doi:10.3772/j.issn.2095-915x.2019.02.001

开放科学
（资源服务）
标识码
（OSID）

图书馆自动问答系统的设计与实现

陆伟[1,2] 戚越[1,2] 胡潇戈[1,2] 黄勇凯[3] 程齐凯[1,2]

1. 武汉大学信息管理学院 武汉 430072；
2. 武汉大学信息检索与知识挖掘研究所 武汉 430072；
3. 武汉大学图书馆 武汉 430072

摘要：数字和智能化发展已经成为高校图书馆建设的重点方向，而参考咨询作为高校图书馆服务的重要组成部分，它的智能化建设有助于更好的服务于师生和学术知识的传播。本研究以武汉大学图书馆为对象，根据现有问答语料特点和武汉大学图书馆特有的业务场景需求，构建了图书馆领域自动问答系统，系统创新性在于多技术模块的融合以及业务和学术知识的抽象和表征。本文从系统的架构设计出发，对各功能模块应用的技术和实现进行了介绍。

关键词：图书馆；自动问答系统；自然语言处理；知识图谱；深度学习
中图分类号：TP311.52 G35

Design and Implementation of Library Automatic Question Answering System

LU Wei[1,2] QI Yue[1,2] HU Xiaoge[1,2] HUANG Yongkai[3] CHENG Qikai[1,2]

1. School of Information Management, Wuhan University, Wuhan 430072, China;
2. Information Retrieval and Knowledge Mining Laboratory, Wuhan University, Wuhan 430072, China;
3. Wuhan University Library, Wuhan 430072, China

Abstract: Digital and intelligent development has become the goal for library construction. As an important part of library services the development of reference consulting helps to better serve teachers and students and to promote the dissemination

基金项目：富媒体数字出版内容组织与知识服务重点实验室开放基金项目（ZD2018-07/04）：“社科类教材语义出版与增强显示研究”。
作者简介：陆伟（1974-），博士，副院长，教授，研究方向：信息检索、知识管理、数据挖掘等，E-mail：weilu@whu.edu.cn；戚越（1995-），硕士研究生，研究方向：信息检索；胡潇戈（1995-），硕士研究生，研究方向：信息检索；黄勇凯（1975-），硕士，副研究馆员，研究方向：数字图书馆建设与服务；程齐凯（1989-），博士，讲师，研究方向：信息检索、文本挖掘与自然语言处理。

doi:10.3772/j.issn.2095-915x.2019.02.001

of academic knowledge. Based on the existing corpus and unique business scenario of Wuhan University library, our research had set up an automatic question answering system. The multi-technology combination and the way we organized business and academic knowledge manifest the research's innovation. From the viewpoint of architecture design, this paper investigated the technology and implementation of each functional module.

Keywords: Library; automatic question answering system; nature language processing; knowledge graph; deep learning

引言

高校图书馆的智能化管理和服务已经成为图书馆发展与建设工作的重要组成部分[1]。一方面，为了给用户提供更高效、智能的咨询服务，各大高校不断加快图书馆转型，充分利用相关线上线下平台，解决传统高校图书馆人工服务效率低，反馈时间较长，找书过程繁琐等使用不便利的现实问题，进一步提升效率并提高师生满意度。另一方面，随着自然语言处理技术的发展，各类自动问答系统的构建日益成熟，且已经在多个领域得到广泛应用。因此，探索问答系统在图书馆领域的应用，给图书馆的智能化发展带来了新的机遇和挑战，能够从根本上提升图书馆管理和服务的效率和质量[2]。

本文首先阐述了自动问答系统的发展和在高校图书馆领域的应用现状，然后系统地介绍武汉大学图书馆智能问答系统的架构与功能模块设计，尤其是根据武汉大学问答语料和业务场景的特点进行的知识存储与获取设计，展示了系统的界面和效果。最后对未来图书馆自动问答系统的研究应用进行思考与讨论，以期能够为高校图书馆自动问答系统的应用和推广提供参考和借鉴。

1 相关研究综述

1.1 自动问答系统发展现状

问答系统最早的实现构想可以追溯到图灵测试。为了测试机器是否具备人类智能，图灵测试要求电脑能在 5 分钟内回答由人类测试者提出的一系列问题，且其超过 30% 的回答让测试者误认为是人类所答。随着人工智能、自然语言处理等相关技术的发展，针对不同的数据形态的变化也衍生出不同种类的问答系统，按年代发展可分为面向限定领域的问答系统，面向开放领域的问答系统和交互式问答系统。

