

面向知识自动化的自动问答研究进展

曾 帅^{1,2} 王 帅^{1,3} 袁 勇^{1,2} 倪晓春^{1,2} 欧阳永基⁴

摘 要 将自动问答系统从基于文本关键词的层面,提升到基于知识的层面,实现个性化、智能化的知识机器人,已成为自动问答系统未来的发展趋势与目标. 本文从知识管理的角度出发,分析和总结自动问答领域的最新研究成果. 按照知识表示方法,对代表性自动问答系统及关键问题进行了描述和分析;并对主流的英文、中文自动问答应用和主要评测方法进行了介绍.

关键词 自动问答系统, 知识机器人, 知识自动化, 知识工程

引用格式 曾帅, 王帅, 袁勇, 倪晓春, 欧阳永基. 面向知识自动化的自动问答研究进展. 自动化学报, 2017, 43(9): 1491–1508

DOI 10.16383/j.aas.2017.c160667

Towards Knowledge Automation: A Survey on Question Answering Systems

ZENG Shuai^{1,2} WANG Shuai^{1,3} YUAN Yong^{1,2} NI Xiao-Chun^{1,2} OUYANG Yong-Ji⁴

Abstract Question answering systems are evolving from the text keywords level to the next knowledge-based level, and thus realizing personalized and intelligent knowledge robots has become the development trend and goal for the future question answering systems. From the viewpoint of knowledge management, this paper analyzes and summarizes the latest research findings in the field of automatic question answering. According to the knowledge representation methods, this paper surveys several representative question answering systems and analyzes their key technologies. Also investigated are some mainstream English and Chinese-language-based question answering applications and the commonly used evaluation methods.

Key words Question answering system, knowledge robots, knowledge automation, knowledge engineering

Citation Zeng Shuai, Wang Shuai, Yuan Yong, Ni Xiao-Chun, Ouyang Yong-Ji. Towards knowledge automation: a survey on question answering systems. *Acta Automatica Sinica*, 2017, 43(9): 1491–1508

在过去的二十年间,搜索引擎产品的出现及其发展,一定程度上满足了用户从海量互联网开源数据中获取信息的需求. 一般而言,搜索引擎基于用户挑选的一个或多个关键词,返回大量按相关程度排序的网页,由用户筛选搜索结果获取信息. 这一过程操作繁琐且具有一定难度. 随着个人携带智能设备的普及和移动互联网的飞速发展,用户对获取精准答案的需求不断增长,自动问答系统 (Question answering (QA) system) 成为新一代互联网的热门应用之一.

对用户来说,自然语言是最佳的人机交互方式. 自动问答系统用自然语言来回答用户的问题,比搜索引擎更友好,更能满足用户操作简单化和知识精准化的需求. 随着大数据和知识管理、智能化深度分析、个人携带智能设备、移动互联网等相关技术与应用的突飞猛进,自动问答的形态和应用场景正在经历深刻变化,也使得用户的行为与需求模式发生了极大的变化. 将问答系统从基于文本关键词的层面,提升到基于知识的层面,实现智能化的知识机器人,已成为问答系统未来的发展趋势与目标.

利用知识自动化的方法,对虚拟空间的大数据进行深度的开发和智力挖掘,才有可能有效应对不确定性、多样性和复杂性问题^[1–2]. 从知识自动化的角度来看,自动问答系统的工作流程可以对应到知识管理生命周期的各个阶段^[3],如图 1 所示.

1) 知识识别: 当接收到问题时,自动问答系统首先对问题进行分析,然后通过知识识别检索封装好的知识,确认系统内部是否存在相关知识.

2) 知识创建: 如果在识别阶段通过知识检索找不到解答问题的相关知识,或者已有知识不够全面,则可能会触发创建或获取新知识的需求. 自动问答系统对数据源进行分析和知识提取,获得新知识,并通过评估确认新知识是否有价值.

收稿日期 2016-09-18 录用日期 2017-05-31

Manuscript received September 18, 2016; accepted May 31, 2017

国家自然科学基金 (61233001, 61533019, 71102117, 71232006, 71402178, 71472174, 71702182) 资助

Supported by National Natural Science Foundation of China (61233001, 61533019, 71102117, 71232006, 71402178, 71472174, 71702182)

本文责任编辑 王飞跃

Recommended by Associate Editor WANG Fei-Yue

1. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室 北京 100190 2. 青岛智能产业技术研究院 青岛 266109 3. 中国科学院大学 北京 100049 4. 解放军 61786 部队 北京 102206

1. State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 2. Qingdao Academy of Intelligent Industries, Qingdao 266109 3. University of Chinese Academy of Sciences, Beijing 100049 4. Troops 61786 of People's Liberation Army, Beijing 102206

答系统的发展趋势与最终目标是实现智能化的知识机器人, 本文着重介绍面向知识自动化的问答系统。

2 自动问答系统的基本架构

自动问答系统的基本架构如图 2 所示, 自动问答系统通过人机交互界面接受用户问题和向用户返回答案, 其中问题和答案均为自然语言形式。按照系统是否限定问题范围, 可分为开放领域自动问答系统和限定领域自动问答系统。

系统的核心模块包括数据源分析、问题分析、知识检索、答案生成 4 部分。此外, 答案评估模块对问题-答案对进行评估, 并将评估结果返回给核心模块。核心模块依据评估结果, 可以通过完善知识体系、提高知识挖掘与检索效率、改善答案生成方法等手段, 以改善用户体验。下面分别对各个核心模块展开描述。

2.1 数据源分析

数据源分析主要是采用词性分析、句法分析、语义分析、依存分析、实体识别等多种方法与手段, 从结构化、半结构化或非结构化数据中, 提取有用的信息与知识, 并将其表示为某种知识表示形式。

2.2 问题分析

问题分析主要是采用包括而不同于词性分析、句法分析、语义分析、对话分析、期望答案类型分

析、问题焦点识别等自然语言处理方法, 对问题进行语义理解和分析。一般而言, 问题可分为事实性问题、列表问题、假设问题、确认问题、因果问题等。问题类型将影响对期望答案类型的判断。例如, 崔恒等^[14] 将问句的预期答案类型分为 16 种基本类型, 如: 人、机构、地点、数量、时间等。期望答案类型的确定可以为知识检索和答案抽取提供参考依据。

经过一系列问题分析之后, 将问题表示为与数据源的知识表示方法相匹配的形式, 如逻辑表示 (如图 3 所示)。

2.3 知识检索

在知识检索模块, 主要是基于问题的知识表示, 在知识库中检索相关知识子集, 涉及数据挖掘、信息检索、知识检索与发现等多种方法与技术。

2.4 答案抽取

在答案抽取模块, 首先对知识子集进行句子相似度计算、依存关系分析、多源数据融合, 实现知识排序与融合, 然后对融合后的知识进行答案检验与评估, 最终生成答案。答案可能是抽取文字或多媒体数据片段, 也可能是生成式答案。

3 历史回顾

根据主要处理的数据类型, 自动问答系统的发展可分为三个阶段。

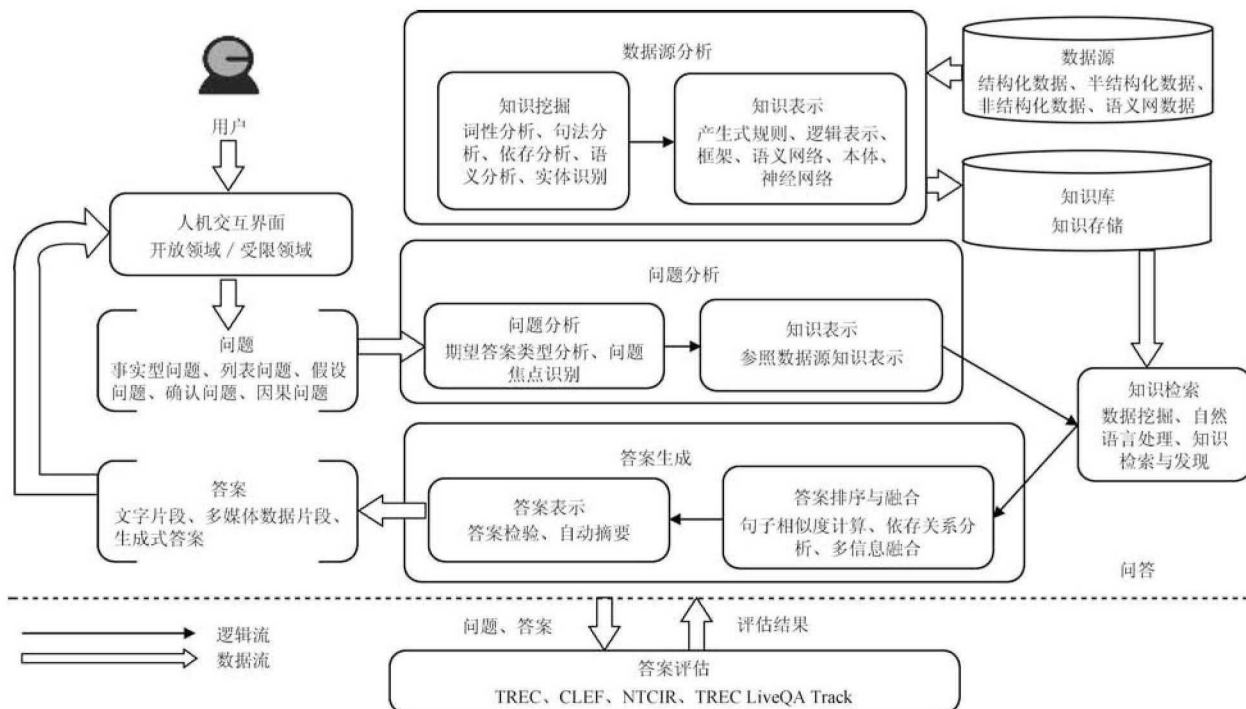


图 2 基本架构

Fig. 2 System architecture

Example:

Heavy selling of Standard & Poor's 500-stock index futures in Chicago relentlessly beat stocks downward.

