

聊天机器人系统研究进展

王浩畅 李 斌

(东北石油大学计算机与信息技术学院 黑龙江 大庆 163318)

摘 要 聊天机器人是自然语言处理和人工智能领域中的一个重要研究方向。近年来,聊天机器人系统的研究受到了广泛关注。综合运用各种人工智能技术,对推动人机对话的发展有重要意义。梳理近些年的研究,从研究背景及其意义,延伸到当前国内外对相关技术的研究现状。同时在聊天机器人的分类、系统框架、构建方法、技术评测等几个方面做详细的介绍。总结现有研究中存在的不足并对未来的发展趋势进行展望。

关键词 聊天机器人 自然语言处理 模板匹配 检索匹配 深度学习

中图分类号 TP391

文献标识码 A

DOI:10.3969/j.issn.1000-386x.2018.12.001

RESEARCH PROGRESS OF CHATBOT SYSTEM

Wang Haochang Li Bin

(School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, Heilongjiang, China)

Abstract Chatbot is an important research direction in the field of natural language processing and artificial intelligence. In recent years, the research of chatbot system has attracted wide attention. The application of AI technology is of great significance to the development of man-machine conversation. In this paper, the research related to chatbot in recent years was reviewed, such as the research background and significance, and the research status of related technologies at home and abroad. The classification, system framework, construction method and technology evaluation of chatbot were introduced in detail. We summarized the shortcomings of the existing research and looked forward to the future development trend.

Keywords Chatbot Natural language processing Template matching Retrieval match Deep learning

0 引 言

随着社会信息化进程的日益加快,人们希望能够用自然语言与计算机进行交流,聊天机器人的市场因此产生。聊天机器人通过学习自然语言来与人类进行对话,属于问答系统的范畴。它不仅能回答用户所提的问题,还能人性化地与用户交流,同时提醒或者安排用户完成一些任务。

20 世纪 90 年代以前的聊天机器人,与用户的交谈仅仅基于谈话技巧和程序技巧,在它们的对话库中存放着很多句型和模板^[1],但缺乏语义分析能力,且几乎没有知识库,因此根本无法真正理解用户所提出问

题,对话质量差强人意。同时由于应用场景的缺失,聊天机器人在过去的十几年间并未得到太多的关注。

随着技术的不断创新和市场需求的持续增大,人工智能正成为数据、科技进入人类生活的重要桥梁。近年来,包括自然语言处理、深度学习、语音识别和模式识别等在内的人工智能技术稳步前进,推动了聊天机器人系统的高度发展,人机交互的方式也发生了很大的变化。目前,很多商业公司纷纷投入到聊天机器人的研发中,陆续推出了相关产品,比如苹果的 Siri、谷歌的 Google Now、微软的 Cortana 与小冰、百度的“度秘”、亚马逊的智能语音助手 Alexa、Facebook 的语音助手 M、以及韩国风靡全球的 SimSimi 等。研究聊天机器人的相关技术,尤其是对话理解问题,对促进人机交

互方式的发展有重要意义^[2]。

1 国内外研究现状

1.1 国外研究现状

国外在聊天机器人领域的研究起步比较早,也比较成熟,一些知名的大学和科研机构都对智能人机对话系统进行了深入的研究和系统开发。

聊天机器人的研究开始于 1950 年,英国知名数学家图灵在《Mind》上发表了一篇经典论文——《计算机器与智能》^[3-4]。在文中,图灵第一次提出“机器智能”的概念,并提出判断计算机是否具备智能的实验方法——“图灵测试”^[5]。“图灵测试”是问答系统的蓝图,也是人工智能的终极目标。

1966 年,麻省理工学院的计算机科学家 Joseph Weizenbaum 开发出了最早也是最著名的聊天机器人系统——Eliza^[6]。该系统的基本原理是在数据库中查找与用户问句中相应的关键词,然后使关键词与相应的模式进行匹配,最后输出与之匹配的答案。Eliza 主要应用于医疗领域,通过这个系统可以引导精神病人与系统进行交互,从而获取治疗需要的信息,能对精神病人进行辅助治疗^[7]。从此,Eliza 系统开启了智能聊天机器人的时代。