早期由于智能技术和领域数据规模的局限性，问答系统主要是面向限定领域的 AI 系统或专家系统，例如解决高中代数相关词汇问题的 STUDENT[3] 系统、可以回答阿波罗月球探测任务相关样本分析结果的 LUNAR[4,5] 系统 等。该时期的问答系统处理的数据类型主要是结构化数据，系统一般是将输入问题转化为数据库查询语句，然后进行数据库检索反馈答案。

面向开放领域的问答系统支持自由文本，例如英文问答式检索系统 Ask Jeeves (http://www.ask.com)、麻省理工学院 (MIT) 计算机科学系人工智能实验室开发的 START(http://

start.csail.mit.edu) 可以回答多个领域的自然语言问题[6]、AnswerBus 多语系自动问答系统，可以回答英语、法语、西班牙语、德语和葡萄牙语的问题[7]。早期的这种问答系统的处理流程主要包括：问题分析、文档及段落检索、候选答案抽取、答案验证。此后，自 1999 年文本检索会议（Text Retrieval Conference，简称 TREC）引入问答系统评测专项（Question Answering Track，简称 QA Track）以来，包括 NTCIR、CLEF 等评测组织的出现，极大推动了自然语言处理技术在问答领域中的研究发展[8,9]。随后网络上出现的社区问答（community question answering, CQA）提供了大规模的用户交互衍生的问题答案对（question-answerpair, QA pair）数据，这为基于问答对的问答系统提供了稳定可靠的问答数据来源。

随着苹果公司 Siri 系统的问世，问答系统进入了智能交互式问答的发展阶段，这种形式的问答系统能够让用户体验更为自然的人机交互过程，并且也使信息服务的相关应用更为方便可行。相对于搜索引擎，问答系统能更好地理解用户提问的真实意图，同时更有效地满足用户的信息需求。

问答系统的分类标准有多种，依据用户问题的所属数据领域，问答系统可分为面向限定域的问答系统、面向开放域的问答系统、以及面向常用问题集（frequent asked questions, FAQ）的问答系统。依据答案的不同数据来源，问答系统可划分为基于结构化数据的问答系统、基于自由文本的问答系统、以及基于问答对的问答系统。此外，按照答案的生成反馈机制划分，问答系统可以分为基于检索式的问答系统和基于生成式的问答系统。

1.2 高校图书馆问答系统发展现状

图书馆的发展是循序渐进的，其发展中的每一次重大变革都伴随着科学技术的创新与突破，从而促使图书馆快速的从传统的人工服务方式转换为现代的自动化方式，“互联网 +”时代下，智能问答机器人入驻图书馆是大势所趋。图书馆自动问答系统很大程度上分担了部分咨询服务工作，成为图书馆与读者沟通互动的桥梁和纽带，这些都是构建数字化图书馆的基础性工程。

国外图书馆智能问答系统的应用最早开始于以德国汉堡大学为代表的一些欧洲图书馆。相比于国外，国内也有很多高校图书馆引入了自动问答系统。清华大学是最早将基于人工智能的问答引擎融入图书馆服务平台的国内高校之一[10]，其研发的“小图”主要基于 AIML（人工智能标记语言）和搜索引擎技术，提供业务咨询和馆藏搜索；南京大学图书馆利用关键字匹配和数据库查询等技术，依托微信平台，构建了问答系统[11]；除南京大学外，包括哈尔滨工业大学图书馆[12]、东南大学图书馆[13]、西安交通大学图书馆[14]等在内的诸多图书馆也基于检索、关键词匹配等技术构建了基于微信平台的问答系统。其他如宁波大学以及成都理工大学图书馆等也先后引入了机器人馆员。除高校图书馆外，部分公共图书馆开始借助问答系统的技术更好地为读者解惑答疑，如上海图书馆、湖北省图书馆、杭州新华书店以及北京图书大厦等。

相比于目前国内高校图书馆应用，武汉大

doi:10.3772/j.issn.2095-915x.2019.02.001

学图书馆自动问答系统有两大特点：从应用场景来说，不仅关注图书馆资源服务方面的业务需求和 FAQ，而且充分利用了图书馆学术领域知识富集的特点，对其做了组织与加工，通过问答形式实现学术知识的搜索和利用；从技术上来说，不局限于单个技术，致力于多技术方案、多功能模块的融合，以适应不同的应用场景需求，利用知识图谱抽象了业务和学术领域的实体、属性和关系，使得相关知识能够在语义网络中得以表征。

