.microsoft.com/>

²⁾ <https://stackoverflow.com/>

询问细节、澄清表述等功能的语句,所以,当前系统的自动回复研究主要集中在响应选择(response selection)和响应生成(response generation)两个方向上。响应选择利用文本匹配模型^[13]从一个候选集中找出与问题最相关的响应,响应生成利用 seq2seq 模型^[14]从头开始构造一个响应。

在响应选择方面,主要的模型有:将对话上下文所有语句作为一段长文本拼接在一起,并通过计算对话上下文与候选响应匹配得分进行选择的 Match-LSTM^[15];首先使用不同方法重构最后一个语句,然后获得该语句与候选响应的循环神经网络^[16](Recurrent Neural Network, RNN)和卷积神经网络^[17](Convolutional Neural Network, CNN)联合表示,由此计算匹配得分的 DL2R^[6];从词序列和语句序列两个角度进行匹配的 Multi-View^[18];先获得每个对话语句与候选响应的匹配矩阵,然后用 CNN 将其转换为向量,最后用 RNN 聚合出匹配得分的 SMN^[19];网络结构与 SMN 类似,但匹配矩阵是由基于 Transformer^[20]的自注意力^[21]表示和交叉注意力表示堆叠得到的 DAM^[22];着重考虑对话语句间关系,并通过深度聚合以获得细粒度对话上下文表示的 DUA^[23];融合对话上下文和候选响应之间单词、n-grams^[24]和语句子序列等多种表示的 MRFN^[25];基于多种表示和层次聚合的 HAMR^[26];通过堆叠将浅层匹配信息扩展到深层匹配的 Io^[27]。上述模型综合应用 RNN、CNN、Transformer、注意力机制等进行对话上下文和候选响应的匹配研究,其目标都是捕捉与放大匹配信号。

上述模型均使用 Ubuntu-v1^[10]和 Ubuntu-v2^[28]语料进行验证,虽然训练集有 500000 个不同的对话,验证集和测试集的对话数也各有 50000 个,但每个对话只有 10 个候选响应,与真实情况相差甚远。为此,Gunasekara 等^[29]提出了 Ubuntu-v3 和 Advising 语料,他们将单个对话的候选集大小扩展到 100,还额外构造了一个大小为 120000 的候选集,模拟从海量候选响应中进行选择的情况。此外,他们还设计了两个测试模型灵敏度的数据集,一个数据集中每个对话包含多个正确响应,另一个数据集中某些对话可能没有正确响应。为了考查模型利用外部知识的能力,他们还提供了一个由 Ubuntu 帮助文档、课程信息和学生个人信息构成的外部知识库。

针对上述任务,Chen 等^[30]提出了序列注意力匹配模型,

Gu 等^[31]提出了序列推理模型,Sun 等^[32]提出了基于记忆网络^[33]的 GSMN 模型,Chiang 等^[34]提出了能捕捉语句级别和对话级别信息的 Highway Recurrent Transformer 模型,Lu 等^[35]提出了基于时空匹配的 STM 模型,Whang 等^[36]提出了基于句子嵌入、语句位置嵌入和用户嵌入的增强序列表示模型。上述模型虽然提出了一些新的解决思路,但未脱离对话上下文和候选响应之间匹配的本质,且越来越多的模型强调多种表示的重要性。

在响应生成方面,大量的工作是从不同角度增加响应的多样性,如引入隐含变量^[37]、主题信息^[38]、动态词汇表^[39]、重排序^[40]等。也有一些工作关注对话上下文建模^[41-42]及其有效性研究^[43]。此外,强化学习同样被应用于响应生成的研究中,Li 等^[44]采用对话中的信息性、连贯性以及易答性 3 个属性作为指标,并使用策略梯度方法对模型进行奖励,实验表明该算法产生了更多的互动响应。在强化学习的基础上,Li 等^[45]进一步引入生成对抗网络,所提方法中包含了两个系统,一个是用于生成响应序列的生成模型,另一个是用于区分人类的对话和机器生成的对话的鉴别器,其中鉴别器的输出作为生成模型的奖励,使用强化学习进行训练。与之相似的研究还有文献^[46]等。还有一些研究并不拘泥于响应选择和生成的界限,而是将它们结合在一起^[47-49],将选择结果和对话上下文一起作为解码器的输入。

在细节问题生成方面,一些研究者做了积极探索。Li 等^[50]使用对话上下文相关的词级注意力来获得更准确的语句表示,以及用问题引导的句级注意力来获得更好的对话上下文表征。利用这些注意力机制,他们的模型能准确地理解何时输出答案,何时生成细节提问。Gao 等^[51]提出使用共指对齐和对话流建模来生成问题。具体地,共指对齐通过关联对话上下文中的指称与生成问题中对应的代名词,使问题与对话上下文联系更加紧密,而对话流建模可以使对话更连贯。Cao 等^[52]引入一个二分类器来控制细节问题生成。区别于前 3 个生成方向的工作,Rao 等^[53]提出使用排序的方法获得要回复的细节问题,排序的依据是它们对用户问题的有用性。Aliannejadi 等^[12]公开了一个细节问题选择数据集 Qulac,并提出一个检索框架,包括问题检索、问题选择和文档检索 3 部分,其中问题选择考虑了用户提问和对话上下文。表 2 列出了无背景交互式问答方法。