1990 年,美国科学家兼慈善家 Hugh G. Loebner 设立了一年一度的人工智能比赛——Loebner Prize(勒布纳奖),目的在于奖励最擅长模仿人类真实对话场景的机器人;Chatterbox Challenge 是 2001 年开始的一年一度的挑战,旨在测试聊天机器人在对话中模拟真人的能力,根据对一系列的问题和整个对话的反应程度来判断该能力。

在这样一些竞赛的推动下,众多科学家和学者纷纷投身到聊天机器人系统研究中,这在推广图灵测试的同时,也推动了人工智能的发展,是聊天机器人的研究的一个黄金期。1995 年, Richard S. Wallace 博士开发了一个基于经验的人工智能聊天机器人——ALICE。它以人工智能标记语言 AIML(Artificial Intelligence Markup Language)作为知识描述语言,凭借其良好的设计和逼真的效果,在 2000 年、2001 年和 2004 年三度获得著名的勒布纳奖,并且在 2004 年获得了 Chatterbox Challenge 比赛的冠军,曾一度被认为是最聪明的聊天机器人。目前,该系统经过改进后已经能够完全支持中文语言,并且在各种中文聊天系统中取得了较好的成绩。

1998 年, Wendell Cowart 用 Java script 和 PERL 语

言编写的 Talk-Bot 作为一个在线聊天系统被发布,在 2001 年、2002 年和 2006 年三次获得了 Chatterbox Challenge 比赛的冠军^[8]。

由 Bruce Wilcox 开发的 ChatScript 在 2010 年、2011 年和 2014 年三次获得了勒布纳奖,促进了聊天机器人领域的快速发展。由于 ChatScript 改进了脚本引擎,使其比 AIML 更具维护性及可读性,并从 AIML 时代基于无状态的对话,进入基于状态、拥有上下文情景及记忆的对话,带来了更高的交互性;同时引入 Ontology,让聊天机器人拥有更丰富的知识^[9]。

新世纪的到来为智能聊天系统创造了更大的发展空间,尤其是深度学习算法的出现,为聊天机器人的研究带来了蓬勃生机^[10]。目前,研究聊天机器人的热度持续上涨,各大厂商也纷纷加入到研制聊天机器人的行列,如苹果语音助手 Siri、韩国的小黄鸡、IBM 沃森系统、谷歌的 Google Now、微软的 Cortana、亚马逊的 Alexa 等。

1.2 国内研究现状

相比之下,国内在聊天机器人领域的研究起步较晚,面临着两大难题:一是中文在信息处理方面的特殊性,往往无法将国外一些成熟技术和研究成果直接利用到中文对话系统中;二是缺乏相应的自然语言处理的相关基础资源,如语料库、知识库以及相应的评价机制等^[1]。

近年来,随着聊天机器人技术的飞速发展,国内的聊天机器人研究事业有了很大的进展。2014 年,微软小冰一时成为热门话题,小冰在与用户的对答中积累了大量的语料基础,同时,微软凭借在大数据、自然语义分析、机器学习和深度神经网络方面的技术积累^[11],大大提升了小冰的对话学习能力。继小冰之后,国内其他厂商紧随其后,各具特色的聊天机器人应运而生,如公子小白、百度度秘、Holoera 等。按照时间的先后顺序将聊天机器人所属的应用场景分类,以及主要功能或核心技术等总结如图 1 所示。

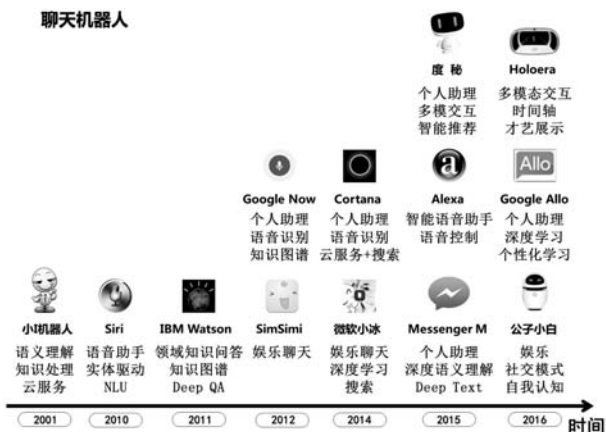


图 1 聊天机器人

2 聊天机器人的分类

聊天机器人种类繁多,且广泛地应用于各个领域。在功能上,可以将聊天机器人划分为以闲聊为主的开放领域和以任务为主的特定领域;而在技术实现上又可以划分为检索式和生成式;本文主要从应用场景上对聊天机器人进行分类,从该方面可以将其划分为娱乐、在线客服、教育、个人助理和智能问答。