2 问答系统架构

本文在已有研究的基础上结合武汉大学图书馆的实际需求，构建图书馆智能问答系统，总体架构如图 1 所示。

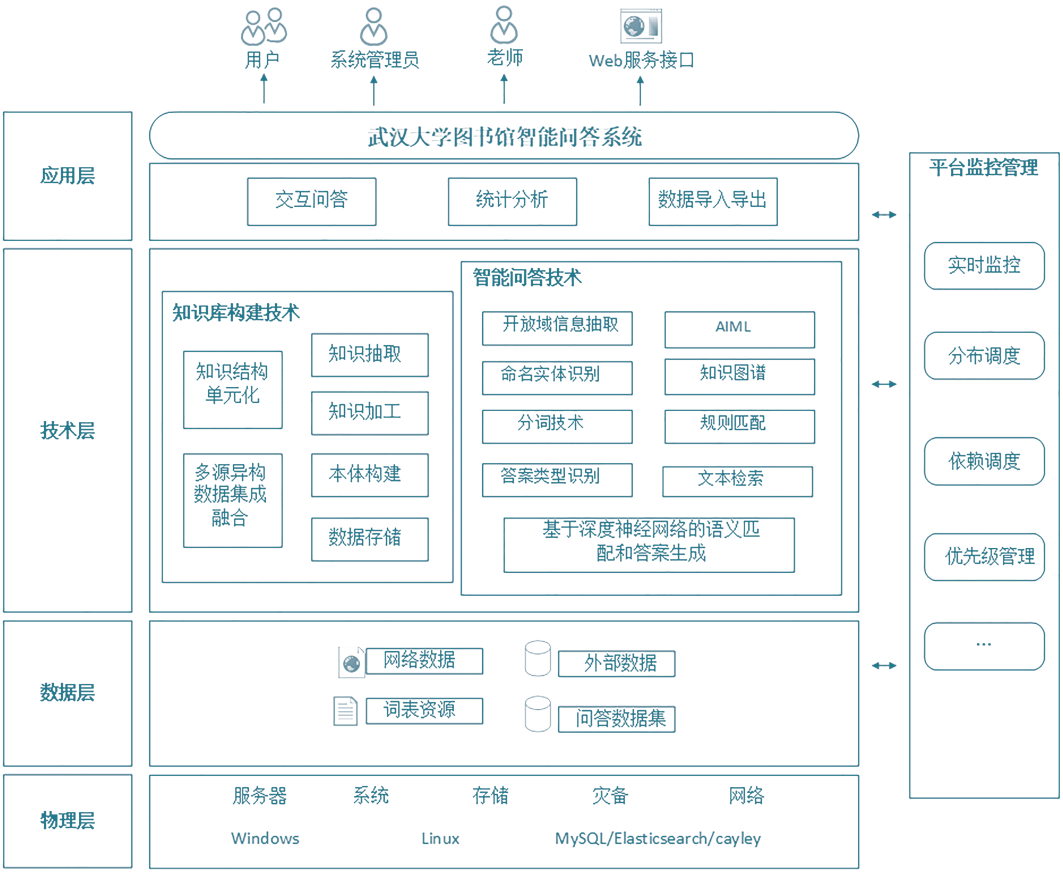

图 1 系统总体架构图

系统包括架构从逻辑上可划分为物理层、数据层、技术层和应用层几部分。其中物理层描述了系统服务器以及存储工具；数据层描述了多源问答数据；在技术层，主要涉及知识库的构建和问答模块，其中问答模块的设计和实现是本研究讨论的核心；在应用层面主要面向全校师生在图书馆相关业务时场景下的辅助参考咨询，除核心咨询服务外会提供后台统计分析和数据导入导出模块供图书馆管理人员查看与使用。接下来对知识库存储与问答模块的设计与实现进行重点说明。

3 知识库存储

本节简要介绍知识库模块问答知识的存储。基于通用图书馆参考咨询问答领域的限定性和学术领域知识的富集性，又结合武汉大学图书馆服务和武汉大学学科领域研究的个性化特点，配合技术选型，本研究设计了几种知识的存储形式，包括原始问答对的直接存储、基于 AIML 格式的数据存储、知识图谱 RDF 三元组形式的存储、问答数据索引形式的存储以及词典形式的存储。