LF:

```
heavy_JJ(x1) & selling_NN(x1) & of_IN(x1, x6) &
Standard_NN(x2) & CC(x13, x2, x3) & Poor_NN(x3)
& 's_POS(x6, x13) & 500-stock_JJ(x6) & index_NN(x4)
& future_NN(x5) & in_NN(x6, x4, x5) & in_IN(x1, x8)
& Chicago_NN(x8) & relentlessly_RB(x12) &
beat_VB(x12, x1, x9) & stocks_NN(x9) & down-
ward_RB(x12)
```

图 3 问题的逻辑表示

Fig. 3 The logic form of an example question

3.1 第一阶段：以结构化数据和数据库检索技术为中心

第一阶段是以结构化数据和数据库检索技术为中心的自动问答系统，始于 20 世纪 60 年代。由于结构化数据一般存储在数据库中，这类 QA 系统也被称为数据库自然语言接口。其目标是允许用户通过自然语言的形式，查询数据库中存储的信息。这类系统得益于人工智能 (Artificial intelligence) 的诞生与兴起^[6]，是自然语言理解与数据库技术的结合。

早期最具代表性的 NLDB 是 1961 年的第一个问答系统 Baseball^[15] 和 1972 年的 LUNAR^[16]。这一阶段的自动问答系统一般需要预先构建领域词典，重度依赖于大量由专家预先制定的复杂规则，并且对问题类型进行了严格的限制，只能回答非常简单的问题。它们使用的数据库一般为领域相关的自定义数据结构，也没有通用的数据库查询语言，很难迁移到其他数据库。

随着面向对象的数据库系统和关系型数据库技术的发展，基于数据库管理系统 (Database management system, DBMS) 的自动问答系统出现了，DBMS 优化了数据库存储与查询方式，使得问答系统可以应用于更复杂的领域。如 LADDER^[17] 采用三层架构：自然语言访问层、智能数据访问层和文件访问管理层，实现了通过自然语言访问分布式数据库的功能。然而由于语法解析器 (Semantic parser) 依赖于特定领域的语义语法信息，因此移植性较差。为了解决这一问题，上世纪 80 年代出现了通用自然语言接口 (Transportable natural language interfaces, transportable NLIs)，这类研究试图实现通用的句法语法解析器和通用的数据库接口，以提高问答系统在不同领域、数据库、数据库管理系统之间的可移植性，如 Chat-80^[18]、ASK^[19]、TEAM^[20]。从 90 年代开始，NLDBs 进入商业化，许多公司开发了商用自动问答系统，如 IBM^[21]、AT&T^[22]。进入本世纪以来，针对准确率的提升^[11, 23]、交互性体验的改善^[24]、非英语的自动问答^[25-26] 等方面，有一些学者展开了研究。

一般而言，与形式化查询语言、图形界面、基于

表单的查询接口等相比，NLDBs 的用户不需要进行形式化语言、表单字段的预学习，而且采用自然语言作为输入，更容易表达用户意图。然而，在可用性和通用性方面，NLDBs 仍然存在一些不足：一方面，关系型数据库与表现力受限的 SQL (Structural query language) 语法限制了数据存储方式和问题内容，因此不适用于大数据 (往往包含大量非结构化数据) 的处理需求，因为“受限的自然语言不再是自然语言 (A subset of natural language is no longer a natural language)^[5]”；另一方面，在句法分析、语法分析方面的能力有限，一般通过对已有数据的简单查询和统计进行问题分析，在学习推理、处理复杂问题方面的表现欠佳。

3.2 第二阶段：以信息和信息检索技术为中心

第二阶段是以信息和信息检索技术为中心的自动问答系统，这类系统从上世纪九十年代起取代 NLDBs 成为了业界和学术界新的关注点。然而早在上世纪六十年代，就已出现了基于文本的自动问答系统，但当时只能处理 10 万到 50 万个词的小规模文本^[4]。

第一个开放领域问答系统 MURAX^[27] 出现于 1993 年，它基于信息检索技术从葛罗里学术百科全书中获得可能包含了答案的文本。同年诞生的还有 Start^[28-29]，第一个基于 Web 的自动问答系统；以及 FAQ (Frequently asked questions) finder^[30]，一个基于 FAQ 数据集的自动问答系统。

TREC (Text retrieval conference) 是文本检索领域最权威的国际评测会议，由美国国家标准技术局 (National Institute of Standards and Technology, NIST) 主持。从 1999 年到 2007 年，TREC 开展了面向英文的开放领域自动问答评测 (TREC question answering track) (详情见第 7.1 节)。TREC 的举办极大地促进了自动问答领域的发展，涌现了一些优秀的问答系统，如集成了许多信息检索和 NLP (Natural language processing) 技术的 Webclopedia^[31]，基于最大熵分类方法进行问题、答案类型预测与命名实体识别的 IBM 统计问答系统^[32]。

此外，还有一些代表性问答系统，如基于搜索引擎 (Google) 的 MULDER^[14, 33]、面向大规模数据集的问答系统^[34]、采用命名实体识别技术的问答系统^[35]、基于关键词语义相似度和 FAQ 库的问答系统^[36-37] 等。

基于信息检索技术的自动问答系统首先从海量信息中检索与问题相关的信息子集，然后从子集中抽取准确的答案。与 NLDBs 相比，这类自动问答系统可以处理更灵活的数据格式，在非结构化大数据

的应用上占有绝对的优势, 尤其是在互联网上海量不断增长的用户生成内容 (User-generate content, UGC) 的利用方面, 取得了突破性的成果. 然而, 信息检索技术一般通过提取问题关键词的逻辑组合表达检索需求, 以关键词为基础进行信息索引、匹配与抽取, 这种做法尽管简单易行, 但是停留在浅层语义分析, 普遍缺乏对信息所蕴含知识的深度理解.

3.3 第三阶段: 以知识和知识自动化为中心

近年来, 随着大数据和知识管理、智能化实时采集、深度分析、个人携带智能设备与移动互联网等相关技术与应用取得了突飞猛进的进展, 问答系统进入了以知识自动化为中心的新阶段. 信息, 受制于表现形式, 是一种素材和起步的讯息. 而知识是不囿于形式的人的认知结果, 是解释、运用和推广信息的途径. 将问答系统从基于文本关键词的层面, 提升到基于知识的层面, 实现智能化的知识机器人, 已成为问答系统未来的发展趋势与目标. 在下一节中, 将展开详细描述.

4 基于知识的自动问答系统

知识表示是贯穿自动问答知识管理生命周期的重要元素, 因此, 下面我们将按照知识表示方式对问答系统进行分类分析. 需要指出的是, 在大部分问答系统中, 知识都不止一种单一的表示方法. 本文依据的是最主要或最具特色的知识表示方法. 根据知识表达方式的不同, 实现各类系统所需的关键技术也有所差异. 下文会一一阐述.

4.1 基于产生式规则 (Production Rule) 的问答系统

4.1.1 描述

产生式规则常用于表示具有因果关系的知识, 其基本形式是 IF a THEN b , 其中 a 代表一组前提或状态, b 代表若干结论或动作. 其含义是如果前提 a 成立, 即可得到结论 b .

4.1.2 代表性系统

产生式规则是概括经验知识的优先采用的方法. 在 TREC-10 测试中, 准确率最高的问答系统^[38] 采用的就是基于规则的方法. 举例而言, 对于出生年份的问题, 可以通过以下规则抽取答案:

```
IF      问题形式为 When (What year) born
      且文本片段形式为  $X(Y-Z)$ 
      且  $X$  为首字母大写的单词
      且  $Y$  和  $Z$  为四位数字
THEN    $Y$  是答案
```

4.1.3 关键技术

在基于产生式规则的问答系统中, 如何进行规

则的自动发现与更新、检验与评估是最关键的问题. 究其原因, 由人工维护的产生式规则需要高水平的、具备丰富的领域知识的专家. 在问答系统所应用的领域较为狭窄时, 这有可能得到满足. 然而, 随着问答系统涉及知识的广度和深度不断提高, 依赖于专家知识对管理规则的难度也大为提高. 这使得开发具有自学习能力的系统逐渐成为研究重点.