表 2 无背景交互式问答方法

Table 2 Interactive question and answer methods without background

任务	方法核心	文献	数据集
响应选择	通过不同方式对上下文的语句、语序以及候选词进行不同表示后聚合匹配得分	文献[6]、文献[15]、文献[19]、文献[22-23]、文献[25-27]	Ubuntu-v1 和 Ubuntu-v2
	使用不同的匹配模型进行增强,如注意力、时空信息以及多参数等	文献[30]、文献[35-36]	Ubuntu-v3 和 Advising
	基于记忆网络、Transformer 或推理模型	文献[31]、文献[33-34]	
响应生成	不同角度增加响应的多样性	文献[37-40]	
	充分考虑上下文信息,利用其建模	文献[41-43]	
选择生成混合	基于强化学习和对抗学习	文献[44-46]	Ubuntu-v1、Ubuntu-v2、Ubuntu-v3 和 Advising
	忽略选择和生成的区别,结合不同方法优势	文献[47-49]	
细节问题生成	通过不同建模方式理解上下文信息,包括注意力、对齐指称以及分类器等,目标是使对话更连贯	文献[50-52]	
	以用户问题有用性为指标进行排序	文献[12]、文献[53]	

2.1 意图理解

Qu等^[54]认为开展交互式问答意图理解研究的目的有3点:第一,只有准确识别用户意图,交互式问答系统才能对语句内容进行有效处理,并以此为基础向用户提供答案或调整之前的答案。这在智能客服机器人系统中极其有用,因为需要根据意图将用户问题划分到不同的处理模块;第二,通过学习不同意图的应用场景,知道何时询问用户细节、如何请求更多信息,交互式问答系统可以模仿人类作答行为。第三,相关的意图理解模型可以用来自动标注更多的对话语句,进行数据分析或完成响应选择等任务。

Yu等^[55]认为意图理解的好处有:一个好的意图理解模型可以减少训练整个交互式问答系统所需的标注样本;对于响应选择式交互式问答系统,意图理解可以缩小搜索空间,加快响应选择。除了交互式问答,其他领域也有意图理解研究,如对话系统^[56-58]、信息检索^[59-60]、社交网络^[61-62]等。这些领域将意图理解看作二分类或多分类问题,而交互式问答因为语句功能多样,其意图理解是一个多标签分类问题,难度也更大。传统的意图理解模型大多基于最大熵^[63]、隐马尔可夫模型^[64]、深度置信网络^[65]、条件随机场^[66]、支持向量机^[67]等技术。

这些模型使用了不同的特征,包括词汇特征、句法特征等。特别地,上下文特征被重点考虑。例如,Venkataraman等^[68]使用的隐马尔可夫模型中隐藏状态是意图标签,它生成可观测的单词序列。观测概率由基于意图特定词的语言模型得到,而基于意图标签的n-gram语言模型给出了各意图标签之间的转移概率。Ji等^[65]和Dielmann等^[69]使用深度置信网络进行序列解码,并检验生成和条件建模方法。条件随机场作为一种强大的序列标注方法,已被广泛用于集成上下文信息的意图分类^[66,70-71]。Ribeiro等^[72]在支持向量机分类器中使用了不同设置来捕捉上下文信息,如n-grams、意图预测值。随着深度学习的兴起,基于卷积神经网络^[73]、循环神经网络^[74]或两者混合^[75]的意图理解模型成为研究的主流。

虽然上下文信息在当前意图理解研究中占据着重要地位,但语句内容之外的辅助信息与意图之间具有很强的关联性,如纳入语句内容之外的发言者类型和语句位置等辅助信息。然而,目前的研究对此涉猎较少,如何有效利用这些信息来提升意图理解能力将是未来研究的方向。

2.2 角色理解

角色理解最初被设计成完形填空式阅读理解(cloze-style reading comprehension)任务,要求机器根据一篇文档回答出问题中被替换掉的单词或实体。角色理解的数据集较多,包括不同领域,具有代表性的有CNN/DailyMail^[76]、Who-did-What^[77](WDW)、Children's Book Test^[78](CBT)、BookTest^[79]和CLOTH^[80]。CNN/DailyMail的文档选自美国有线电视新闻网(CNN)和《每日邮报》新闻,问题生成自新闻末尾附带的要点总结,答案为要点总结中的实体。

WDW的文档也来源于新闻语料,但文档和问题是来自

同一事件的两篇不同的报道,以此增加机器理解的难度,扩大其与人类水平之间的差距。另一方面,WDW中的人名实体未做匿名处理,因而可以利用外部资源来提升模型性能。

CBT选取儿童读本中任意连续的21句话,将前20句话作为文档,将去除某一单词的第21句话作为问题。BookTest与CBT类似,但规模比CBT大60倍。因为训练样本更多,在不改变网络结构的前提下,BookTest模型也能获得较大的性能提升。相比前面几个自动生成的数据集,CLOTH则构造自老师精心设计的初、高中英语完形填空试题,选项之间的混淆性更大,对语言理解的要求也更高。

Multiparty-Dialog-RC^[81]数据集与上述几个数据集最大的区别在于将口语对话作为文档。相比书面文本,口语对话有其独特的结构和语言现象。通常,一个口语对话有两个或多个发言者参与其中,而每个发言者有自己的表达习惯。对话过程中,这些发言者围绕同一个目的或主题陈述自己的想法,每次发言可能是一句话,也可能是多句。因此,口语对话中的信息往往更加分散,有时多句话或多次发言才能完整地表达某一信息,对应到书面文本,可能只是一句话。此外,处理书面文本时,只需要关注内容本身,而口语对话除了对内容的理解之外,还需要明确发言者和内容之间的关系。如果不能做到这一点,对话中的指代消解就无从谈起,对细节的理解也会发生错误,亦对整个对话主旨的掌握产生消极影响。

综上,口语对话理解更具挑战性。以口语对话为主体,研究者还提出了面向抽取式阅读理解(extractive reading comprehension)任务的FriendsQA数据集^[82],以及面向有背景交互式问答的Molweni数据集^[83]。

角色理解的模型方面,Chen等^[84]提出了基于人工提取特征及传统机器学习的EntityCentric模型,以及基于自注意力的双向LSTM^[85]模型。Trischler等^[86]提出了结合推理器和抽取器的EpiReader模型,其能够同时使用CNN和RNN编码文档和问题。Dhingra等^[80]提出了关注文档和问题乘法交互的Gated-Attention模型。Cui等^[87]引入Attention-over-Attention模型,分别计算文档到问题的注意力和问题到文档的注意力。