武汉大学图书馆智能问答应用可分为三个场景，分别是图书馆业务咨询、学术搜索咨询和通用问答。不同知识存储形式被设计用来支持不同业务场景下的咨询问答，业务咨询主要产生于师生在使用图书馆提供的资源服务，比如查新、查重、检索、馆藏资源的获取和其他资源的利用等，原始问答对、基于索引形式的问答数据以及知识图谱均适用于这种业务场景；学术领域咨询主要产生于对学术知识的获取，由于高校在学科领域研究的深入性，会涉及对学术领域知识如领域期刊、作者、文献、数据库等信息的获取，本研究主要利用知识图谱来构建学术领域的语义网络，表现为 RDF 三元组形式。通用问答场景不受领域限制，典型数据如通用问候，以问答对和 AIML 格式存储。

4 问答模块设计与实现

4.1 系统问答流程

核心问答模块按照回答场景的不同选择不同的技术选型，又可细分为文本硬匹配、AIML、知识图谱、检索匹配和深度学习等子功能模块。流程框架如图 2 所示：

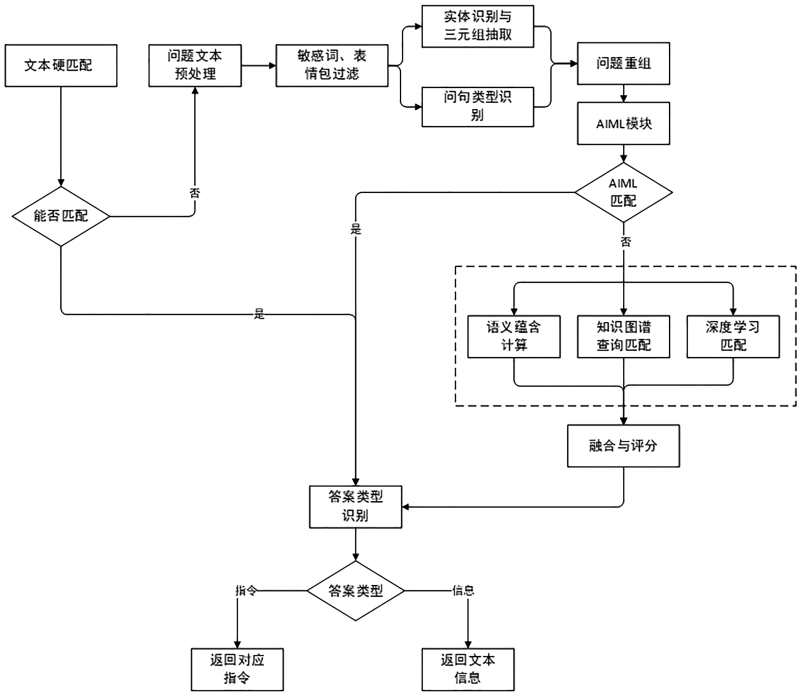

图 2 系统问答模块流程架构图

doi:10.3772/j.issn.2095-915x.2019.02.001

整个问答流程采用责任链模式，基于一个调度中心，调度多个问答子模块的形式。用户输入问句后，系统首先做文本硬匹配，当文本硬匹配失败时，进入问句语义识别模块，识别用户的查询意图，然后将问题重组，推给 AIML 模块进行规则和模板匹配，若匹配失败，进入知识图谱和深度学习模块进行语义匹配，融合计算答案得分并进行置信度校验最终返回。在每个模块中除了针对单个问题返回系统认为最优答案外，系统根据每个提问进行了相似问题推荐。当用户选择人工回复时，系统会直接把问题推给专家坐席，由相关老师人工回复。

4.2 功能模块

4.2.1 文本硬匹配

本研究中的文本硬匹配指的是问句文本的完全匹配，利用已有的图书馆咨询问答历史数据，构建问答对存储结构，存储在缓存中。当用户再次提问历史提问过的完全相同的问题时，直接返回该问题对应的答案。

文本硬匹配作为本系统处理问题的第一个子模块，应用场景多在于解决 FAQ 问题，即对于用户最常见的问题，直接从系统缓存中取出返回对应答案，可以进行快速回复，而不需要经过系统内复杂的问题处理逻辑包括：对文本进行预处理、语义识别、选择合适的子模块返回答案以及置信度校验等，有效提升了系统的处理和回复效率。