Lin 等^[39] 提出了一种从文本中发现推理规则的非监督学习方法. 它基于哈里斯 (Harris) 的分布假设 (Distributional hypothesis), 该假设认为共现率越高的词, 相似度越高, 在由文档集生成的依赖树中, 搜索连接到相同词集 (Word set) 的路径, 这些路径所代表的句子模式具有相似的含义. 例如, 通过该方法可以发现 “ a writes b ” 和 “ a is the author of b ” 是相似的句子模式. Mollá^[40] 将问题和答案句子表示为图, 采用图处理方法自动学习问答规则. 如图 4 所示为一个图结构表示的产生式规则, 其含义为年份的问题, 可以通过以下规则抽取答案:

```
IF      问题包含结构  $X(\text{obj}) - Y(\text{det}) - \text{what}$ 
      且文本片段与问题重叠的子图包含了 answer 节点
THEN   answer 节点的子节点是答案
```

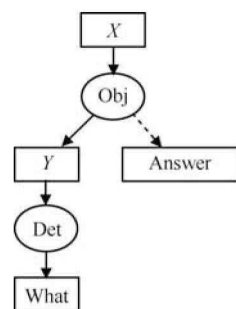


图 4 图结构的问答规则^[40]

Fig. 4 An example QA rule represented as a graph^[40]

4.2 基于逻辑表示 (Logical Representation) 的问答系统

4.2.1 描述

逻辑表示是知识表示的主要方法之一, 以逻辑公式来描述对象、性质、状况和关系. 例如 “中国的首都是北京” 可以描述为: Capital (China, Beijing). 这里 “Capital” 是谓词, “China” 表示中国, “Beijing” 表示北京. 在逻辑表示方法中, 知识可以看成一组逻辑公式的集合, 通过增加、删除或修改逻辑公式对知识进行更新.

4.2.2 代表性系统

Moldovan 等^[41] 首先将问题和答案表示为逻辑形式 (Logic form, LF), 然后基于 WordNet 构建词汇链, 将问题中的概念和答案中的概念进行关联, 最

后通过逻辑证明器 (Logic prover) 排序、抽取和验证答案. 具体而言, 逻辑形式是基于语法解析和深层语义形式之间的中间步骤, 在逻辑形式中体现了基于语法的关系, 如主语、宾语、介词短语、复合名词、形容词/状语等.

图 3 为一个逻辑表示的示例. 句子中的单词和逻辑表示中的谓词是一一映射的关系, 谓词命名由单词本身和词性标记构成, 名词还会有一个参数作为标识符, 可以作为其他谓词的参数. 例如, 单词 “selling” 被映射为谓词 “selling_NN(x1)”, 其中词性标记 “NN” 表示 “单数或不可数名词 (Noun, singular or mass)”, 参数 “x1” 为标识符. 谓词 “of_IN(x1, x6)” 表示从属连词 “of” 连接了 “x1” 和 “x6”. 由于动词谓词的固定位置分配机制, 使得逻辑证明器可以很容易分辨动作的主语和对象.

Tang 等^[42] 利用归纳逻辑编程 (Inductive logic programming, ILP) 中的多子句构造器 (Multiple clause constructors) 将多种不同的学习策略融合在一个统一的假设检验框架中. 相比单一的学习方法, 在样本较小的领域问答系统中取得了更好的结果.

4.2.3 关键技术

逻辑具有两个重要的相互关联的部分: 公理系统和演绎结构. 前者说明什么样的关系和蕴含可以形式化, 后者是推理规则集合, 它能从公理集合中推导出定理. 在一个完整的逻辑架构下, 所有的演绎推理都保证正确. 因此, 如何将自然语言自动地转换为逻辑表示, 是这类问答系统的关键技术之一.

此外, 由于逻辑不擅长表示不确定性知识和启发性知识, 而在自然语言的分析处理中, 不可避免的一点就是语义不确定性. Zadeh^[43] 提出的模糊集合

论、模糊逻辑、模糊推理等, 在自然语言的知识表示方面作出了巨大的贡献, 也为基于逻辑的问答系统提出了新的发展方向.

4.3 基于框架 (Frame) 表示的问答系统

4.3.1 描述

框架是把某一特殊事件或对象的所有知识存储在一起的一种复杂数据结构. 每个框架由若干个槽 (Slot) 构成, 每个槽都由特定的值来填充, 值本身可以是框架, 也可以指向其他框架.

4.3.2 代表性系统

Clark 等^[44] 介绍了一个面向特定领域的问答系统, 该领域的每一个概念 (Concept) 都是由框架来表示. 图 5 所示为针对 Binding-event 这一概念的框架表示. 当针对 Binding-event 的某个子事件查询请求发生时, 系统从 Binding-event 框架以及其他相关框架中抽取信息. 每一个被发掘出来的子事件被称为一个实例 (Instance), 根据实例之间的关系最终形成一个实例图 (Instance graph), 用于生成答案.

Barker 等^[45] 提出了一个化学领域的自动问答系统, 其知识库应用的也是基于框架的知识表示. 每个框架代表一种类 (Class) 或者一个实例 (Instance), 类被组织成半格 (Semi-lattice). 框架拥有多个槽, 槽由相应槽值的定理 (Axiom) 来填充. 这些定理拥有陈述式的或者过程式的语义, 从而使得答案推理得以进行. 该问答系统知识库的核心知识结构有两个: 化学名词和定律 (Chemistry terms and laws). 前者是化学名词的概念定义; 后者是解决问题的方法以及何时用何种方式解决该问题的说明.

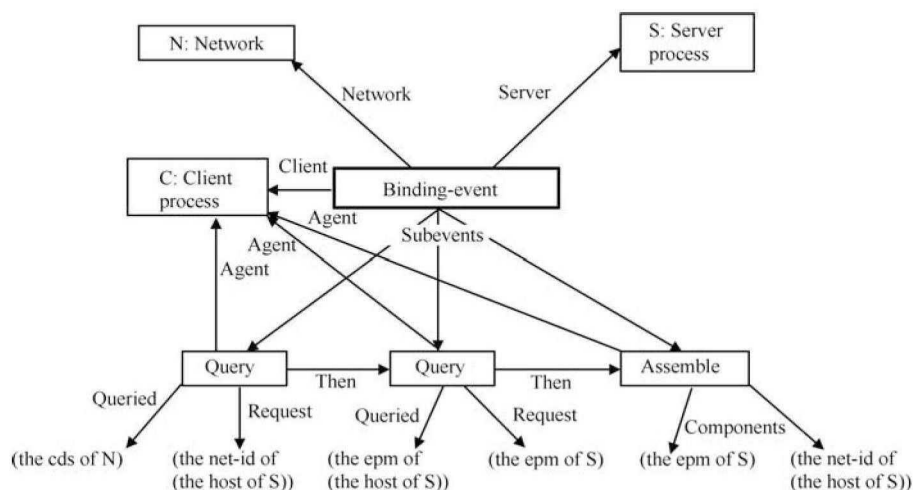


图 5 概念 Binding-event 的框架表示^[44]

Fig. 5 The frame of Binding-event^[44]

汉语框架网 (Chinese frame net, CFN)^[46] 是一个以基于 Fillmore 的框架语义学 (Frame semantics) 为理论基础、以英文 FrameNet 为参考、以汉语语料事实为依据的供计算机使用的汉语词汇词义数据库, 由框架库、句子库和词元库三部分组成. 王智强等^[47] 以 CFN 为依据, 在句子级框架语义分析基础上, 利用框架关系、有定零形式线索和篇章框架关系, 提出了一种篇章框架语义分析方法, 用于答案句检索和答案抽取方法. 但由于 CFN 语义资源覆盖率低, 导致该方法所涉及的问答数据也比较有限.

4.3.3 关键技术

在基于框架表示的问答系统中, 高质量的框架结构是系统高性能的必要条件. 此外, 从列举的代表性系统可以看出, 一般而言, 框架表示是面向特定领域的知识表示方法, 其构建依赖于领域专家知识, 具有领域受限、静态等特征. 因此, 如何自动构建通用框架、实现对框架的评估与动态更新, 是这类问答系统要面对的关键问题.

4.4 基于语义网络 (Semantic Networks) 的问答系统

4.4.1 描述

语义网络是通过节点及带标记的边来表达知识的一种有向图, 一般节点代表概念, 边代表语义关系. 语义网络是知识表示中最重要方法之一, 是一种表达能力强而且灵活的知识表示方法. 因此, 许多知名的知识库都采用语义网络而构建, 如 Freebase 知识库^[48]、YAGO 知识库^[49-50] 等. 它们在自动问答领域得到了广泛的应用.

4.4.2 代表性系统

Jacana-freebase^[51] 是一个基于 Freebase 知识库的问答系统. 该系统将知识库看作是由知识条目相互连接构成的网络, 为了在知识网络中获得答案, 该系统在问题特征和 Freebase 表示的答案之间直接建立匹配模式, 并且基于 ClueWeb 建立知识条目关系与自然语言文本的映射关系, 以提高匹配准确率.