娱乐聊天机器人的主要功能是同用户进行开放主题的对话。代表性的系统有微软“小冰”、微信“小微”、“小黄鸡”等。此类应用不仅能够与用户进行闲聊,实现情感抚慰和精神陪伴等;还能满足日常生活需求,如新闻播报和天气预报等。

在线客服聊天机器人的主要功能是针对用户提出的有关产品或服务的相关问题进行回答。代表性的系统有小i机器人、京东的JIMI客服机器人等。一般而言,客服机器人具备一定的拒识能力,即当知道自己不能回答用户的问题时,就会提供某种机制将用户转向人工客服。

教育场景下的聊天机器人主要功能包括帮助用户学习新语种;辅助用户在特定年龄阶段对某种知识的学习;指导用户对某项专业技能的学习等。代表性的系统有科大讯飞公司的开心熊宝和微软亚洲研究院开发的微软小英。

个人助理类聊天机器人的主要功能侧重于实现个人事务,如信息查询及代办其他事件等功能,目的是更便捷地辅助用户的日常事务处理。代表性的系统有苹果Siri、百度度秘、微软Cortana和GoogleNow等。

智能问答类的聊天机器人主要功能是理解并回答用户提出的问题,这些问题比较侧重于事实性问题或者是需要计算和逻辑推理型的问题。代表性的系统有IBM Watson、Wolfram Alpha等。

3 聊天机器人系统框架

图2显示了一个通用的聊天机器人系统框架,其中包含五个主要的功能模块。基本的流程为:用户通过文字形式或语音形式输入之后进行预处理,转化成文本形式进行自然语言理解,然后再通过语义表示和上下文进入对话管理,接着对当前对话模型进行答案提取,最后将生成的回复文本进行合成输出给用户。

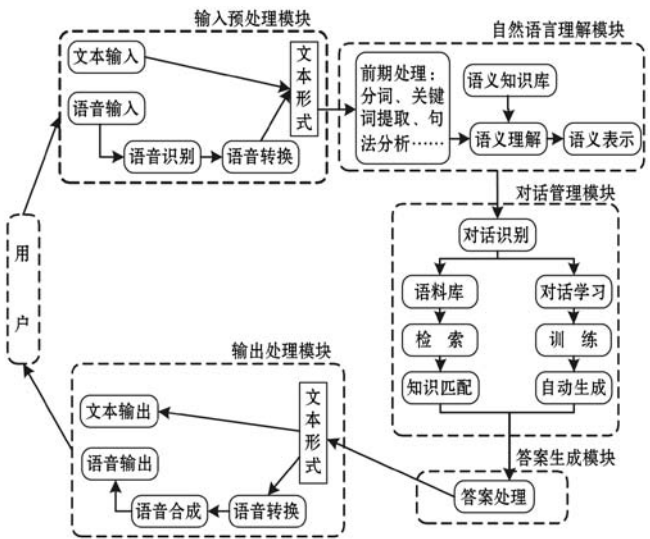


图2 聊天机器人系统框架

3.1 输入预处理

输入预处理模块主要负责对用户输入的问句进行前期的处理,如果是文字输入,就直接交给自然语言理解模块进行处理;如果是语音输入,就需要先将语音转换成文字形式。语音输入涉及到语音识别技术,语音识别技术主要包括特征提取技术、模式匹配准则以及模型训练技术3个方面^[12]。

3.2 自然语言理解

自然语言理解模块主要负责从完整的语句中提取出需要的信息,并且理解用户输入问句的语义,产生一个特定的语义表达式。聊天机器人系统中的自然语言理解功能包括用户意图识别、用户情感识别、指代消解、省略恢复、回复确认及拒识判断等技术^[13]。并且语法分析、语义分析、关键词提取、相似度计算等自然语言处理技术也是必不可少的。