4.2.2 问句语义识别

问句语义识别模块即对问句的语义解析，包括文本预处理、问句类型识别等。此模块的核心是理解用户查询意图，是后续其他子模块的基础。

首先对输入文本进行规范化处理，包括敏感词过滤如粗口、色情等不适合呈现的内容，表情包过滤，繁体转简体和中英文符号统一化等。其次，对文本进行分词和命名实体识别，采用词典和 CRF 结合的技术方案，本研究根据图书馆咨询场景构建了业务咨询和学术搜索领域专有词典。再结合规则匹配和模板匹配，对输入问句进行知识图谱查询匹配所需的三元组等关键信息的抽取。

为了解决抽取出的实体过少或冗余的情况，本研究还结合了问句类型识别来提升用户意图理解的效果。问句类型识别指的是系统在解析一个问句时，依据其所期望的答案的类型来给问句打标签分类的过程，其中答案类型指的是答案所属的类别或者它的本体。问句类型识别被证明在问答系统中起着重要的作用，尤其是对于事实型问题 [15]。本研究的图书馆智能问答系统参考了 BBN 对问句本体，即答案类型的定义 [16]，并根据图书馆实际业务咨询和学术搜索语料所涉及的本体定义了答案类型 answer-Type，如表 1 所示。

对于像高校图书馆咨询这样的限定领域，以上定义的问句类型可以覆盖大部分图书馆咨询问题。

4.2.3 AIML

AIML 可以理解为基于 XML 解析的问答技术，引入该模块的原因是，本研究在分析了图书馆咨询问答领域语料后发现，有相当一部分问题可以用基于规则的方式去解决；而且 AIML 提供了丰富的语义标签，增加了规则的

灵活性，例如相同候选答案的随机回复、同义问句标签以及会话的支持很适合咨询问题产生歧义时，上下位问题深度通常为 2 层的情况，有助于提升系统与用户交互的友好性。

表 1 问句本体（答案类型）定义

answerType	类型定义	样例
Entity	实体	武汉大学图书馆智能问答系统是什么?
description	描述	帮介绍下中国知网吧?
Date	日期	豁免日是哪一天呀?
Time	时间	图书馆开馆时间是什么时候?
person	人	请问我想去咨询下图书馆志愿者应该找哪位老师呀?
location	地点	图书馆自助打印机在哪里?
reason	原因	为什么我的校园卡被禁?
behavior	行为方式	怎么座位预约?
boolean	布尔	违约三次还可以借书吗?
numeric	数量	违约几次不能再借书了呀?
money	金钱	图书馆 3d 打印一次要多少钱呀?

在本研究开发的图书馆智能问答系统中 AIML 模块具体应用在以下几个场景：独立于图书馆业务咨询和学术搜索领域之外的通用问答，比如通用问候、许多不相干业务的问答；问题和答案定义较明确的场景，如有关武汉大学一些事实信息的描述，比如“武大校长是谁？”、“简介下武汉大学图书馆智能问答系统”；对于一些特定问题需要进行意图确认的多轮对话场景，比如投诉类问题和部分意图不明确的问题：对投诉类问题的回复具有一致性，多是如“记录投诉信息并反馈给相关老师”，对部分意图不明确的特定问题，比如“预约”，在图书馆业务咨询领域中分为“座位预约”和“图书预约”，可以利用 AIML 的 <that> 标签进行投诉信息确认和提问意图的补全和完善。

此外，由于 AIML 最初设计应用在英文聊天机器人，只支持英文的特点，本系统在对问题进行语义解析后，需要对问题关键信息进行重组，来达到中文应答效果。

4.2.4 知识图谱

知识图谱最早由 Google 提出，从本质上看，知识图谱是一种是结构化的语义知识库 [17]，是一种大规模的语义网络 [18]，它以结构化的形式描述客观世界中的概念、实体及关系。知识图谱的基本组成单位是 RDF 三元组，即“实体 - 关系 - 实体”或者“实体 - 属性 - 属性值”。

知识图谱应用于自动问答时不同于 web 文档搜索，在自然语言问题能够很好的被理解并精确地映射到知识库的结构化查询时，基于知识库的问答（KBQA, Question Answering over Knowledge Bases）可以给出准确且简洁的答案[19]。

doi:10.3772/j.issn.2095-915x.2019.02.001

知识图谱子模块主要利用了问句语义识别模块的实体和三元组抽取结果，即候选的实体、属性 / 关系词，以及问句类型。因为知识库中存储的是标准的消歧后的语义项，所以需要将抽取出的实体和关系 / 属性进行消歧和标准化映射 [20]。