Yih 等^[52] 也提出了一个基于 Freebase 知识库^[48] 的问答系统. 图 6 所示为该系统表示电视节目 Family Guy 的子图. 图 6 中的结点代表实体, 两个实体间的关系由连线上的谓词表示.

该问答系统提出一种新的语义解析方法 (Semantic parsing), 将一个自然语言问题映射到逻辑形式的查询 (Logical form query), 接着在知识库中执行查询, 通过检索得到答案. 核心是将知识库和查询均以图的形态表示, 将语义解析简化成问题到查询图 (Query graphs) 的映射, 从而提高了检索效率. 图 7 即为针对上述电视节目 Family Guy 的一

个问题的查询图 (问题为: Who first voiced meg on Family Guy?).

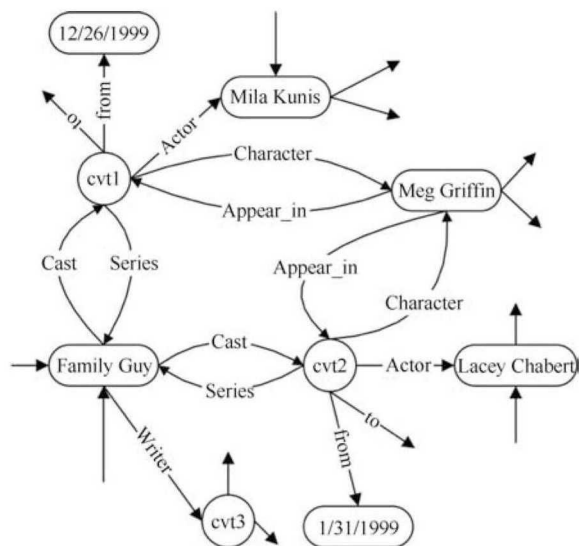


图 6 语义网络示例^[52]

Fig. 6 An example of semantic networks^[52]

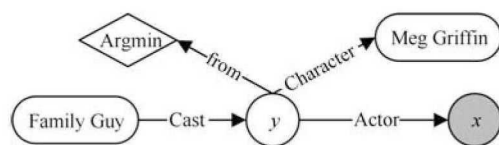


图 7 问题查询示例^[52]

Fig. 7 An example of logic form query^[52]

两个实体 Family Guy 以及 Meg Griffin 用圆角矩形结点表示, 圆形结点 y 表示存在某个实体代表角色选派关系 (比如演员和时间等), 阴影圆形结点 x 为答案结点, 菱形的结点 $argmin$ 将答案限制为最早出演这个角色的演员. 如果不考虑 $argmin$ 操作, 这个查询可以表示为

$$\lambda x. \exists y. \text{cast}(\text{FamilyGuy}, y) \wedge \text{actor}(y, x) \wedge \text{character}(y, \text{MegGriffin}) \quad (1)$$

在图 6 中执行该查询可得到 Lacey Chabert 和 Mila Kunis 两个答案, 再比较两个节点的 $from$ 属性, 其中最小者 Lacey Chabert 为最终的正确答案. 该问答系统在 Web Questions 数据集上的 F1 值达到 52.5%, 比以往最好方法提高 7.2%.

4.4.3 关键技术

知识库是基于语义网络的问答系统中最重要组成模块, 其关键技术是语义解析、知识库的构建与更新. 其中, 语义解析是对自然语言形式的问题进行分析和处理, 最终转换为对知识库的查询语句^[52].

知识库的构建以 YAGO 知识库为例. YAGO 知识库^[49] 基于维基百科而自动建立. 将维基百科

中的每篇文章自动映射为知识库中的一个实体,并根据维基百科中的分类信息对 WordNet 中的分类体系进行扩展. YAGO 将 Fact 表示为主体 (S), 谓语 (P) 和对象 (O) 的三元组, 和 RDF 数据模型兼容. 每一个实例 (SOP 三元组) 被赋予一个标识符, 标识符可以作为其他实例的主语或对象. 与 RDF 不同的是, YAGO 只能对知识库中已存在的 Fact 进行实例化. 因此, YAGO 避免了不可判定性的问题. YAGO 知识库诞生之后, 在 QA 领域引起了广泛关注. 它的进阶版本 YAGO2 知识库^[50] 用陈述规则集 (Declarative rules) 代替了原来的硬编码规则 (Hard-wired rules), 减少了代码量, 更容易扩展. 这些规则也是 SPO 三元组形式, 它们可以看作是 YAGO2 知识库的附加实例, 包括事实性规则、关联规则、替换规则、抽取规则. 此外, YAGO2 还基于 GeoNames, 为实例和实体扩展了时间、空间维度.

知识库所涵盖知识的广度和深度往往是有限的. 对于知识不完备的情况, West 等^[53] 提出了一种基于问答系统完善知识的方法, 利用基于 Web 的问答系统获取最新的相关文本、筛选候选答案, 根据知识库中的实体信息对候选答案进行排序, 最终得到“主语-关系对”相应的宾语.

4.5 基于本体 (Ontology) 的问答系统

4.5.1 描述

本体层级体系由 5 个基本建模元语构成, 分别为: 类、关系、函数、公理和实例. 基于本体的知识表示能够以一种显式、形式化的方式来表示语义, 提高异构系统之间的互操作性, 促进知识共享, 因此得到了广泛的应用.

4.5.2 代表性系统

Pythia^[54] 是一个基于 LexInfo¹ 的问答系统. 如图 8 所示, 它通过对 LexInfo 模型语言信息的丰富, 生成基于本体的语法条目. 用户输入问题经过该语法条目及本体无关的语法条目的处理, 构建为形式化查询语句. 它依赖于深度语言学分析, 可以处理比较复杂的自然语言问题, 如定量问题、最高级问题 (如: 哪座山是世界上的第二高峰?) 等.

周永梅等^[55] 对专业词语采用基于本体的概念相似度方法, 综合考虑语义重合度、语义距离、本体结构三方面的相似度进行计算:

$$\begin{aligned} Sim(W_1, W_2) = & \sqrt[3]{Sim(Contraction)} \times \\ & \sqrt[3]{Sim(Distance)} \times \sqrt[3]{Sim(Structure)} \quad (2) \\ \text{s.t. } & a + b + c = 1 \end{aligned}$$

¹http://lexinfo.net/, 一个语言学领域的本体模型

其中, 参数 a 、 b 、 c 用于调整语义重合度相似度 $Sim(Contraction)$ 、语义距离相似度 $Sim(Distance)$ 、本体结构相似度 $Sim(Structure)$ 三者的权重.

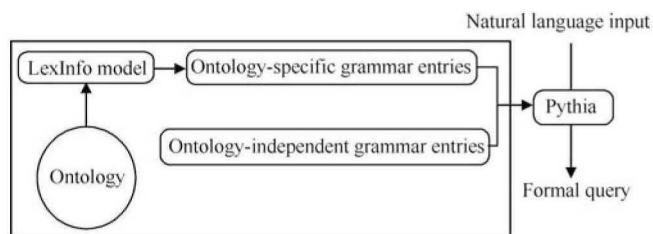


图 8 基于本体的问答系统: Pythia^[54]

Fig. 8 An example QA system based on ontology: Pythia^[54]

4.5.3 关键技术

本体的构建, 是实现这类问答系统的关键技术.

本体可以自顶向下或自底向上而建立. 目前主要的本体构建方法有 IDEF5、骨架法、TOVE 法、METHONTOLOGY 法、KACTUS 法、七步法等^[56]. 这些方法没有绝对的优劣, 最佳的方法需要综合领域特点、应用软件及扩展性而确定. 然而, 本体的开发和完善是一个反复叠加的过程, 需要领域专家的持续参与. 为了加快领域本体建设速度、并减轻对领域专家的依赖, 魏顺平等^[57] 提出了一种基于文本挖掘的领域本体半自动构建方法, 借助于文本挖掘技术自动提炼概念、概念定义和概念之间的关系, 并辅之以学科专家和知识工程师的参与.

4.6 基于神经网络的问答系统

4.6.1 描述

基于神经网络的方法一般将问题和答案都表示为语义向量, 从而将复杂的问答任务转化为在语义空间中, 计算输入问题和候选答案之间相似度的问题. 具有最高相似度的候选答案被认为是最终的答案. 由于神经网络自身的适应性和稳健性, 基于神经网络的方法得到了越来越多的关注.

4.6.2 代表性系统

QANTA^[58] 是一个基于递归神经网络的问答系统, 利用问题句子的依存解析树, 从词和短语层面对问题句子进行表示, 在机智问答等事实性问题的自动问答方面取得了很好的效果.