上述模型有一个共同特点,即直接匹配问题和对话。但将这些模型直接应用在角色理解任务上不太合适,因为它们忽视了角色在匹配中的中心地位,并且无法捕捉角色之间的差异性。

3 有背景交互式问答

背景信息包含了领域知识,是理解对话的重要信息。经过整理后,我们总结了目前研究中较为常见的公开数据集,具体如表3所列。其中,ShARC^[88]数据集的背景为政府网站发布的规则文本,任务形式类似无背景交互式问答(回复有两种,一种是询问细节的问题,另一种是答案Yes/No);QBLink^[89]数据集的形式类似智力竞赛,竞赛的答案则为维基百科词条;DREAM^[90]数据集的背景为日常对话,其任务形

式为多项选择的;RC2^[91]数据集的背景为商品评论,且答案是开放的;VisualDialog^[92]数据集的背景为图片;SQA^[93]数据集的背景为表格;背景为知识图谱的数据集有CSQA^[94],Dialog-to-Action^[95],ConvQuestions^[96]等。

表3 包含背景问答数据集

Table 3 Question and answer data sets with background information

数据集	背景信息
CoQA	多领域背景
QuAC	维基百科
ShARC	政府发布文本
QBLink	智力竞赛
DREAM	日常对话
RC2	电商评论
VisualDialog	图片
SQA	表格
CSQA	知识图谱
Dialog-to-Action	知识图谱
ConvQuestions	知识图谱

有背景交互式问答可以看作机器阅读理解^[97](machine reading comprehension)任务的延续,将一问一答扩展为多轮问答。典型的机器阅读理解数据集有SQuAD^[98-99],MS MARCO^[100]等。借鉴SQuAD的构建和评价方法,研究者提出了两个有背景交互式问答数据集:CoQA^[2]和QuAC^[101]。它们的任务形式基本相同,只在某些方面有所区别:1)CoQA背景对问答双方都可见,而QuAC背景仅对回答者可见,因此QuAC对话主题切换得更加频繁,也包括更多的细节问题;2)CoQA允许答案是自由文本(freetext),而QuAC要求答案必须是背景文本的一个片段(span),这导致QuAC答案的平均长度是CoQA的5倍;3)CoQA背景体裁多样,包含维基百科、标准化考试阅读理解、儿童故事书、新闻报道等多个领域,而QuAC背景只来源于维基百科。根据我们对不同数据集研究的基数统计,本文主要介绍基于CoQA和QuAC这两个数据集的研究成果。

下面对两类背景理解模型进行概述。一类是传统模型,它们大多使用了RNN和注意力机制;另一类是基于BERT^[102]等的预训练模型^[103]。Ma等^[104]从对话上下文建模方法、推理策略、响应生成、评价指标等角度对这些模型进行了系统研究。

3.1 传统背景理解模型

在CoQA数据集上,Reddy等^[2]提出了一个结合机器阅读理解和seq2seq的混合模型:DrQA^[105]+PGNet^[106-107]。他们首先将历史问答对和当前问题按语句顺序拼接在一起,将其看作机器阅读理解中的问题序列,然后由DrQA从背景文本中抽取片段形式的答案依据(rationale),以此为PGNet输入,生成自由文本形式的答案。Zhu等^[108]提出了SDNet模型,利用互注意力和自注意力来更有效地理解背景文本和对话上下文。在处理历史问答对和当前问题之间的关系上,Zhu等发现拼接2轮历史问答对效果最好。Su等^[109]针对一般疑问句和特殊疑问句,提出了一个自适应框架,以不同的方式来处理这两种类型的问题。他们首先抽取答案依据,然后

利用一个门控(gated)组件来判断问题类型,最后由MatchNet^[110]和DistillNet^[107]分别回答一般疑问句和特殊疑问句。Gu等^[111]提出了TT-Net模型,利用时间卷积网络^[112]捕捉问答过程中的话题迁移特征。Mandya等^[113]提出了QANet模型,检验共指消解对不同领域文本和对话长度的影响。

在QuAC数据集上,Choi等^[101]提出了BiDAF+w/n-ctx模型,在BiDAF^[114]的基础上增加了自注意力^[115]和ELMo^[116]上下文嵌入,并且每个问题之前拼接了 n 轮历史问答对($n=2$ 时效果最好)。Yatskar^[117]还将BiDAF+w/n-ctx应用在CoQA数据集上,取得了非常好的效果($n=3$)。由于CoQA中97.8%的答案内容与背景文本重叠,有一些工作将CoQA和QuAC放在一起进行研究。如Huang等^[118]提出的FlowQA模型,引入流机制,通过交替并行处理结构,将回答之前问题得到的中间表示集成到当前问题的处理上来。他们认为之前模型采用的拼接方法忽略了历史问答对的推理过程,而流机制通过中间表示捕获短语、事实等相关信息,可以弥补这一缺陷,为回答当前问题提供了额外线索。又如Chen等^[119]提出的GraphFlow模型通过动态构建稀疏图来捕获对话流,并以此为基础进行显式的答案推理。

3.2 基于预训练的背景理解

在CoQA和QuAC数据集上,Ohsugi等^[120]为突破BERT对输入序列长度的限制,提出了BERTw/n-ctx模型,分别编码背景、问题和答案,获得基于BERT的上下文表示,然后拼接在一起进行预测。Yeh等^[121]扩展FlowQA并提出了BERT-FlowDelta模型,通过对问答推理中的信息增益进行建模,让模型获得更多推理线索。