本研究综合利用了同义词词典并计算了原始语义项和规范化语义项之间的相似度，返回相似度得分最高的语义项作为最终结果，实现 mention 到 entity 和 relation 到 property 的转换。此外，为了弥补信息抽取时存在冗余和识别错误的问题以及提升知识图谱查询匹配效率的问题，本研究在规范化映射和查询匹配之间增加了关联关系匹配，即专门构建了实体 - 属性或实体 - 关系关联表，利用该表来过滤与用户意图无关的实体或属性，筛选出具有主谓关系的实体 - 属性或实体 - 关系对。然后利用具有主谓关系的语义项和问句类型属性，构建相应的结构化查询语句，查询语言选择 cayley 提供的基于 Gremlin 的 gizmo 语言，在图数据库中进行查询和子图匹配，进行子图匹配时计算了相似度得分，返回最高得分的结果作为最终的答案。

图数据库存储示例如图 3 所示。

图 3 “图书馆”实体在图数据库中的存储

图 3 展示了“图书馆”实体及其相关知识在图数据库中的存储。其中，蓝色点代表“图书馆”实体，边代表图书馆的属性或关系，黄色点代表图书馆的属性值或者是与图书馆有关联关系的实体。具体信息为：图书馆 - 开馆时间 - 代表开馆时间的 URI（Uniform Resource Identifier，统一资源标识符）。

利用图数据库进行查询时，例如询问“图

书馆开馆时间是什么时候？”，转换成结构化的查询语句为：

g.V(“图书馆”).LabelContext(“<service_graph>”)

.Out(“开馆时间”)

.LabelContext(null).Save(“value”,”value”)

.And(g.V().Has(“answerType”,“time”))

.All()

查询语句信息为：查询属于业务咨询 <service> 子图下的图书馆开馆时间的属性值，而且问句（答案）类型为 time（时间）类型。

4.2.5 多轮对话

本研究系统的多轮对话模块基于知识图谱模块，是对知识图谱模块的补充。知识图谱存储依赖于 RDF 三元组形式，使得其更适合回答较明确的主谓宾形式的事实型问题 [21]，即根据实体属性获得属性值或者根据实体和关系获得实体，但实际用户提问具有多样性，往往会出现主谓角色缺失和角色歧义的情况。

本研究设计了多轮对话来进行用户意图的确认，轮数限制在 2~3 轮以内，一定程度上增加了 KBQA 问答模块的鲁棒性。对于角色缺失的情况，当只识别出主谓其中之一时，先去知识图谱中查询该角色所对应的另一种角色的候选值，然后返回给用户进行角色值的确认；对于角色歧义的情况，通常涉及上下位类查询，系统同样通过返回属性列表来确认唯一的下位类主语。多轮对话的设计目的是通过与用户的交互，构建完整的无歧义的结构化查询，然后将查询匹配出属性值或实体作为答案返回。

4.2.6 深度学习

前几个模块严格依赖语义分析准确程度、规则和模板定义的复杂程度以及结构化知识库的完善程度，特点在于准确率较高。深度学习模块基于问答语料检索匹配，并致力于优化排序，对于图书馆咨询这样的限定领域，尤其是业务咨询，能很好的解决其他模块没有覆盖的问题，提升召回率。

基于深度学习的文本匹配相比于传统方法，模型能够从文本中提取出词语间的关系和结构信息，更准确地表征文本匹配问题。

该模块基于利用 Elasticsearch 构建的问答句对索引，先利用 Elasticsearch 根据问题检索出候选答案集，然后以用户输入的问题和检索出的答案候选集为输入，利用训练好的模型进行优化排序，输出匹配度最高的答案。本研究采用基于注意力机制的 BiGRU 模型构建网络结构，模型网络结构如图 4 所示。

基于注意力机制的双向 BiGRU 模型首先基于词嵌入（Word Embedding）技术，把问题或答案句子中的词表示为词向量；其次，利用双向循环神经网络（Bidirectional Recurrent Neural Network, Bi-RNN）对词向量进行编码，利用前向、后向最后时刻的状态，计算注意力权重，利用答案的注意力权重对问题编码向量加权，以及利用问题的注意力权重对答案编码向量加权，将加权后的问题和答案向量进行拼接，再利用全连接网络计算问题和候选答案的相似度，得到 matching 值；最后对 matching 值进行排序，得到最佳匹配的答案，作为该模块最终输出。