Yih 等^[59] 提出了一个基于卷积神经网络 (Convolutional neural network, CNN) 训练实体与关系的语义特征向量, 其主要思路是利用语义相似度建立实体题集与实体关系的匹配关系, 用于对单关系问题的自动问答. 该模型由词散列层、一个三层的神经网络和语义层构成, 如图 9 所示. 在词

散列层中, 为每个词 (如 “cat”), 添加首尾分隔符 (如 “#cat#”), 然后根据词中字母三元组 (如 “#ca, cat, at#”), 将词散列化为词向量, 作为神经网络的输入. 神经网络由卷积层、Max pooling 层和前馈层构成. 首先, 利用词在句子中的局部上下文信息, 通过卷积层的处理得到一组局部上下文特征向量. 在局部上下文特征空间中, 语义相似的 N 元词组 (Word- n -grams) 将被映射为相似的向量. 考虑到句子的含义往往取决于少数几个关键词, 采用 Max pooling 层抽取最显著的局部特征向量, 从而得到固定长度的全局特征向量. 最后, 通过前馈神经网络, 从全局特征向量中抽取非线性语义特征. 语义层实际上也是一个神经网络, 其激活函数为 Softmax, 计算语义特征相似度, 从而得到问题实体指代 (Entity mentions) 和知识库实体相似度, 以及关系模式和知识库关系的相似度.

在大部分基于神经网络的问答系统中, 对于与不同答案的相似度计算, 问题都被表示为相同的向量. 然而, 有研究者^[60] 认为应该根据候选答案的焦点, 基于注意力模型 (Attention model, AM) 动态计算问题的向量表示. 如图 10 所示. 系统将问题中的每个词表示为词向量 (Word embedding), 并经过长短时记忆模型 (Long short-term memory, LSTM) 获得问题的句向量; 将答案从答案实体 a_e 、答案关系 a_r 、答案类型 a_t 、答案上下文 a_c 等方面基于知识库分别表示为答案向量 e_e 、 e_r 、 e_t 、 e_c ; 最后以句向量和答案向量为输入, 经过注意力模型, 获得

最终的问题向量表示.

4.6.3 关键技术

迄今为止, 神经网络技术的发展经过了两次飞跃: 1) 1974 年, Werbos^[61] 的博士论文提出了多层感知机和反向传播 (Back propagation) 算法的提出; 2) 2006 年, Hinton 等^[62] 提出了快速训练深度信任网络 (Deep belief network) 的方法, 后者标志着神经网络进入深度学习阶段. 2016 年初, AlphaGo^[63] 的成功是深度学习研究的又一个标志性事件. 在自动问答领域, 卷积神经网络、递归神经网络、Max-out 网络、注意力模型等最流行的深度

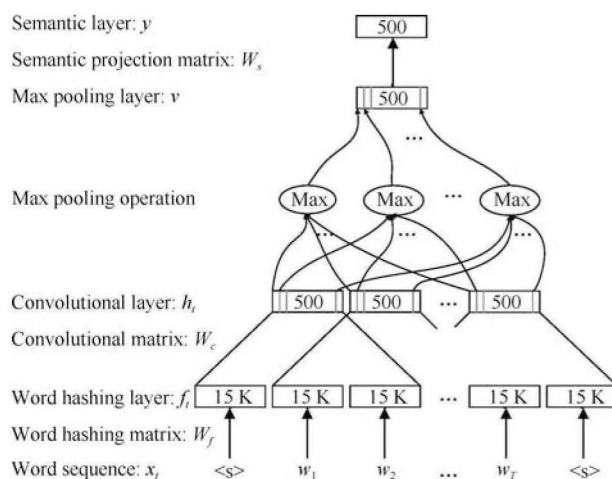


图 9 卷积神经网络模型^[59]

Fig. 9 An example CNN model^[59]

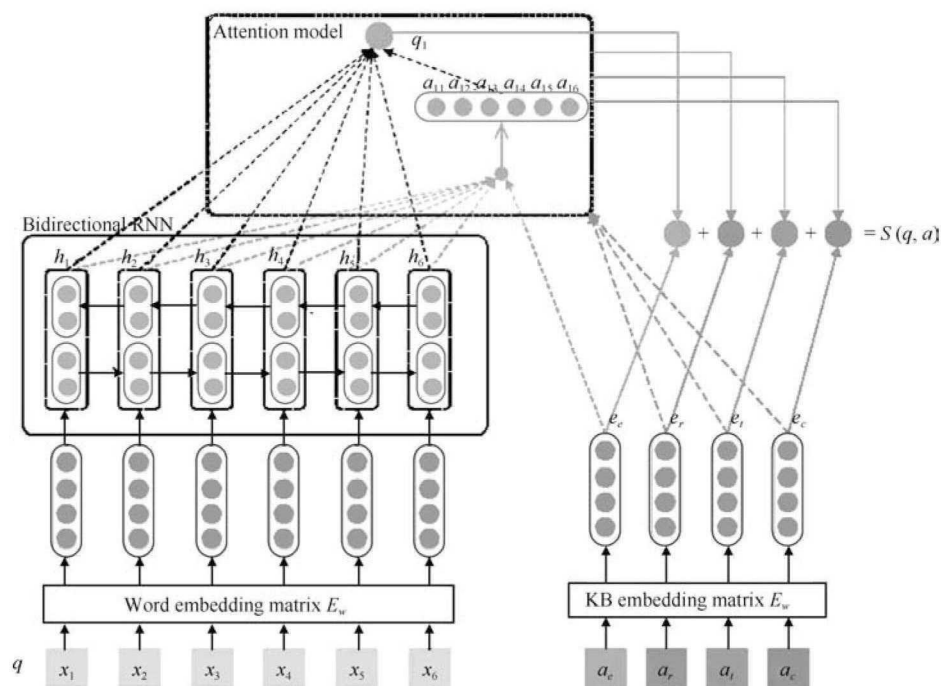


图 10 注意力模型^[60]

Fig. 10 An example attention model^[60]

学习模型都已投入应用, 其实现效果也得到了广泛认可. 然而, 如何构建适合于实际应用场景的神经网络, 尤其是自适应调整网络参数甚至结构, 依然是这类问答系统面临的关键问题和主要挑战.

此外, 一些研究通过以深度学习为代表的表示学习, 将知识库中的实体、关系以及问题文本转换为一个低维语义空间中的数值向量, 进而利用数值计算, 直接匹配与问题语义最相似的答案. 从目前的结果看, 基于表示学习的知识库问答系统在性能上已经超过传统知识库问答方法^[64].

5 典型英文系统应用

5.1 START

START 是世界上第一个基于 Web 的英文自动问答系统, 于 1993 年开始上线并运营至今. 不同于搜索引擎这样传统的信息检索系统, START 目标是为用户反馈准确凝练的答案. START 目前可以回答有关地理、人物、艺术与娱乐、历史和文化等事实型 (Factoid) 和定义型 (Definition) 问题.

START^[28-29] 采用基于知识库和基于信息检索的混合模式. 如果用户提出的问题属于自有知识库的范畴, 则 START 直接返回精确答案. 反之, 系统通过问题分析得到关键字集, 然后将关键字集提交给信息检索模块 (此处是搜索引擎) 来查找相关的网页或者文档, 最后答案抽取模块从相关的网页和文档中提取出答案. START 所利用的核心技术称为自然语言注释 (Natural language annotations), 该技术利用自然语言语句以及短语-注释 (Phrases-

annotations) 作为不同分割粒度下信息片段的内容描述. 注释指的是用于描述不同信息片段所包含内容的自然语言句子和短语, START 将这些注释进行语法分析并将其语法结构存储成三元组的形式, 问答阶段将问句的三元组与注释三元组进行匹配, 若匹配成功, 则返回注释三元组所对应的信息片段作为答案.

5.2 AnswerBus

AnswerBus 是由美国密歇根大学开发的一款面向开放领域句子等级的互联网信息检索自动问答系统 (Sentence level web information retrieval QA)^[65]. 可以接受用户用英语、西班牙语、德语、意大利语、葡萄牙语的提问 (若用非英语提问, 则第一步先将问题翻译成英语), 最终用英语返回答案. 当输入问句后, AnswerBus 利用 5 个搜索引擎 (Google、Yahoo、WiseNut、AltaVista 以及 Yahoo News) 检索与用户所提问题相关的互联网页面, 之后 AnswerBus 从页面中抽取可能包含问题答案的若干个语句并对其进行打分, 最后返回分数靠前的语句作为问题答案.

5.3 Evi

Evi 是亚马逊旗下的一款面向开放领域的问答软件, 能够回答用户关于书籍、音乐、电影、历史、人物以及地点等方面的自然语言提问. Evi 本身拥有一个结构化数据库, 此外还能利用来自 Yelp 和第三方网站的数据和 API 进行查询, 从而为用户提供精准全面的答案. 其基本架构^[66] 为:

表 1 典型英文自动问答系统
Table 1 A list of English QA systems

问答系统	问题类型	数据源	答案形式	相关技术
START	What, Who, When 等开头的事实型或者定义型问题	START KB、Internet Public Library	一句话或者一段文字	自然语言注释 (Natural language annotations)、句子级别的自然语言处理 (Sentence-level NLP)
AnswerBus	开放领域问答系统	互联网	按照相关程度返回若干个可能的候选答案语句	命名实体抽取 (Named entities extraction)
Evi	开放领域问答系统	自有结构化知识库 (Structured knowledge base), Yelp 和第三方网站的数据和 API	类似人类语言风格的简明回答	知识表示
AskJeeves	开放领域问答系统	自有问答数据库、互联网	文本、文档链接以及内容摘要	自然语言检索技术 (NLP)、人工操作目录索引
Wolfram Alpha	开放领域问答系统	内置的结构化知识库	包含答案信息的各种数据和图表	计算知识引擎 (Computational knowledge)
Watson	开放领域问答系统	定义了自身的知识框架, 并从海量结构化和半结构化资料中抽取知识构建知识体系	针对用户提问的精准回答	统计机器学习、句法分析、主题分析、信息抽取、知识库集成和知识推理

Evi 拥有一个包含世界上各类知识的庞大并且不断拓展的知识库(知识库中包含常识型知识、事实型知识以及词汇型知识), 知识库中的各类条目均以结构化方式存储. 当收到用户用自然语言的提问后, 自然语言转换系统将用户的提问转换为一个与语言无关的查询, 查询结果可以直接从系统本身的数据库得到, 也可以通过推理系统从外部资源中生成.