Qu等^[122-123]先后提出了BERT+HAE和HAM模型。BERT+HAE引入了一个语句选择模块来获取对话上下文中与当前问题最相关的一些语句,并且在BERT输入表征中增加了一个历史答案嵌入(History Answer Embedding, HAE),用于标记背景文本中作为之前问题答案的词汇。HAM从3个方面完善了BERT+HAE:改进语句选择为基于注意力的方法;在历史答案嵌入中加入语句位置信息;增加两个对话行为预测辅助任务。区别于前两个工作,他们只在QuAC上进行了模型验证。

Ju等^[124]面向CoQA提出了RoBERTa+AT+KD模型,以更先进的RoBERTa^[125]作为预训练模型,并结合对抗训练^[126](Adversarial Training, AT)和知识蒸馏^[127](Knowledge Distillation, KD),将性能良好的模型作为额外的微调信号。此外,他们还设计了一个答案依据标注(rationale tagging)任务,用于预测每个词是否为答案依据。

综上,早期有背景交互式问答研究主要使用传统模型。随着预训练模型的兴起,研究者纷纷改用微调预训练模型的方法进行训练,并且通过调整输出层结构、增加词嵌入信息、改进训练目标等方式提升模型性能。表4列出了有背景交互式问答方法。

表4 有背景交互式问答方法

Table 4 Interactive question and answer methods with background

任务	方法核心	文献
传统背景理解	基于 seq2seq 的架构,设计不同的编码器和解码器模型,通过不同方式的文本拼接进一步表示上下文的关系	文献[2]、文献[108]、文献[111]、文献[113]
	对一般疑问句和特殊疑问句,提出了一个自适应框架,以不同的方式来处理这两种类型的问题	文献[109]
	对 BiDAF+w/n-ctx 进行优化	文献[101]、文献[117]
	引入流机制,通过不同的方式表示对话流,如交替并行处理结构、动态构建稀疏图等	文献[118-119]
基于预训练	基于 BERT 的预训练模型,根据对话的特性对模型进行改进,如获取与问题相关的语句以及背景信息等	文献[120-123]
	结合对抗学习以及知识蒸馏技术进行微调	文献[124]

4 背景迁移的交互式问答

最近,不少研究开始将迁移学习^[128]应用于交互式问答。根据目前研究的任务性质、领域和学习顺序的不同,大致可以将研究分为领域适应(domain adaptation)、跨语言学习(cross-lingual learning)、多任务学习(multi-task learning)和序列迁移学习(sequential transfer learning)4种。

除跨语言学习之外,其他3种在交互式问答中都有应用。例如,Qiu等^[129]使用领域适应将资源丰富领域中的交互式问答知识迁移到电子商务领域,解决标注数据不足的问题。Whang等^[130]结合领域适应和顺序迁移学习,提出了一个适用于交互式问答响应选择的后训练(post-training)方法。Ubuntu-v1和Advising两个数据集上的验证实验表明,后训练方法能有效提升特定领域模型的性能。Yan等^[131]提出联合学习响应选择和用户输入建议两个任务,使对话更主动并保持用户参与的积极性。

Tao等^[132]为解决直接集成ELMo到响应选择模型效果不佳的问题,提出利用大规模响应生成语料训练一个层次Encoder-Decoder^[133]网络,让其学习语句级别和对话级别的上下文词表示,然后将这两种表示分别集成到响应选择模型的输入层和输出层,实现响应生成知识到响应选择知识的迁移。Wolf等^[134]提出了一个微调GPT^[115]的响应生成模型,微调过程中引入了语句分类和语言模型学习两个辅助任务。

类似地,Qu等^[123]在微调BERT时引入了两个对话行为预测辅助任务;Ju等^[124]在微调RoBERTa时引入了一个答案依据标注辅助任务。Yatskar^[117]应用顺序迁移学习先在SQuAD语料上训练,再在CoQA和QuAC上微调,以此提升有背景交互式问答模型的性能。

5 总结与展望

5.1 总结

总体上,随着深度学习技术的发展以及标记数据集的不断丰富,交互式问答的研究得到了突飞猛进的发展。

在无背景交互式问答研究中,由于对话可能回复答辩或询问细节、澄清表述,所以研究主要分为响应选择和响应生成两个方向。对于响应选择,主要是通过不同方式对上下文的语句、语序以及候选词进行不同表示后,聚合匹配得分;对于响应生成,上下文信息是多数研究的重要关注点,就效果而言,强化学习和对抗学习的加入可以明显提升对话系统的性能。

在有背景交互式问答研究中,技术上的进展可以分为两个阶段:第一阶段主要基于seq2seq架构,设计不同的编码器和解码器模型,通过不同方式的文本拼接进一步表示上下文的关系;第二阶段则主要基于BERT等预训练模型,根据对话的特性对模型进行改进,如获取与问题相关的语句以及背景信息等。总体上,基于预训练模型的方法提高了对话系统的性能。

背景迁移交互式问答旨在融合不同领域的数据集特征,以达到跨领域、跨语言对话的目的。其方法包括预训练模型、知识蒸馏以及多任务学习等技术的综合应用。背景迁移交互式问答属于新兴的研究方向,就目前的实验结果而言,依旧有较大的进步空间。

5.2 展望

虽然交互式问答研究方兴未艾,亦取得了阶段性的成果,但仍有以下问题亟待解决:1)意图识别不准。现有意图理解模型强调语句上下文的重要性,忽视了语句内容之外的发言者类型和语句位置等辅助信息。经分析,这些辅助信息与意图有着很强的关联性,影响着意图识别的准确率。因此有必要研究如何利用这些辅助信息来提升意图理解性能,让交互式问答系统更加智能。2)角色辨析不清。现有角色理解模型直接匹配问题和对话的做法,忽视了角色在匹配中的中心地位,也没有考虑角色之间的差异性,导致容易预测出与对话内容背离或使问题前后矛盾的角色。因此有必要改进匹配思路,以更好地辨识对话中出现的不同角色,让交互式问答系统更具个性化。3)模型过大难以部署。预训练模型凭借优异的自然语言理解能力,已成为当前交互式问答研究的主流。但因其参数量巨大和计算复杂度高,使得模型训练和预测时间成倍增长,资源消耗也越来越大。因此有必要提出更轻量的、适用于交互式问答的预训练模型,使其在低资源消耗的情况下也能有优异的性能表现,让交互式问答系统的运行更加高效。

