

网络答疑系统现状分析

刘亚军

(西昌学院,四川 西昌 615013)

【摘要】网络答疑系统在远程教育中起着非常大的作用,它能够促进师生交流、帮助学生找到问题的答案,提高学生的学习效率,同时也可以避免教师对同样的问题做出反复的解答,有利于教学质量的提高。本文对国内外网络答疑系统的现状进行了分析,指出了网络答疑系统目前存在的不足,并指出了网络答疑系统发展的方向。

【关键词】网络答疑系统;自然语言;人工智能

【中图分类号】TP393 **【文献标识码】**A **【文章编号】**1673-1891(2010)03-0068-03

一 前言

现代远程教育使教育资源实现了跨时间和空间的传递,学生不仅可以从教师那里得到知识,还能从网络上获取广泛而丰富的教学知识和内容。在整个远程教育过程中,除了网络课堂教学之外,学生自主学习的时间占了很大一部分。然而,由于受到地域和时间的限制,在自主学习阶段,学生遇到困难,不能与教师实时互动,网络答疑系统的出现使得这个问题得到了解决,它有机地把学生经常遇到的问题和老师的解答放到问题库中,并利用自然语言处理技术语义分析学生的疑问同时自动匹配答案来解决学生的疑问。因此网络答疑系统在远程教育中起着很大的作用,它可以增强教师和学生的沟通,找到解决难题的答案,提高学生的学习效率,也可以避免教师反复解答同一问题,改善教学质量。

随着网络和信息技术的迅猛发展,网络答疑系统已经成为国内外大多数远程教育系统的一个重要模块,并且在实际的教学过程中得到了广泛的应用;另一方面,网络答疑系统的使用在网络教学中也带来新的问题。

二 国内外对网络答疑系统的研究

(一) 国外研究概况

国外的远程教育网站功能比较全面,一般都具有良好的交互,自动答疑、及时反馈的功能。相应的这些网站上都有比较成熟的适用于各类问题的网络答疑系统。比较典型的有以下几类:

1. Start地理学知识专业答疑系统

Start系统是麻省理工学院于1993年开发的,具体地址是: <http://www.ai.mit.edu/projects/infolab/>。Start系统是世界上第一个面向因特网的自然语言答疑系统,专门针对麻省理工学院信息实验室的地理学专业的用户。例如:问题“What is the longest

river in the world?”,Start的回答是“With a length of 4,180 miles, the Nille River is the longest river in the world.”。此外,Start系统的答案不仅仅是文本,还有图片、声音或者动画等多媒体内容。由于Start是最早的答疑系统,所以回答问题的能力不强,系统的精确性和实时性都较差。另外,Start系统只能回答以英语提交的问题,对其他语种提出的问题无能为力。

2. AskJeeves自然语言答疑系统

和Start系统相比,美国AskJeeves公司开发的自然语言答疑系统是较先进的网络答疑系统,它的地址是: <http://www.askjeeves.com/>。AskJeeves最大的优点是用户可以用自然语言句型提问:当用户提问时,AskJeeves系统会自动分析用户的问题,然后给出一个反问,采用人机交互的方式,准确地识别用户的意图,因此,用户能够充分实现自己的检索目的。通过这种检索方式检索到的网页比单纯通过关键词搜索到的网页更能接近用户的要求。但是,AskJeeves返回的结果仍然是网页,并不是用户最想要的直接答案。

3. AnswerBus基于搜索引擎的答疑系统

AnswerBus系统是一个比较成熟的网络答疑系统,地址是: <http://misshoover.si.umich.edu/~zzheng/qa-new/>。AnswerBus是一个内核集成了Google, Yahoo, YahooNews, Alta Vista 和 Wisenu 等搜索引擎的网络答疑系统,它支持多语种问题,不仅可以回答英语提出的问题,还可以回答法、西、德、意和葡语等语言提出的问题。当用户提交相关问题时,AnswerBus系统会根据用户的提问选择一个或多个搜索引擎,查询不同的站点,提取页面中的句子并把结果反馈给用户。

4. FAQFind答疑系统

FAQFind是目前国际上比较先进的采用语义网技术构建的网络答疑系统,由芝加哥大学人工智能

实验室研制而成。该系统采用的原理是:预先建立一个常见问题库(FAQ库),通过使用语义网分析与概念匹配技术,采用基于向量的搜索引擎从问答对文件中抽取答案,最后得到问题答案。由于语义网技术的运用现在还未成熟,FAQFind答疑系统的功能有待实际运用检验。