3.3 对话管理

对话管理模块是整个系统的核心,主要功能是协调聊天机器人的各个部分,提供人机交互的对话策略,帮助用户高效自然地完成任务,并维护和更新整个对话的状态^[14]。其中涉及到对话行为识别、对话状态跟踪、动作选择、对话信息记录、策略优化以及对话奖励等关键技术。对话管理模块首先要识别出用户输入问题的形式是属于封闭式还是开放式。封闭式是指有明确的服务目标和服务对象,仅处理特定领域中的问题,给对话的主题进行了限制。开放式则没有一个清晰的目标和对象,对话涉及的主题很广,所需要的知识量也是巨大的。接着进行知识匹配,最后得到可以表示或者指向答案的模型。对话管理模块管理着对话的上下文和后端的知识库,以及对于生成最终的对话答案产生了直接的影响^[15],在整个系统中是非常重要的。

3.4 答案生成

答案生成模块主要负责根据当前对话的模型进行答案的提取。答案生成通常根据对话管理部分产生的文本信息,自动生成面向用户的自然语言反馈^[13],整个答案生成的过程包括了内容选择、文本规划、语句合成、指代表达生成、表层实现等阶段。目前,聊天机器人的对话生成技术主要包括检索式和生成式两种。检索式是指在对话库中以搜索匹配的方式找到适合用户输入语句的最佳答案予以回复。生成式是指采用一定的技术手段(如深度学习技术)自动生成新的回复内容。

3.5 输出处理

输出处理模块主要负责了解用户是以何种形式进行提问。提问时采用的是文字,就直接将文字输出给用户;提问时采用的是语音,就先将语音转化成文字。

4 聊天机器人的构建方法

对于聊天机器人的技术而言,常见的几种主流技术构建方法有:基于人工模板的聊天机器人、基于检索的聊天机器人和基于深度学习的聊天机器人。

4.1 基于人工模板的聊天机器人

该方法是根据用户输入的句子,在模板库中找到匹配的问句模板,然后按照对应的应答模板生成答案,返回给用户,如 ALICE、ChatScript 等。ALICE^[16] 采用 AIML^[17] 语言来描述知识库,AIML 语言采用 XML 语法来存储数据^[18],图 3 所示为 AIML 规则实例。这种方法的优点是精确,缺点是需要大量的人力,并且缺乏灵活性,可扩展性差。



图 3 AIML 规则

4.2 基于检索的聊天机器人

基于检索技术的聊天机器人根据用户输入的句子,在对话库中以搜索匹配的方式进行匹配,从中找到最合适的应答内容,将答案提取出来返回给用户,其工作流程图如图 4 所示。这种方法对对话库的要求很高,并且需要足够大,但是它的优点是回答质量高,表

达比较自然。

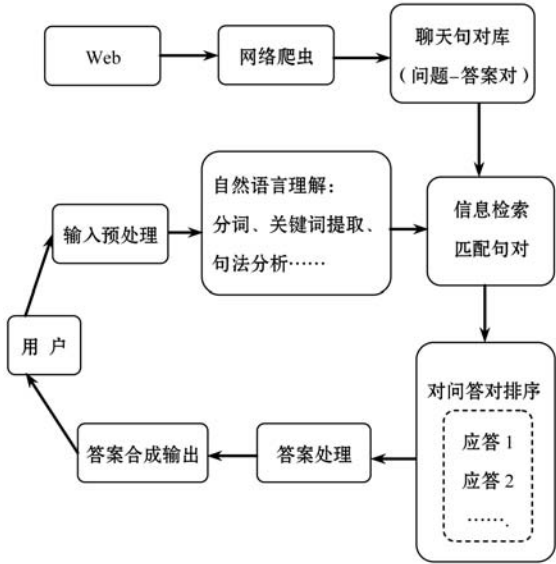


图 4 基于检索的聊天系统工作流程

4.3 基于深度学习的聊天机器人

基于深度学习的聊天机器人根据用户输入的句子,利用模型逐词或逐字生成答案,然后将答案回复给用户。其中多数技术采用了 Encoder-Decoder^[19-21] 模型,即编码-解码模型,其框架技术原理如图 5 所示。Encoder 就是对输入序列进行编码,通过一系列非线性变换转化成带有语义固定长度的向量;Decoder 就是根据之前生成的固定向量再转化成输出序列。实现聊天系统的时候,Encoder 和 Decoder 都不是固定的,可选的模型有 RNN^[22]、LSTM^[23]、GRU^[24] 等,可以自由组合。这种方法思路简单,可扩展,能够更好地理解上下文,但是模型很难训练,回复时经常存在一些语法错误。