参考文献

- [1] SARIKAYA R. The technology behind personal digital assistants: An overview of the system architecture and key components [J]. IEEE Signal Processing Magazine, 2017, 34(1): 67-81.
- [2] REDDY S, CHEN D, MANNING C D. CoQA: A Conversational Question Answering Challenge [J]. Transactions of the Associa-

- tion for Computational Linguistics, 2019, 7:249-266.
- [3] LI X, CHEN Y N, LI L, et al. End-to-End Task-Completion Neural Dialogue Systems[C]//Proceedings of the Eighth International Joint Conference on Natural Language Processing. 2017:733-743.
- [4] WEI Z, LIU Q, PENG B, et al. Task-oriented dialogue system for automatic diagnosis[C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. 2018: 201-207.
- [5] SERBAN I V, SORDONI A, BENGIO Y, et al. Building end-to-end dialogue systems using generative hierarchical neural network models[C]// Thirtieth AAAI Conference on Artificial Intelligence. 2016.
- [6] YAN R, SONG Y, WU H. Learning to respond with deep neural networks for retrieval-based human-computer conversation system[C]// Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2016:55-64.
- [7] WANG B X. Research on semantic mining for online community Q&A pairs [D]. Harbin: Harbin Institute of Technology, 2017.
- [8] ZHOU X Q. Research on interactive question answering technology based on deep learning [D]. Harbin: Harbin Institute of Technology, 2017.
- [9] MITCHELL T M, BETTERIDGE J, CARLSON A, et al. Populating the semantic web by macro-reading internet text[C]// International Semantic Web Conference. 2009:998-1002.
- [10] LOWE R, POW N, SERBAN I V, et al. The Ubuntu Dialogue Corpus: A Large Dataset for Research in Unstructured Multi-Turn Dialogue Systems[C]// Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue. 2015:285-294.
- [11] QU C, YANG L, CROFT W B, et al. Analyzing and characterizing user intent in information seeking conversations[C]// The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. 2018:989-992.
- [12] ALIANNEJADI M, ZAMANI H, CRESTANI F, et al. Asking clarifying questions in open-domain information-seeking conversations[C]// Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019:475-484.
- [13] WANG S, JIANG J. A compare-aggregate model for matching text sequences[C]// 5th International Conference on Learning Representations(ICLR). 2019.
- [14] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks[C]// Advances in Neural Information Processing Systems. 2014:3104-3112.
- [15] WANG S, JIANG J. Learning Natural Language Inference with LSTM[C]// Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics; Human Language Technologies. 2016:1442-1451.
- [16] SCHUSTER M, PALIWAL K K. Bidirectional recurrent neural networks[J]. IEEE Transactions on Signal Processing, 1997, 45(11):2673-2681.
- [17] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1(4):541-551.
- [18] ZHOU X, DONG D, WU H, et al. Multi-view response selection for human-computer conversation[C]// Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016:372-381.
- [19] WU Y, WU W, XING C, et al. Sequential Matching Network: A New Architecture for Multiturn Response Selection in Retrieval-Based Chatbots[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. 2017:496-505.
- [20] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// Advances in Neural Information Processing Systems. 2017:5998-6008.
- [21] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[C]// 3rd International Conference on Learning Representations(ICLR). 2015.
- [22] ZHOU X, LI L, DONG D, et al. Multi-turn response selection for chatbots with deep attention matching network[C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. 2018:1118-1127.
- [23] ZHANG Z, LI J, ZHU P, et al. Modeling Multi-turn Conversation with Deep Utterance Aggregation[C]// Proceedings of the 27th International Conference on Computational Linguistics. 2018:3740-3752.
- [24] ZONG C Q. Statistical natural language processing(Second Edition) [M]. Beijing: Tsinghua University Press, 2013:83-85.
- [25] TAO C, WU W, XU C, et al. Multi-representation fusion network for multi-turn response selection in retrieval-based chatbots[C]// Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining. 2019:267-275.
- [26] MAO G, SU J, YU S, et al. Multi-Turn Response Selection for Chatbots With Hierarchical Aggregation Network of Multi-Representation[J]. IEEE Access, 2019, 7:111736-111745.
- [27] TAO C, WU W, XU C, et al. One time of interaction may not be enough: Go deep with an interaction-over-interaction network for response selection in dialogues[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. 2019:1-11.
- [28] LOWE R T, POW N, SERBAN I V, et al. Training end-to-end dialogue systems with the ubuntu dialogue corpus[J]. Dialogue & Discourse, 2017, 8(1):31-65.
- [29] GUNASEKARA C, KUMMERFELD J K, POLYMENAKOS L, et al. Dstc7 task 1: Noetic end-to-end response selection [C]// Proceedings of the First Workshop on NLP for Conversational AI. 2019:60-67.
- [30] CHEN Q, WANG W. Sequential attention-based network for noetic end-to-end response selection [J]. arXiv:1901.02609, 2019.
- [31] GU J C, LING Z H, RUAN Y P, et al. Building sequential inference models for end-to-end response selection[J]. arXiv:1812.00686, 2018.
- [32] SUN S, TAM Y C, CAO J, et al. End-to-end Gated Self-attentive Memory Network for Dialog Response Selection[C]// Workshop on DSTC7. 2019.