5.4 AskJeeves

一款人工操作目录索引的搜索引擎, 同时也是一款问答系统. 对于用户的自然语言提问, AskJeeves 首先利用语义分析技术对用户的查询问题和数据库中已有的问题做精确匹配(其数据库存储了超过 1000 万个问题的答案), 若该问题存在, 则直接返回答案, 否则采用搜索引擎的方式为用户提供一系列外部文档链接及其内容摘要.

该问答系统中的问题分析模块是依赖手工完成的, 为了能够正确理解用户的查询意图, AskJeeves 雇佣了数百名专职人员构造问题模板. 系统的问题模板虽然能够细化和明确用户的需求, 但由于需要人工产生和维护, 工作量非常大.

5.5 Wolfram Alpha

由 Wolfram Research 公司推出的一款在线自动问答系统. 其数据来源包括学术网站和出版物、商业网站和公司、科学机构(例如中央情报局出版物《世界概况》、道琼斯公司、美国地质调查局等), 从而发掘、建立起一个异常庞大的且经过组织的数据库. 此外, 基于 Wolfram 早期产品 Mathematica (其囊括了计算机代数、符号和数值计算、可视化和统计功能), Wolfram Alpha 可以完成数学、统计学、物理、经济学、音乐、体育和天气等各个领域的查询、计算和分析. 例如, 输入“China GDP”, 问答系统返回的不是一堆网页, 而是直观的数据和图表. 包括: 中国 GDP 最新情况, 从 1970 年至今的中国 GDP 增长情况、中国通货膨胀率、失业人口率等.

5.6 Watson

Watson 通过加载数以百万计的文件, 包括字典、百科全书、网页主题分类、宗教典籍、小说、戏剧和其他资料, 来构建它的知识体系. 在 Watson 分析问题并确定最佳答案的过程中, 利用其核心的 IBM DeepQA 技术来生成假设、收集大量证据并进行分析和评估.

在问答阶段, Watson 问答系统^[67-70]通过使用多种不同的算法来搜索问题的潜在候选答案, 并为每个候选答案进行打分(这里采用假设生成技术(Hypothesis generation)将散落在各处的知识片段连接起来, 进行推理、分析、对比、归纳、总结和论

证, 从而获取决策的证据), 同时为每个候选答案收集其他支持材料, 利用自然语言处理对所搜集到的相关材料进行深度评估. 当越来越多算法运算得到的结果聚焦于某一个答案时, 这个答案的可信度就会越高. 此外, Watson 还会衡量每个候选答案的支持证据, 进一步确认最佳的答案选择及其可信度. 当该答案的可信度达到一定的水平时, Watson 就会将其作为最佳答案呈现出来.

6 典型中文系统应用

6.1 微软小冰

微软小冰是一款智能聊天娱乐机器人. 除了人机互动聊天之外, 还兼具提醒、天气、星座、交通指南、餐饮点评等实用功能. 微软小冰集合了中国近 7 亿网民多年来积累的全部公开聊天记录, 并从中精炼出 1500 万条语料库, 基于数据挖掘与智能搜索, 通过理解对话的语境和语义而实现人机问答的自然交互. 在 2016 年 8 月推出的第四代产品中, 在自然语义的长时学习和理解的基础上, 微软小冰又增加了情感计算、自主知识学习(Doc chat)以及意图对接对话引擎(Intention-X engine), 从而实现了更精准的用户意图识别和互动交流.

数据显示, 小冰自 2014 年 6 月发布至今已与 4200 万用户进行了逾 220 亿次对话, 平均对话轮数达到了创纪录的 23 轮, 语音自然度达 4.32(介乎于“自然”和“比较自然”两个分值之间), 微软小冰已经成为全球科技史上最大规模的一次图灵测试.

6.2 京东 JIMI

京东 JIMI 是京东自主研发的一款基于大数据的智能客服问答机器人, 涵盖售前咨询、售后服务等电子商务各个环节, 旨在为用户提供更好的导购和咨询体验.

JIMI 智能客服的核心技术是融合了深度神经网络的自然语言处理体系. 在用户输入问题后, 首先利用意图识别和命名实体识别对用户提问进行分析, 了解客户真实意图并抽取用户问题中的关键信息(如商品属性、订单状态等). 面对一些语义模糊的问题, 则会采用深度神经网络模型对用户意图进行识别, 相对于传统的分类算法, 深度学习识别准确率提升了 4% 至 5%. 在确定问题种类后, JIMI 会快速准确地从海量候选集里进行答案匹配以及候选答案抽取和排序, 最后反馈给用户最佳答案和建议.

此外, 通过与用户的交互以及对互联网数据挖掘开展的学习, JIMI 能够自动完善自身的知识结构, 并随着时间推移变得越来越智能和强大.

6.3 度秘

度秘是百度在 2015 年推出的一款旨在为用户提供秘书化搜索服务的机器人助理。除具备基础问答功能外,度秘还在广泛索引真实世界服务和信息的基础上,依托其强大的搜索及智能交互技术,以自然问答方式为用户提供电影、美食以及生活服务等各种推荐服务,还可以通过多轮对话不断揣摩用户的真实所想。

对每一项接入的服务,百度后台通过全网数据挖掘和机器学习的方式,为服务贴上标签,建立丰富的索引维度,从而让用户在提出个性化的需求时能够及时为其推送精准的服务,实现用户需求与商家服务的精确匹配。

6.4 小 i 机器人

小 i 机器人是目前国内比较成熟的一款智能客服助理,能够以拟人化方式与用户进行实时交互,实现智能客服咨询和产品营销推广等功能。

小 i 机器人强大的交互能力得益于其积累的完整的知识模型和庞大的中文知识库。小 i 拥有全球最大的领域知识库和语言语义库,其包括:上百万规模的词类、通用对话库(包括 16 大类,135 个小类,近二十万语料),以及数十个专业领域知识本体基类以及海量的通用语言知识库和对话日志,涉及电信、金融、政府、电子商务等诸多领域。

在自动问答方面,小 i 整合了自然语言处理和人机交互技术,不断适应客户的提问方式;提供基于语义的智能搜索和高级搜索,满足用户不同的搜索需求;利用基于语义模型的句子相似度计算技术推算关联知识;提供可自定义的知识对比方案,提供对知识不同纬度的对比。此外,小 i 综合了语言学、统计学,具有场景反问和上下文关联处理,其智能交互算法经过每月 10 亿多次交互的验证,确保良好的用户

体验和应答准确率,支持高并发和平行扩展服务能力。

6.5 阿里小蜜

阿里小蜜是阿里巴巴集团在 2015 年 7 月推出的一款人工智能购物助理虚拟机器人。依托于客户真实的需求,通过智能 + 人工的方式为客户提供购物体验服务。

阿里小蜜集合了阿里巴巴集团淘宝网、天猫商城、支付宝等平台日常使用规范、交易规则、平台公告等信息,凭借阿里在大数据、自然语义分析、机器学习方面技术积累,精炼为几千万条真实而有趣,并且实用的语料库(此后每天净增 0.1%),通过理解对话的语境与语义,实现了超越简单人机问答的自然交互,最终成长为用户的私人购物助理。

区别于传统的基于自定义知识库的问答型智能客服产品,阿里小蜜采用人工智能 + 知识图谱的方式。目前人机对话语义意图的精确匹配率提升到了 93%,服务满意度比传统的自助服务提升了一倍。

7 评估方法

通常意义下的文本检索输入的查询是关键词,返回的是相关文本。而对于自动问答,输入的查询是自然语言提出的问题,希望返回的不是整篇文本,而是在给定的语料库中检索并返回问题的直接答案。

在过去的问答系统研究中,所有的研究都是在各自假设下进行,加上系统复杂度高,不同单位的研究成果很难拿来作客观的评估与比较。除此之外,这类系统的评估是非常消耗人力的,事前的准备包含要产生足够多且合适的问题题目,同时每一题可能出现的答案都必须以人工方式从比赛语料中挑选出来。以上所述对问答系统的研究发展非常不利。有鉴于此,由单一组织举办、多个研究单位共同参与的问答系统比赛应运而生。

表 2 典型中文自动问答系统
Table 2 A list of Chinese QA systems

问答系统	问题类型	数据源	答案形式	相关技术
微软小冰	日常聊天伴侣	海量网民聊天语料库	拟人化回答	情感计算、自主知识学习、意图对接对话引擎
京东 JIMI	电商售前、售后咨询	自有问答库	文本	深度神经网络、意图识别、命名实体识别
小 i 机器人	业务咨询	语言知识库以及业务知识库	文本	知识表示、本体理论、分领域的语义网络
度秘	生活服务类咨询	互联网	服务推荐(如餐厅、影院)	全网数据挖掘和聚合
阿里小蜜	导购咨询	自有语料库	文本、语音、网页链接等	知识图谱、语义理解、个性化推荐、深度学习

7.1 The TREC Question Answering Track

TREC 是文本检索领域最权威的国际评测会议, 由美国国家标准技术局 (NIST) 主持. 该会议开始于 1992 年, 每年举办一届, 迄今已举办了 20 多届. TREC 会议的评测项目一开始只有随机检索 (Ad hoc) 和分流 (Routing) 两项, 之后随着文本检索技术的不断发展, 逐渐转向自动问答领域. 早在 1999 年, TREC 就首次开展了面向英文的开放领域自动问答评测 (TREC question answering track). 自此以后, QA track 一直都是该领域关注度最高、参与机构最多的评测项目之一.

