

智能问答系统研究

文/王瑛 何启涛

摘要

问答系统是一个用自然语言与人们进行互动的计算机系统，近些年来机器学习、强化学习，尤其是深度学习等技术的革命性发展给问答系统带来了长足的进步。本文将从问答系统的定义入手展开文章，再介绍了问答系统的发展历程，以及给出问答系统的一般处理框架，再从面向任务型和非任务型两个方面给出问答系统的分类及其相应技术，并给出问答系统的常见的评价标准方法。最后，在文章结尾处给出当前问答系统的一些挑战以及未来的可能的研究方向。

【关键词】问答系统 处理框架 分类 深度学习

1 引言

随着互联网的发展，互联网上的信息越来越繁杂，已呈现出爆炸式增长的状态，面对信息洪灾，人们越来越难以在网上便捷获取有效信息，获取有效信息的成本也越来越高。过去人们通常根据搜索引擎输入关键词，搜索服务提供商运用检索算法根据相关性从高到低给出相关信息。但随着信息越来越繁杂，传统搜索算法愈来愈显得力不从心，其原因在于传统算法未能获取语言更深层次的语义信息，此外，相似信息太多，甚至包含了很多虚假信息。因此也不断催生着新技术的诞生，尤其是计算语言学的发展和人工智能的进步。

而随着深度学习技术的不断进步，问答系统不仅仅可以获取语言表面上的信息，还可以获取更深层次的语义信息，而这种进步恰好也给了问答系统生存的土壤与营养。问答系统所能带来的变革与收益，诸如在取代人工客服，聊天机器人，智能家居等领域的应用，进一步引起了学术界和工业界的广泛关注。当前，问答系统已成为自然语言处理领域研究的一大热点。问答系统的成熟应用会将人类从大量重复性的劳动中释放出来，并且会改变人类生产力

的生产方式，对人类的社会进步是不可估量的。同时，智能的问答系统也是构建下一代智能世界不可或缺的一个重要角色。在本文中，我们将从不同的角度对问答系统进行概述，并在文章结尾处讨论了一些当前问答系统所面临的挑战与未来可能的发展方向。

2 问答系统

2.1 问答系统的定义

问答系统是一种以自然语言或语音和用户进行自由问答交流的计算机程序，它在用户和基于计算机的应用程序之间提供了一个接口，该接口允许以一种相对自然的方式与应用程序进行交互。目前，问答系统正以文本、图形、语音等多模态的形式发展。

2.2 问答系统的组成

问答系统通常由四个部分组成：

(1) 自然语言理解 (NLU)，将自然语言信息转换成语义槽，通俗来说就是将文本语言转换为计算机可以表示并理解的信息；

(2) 问答状态跟踪，即问答管理，这一阶段系统根据历史问答和当前用户的输入，来产生当前的问答状态，即输出当前的状态所采取的动作；

(3) 策略学习，策略学习的目标是生成下一状态，这也是得益于强化学习的引入，使得系统不仅仅是考虑历史的问答记录，还应考虑未来可能的问答状态；

(4) 自然语言生成 (NLG)，将计算机的语言理解表示映射为人类所熟悉的自然语言。当然，有些问答系统的输入输出并非自然语言，也可能是语音，那么在输入时还需要将语音转换为自然语言，在输出时将自然语言转换为语音。

3 问答系统的发展历程

根据技术的变革发展，我们可以将问答系统分为三个大的阶段：

3.1 基于符号规则的第一代

二十世纪 80 年代末是符号规则技术的高潮期，也是最早一代问答系统的诞生期，其主要技术是基于专家设计符号规则，来产生指定的回答系统，它的缺点当然很明显：

(1) 很难扩展；

(2) 过于依赖专家，其成本太高；

(3) 系统只是基于规则来产生回复，并没有涉及学习。

3.2 基于数据统计的第二代

二十一世纪以后由于互联网的快速发展，产生了大量的文本信息，尤其是社区问答信息，这为问答系统进入数据驱动时代铺垫了良好的基础，此时问答系统已进入开放领域。基于统计学的问答系统大放异彩，问答系统已经开始出现浅层学习技术，但是这种基于浅层学习的问答系统虽说有一定的学习能力，但由于这种学习能力不强，也不易于理解，在很多方面难以给出解释，这些因素限制了当时问答系统的进一步发展。

3.3 基于深度学习的第三代

互联网信息的爆炸式产生和深度学习技术的快速发展，大量的数据和革新的技术给新一代问答系统提供了发展的客观条件。此时，深层的神经网络蓬勃发展，用更深层的学习来获取更深层次的语义表征，使得深层神经网络具有强大的学习能力。正由于其强大学习能力这也使得问答系统得到了空前的关注。虽说第三代问答系统的学习能力远超前两代的问答系统，但仍具有一些局限性，例如仍然有一些现象很难解释，很难扩展到其他领域，无法做到通用问答系统，并且市场上并没有完全成功的商用问答系统。

4 问答系统处理框架

不同类型的问答系统在数据处理的方式上有所不同。虽然不同的问答系统面对不同的任务有着各自的架构体系，但根据数据的流动

●项目基金：广东省科技项目：消防应急救援车辆关键技术研发（2015B090923001）。

方式,一般可以分为三个部分,包括问题理解,信息检索和答案生成。

问答系统研究包含三个基本问题:如何将自然语言在计算机中表示出来,让计算机理解其语义,如何选取最优的答句,以及将答案映射为自然语言表示出来。

(1) 提问处理模块:负责处理用户的问题,生成查询关键词,确定问题答案的类型以及问题的语义表示。

(2) 检索模块:根据提问处理模块所转换的语义表示,使用强化学习和检索方式筛选出候选答案。

(3) 答案抽取模块:依据某种打分规则,对检索模块的候选答案进行打分,返回得分最高的那个候选项。

5 问答系统的分类与技术

按照问答系统的任务类型,大致可分为两种类型:

(1) 面向任务型问答系统;

(2) 面向非任务型问答系统。

面向任务型问答系统的目的是完成具体的任务,例如查询酒店,订餐等。面向非任务型问答系统的主要目的是和用户进行自由交流,很典型的就是当前流行的聊天机器人。

面向非任务型问答系统的主要有三种方法:

(1) 基于检索的方法,从事先定义好的数据库中,根据某种规则,选出最合适的答案。优点是通俗易懂,易于解释,但缺点也很明显,非常依赖事先定义的数据库和检索算法,如果数据质量不高,那么很可能效果就很差。

(2) 基于生成的方法,这是当前研究领域的一个主流热点,它是通过理解提问的问题,来生成适当的回复,典型的生成模型有 seq2seq, transformer 等。它的优点是不依赖定义好的数据库,可以灵活回复,缺点是容易生成一些无意义的回复和重复的回复等。

(3) 基于检索和生成混合的方法,结合检索回复较为精确和生成回复较为灵活的优点,可以大幅提升问答系统的回复效率和性能。这也是问答系统实际商业应用的主流方法。

6 问答系统的评价

如何评价一个问答系统也是问答系统的一个重要方面,然而由于回复的多样性,当前业界很难有统一的评价标准。

面向任务型问答系统可以用任务的完成情况来给出评价。

面向非任务型问答系统通常有以下几种评价方法:

(1) 计算嵌入距离,可分为三种情况,一种是直接求和再取平均,一种是先取绝对值,然后求平均,还有一种是贪心匹配。

(2) 评测多样性,计算 distinct-ngram 的数量和熵值。

(3) 图灵测试,用 retrieval 的 discriminator 来评估问答系统所生成的回复。

7 问答系统的挑战与未来展望

深度学习的革命性发展给问答系统带来了长足的进步,从序列到序列的模型,到端到端的模型,再到最近很火的预训练,给问答系统留下无限的发展空间。尽管深度学习技术获得了快速的发展,但目前仍达不到商业普及应用的要求,问答系统仍然有着一些挑战,如语义嵌入尚不完善,如何利用大量的无标注数据,如何将常识库加入问答系统等等。接下来,我们将讨论一些可能的发展方向。

7.1 预训练

谷歌的 BERT 刷新了多项 nlp 记录,也为问答系统开辟了新的道路。充分利用大量无标注的文本数据,将常识引入问答系统,是问答系统走向通用智能的一个可能的趋势。

7.2 通用模型

由于各个领域的知识数据结构的不同,每个领域的问答系统框架也不同,导致当前的问答系统仍主要用于单一领域。而处理多领域多语言的问答系统是至关重要的,迁移学习和主动学习则是解决此问题的一种趋势。

7.3 深度推理

当前的问答系统大都基于标注好的数据和事先定义的数据库,在回复生成上仍缺少多样性,其根本原因在于当前的问答系统仍无法深刻的理解真实世界,无法有效理解自然语言的丰富信息。语言的本质是一个个符号,符号与符号之间的关系神经网络可以学习到,但符号背后的物理世界则是计算机所无法理解的,如何让计算机有效获取物理世界的信息是极其重要的一步。

参考文献

- [1] 李沛晏,朱露,吴多胜. 问答系统综述 [J]. 数字技术与应用, 2015 (04): 69+71.
- [2] 毛先领,李晓明. 问答系统研究综述 [J]. 计算机科学与探索, 2012, 6 (03): 193-207.
- [3] 刘里,曾庆田. 自动问答系统研究综述 [J]. 山东科技大学学报 (自然科学版), 2007 (04): 73-76.
- [4] Chen H, Liu X, Yin D, et al. A Survey on Dialogue Systems [J]. ACM SIGKDD Explorations Newsletter, 2017, 19 (2): 25-35.
- [5] Arora S, Batra K, Singh S. Dialogue system: A brief review [J]. arXiv preprint arXiv:1306.4134, 2013.
- [6] Devlin J, Chang M W, Lee K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding [J]. arXiv preprint arXiv:1810.04805, 2018.
- [8] 王东升,王卫民,王石,符建辉,诸峰. 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 (08): 1-8+41.
- [9] 李舟军,李水华. 基于 Web 的问答系统综述 [J]. 计算机科学, 2017, 44 (06): 1-7+42.
- [10] 何靖. 开放域问答系统研究综述 [A]. 中国中文信息学会信息检索与内容安全专业委员会. 第六届全国信息检索学术讨论论文集 [C]. 中国中文信息学会信息检索与内容安全专业委员会: 中国中文信息学会, 2010: 8.

作者简介

王瑛 (1970-), 女, 湖南省长沙市人。高级实验师, 广东工业大学计算机学院。研究方向为云计算、大数据、知识工程。

何启涛 (1993-), 男, CCF 会员, 广东工业大学计算机学院在读硕士。研究方向为自然语言处理, 问答系统。

作者单位

广东工业大学计算机学院 广东省广州市 510006