(二) 国内研究概况

和国外比较,由于起步晚,技术落后,国内的答疑系统功能比较简单,基本上作为远程教育网站的一个附属部分,独立的具有答疑功能的网络答疑系统非常少,但近几年发展很快,呈追赶之势。国内网络答疑系统的功能主要包括基本的问题答疑、浏览信息、知识管理、问题统计等,此外,少部分答疑系统增加了一些附加功能,比如用户个性化定制以增强系统的特点。国内答疑系统功能的实现通过两种方式:人工答疑和自动答疑。

目前,国内具有代表性的网络答疑系统,有下面几类:

1. 以E-mail、bbs方式答疑的答疑系统

在系统中,E-mail、bbs是最主要的答疑方式。师生之间的信息交流仅限于电子邮件联系,或者利用bbs发布公告信息。这类典型系统有:北京邮电大学现代远程教育系统;华南理工大学远程教育系统;浙江大学现代远程教育系统和大部分的面向中小学校的远程教育系统。

2. 部分具有答疑功能的答疑系统

该类系统的特点是:学生可以在老师主持下或者独立借助于bbs以聊天的方式进行模拟上课,答疑的方式采用实时或者非实时。此类系统有:北京大学现代远程教育系统、清华大学远程教育系统、湖南大学多媒体信息教育学院的远程教育系统。

3. 问题库(FAQ)智能答疑系统

FAQ系统一般有一个存储问题和答案的数据库,初步具备了智能化答疑系统的特点。用户可以在数据库中输入关键词寻找相关的资料,如果遇到新问题,可以把它增加到数据库中,最后由专家做出解答。

这类系统的典型代表是上海交通大学远程教育设计中心开发的基于网络的答疑系统。在该系统中,用户可能提出的问答对放在FAQ库里,当用户输入自然语言语句时,系统会自动抽取其中的关键词并和库中存储的问题进行关键词匹配,同时把最适合的问题答案返回给用户。总的来说,这类系统并没有涉及语义理解方面的复杂技术,仅仅采用了基于关键词的简单匹配技术。

4. 采用全文检索方式的智能答疑系统

这类系统主要基于信息检索技术来实现问题答疑。它的典型特点是知识库是相关文档库而不是现成的问题答案对。系统首先对用户提问进行自然语言理解,同时利用信息检索技术对文档库中的文档进行检索,然后输出按查询相关度排序的文档,最后采用答案抽取技术对相关度比较高的文档进行答案抽取后返回给用户。这种答疑系统在国内外比较少,华南理工大学曾经开发过类似的系统。

三 当前网络答疑系统的不足

在实际的使用过程中,大多数网络答疑系统还存在着诸多不足,具体表现在以下几个方面:

(一) 系统交互性差

多数网络答疑系统交互性都比较差,当学生提问后,系统不能及时反馈,缺少交互,滞后期较长,这些缺点降低了学生的学习兴趣,同时他们的学习进度也受到严重的影响:旧的问题得不到及时解答,新的问题又产生。即使是较为先进的智能答疑系统,提供给学生的答案也是盲目无序的,真正的答案需要学生自己去筛选,他们的学习需要无法满足,这使得网络教育的优点根本没有体现出来。

(二) 答疑方式过于简单

网络答疑系统答疑方式分为同步和异步。同步方式响应速度快,比如在线答疑;异步方式能够针对问题进行详细的解答,比如电子公告板。现在的答疑系统一般都只有一种答疑方式,答疑方式过于简单,没有很好的把两种方式的优点结合起来,不能有效的体现网络教育的优越性。

(三) 系统对自然语言的提问支持不够

当前网络答疑系统的查询方式一般采取通过关键词或关键词的逻辑组合,对学生的自然语言提问不能很好地支持。对于使用答疑系统的学生来说,他们需要具备相当的抽词能力和逻辑组合能力,因此在使用上造成一定的困难;同时,由于答疑系统对学生提问的意图无法真正理解,造成查询结果太多而且无序,从而无法适应应用的需要。