从 1999 年的 TREC-8 到 2007 年最后一届 TREC-2007, QA track 已经成功举办了 9 届. 每届的评测任务和评测指标都略有调整. 下面以最后一届 TREC 2007 QA track^[71-72] 为例, 介绍一下该评测的任务及评测指标.

7.1.1 问题类型

QA track 的测试集包含若干个问题序列 (Question series), 每个问题序列也称作一个目标 (Target). 每一个问题序列包含若干事实型问题 (Factoid questions)、1~2 个列表型问题 (List questions) 以及一个称之为“其他 (Other)”的问题.

QA track 2007 总共包含 70 个问题序列. 其中包括 19 个人物问题, 17 个组织问题, 15 个事件问题以及 19 个物品问题.

7.1.2 问答文档集 (Document Set)

测试集中所有问题的答案都要求从会议主办方提供的官方文档集中提取 (如果某个问题在文档集中确实没有相应答案, 则返回 “NIL”).

QA track 2007 官方文档集包含两部分, 分别为 AQUAINT-2 文档集以及 Blog06 文档集. AQUAINT-2 文档集包含了从 2004 年 10 月到 2006 年 3 月期间, 世界各大主流新闻媒体 (包括法新社、华盛顿邮报、纽约时报等) 的英文新闻报道; Blog06 文档集包含了 2005 年 12 月到 2006 年 2 月期间摘取的部分英文博客内容以及相应的评论数据.

7.1.3 评测标准

除了答案之外, QA track 也要求给出证据, 即答案所在的文档编号. 对于所有类型的问题, 其分数范围均在 [0.0, 1.0] 之间.

1) Factoid questions (事实型问题): 对于每个事实型问题, 都由两位人工评委独立给出分数. 当两位评委意见不一致时, 由第三位评委给出最终得分.

评委根据如下标准从 [0.0, 1.0] 给出分数:

a) Incorrect: 回答错误或者答非所问;

b) Unsupported: 不支持 (答案正确, 但给出的证据不支持);

c) Non-exact: 不确切 (与正确答案有交集);

d) Locally correct: 局部正确 (文档集中存在该答案, 但该答案并非是整个文档集中的最佳答案);

e) Globally correct: 全局正确.

2) 列表型问题: 此类问题返回的是一个实例. 对于列表类问题, TREC 提出了实例查准率 IP (Instance precision)、实例查全率 IR (Instance recall) 和 F 值等指标. 对于某个列表类问题, 实例查准率 IP 指的是问答系统给出的正确答案占给出的全部答案的比例, 实例查全率 IR 指的是问答系统给出的正确答案的数量占所有正确答案的比例. 以 S 代表全部正确答案的数量, D 代表问答系统给出的正确答案数量, N 代表问答系统给出的总的答案数量, 则 $IP = D/N$, $IR = D/S$. F 值为实例查准率和查全率的调和平均值, 即: $F = 2 \times IP \times IR / (IP + IR)$.

3) 其他问题: 其他问题的前身是定义问题 (Definition questions). TREC 2003 之后, 为了增加比赛难度, 将其改为现类型. 由于需要返回关于问题目标的其他有用或者有趣的信息, 而且不能包含事实型、列表型问题中已经回答过的内容, 所以评判主观性较强且难度较大. 目前的评判方法为

a) 首先由评委从官方文档集以及所有参赛系统提交的答案中汇总关于某个问题目标的全部有用信息, 构成标准答案集合. 然后, 评委对标准答案集合中的每一个条目根据其与该问题的相关程度, 打上 “Vital” 或者 “Okay” 的标签.

b) 若干位评委对问答系统返回答案的每一个条目进行评估, 标注出与标准答案集合中标签为 “Vital” 的条目类似的条目. 越多的评委认为该条目为 “Vital”, 则其对应的分数越高. 最后每个条目的分数都标准化在 [0.0, 1.0] 之间, 记做 Weight.

c) 最终分数按如下式子得到:

$$F(\beta) = \frac{(\beta^2 + 1) \times NP \times NR}{\beta^2 \times NP + NR} \quad (3)$$

其中, β 用于调整 NR (条目查全率) 和 NP (条目查准率) 的权重. NR 是问答系统返回的各个条目的 Weight 值总和与标准答案集合中所有条目 Weight 值总和的比值. NP 用返回条目的非空格字符数来近似计算, 记 length 为非空字符个数, allowance 为参考值 (如 100), 则

NP =

$$\begin{cases} 1, & \text{若 length} < \text{allowance} \\ 1 - \frac{\text{length} - \text{allowance}}{\text{length}}, & \text{否则} \end{cases} \quad (4)$$

问答系统的最终得分为上述三种类型问题分数的平

均值.

7.2 CLEF (Cross-language Education and Function)

CLEF² 是由欧盟资助的一个评测会议系列^[72-73], 其与 QA track 的最大区别是 CLEF 针对的是欧洲范围内跨语言检索问题的评测, 参赛语言高达 9 种. CLEF 与 QA track 类似, 也是要求问答系统针对某个问题返回精确回答以及支持该答案的文档. 在问题设置方面, CLEF 除事实性问题、定义问题外, 还包括推理 (Reason)、动机 (Purpose)、过程 (Procedure) 三类问题:

1) 推理问题: 要求回答某事发生的原因或者动机. 如 Why did a commission expert conduct an inspection visit to Uruguay?

2) 动机问题: 要求回答某事的目标或者目的. 如: What is the overall objective of the eco-label?

3) 过程问题: 要求回答完成某任务时需要进行的一系列具体步骤. 如: What is the common practice with shoots when packing them?

每个问题都由人工评委进行评判, 评判标准为

R (Right): 回答正确

W (Wrong): 回答错误

U (Unanswered): 没有回答

问答系统的最终得分为

$$F = \frac{n_R + \frac{n_U \times n_R}{n}}{n} \quad (5)$$

其中, n_R 表示回答正确的答案数量, n_U 表示未回答的答案个数, n 表示问题的总数.

7.3 NTCIR

NTCIR 评测会议³ 是由日本国立情报学研究所 (National Institute of Informatics, NII) 和日本学术振兴会 (Japan Society for Promotion of Science) 共同主办. 它的目的是建立一个日文标准测试集, 同时它还研究亚洲语言的跨语言检索问题, 比如日语、中文、朝鲜语等. 此外, 该会议还包含其他的评测项目, 如专利检索、问答检索、自动摘要等.

7.4 TREC 2015 LiveQA Track

从 2015 年开始, TREC 开始引入针对真实用户提问的实时问答系统评测项目 (LiveQA track)^[74]. 所有的评测问题都是 Yahoo Answers 上由用户提交但尚未得到解答的问题, 要求问答系统给出实时解答 (响应时间限制在一分钟之内, 系统也可以选择 not answer 某个问题, 返回 NULL). 与以往的 QA track

中常用到的事实型问题或者定义型问题不同, Yahoo Answers 用户的提问更加复杂和多样化, 例如观点类问题、建议类问题等. 例如, Can lazy eyes fix themselves? 以及 Is the ability to play an epic guitar solo attractive in a woman? 此外, 由于用户个体的语言风格不同, 致使该评测更加贴近实际, 难度也更大.

1) 训练数据集

由于第一次举办该评测, 没有现成的数据集. 为此, Yahoo Answers 提供了公开问答数据集. 除此之外, 比赛组织者还提供了 1000 个 Yahoo Answers 上的提问以及一个网络爬虫程序用以爬取问答对. 以上数据用于赛前对问答系统进行训练.

2) 评测方法评测期间, 评测系统每一分钟向参赛系统发送一个问题, 要求系统在一分钟之内返回答案, 并且要求答案字符数小于 1000.