- [33] SUKHBAAATAR S, WESTON J, FERGUS R, et al. End-to-end memory networks[C]// *Advances in Neural Information Processing Systems*. 2015;2440-2448.
- [34] CHIANG T R, HUANG C W, SU S Y, et al. Learning Multi-Level Information for Dialogue Response Selection by Highway Recurrent Transformer[J]. arXiv:1903.08953, 2019.
- [35] LU J, XIE Z, LING G, et al. Spatio-Temporal Matching Network for Multi-Turn Responses Selection in Retrieval-Based Chatbots[C]// *Workshop on DSTC7*. 2019.
- [36] WHANG T, LEE D, LEE C, et al. Enhanced Sequential Representation Augmented with Utterance-level Attention for Response Selection[C]// *Workshop on DSTC7*. 2019.
- [37] SERBAN I V, SORDONI A, LOWE R, et al. A hierarchical latent variable encoder-decoder model for generating dialogues [C]// *Thirty-First AAAI Conference on Artificial Intelligence*. 2017.
- [38] XING C, WU W, WU Y, et al. Topic aware neural response generation[C]// *Thirty-First AAAI Conference on Artificial Intelligence*. 2017.
- [39] WU Y, WU W, YANG D, et al. Neural response generation with dynamic vocabularies[C]// *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.
- [40] LI J, GALLEY M, BROCKETT C, et al. A Diversity-Promoting Objective Function for Neural Conversation Models[C]// *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2016;110-119.
- [41] TIAN Z, YAN R, MOU L, et al. How to make context more useful? an empirical study on context-aware neural conversational models[C]// *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*. 2017;231-236.
- [42] ZHANG W, CUI Y, WANG Y, et al. Context-sensitive generation of open-domain conversational responses[C]// *Proceedings of the 27th International Conference on Computational Linguistics*. 2018;2437-2447.
- [43] SANKAR C, SUBRAMANIAN S, PAL C, et al. Do Neural Dialog Systems Use the Conversation History Effectively? An Empirical Study[C]// *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 2019;32-37.
- [44] LI J, MONROE W, RITTER A, et al. Deep Reinforcement Learning for Dialogue Generation[C]// *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. 2016;1192-1202.
- [45] LI J, MONROE W, SHI T, et al. Adversarial Learning for Neural Dialogue Generation[C]// *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 2017;2157-2169.
- [46] XU Z, LIU B, WANG B, et al. Neural response generation via gan with an approximate embedding layer[C]// *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 2017;617-626.
- [47] SONG Y, LI C T, NIE J Y, et al. An ensemble of retrieval-based and generation-based human computer conversation systems [C]// *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. 2018;4382-4388.
- [48] YAN R, ZHAO D. Coupled context modeling for deep chat-chat: towards conversations between human and computer[C]// *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2018;2574-2583.
- [49] WU Y, WEI F, HUANG S, et al. Response generation by context-aware prototype editing [C] // *Proceedings of the AAAI Conference on Artificial Intelligence*. 2019;33;7281-7288.
- [50] LI H, MIN M R, GE Y, et al. A context-aware attention network for interactive question answering[C]// *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2017;927-935.
- [51] GAO Y, LI P, KING I, et al. Interconnected Question Generation with Coreference Alignment and Conversation Flow Modeling[C]// *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 2019;4853-4862.
- [52] CAO Y T, RAO S, DAUMÉ III H. Controlling the Specificity of Clarification Question Generation[C]// *Proceedings of the 2019 Workshop on Widening NLP*. 2019;53-56.
- [53] RAO S, DAUMÉ III H. Learning to Ask Good Questions: Ranking Clarification Questions using Neural Expected Value of Perfect Information[C]// *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*. 2018;2737-2746.
- [54] QU C, YANG L, CROFT W B, et al. User intent prediction in information-seeking conversations[C]// *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*. 2019;25-33.
- [55] YU Y, PENG S, YANG G H. Modeling Long-Range Context for Concurrent Dialogue Acts Recognition[C]// *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. 2019;2277-2280.
- [56] JUDY D, RON Z. Pragmatic determinants of intonation contours for dialogue systems[J]. *International Journal of Speech Technology*, 1997, 1(2): 109-120.
- [57] STOLCKE A, RIES K, COCCARO N, et al. Dialogue act modeling for automatic tagging and recognition of conversational speech[J]. *Computational linguistics*, 2000, 26(3): 339-373.
- [58] PARK S, KIM D, OH A. Conversation Model Fine-Tuning for Classifying Client Utterances in Counseling Dialogues[C]// *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2019;1448-1459.
- [59] SONG W. Research on topic based query intention recognition [D]. Harbin: Harbin Institute of Technology, 2013.
- [60] LU W, ZHOU H X, ZHANG X J. Review of Research on Query Intent[J]. *Journal of Library Science in China*, 2013, 39(1): 100-111.
- [61] CHEN H C. Mining consumption intention based on microblog [D]. Harbin: Harbin Institute of Technology, 2014.
- [62] JIA Y L, HAN D L, LIN H Y, et al. Consumption Intention Recognition Algorithm for Weibo Users[J]. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2020, 56(1): 68-74.
- [63] ANG J, LIU Y, SHRIBERG E. Automatic dialog act segmenta-