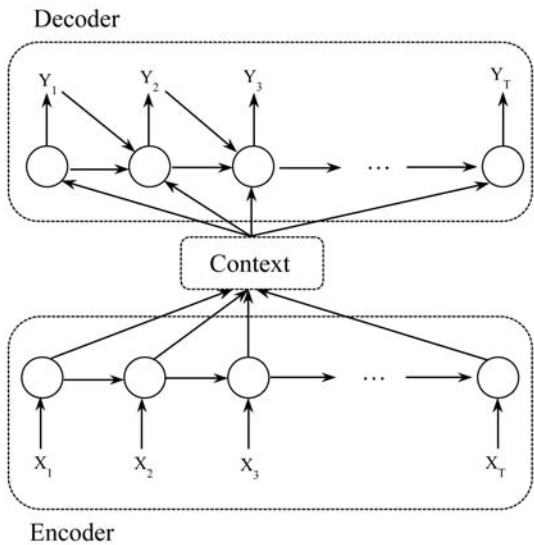


图 5 Encoder-Decoder 框架

目前,在构建系统时都是将多种方法融合。根据三种构建方法的不同,将聊天机器人构建方法进行对

比,如表 1 所示。

表 1 三种构建方法的比较

构建方法	基于人工模板	基于检索	基于深度学习
聊天语料库	问答模板	问答句对	问答句对
人工工作量	大	中等	低
语法准确性	准确	准确	不能保证
应答可控性	高	中等	低
可扩展性	差	中等	好
技术难度	容易	中等	复杂
适用场景	高频聊天请求 特定聊天场景	中等频率聊天请求 垂直领域的聊天	长尾聊天 请求

5 聊天机器人技术评测及评测指标

5.1 聊天机器人技术评测

技术评测在聊天机器人(人机对话)领域非常重要,不仅对人工智能的技术进步起到了巨大的推动作用,还能够促进整个科学研究领域的发展。近年来,出现了一些专门针对人机对话技术的评测,如 NIPS、DSTC、SMP2017-ECDT 等。

神经信息处理系统大会 NIPS(Advances in Neural Information Processing Systems)是由连接学派(connectionist)神经网络的学者于 1987 年在加拿大创办,是一个关于机器学习和神经计算领域的顶级会议,该会议于每年 12 月由 NIPS 基金会主办。2017 年 NIPS 在美国西岸加州南部城市长滩市举办,竞赛主要有 5 项,包括人机问答、学习跑步、会话智能挑战等。

对话状态跟踪挑战 DSTC(Dialog State Tracking Challenge)是从 2013 年开始进行的系列研究挑战任务,每年一次,到 2017 年已经组织了五次。该评测是一种任务驱动多轮对话评测标准,其重点在于改善跟踪对话系统的技术水平。状态跟踪是指在对话过程中准确估计用户的目标,该挑战为对话状态跟踪提供了第一个通用的测试平台和评估标准。

全国社会媒体处理大会 SMP(National Conference of Social Media Processing)从 2012 年开始,每年举行一次,现已成为社会媒体处理的重要学术活动。第六届 SMP 于 2017 年 9 月在北京召开,此次会议首次组织举办了中文人机对话技术评测(The Evaluation of Chinese Human-Computer Dialogue Technology,SMP2017-ECDT),此次评测主要为了促进中文人机对话系统相关研究的发展,为人机对话技术相关的学术研究人员和产业界从业人员提供一个良好的沟通平台。表 2 列出了近 5

年评测任务及评测指标。

5.2 聊天机器人系统评测指标

聊天机器人系统的评测指标^[30-34]采用国际上通用的平均排序倒数 MRR(Mean Reciprocal Rank)、信赖加权评分 CWS(Confidence Weighted Score)、准确率 P(Precision)、召回率 R(Recall)和 F-Measure(F)几个评测指标。

$$MRR = \frac{1}{N} \sum_{i=1}^N \frac{1}{M}$$

(1)