该模块作为整个系统责任链的最后一个子功能模块，引入了置信度校验，当答案低于阈

doi:10.3772/j.issn.2095-915x.2019.02.001

值时会返回系统默认应答指令。虽然依赖于语料库的大小，但利用了图书馆咨询领域问题重复率较高的特点，对于问句解析阶段语义项过少或过多的问题应答效果较好。

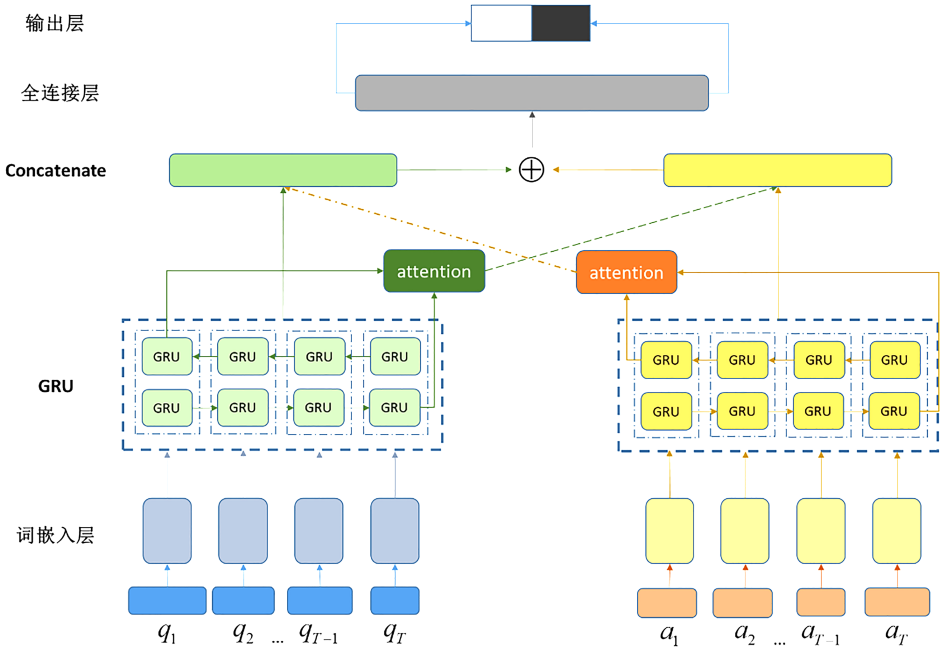

图 4 基于注意力机制的 BiGRU 模型结构示意图

5 系统实现及实验

5.1 系统功能

本系统核心功能是图书馆辅助参考咨询的自动问答，辅助图书馆相关管理人员解决师生在获取和使用图书馆资源或者学术知识获取场景下遇到的疑问。不仅如此，在基础智能问答的基础上，系统支持热点问题推荐和相关问题推荐，热点问题即师生常见问题，相关问题是与上一轮用户问题相似的问题。系统同时支持置信度校验，低于置信度的答案系统不采纳。为弥补自动回复能力的不足，系统提供了转人工服务，用户可以通过相应关键字开关，开启人工服务，问题将直接转到专家坐席，由相关老师或职工来回答，结束后可通过开关回到自动模式。此外，系统提供了语音输入功能，支持语音转文字交互。

5.2 系统界面截图

“武汉大学图书馆智能问答系统”嵌入平台包括 web 和微信公众号两个，用户交互界面如图 5 和图 6 所示。

图 5 展示了系统 web 端用户交互界面，5A 展示了“武汉大学图书馆自动问答系统”初始界面，5B 和 5C 是系统的问答主界面，可以看出，系统不仅实现了问题的自动回复，而且提供了常用问题推荐并按标号回复以及语音输入功能。图 6 展示了系统的微信公众号交互界面，微信公众号交互方式依托微信服务器，基本功能和 web 端一致，删除了问题推荐功能。

图 5 5A，5B，5C“武汉大学图书馆智能问答系统”web 端界面

图 6 “武汉大学图书馆智能问答系统”微信公众号界面

5.3 系统测试

本研究针对图书馆常见问题对系统做了测试，测试问题数量为 200，主题包括通用问题、电子资源获取、纸质资源获取、入馆指南、证卡服务、投诉及建议等多个主题，测试结果准确率达到 74.3%。实验证明，本研究开发的自动问答系统可以达到一定程度的替代人工咨询服务的要求。