- tion and classification in multiparty meetings[C]//Proceedings. (ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing. 2005.
- [64] SURENDRAN D, LEVOW G A. Dialog act tagging with support vector machines and hidden Markov models[C]//Ninth International Conference on Spoken Language Processing. 2006.
- [65] JI G, BILMES J. Dialog act tagging using graphical models[C]//Proceedings. (ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing. 2005.
- [66] KIM S N, CAVEDON L, BALDWIN T. Classifying dialogue acts in one-on-one live chats[C]//Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing. 2010;862-871.
- [67] FERNANDEZ R, PICARD R W. Dialog Act Classification from Prosodic Features Using Support Vector Machines[C]//Speech Prosody 2002, International Conference. 2002.
- [68] VENKATARAMAN A, FERRER L, STOLCKE A, et al. Training a prosody-based dialog act tagger from unlabeled data [C]//2003 IEEE International Conference on Acoustics, Speech, and Signal Processing. 2003.
- [69] DIELMANN A, RENALS S. Recognition of dialogue acts in multiparty meetings using a switching DBN[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2008, 16(7): 1303-1314.
- [70] QUARTERONI S, IVANOV A V, RICCARDI G. Simultaneous dialog act segmentation and classification from human-human spoken conversations[C]//2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2011: 5596-5599.
- [71] CHEN L, DI EUGENIO B. Multimodality and dialogue act classification in the RoboHelper project[C]//Proceedings of the SIGDIAL 2013 Conference. 2013;183-192.
- [72] RIBEIRO E, RIBEIRO R, DE MATOS D M. The influence of context on dialogue act recognition[J]. arXiv:1506.00839, 2015.
- [73] BARAHONA L M R, GASIC M, MRKI N, et al. Exploiting Sentence and Context Representations in Deep Neural Models for Spoken Language Understanding[C]//Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. 2016;258-267.
- [74] KHANPOUR H, GUNTAKANDLA N, NIELSEN R. Dialogue act classification in domain-independent conversations using a deep recurrent neural network[C]//Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. 2016;2012-2021.
- [75] LIU Y, HAN K, TAN Z, et al. Using context information for dialog act classification in dnn framework[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017;2170-2178.
- [76] HERMANN K M, KOCISKY T, GREFFENSTETTE E, et al. Teaching machines to read and comprehend[C]//Advances in Neural Information Processing Systems. 2015;1693-1701.
- [77] ONISHI T, WANG H, BANSAL M, et al. Who did What: A Large-Scale Person-Centered Cloze Dataset[C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016;2230-2235.
- [78] HILL F, BORDES A, CHOPRA S, et al. The goldilocks principle: Reading children's books with explicit memory representations[C]//Proceedings of the Forth International Conference on Learning Representations. 2016.
- [79] BAJGAR O, KADLEC R, KLEINDIENST J. Embracing data abundance: Booktest dataset for reading comprehension[J]. arXiv:1610.00956, 2016.
- [80] DHINGRA B, LIU H, YANG Z, et al. Gated-Attention Readers for Text Comprehension[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. 2017: 1832-1846.
- [81] MA K, JURCZYK T, CHOI J D. Challenging Reading Comprehension on Daily Conversation: Passage Completion on Multiparty Dialog[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2018;2039-2048.
- [82] YANG Z, CHOI J D. FriendsQA: Open-domain question answering on TV show transcripts[C]//Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue. 2019;188-197.
- [83] LI J, LIU M, KAN M Y, et al. Molweni: A Challenge Multiparty Dialogues-based Machine Reading Comprehension Dataset with Discourse Structure[J]. arXiv:2004.05080, 2020.
- [84] CHEN D, BOLTON J, MANNING C D. A Thorough Examination of the CNN/Daily Mail Reading Comprehension Task[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. 2016;2358-2367.
- [85] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [86] TRISCHLER A, YE Z, YUAN X, et al. Natural Language Comprehension with the EpiReader[C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016;128-137.
- [87] CUI Y, CHEN Z, WEI S, et al. Attention-over-Attention Neural Networks for Reading Comprehension[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. 2017;593-602.
- [88] SAEIDI M, BARTOLO M, LEWIS P, et al. Interpretation of Natural Language Rules in Conversational Machine Reading [C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018;2087-2097.
- [89] ELGOHARY A, ZHAO C, BOYD-GRABER J. A dataset and baselines for sequential open-domain question answering[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018;1077-1083.
- [90] SUN K, YU D, CHEN J, et al. DREAM: A Challenge Data Set and Models for Dialogue-Based Reading Comprehension[J]. Transactions of the Association for Computational Linguistics, 2019, 7;217-231.
- [91] XU H, LIU B, SHU L, et al. Review conversational reading comprehension[J]. arXiv:1902.00821, 2019.
- [92] DAS A, KOTTUR S, GUPTA K, et al. Visual dialog[C]//Pro-

- ceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017;326-335.
- [93] IYYER M, YIH W T, CHANG M W. Search-based neural structured learning for sequential question answering[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. 2017;1821-1831.
- [94] SAHA A, PAHUJA V, KHAPRA M M, et al. Complex sequential question answering; Towards learning to converse over linked question answer pairs with a knowledge graph[C]// Thirty-Second AAAI Conference on Artificial Intelligence. 2018.
- [95] GUO D, TANG D, DUAN N, et al. Dialog-to-action: Conversational question answering over a large-scale knowledge base [C]// Advances in Neural Information Processing Systems. 2018;2942-2951.
- [96] CHRISTMANN P, SAHA ROY R, ABUJABAL A, et al. Look before you hop: Conversational question answering over knowledge graphs using judicious context expansion[C]// Proceedings of the 28th ACM International Conference on Information and Knowledge Management. 2019;729-738.
- [97] LI Z J, WANG C B. Survey on Deep-learning-based Machine Reading Comprehension[J]. Computer Science, 2019, 46(7): 7-12.
- [98] RAJPURKAR P, ZHANG J, LOPYREV K, et al. SQuAD: 100000+ Questions for Machine Comprehension of Text[C]// Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016;2383-2392.
- [99] RAJPURKAR P, JIA R, LIANG P. Know What You Don't Know: Unanswerable Questions for SQuAD[C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. 2018;784-789.
- [100] NGUYEN T, ROSENBERG M, SONG X, et al. MS MARCO: A human generated machine reading comprehension dataset[C]// CoCo 2016-Proceedings of the Workshop on Cognitive Computation: Integrating Neural and Symbolic Approaches 2016, co-located with the 30th Annual Conference on Neural Information Processing Systems, NIPS 2016. 2016.
- [101] CHOI E, HE H, IYYER M, et al. QuAC: Question Answering in Context[C]// Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018;2174-2184.
- [102] DEVLIN J, CHANG M W, LEE K, 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. 2019.
- [103] QIU X, SUN T, XU Y, et al. Pre-trained Models for Natural Language Processing: A Survey[J]. arXiv;2003. 08271, 2020.
- [104] MA L, ZHANG W N, LI M, et al. A Survey of Document Grounded Dialogue Systems (DGDS) [J]. arXiv; 2004. 13818, 2020.
- [105] CHEN D, FISCH A, WESTON J, et al. Reading Wikipedia to Answer Open-Domain Questions[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. 2017;1870-1879.
- [106] GU J, LU Z, LI H, et al. Incorporating Copying Mechanism in Sequence-to-Sequence Learning[C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. 2016;1631-1640.
- [107] SEE A, LIU P J, MANNING C D. Get To The Point: Summarization with Pointer-Generator Networks[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. 2017;1073-1083.
- [108] ZHU C, ZENG M, HUANG X. SDNet: Contextualized Attention-based Deep Network for Conversational Question Answering[J]. arXiv;1812. 03593, 2018.
- [109] SU L, GUO J, FAN Y, et al. An Adaptive Framework for Conversational Question Answering[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2019;10041-10042.
- [110] MIN S, ZHONG V, SOCHER R, et al. Efficient and Robust Question Answering from Minimal Context over Documents [C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. 2018;1725-1735.
- [111] GU Y, GUI X, LI D. TT-Net: Topic Transfer-Based Neural Network for Conversational Reading Comprehension[J]. IEEE Access, 2019, 7;116696-116705.
- [112] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[J]. arXiv;1803. 01271, 2018.
- [113] MANDYA A, BOLLEGALA D, COENEN F. Evaluating Co-reference Chains based Conversation History in Conversational Question Answering[C]// International Conference of the Pacific Association for Computational Linguistics. Springer, Singapore. 2019;280-292.
- [114] SEO M, KEMBHAVI A, FARHADI A, et al. Bidirectional attention flow for machine comprehension[C]// 5th International Conference on Learning Representations (ICLR). 2017.
- [115] CLARK C, GARDNER M. Simple and Effective Multi-Paragraph Reading Comprehension [C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. 2018;845-855.
- [116] PETERS M, NEUMANN M, IYYER M, 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. 2018;2227-2237.
- [117] YATSKAR M. A Qualitative Comparison of CoQA, SQuAD2.0 and QuAC [C]// Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2019;2318-2323.
- [118] HUANG H Y, CHOI E, YIH W. FlowQA: Grasping Flow in History for Conversational Machine Comprehension [C]// 7th International Conference on Learning Representations (ICLR). 2019.
- [119] CHEN Y, WU L, ZAKI M J. Graphflow: Exploiting conversation flow with graph neural networks for conversational machine comprehension[J]. arXiv;1908. 00059, 2019.
- [120] OHSUGI Y, SAITO I, NISHIDA K, et al. A Simple but Effective Method to Incorporate Multi-turn Context with BERT for Conversational Machine Comprehension[C]// Proceedings of the

- First Workshop on NLP for Conversational AI. 2019;11-17.
- [121] YE H Y T, CHEN Y N. FlowDelta: Modeling Flow Information Gain in Reasoning for Conversational Machine Comprehension [C] // Proceedings of the 2nd Workshop on Machine Reading for Question Answering. 2019;86-90.
- [122] QU C, YANG L, QIU M, et al. BERT with History Answer Embedding for Conversational Question Answering [C] // Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019;1133-1136.
- [123] QU C, YANG L, QIU M, et al. Attentive History Selection for Conversational Question Answering [C] // Proceedings of the 28th ACM International Conference on Information and Knowledge Management. 2019;1391-1400.
- [124] JU Y, ZHAO F, CHEN S, et al. Technical report on Conversational Question Answering [J]. arXiv:1909.10772, 2019.
- [125] LIU Y, OTT M, GOYAL N, et al. RoBERTa: A Robustly Optimized BERT Pretraining Approach [J]. arXiv:1907.11692, 2019.
- [126] GOODFELLOW I J, SHLENS J, SZEGEDY C. Explaining and harnessing adversarial examples [J]. arXiv:1412.6572, 2014.
- [127] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network [J]. arXiv:1503.02531, 2015.
- [128] RUDER S. Neural Transfer Learning for Natural Language Processing [D]. Galway, Ireland: National University of Ireland, 2019.
- [129] QIU M, YANG L, JI F, et al. Transfer Learning for Context-Aware Question Matching in Information-seeking Conversations in E-commerce [C] // Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. 2018;208-213.
- [130] WHANG T, LEE D, LEE C, et al. Domain adaptive training bert for response selection [J]. arXiv:1908.04812, 2019.
- [131] YAN R, ZHAO D, E W. Joint learning of response ranking and next utterance suggestion in human-computer conversation system [C] // Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2017;685-694.
- [132] TAO C, WU W, XU C, et al. Improving matching models with contextualized word representations for multiturn response selection in retrieval-based chatbots [J]. arXiv:1808.07244, 2018.
- [133] CHO K, VAN MERRIËNBOER B, GULCEHRE C, et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation [C] // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2014;1724-1734.
- [134] WOLF T, SANH V, CHAUMOND J, et al. TransferTransfo: A transfer learning approach for neural network based conversational agents [J]. arXiv:1901.08149, 2019.



HUANG Xin, born in 1984, lecturer, is a member of China Computer Federation. His main research interests include machine learning, natural language processing and biological information.



LEI Gang, born in 1974, associate professor, master tutor. His main research interests include natural language processing, knowledge discovery and machine learning.