如果标准答案存在于系统给出的排序结果中的多个位置,以排序最高的位置计算;如果标准答案不在系统给出的排序结果中,得 0 分。式(1)中的 M 表示标准答案在系统给出的排序结果中的位置, N 表示测试集中的提问个数。

表 2 人机对话技术评测

名称	时间	评测任务	评测指标
DSTC 1 ^[25]	2013	公共汽车时刻表领域	(1) false-accepts(FA) (2) false-rejects(FR) (3) correct-accepts(CA)
DSTC 2 ^[26]	2014	评估用户通过指定约束来搜索餐厅的对话的状态跟踪	(1) Accuracy (2) L2 norm (3) ROC V2 CA 5
DSTC 3 ^[27]	2014	餐厅或酒店信息领域	(1) Accuracy (2) L2 norm (3) ROC V2 CA 5
DSTC 4 ^[28]	2015	主要任务:评估在旅游场景下人与人对话中的对话状态建模 子任务: (1) 口语理解(SLU) (2) 言语行为预测(SAP) (3) 口语生成(SLG) (4) 端到端系统(EES)	(1) Accuracy (2) 准确率(P)、召回率(R)和 F-score
DSTC 5 ^[29]	2016	主要任务:评估在旅游场景下人与人对话中的对话状态建模,并提出通过机器翻译实现跨语言对话建模 子任务: (1) 口语理解(SLU) (2) 言语行为预测(SAP) (3) 口语生成(SLG) (4) 端到端系统(EES)	(1) Accuracy (2) 准确率(P)、召回率(R)和 F-score
DSTC 6	2017	(1) 端到端的面向目标的对话学习 (2) 端到端对话建模 (3) 对话框故障检测	
NIPS	2017	给定特定的文章(新闻或维基百科文章)的人机对话评测	(1) 整个对话的质量(以广度和参与度评估) (2) 回答内容的自然度

续表 2

名称	时间	评测任务	评测指标
SMP	2017	任务 1: 用户意图领域分类(包括封闭式和开放式) 任务 2: 特定域任务型人机对话在线评测	任务 1: 准确率(P);召回率(R); F-score 任务 2: (1) 任务完成率 (2) 用户满意度 (3) 回复语言的自然度 (4) 对话轮数 (5) 静态的数据库资源未覆盖情况的引导能力

$$CWS = \frac{1}{N} \sum_{i=1}^N \frac{C}{i}$$
 (2)

式中: C 表示前 i 个提问中被正确回答的提问数, N 表示测试集中的提问个数。

$$P = \frac{\text{查询出来的准确答案个数}}{\text{查询出来的所有答案个数}}$$
 (3)

$$R = \frac{\text{查询出来的准确答案个数}}{\text{数据库中的所有准确答案个数}}$$
 (4)

$$F = \frac{2PR}{(P + R)}$$
 (5)

6 结 语

目前, 聊天机器人的研究仍存在许多不足, 这些不足为我们指出了未来的发展趋势。

(1) 理解提问者的问题及语言。目前有些聊天机器人忽略了最基础的语言理解能力。当用户所提问题较短时, 系统不能有效的理解问题^[35]。谷歌和 Facebook 就对这一问题进行了深入的研究, 他们使用深度神经网络来推进聊天机器人的自然语言处理能力, 希望通过分析海量数据来学会理解语言, 并以自然的方式回应人类语言。

(2) 学习能力。学习应当是一个优化的过程, 聊天机器人应该要建立自己的道德标准体系, 识别哪些是应该学习的。除此之外, 还要有主动学习的能力。现在聊天机器人开放式学习的能力依旧是有限的, 仍需要人类教会它们如何获取及处理数据, 这背后需要一个庞大的团队。如果可以通过互联网自行搜索信息, 并进行自动学习和获取知识, 就能减少大量的人力, 这也是走向人工智能的重要一步。

(3) 推理判断能力。聊天机器人如果没有推理判断能力, 就无法分辨出自己的回答是否正确、合理。因此需要建立自己的评价体系, 对所有回答进行判断, 并找出最优答案。目前已有一些研究人员和机构将深度学习和知识图谱技术应用到聊天机器人中, 使得聊天