6 结语

随着深度学习与自然语言处理技术的不断发展，自动问答系统已经成熟应用于诸多行业，尤其是在智能客服领域的广泛应用，为图书馆的智能化发展提供了全新的机会。图书馆在知识一体化基础上，借助自动问答系统能够减少图书馆工作人员在咨询任务上的工作量，更加有力地保障咨询服务的提供，同时，通过学术领域知识的存储与服务，能够使得高校学术知识得到更好的传播与应用。

在未来，图书馆自动问答系统具有一定的现实需求和研究应用前景。图书馆专业语料库质量的提升，会使问答系统知识库中的知识更加丰富，从而让问答系统返回的答案更加精确和专业，自动问答系统会将图书馆的智能化服务推向一个新的高度。人工智能与人机交互的发展，更有利于图书馆自动问答系统管理和服务体系的形成，为用户提供更加智能和精确的书籍检索、学术搜索结果以及信息咨询等服务，从而全面提升图书馆的服务质量，吸引更多的

doi:10.3772/j.issn.2095-915x.2019.02.001

师生使用图书馆资源和服务。

参考文献

[1] 张丽娟，陈越，李丽萍．高校图书馆的智能化管理与服务——北卡罗来纳州立大学图书馆带来的启示 [J]. 大学图书馆学报，2015, 33(2):26-29.

[2] 曾帅，王帅，袁勇，等．面向知识自动化的自动问答研究进展 [J]. 自动化学报，2017,43(9): 1491-1508.

[3] Bobrow D G. A Question-Answering System for High School Algebra Word Problems[C]. October, Fall Joint Computer Conference. ACM, 1964.

[4] Wood J A, Dickey J S, Marvin U B, et al. Lunar anorthosites and a geophysical model of the moon[J]. Geochimica Et Cosmochimica Acta Supplement, 1970, 1.

[5] Woods W A. Lunar rocks in natural English: explorations in natural language question answering[J]. Linguistic Structures Processing, 1977(5): 521−569.

[6] Boris K, Gregory M, Gary B, et al. The START Natural Language Question Answering System [EB/OL]. [2006-12-16]. http://start.csail.mit.edu.

[7] Zhiping Z. AnswerBus Question Answering System [EB/OL]. [2006-12-16]. http://answerbus.coli.uni-saarland.de/index.shtml.

[8] Lin J, Demner-fushman D. Will pyramids built of nuggets topple over?[C]. Proceedings of the Main Conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics (HLT-NAACL '06), Morristown, NJ, USA, 2006. Stroudsburg, PA, USA: Association for Computational Linguistics, 2006: 383–390.

[9] Lin J, Demner-fushman D. Automatically evaluating answers to definition questions[C]. Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT '05). Stroudsburg, PA, USA: Association for Computational Linguistics, 2005: 931−938.

[10] 姚飞，纪磊，张成昱，等．实时虚拟参考咨询服务新尝试——清华大学图书馆智能聊天机器人 [J]. 数据分析与知识发现，2011, 27(4):77-81.

[11] 沈奎林，邵波，赵华．利用微信构建图书馆智能问答系统 [J]. 图书馆学研究，2015(8):75-80.

[12] 李雪婷，李莘．图书馆微信平台自动问答机器人语言体系研究 [J]. 现代情报，2016, 36(10):99-101,122.

[13] 罗涛，朱莹．图书馆 MSN 咨询机器人的研究与实现 [J]. 图书情报工作，2012, 56(19):105-108.

[14] 李丹．图书馆微信平台建设实践与思考 [J]. 现代图书情报技术，2016(4):104-110.

[15] Metzler D, Croft W B. Analysis of Statistical Question Classification for Fact-Based Questions[J]. Information Retrieval, 2005, 8(3):481-504.

[16] Bikel D M. An algorithm that learns what's in a name[J]. Machine Learning, 1999, 34.

[17] 刘峤，李杨，段宏，等．知识图谱构建技术综述 [J]. 计算机研究与发展，2016, 53(3):582-600.

[18] 李薇，肖仰华，汪卫．基于中文知识图谱的人物实体识别 [J]. 计算机工程，2017, 43(3):225-231.

[19] Cui W, Xiao Y, Wang H, et al. KBQA: learning question answering over QA corpora and knowledge bases[J]. Proceedings of the VLDB Endowment, 2017, 10(5):565-576.

[20] Zhang J, Liu S, Li M, et al. Question Answering over Linked Data Using First Order Logic[C]. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2014.

[21] Cui W, Xiao Y, Wang W. KBQA: an online template based question answering system over freebase[C]. International Joint Conference on Artificial Intelligence. AAAI Press, 2016.