(四) 中文信息处理技术较为落后

对当前国内网络答疑系统智能化影响最大的是中文信息处理技术。虽然经过了几十年,但是我国对自然语言的很多研究仍然没有取得突破性的进展。目前国内还没有一个各方面认可的“现代汉语语料库”或者“国家标准现代汉语语料库”,自然语言处理技术的发展受到严重的限制,这也决定了目前国内网络答疑系统的低准确率,自然语言处理技术的落后成为制约网络答疑系统的最主要原因。

四 网络答疑系统的发展趋势

虽然网络答疑系统已经在全世界广泛应用,但它的发展才刚起步,真正集实用、智能于一体的答疑系统还很少。一旦自然语言处理技术、语义分析技术、数据挖掘等技术取得突破,网络答疑系统在查全率和查准率上将会有质的提高。笔者认为未来的网络答疑系统将具有智能型(最重要的特点)、开放性、独立性、方便性等重要的特点。

(一) 智能型

未来的网络答疑系统应能通过友好的交互接口和问题逻辑推理部件更容易接受和理解用户的问题,并根据一定计算方法和推理搜索已有的知识库和信息库找到问题的可能答案或相应资源,找到的信息通过答案的解释部件呈现给用户尽可能简单、易懂的问题答案;系统还支持多种问题求解途径包括提供具有自动特征的人际交互渠道;该系统的知识库和信息库具有自动学习和更新的能力;系统的问题提交和答案呈现模块支持多媒体的呈现方式;系统同时提供对答疑行为和情况相关数据的计算和统计能力,从而优化系统知识库和信息库的结构以及提供有关数据输出;系统提供自动化和可操作的管理功能。

(二) 开放性

网络答疑系统利用资源的范围不仅仅局限于本地问题库,还应包含互联网上已存在的公共数据资源,这样不仅可以充分利用网上已有的优秀教学资源,同时避免了资源的重复建设。

(三) 独立性

目前网络答疑系统仅限于教育领域,当它发展成熟后,最终将走出教育领域,发展成为一个独立的信息咨询平台。

(四) 多媒体性

当前的网络答疑系统基本上采用的是文字答疑的方式,随着多媒体信息的大量运用,网络带宽的迅速扩大,未来的答疑系统答疑方式应该是声、图、文和动画并茂的,这将使用户使用答疑系统更方便。

五 结束语

网络答疑系统已经成为远程教育网站不可缺少的组成部分,它对提高远程教育的质量起着至关重要的作用。但目前的网络答疑系统功能比较简单、答疑效率低、智能性不高。随着人工智能和自然语言处理等相关技术的发展,未来的网络答疑系统将会在智能性、开放性、独立性、多媒体性上取得突破性的进展,从而使网络答疑系统真正给远程教育带来质的提高。

注释及参考文献:

- [1]C.Kwok, O.Etzioni and D.Weld. Scaling question answering to the Web[C]. In Proceedings of the 10th World Wide Web Conference(WWW'10), 2001.
- [2]Harabagiu, S., Maiorano, S., and Pasca, M. 2000. Experiments with Open-Domain Textual Question Answering[C]. In Proceedings of COLING-2000, 2000.
- [3]Dragomir Radev, Weiguo Fan, Hong Qi et al. Probabilistic Question Answering on the Web[C]. WWW2002, 2002.
- [4]柳泉波, 黄荣怀, 何克抗. 智能答疑系统的设计与实现[J]. 中国远程教育, 2000(8): 43-44.
- [5]梁南元. 书面汉语自动分词系统—CDWS[J]. 中文信息学报, 1987, 2(2): 44-52.
- [6]李广健, 张蕾. 网上搜索引擎的几个理论问题[J]. 情报科学, 1999, 17(4): 366-367.
- [7]戴莉. 基于Web的答疑系统关键技术研究[D]. 北京师范大学, 2008.

Analysis on the Status of Question-answering System Based on Network

LIU Ya-jun

(Xichang College, Xichang, Sichuan 615013)

Abstract: In the remote education question-answering system based on network plays an important role; it can promote the communication between teachers and students, help students to find the answer to the question, improve their learning efficiency, and also can avoid making the same problem solutions, and improve the quality of teaching. This article reviews the status of the question-answering system based on network, points out the shortcomings of the current answering system, and points out the development direction of the question-answering system based on network.

Key words: Question-answering system based on network; Natural language; Artificial intelligence