机器人具有一定的推理能力, 且对话更加流畅自然。

(4) 记忆能力。目前的聊天机器人只是回答一些简单的问题, 几乎没有记忆能力。在一段对话中, 人们经常有所省略, 用一个或几个字来代指上文中所说的某一件事, 对于这种情况, 聊天机器人基本上无法根据上下文来回答。这就需要它们通过深度学习, 按照语义的逻辑关系正确代入指代对象^[36]。

(5) 语音识别技术。即使聊天机器人已拥有先进的语音识别技术, 但口音的差异仍是一大难题。由于语言的差异及各种口音的存在, 聊天机器人必须学习适应语种转换和音调变化。目前, Alexa、Siri、Cortana 等个人助理正试图利用现有数据补充语音功能, 以增强语音识别技术, 并改善对复杂语音模式的理解, 使用数据挖掘和神经网络来分析和模仿人类语言。

一个好的聊天机器人不仅需要用到各种信息处理技术, 还要考虑各种知识的融合, 而现在的人工智能技术也正朝着这个方向发展。目前的聊天机器人虽然离我们理想的目标还有一段距离, 功能不够全面, 产业和技术发展仍存在着很多局限性, 但是随着研究的不断深入, 聊天机器人的智能化水平将不断提高, 聊天机器人也在逐步走向成熟。如今, 聊天机器人技术正处于蓬勃发展的时期, 相信在不久的将来聊天机器人会取得一次次重大的突破。

参 考 文 献

[1] 宁长英. 智能聊天机器人的关键技术研究[D]. 杭州: 杭州电子科技大学, 2011.

[2] 冯升. 聊天机器人问答系统现状与发展[J]. 机器人技术与应用, 2016(4): 34-36.

[3] Turing A M. Computing machinery and intelligence[M]//Computation & intelligence. American Association for Artificial Intelligence, 1950: 433-460.

[4] 冯晓波. 中文问答系统中问题分类和关键词扩展的研究[D]. 北京: 北京邮电大学, 2011.

[5] 刘芳, 于斐. 面向医疗行业的智能问答系统研究与实现[J]. 微电子学与计算机, 2012, 29(11): 95-98.

[6] Weizenbaum J. ELIZA—a computer program for the study of natural language communication between man and machine[J]. Communications of the Acm, 1966, 9(1): 36-45.

[7] 王肖磊. 自动问答系统的研究与应用[D]. 北京: 北京邮电大学, 2015.

[8] 冯德虎. 基于 ALICE 的研究生招生咨询智能聊天机器人研究与实现[D]. 成都: 西南交通大学, 2013.

[9] 杨凯翔. 聊天机器人知识挖掘方法研究[D]. 广州: 华南理工大学, 2015.

- trum, 2012, 34(2):407-427.
- [11] Chien C, Tseng F, Chen C. An evolutionary approach to rehabilitation patient scheduling: a case study[J]. *European Journal of Operational Research*, 2008, 189(3):1234-1253.
- [12] Lebowitz P. Schedule the short procedure first to improve or efficiency[J]. *Aorn Journal*, 2003, 78(4):651-654.
- [13] Lee S, Yih Y. Reducing patient-flow delays in surgical suites through determining start-times of surgical cases[J]. *European Journal of Operational Research*, 2014, 238(2):620-629.
- [14] 郝志刚. 考虑医患满意度和手术成本的手术排程方法[J]. *工业工程*, 2017, 20(4):49-71.
- [15] 刘洪伟, 张天奕, 许丹. 多科室择期手术日排程与医生指派研究[J]. *工业工程与管理*, 2016, 21(6):10-16.
- [16] Hsu V N, Matta R D, Lee C Y. Scheduling patients in an ambulatory surgical center[J]. *Naval Research Logistics*, 2003, 50(3):218-238.
- ~~~~~
- (上接第6页)
- [10] 曹东岩. 基于强化学习的开放领域聊天机器人对话生成算法[D]. 哈尔滨:哈尔滨工业大学, 2017.
- [11] 孙冰. 微软“小冰”的爆红与暴毙[J]. *中国经济周刊*, 2014(22):64-65.
- [12] 李伶俐. 一种基于MFCC和SVM的语音识别方法[J]. *软件导刊*, 2012, 11(3):153-154.
- [13] 元传伟. 基于Bot Framework的高职自主学习智能平台[J]. *濮阳职业技术学院学报*, 2017, 30(3):66-68.
- [14] Jurafsky D, Martin J H. *Speech and language processing* [M]. 2nd ed. Prentice-Hall, Inc. 2009.
- [15] 冯升. 聊天机器人系统的对话理解研究与开发[D]. 北京:北京邮电大学, 2014.
- [16] 夏天, 樊孝忠, 刘林. ALICE 机理分析与应用研究[J]. *计算机应用*, 2003, 23(9):1-5.
- [17] Zhang H, Kishore R, Sharman R, et al. Agile Integration Modeling Language (AIML): A conceptual modeling grammar for agile integrative business information systems[J]. *Decision Support Systems*, 2008, 44(1):266-284.
- [18] 李臻贤. 中文问答系统知识库的自动构建问题研究[D]. 济南:山东财经大学, 2015.
- [19] Sutskever I, Vinyals O, Le Q V. Sequence to sequence learning with neural networks[C]//*Proceedings of Advances in neural information processing systems*. 2014: 3104-3112.
- [20] Cho K, Merriënboer B V, Gulcehre C, et al. Learning phrase representations using RNN Encoder-Decoder for statistical machine translation[EB]. *arXiv preprint arXiv:1406.1078*, 2014.
- [21] Vinyals O, Le Q. A neural conversational model[EB]. *arXiv preprint arXiv:1506.05869*, 2015.
- [22] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors[J]. *Nature*, 1986, 323:533-536.
- [23] Hochreiter S, Schmidhuber J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8):1735-1780.
- [24] Cho K, Merriënboer B V, Bahdanau D, et al. On the properties of neural machine translation: Encoder-Decoder approaches[EB]. *arXiv preprint arXiv:1409.1259*, 2014.
- [25] Williams J, Raux A, Ramachandran D, et al. The dialog state tracking challenge[C]//*Proceedings of the SIGDIAL 2013 Conference*. 2013:404-413.
- [26] Henderson M, Thomson B, Williams J D. The second dialog state tracking challenge[C]//*Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*. 2014:263-272.
- [27] Henderson M, Thomson B, Williams J D. The third dialog state tracking challenge[C]//*Proceedings of IEEE Spoken Language Technology Workshop (SLT)*. IEEE, 2014:324-329.
- [28] Kim S, D'Haro L F, Banchs R E, et al. The fourth dialog state tracking challenge[C]//*Proceedings of the 7th International Workshop on Spoken Dialogue Systems (IWSDS)*. 2016:1-14.
- [29] Kim S, D'Haro L F, Banchs R E, et al. The fifth dialog state tracking challenge[C]//*Proceedings of Spoken Language Technology Workshop*. IEEE, 2016:511-517.
- [30] 吴友政, 赵军, 段湘煜, 等. 构建汉语问答系统评测平台[C]//*第一届全国信息检索与内容安全学术会议*. 2004: 315-323.
- [31] 王东升, 王卫民, 王石, 等. 面向限定领域问答系统的自然语言理解方法综述[J]. *计算机科学*, 2017, 44(8):1-8.
- [32] Voorhees E M, Dang H T. Overview of the TREC 2005 question answering track[C]//*Proceedings of the Fourteenth Text REtrieval Conference (TREC 2005)*. New York, 2005: 1-15.
- [33] 王恒. 中文问答系统的研究与实现[D]. 哈尔滨:哈尔滨工业大学, 2008.
- [34] Noguera E, Llopis F, Ferrandez A, et al. Evaluation of Open-Domain Question Answering systems within a time constraint[C]//*International Conference on Advanced Information Networking and Applications Workshops*. IEEE Computer Society, 2007:260-265.
- [35] 王正华. 自动问答系统的研究与实现[D]. 绵阳:西南科技大学, 2015.
- [36] 夏斌. 聊天机器人会成“键盘终结者”吗[N]. *解放日报*, 2016-08-01(12).