2015 LiveQA track 中, 共有 1340 个实时问题提交给问答系统, 总共返回 27369 个有效答案 (2200 个答案被排除, 原因是超时或者没有返回答案). 问答系统平均响应时间为 21.35 秒.

最后由人工评委对问答系统返回的答案进行打分. 评分标准为:

0 分	unanswered
1 分	poor
2 分	fair
3 分	good
4 分	excellent

对于某问答系统, 其评价标准为

avgscore: 平均分

succ@i+: 分数在 i 分到 4 分的答案数量与全部问题数量的比值. 例如, succ@2+ 代表答案质量在 fair (2 分) 及以上的答案所占比率.

prec@i+: 分数在 i 分到 4 分的答案数量与该问答系统回答的问题总量的比值. 该指标用于评测问答系统的精度.

8 总结与趋势展望

本文回顾了自动问答系统的发展历史, 从知识自动化的角度出发, 将当前自动问答系统按照知识表示方法进行了分析和总结. 此外, 本文还分析和比较了实际投入应用的英文系统和中文系统, 介绍了自动问答系统的评测方法.

移动智能终端的广泛应用加速了用户对精准答案需求的提升, 同时也使得根据用户的个性化需求提供答案成为可能. 在自动问答系统发展的前两个阶段 (即, 数据库自然语言接口、以信息检索为中心

²CLEF Homepage: <http://clef.isti.cnr.it>

³NTCIR Homepage: <http://research.nii.ac.jp/ntcir>

的阶段), 其本质是对已有信息的检索, 因此回答问题的能力受限于信息的完备性. 而将信息升华为知识, 实现知识自动化才可以赋予自动问答系统像人类一样的思考能力. 虽然当前还没有自动问答系统完成达到用户理想的状态, 它已经在朝着智能化、个性化的知识机器人发展. 深度学习、计算语言学、知识工程等领域的研究成果, 将推动自动问答系统的急剧进步.

References

- 1 Wang Fei-Yue. Software-defined systems and knowledge automation: a parallel paradigm shift from newton to merton. *Acta Automatica Sinica*, 2013, **41**(1): 1–8
(王飞跃. 软件定义的系统与知识自动化: 从牛顿到默顿的平行升华. 自动化学报, 2013, **41**(1): 1–8)
- 2 Wang Fei-Yue. On future development of robotics: from industrial automation to knowledge automation. *Science & Technology Review*, 2015, **33**(21): 39–44
(王飞跃. 机器人的未来发展: 从工业自动化到知识自动化. 科技导报, 2015, **33**(21): 39–44)
- 3 Bidian C, Evans M M, Dalkir K. A holistic view of the knowledge life cycle: The Knowledge Management Cycle (KMC) model. *Electronic Journal of Knowledge Management*, 2014, **12**: 85–97
- 4 Simmons R. Answering English questions by computer: a survey. *Communications of the ACM*, 1965, **8**(1): 53–70
- 5 Androutsopoulos I, Ritchie G D, Thanisch P. Natural language interfaces to databases — an introduction. *Natural Language Engineering*, 1995, **1**(1): 29–81
- 6 Indurkha N, Damerau F J. *Handbook of Natural Language Processing* (Second Edition). Florida: CRC Press, 2010
- 7 Zheng Shi-Fu, Liu Ting, Qin Bing, Li Sheng. Overview of question-answering. *Journal of Chinese Information Processing*, 2002, **16**(6): 46–52
(郑实福, 刘挺, 秦兵, 李生. 自动问答综述. 中文信息学报, 2002, **16**(6): 46–52)
- 8 Tang Yong, Lin Lu-Xian, Luo Ye-Min, Pan Yan. Survey on information retrieval system based on question answering system. *Computer Applications*, 2008, **28**(11): 2745–2748
(汤庸, 林鹭贤, 罗烨敏, 潘炎. 基于自动问答系统的信息检索技术研究进展. 计算机应用, 2008, **28**(11): 2745–2748)
- 9 Bouziane A, Bouchiha D, Doumi N, Malki M. Question answering systems: survey and trends. *Procedia Computer Science*, 2015, **73**: 366–375
- 10 Mishra A, Jain S K. A survey on question answering systems with classification. *Journal of King Saud University-Computer and Information Sciences*, 2016, **28** (3): 345–361
- 11 Burger J, Cardie C, Chaudhri V, Gaizauskas R, Harabagiu S, Israel D, Jacquemin C, Lin C Y, Maiorano S, Miller G, Moldovan D, Ogden B, Prager J, Riloff E, Singhal A, Shrihari R, Strzalkowski T, Voorhees E M, Weischedel R. Issues, tasks and program structures to roadmap research in question & answering (Q&A). *Document Understanding Conferences Roadmapping Documents*, 2001. 1–35
- 12 Huang Chang-Ning. Thinking about DeepQA beating human champions. *Journal of Chinese Information Processing*, 2011, **25**(6): 21–25
(黄昌宁. 从 IBM 深度问答系统战胜顶尖人类选手所想到的. 中文信息学报, 2011, **25**(6): 21–25)
- 13 Mao Xian-Ling, Li Xiao-Ming. A survey on question and answering systems. *Journal of Frontiers of Computer Science and Technology*, 2012, **6**(3): 193–207
(毛先领, 李晓明. 问答系统研究综述. 计算机科学与探索, 2012, **6**(3): 193–207)
- 14 Cui Huan, Cai Dong-Feng, Miao Xue-Lei. Research on web-based Chinese question answering system and answer extraction. *Journal of Chinese Information Processing*, 2004, **18**(3): 24–31
(崔桓, 蔡东风, 苗雪雷. 基于网络的中文问答系统及信息抽取算法研究. 中文信息学报, 2004, **18**(3): 24–31)
- 15 Green B F, Wolf A K, Chomsky C, Laughery K. Baseball, an automatic question-answerer. In: *Proceedings of the Western Joint IRE-AIEE-ACM Computer Conference*. Los Angeles, California, USA: ACM, 1961. 219–224
- 16 Woods W A, Kaplan A M, Nash-Webber B. The lunar sciences natural language information system. *Journal of Neuroimmunology*, 1972, **174**(1–2): 32–38
- 17 Hendrix G G, Sacerdoti E D, Sagalowicz D, Slocum J. Developing a natural language interface to complex data. *ACM Transactions on Database Systems*, 1978, **3**(2): 105–147
- 18 Warren D H D, Pereira F C N. An efficient easily adaptable system for interpreting natural language queries. *Computational Linguistics*, 1982, **8**(3–4): 110–122
- 19 Thompson B H, Thompson F B. Introducing ask, a simple knowledgeable system. In: *Proceedings of the 1st Conference on Applied Natural Language Processing*. Santa Monica, USA: ACL, 1983. 17–24
- 20 Grosz B J, Appelt D E, Martin P A, Pereira F C N. Team: an experiment in the design of transportable natural-language interfaces. *Artificial Intelligence*, 1987, **32**(2): 173–243
- 21 Ott N. Aspects of the automatic generation of SQL statements in a natural language query interface. *Information Systems*, 1992, **17**(2): 147–159
- 22 Hindle D. An analogical parser for restricted domains. In: *Proceedings of the Workshop on Speech and Natural Language*. New York, USA: ACL, 1992. 150–154
- 23 Popescu A M, Armanasu A, Etzioni O, Ko D, Yates A. Modern natural language interfaces to databases: composing statistical parsing with semantic tractability. In: *Proceedings of the 20th international conference on Computational Linguistics*. Geneva, Switzerland: ACL, 2004. Article No. 141

- 24 Li F, Jagadish H V. Constructing an interactive natural language interface for relational databases. *Proceedings of the VLDB Endowment*, 2014, **8**(1): 73–84
- 25 Llopis M, Ferrández A. How to make a natural language interface to query databases accessible to everyone: an example. *Computer Standards and Interfaces*, 2013, **35**(5): 470–481
- 26 Wang S, Meng X F, Liu S. Nchiql: a Chinese natural language query system to databases. In: *Proceedings of the 1999 International Symposium on Database Applications in Non-Traditional Environments*. Kyoto, Japan: IEEE, 1999.
- 27 Kupiec J. Murax: a robust linguistic approach for question answering using an on-line encyclopedia. In: *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*. New Orleans, USA: ACM, 2001. 181–190
- 28 Katz B. Annotating the World Wide Web using natural language. In: *Proceedings of the 5th RIAO Conference on Computer Assisted Information Searching on the Internet*. Montreal, Quebec, Canada: ACM, 1997. 136–155
- 29 Katz B, Borchardt G C, Felshin S. Natural language annotations for question answering. In: *Proceedings of the 19th International Florida Artificial Intelligence Research Society Conference*. Florida, USA: AAAI, 2006. 303–306
- 30 Burke R D, Hammond K J, Kulyukin V A, Lytinen S L, Tomuro N, Schoenberg S. Question answering from frequently asked question files: experiences with the FAQ finder system. *AI Magazine*, 1997, **18**(2): 57–66
- 31 Hovy E, Gerber L, Hermjakob U, Junk M, Lin C Y. Question answering in webclopedia. In: *Proceedings of the TREC–9 Conference*. Gaithersburg, USA: NIST, 2000. 655
- 32 Ittycheriah A, Franz M, Zhu W J, Ratnaparkhi A, Mammon R J. Ibm's statistical question answering system. *Experimental Techniques*, 2000, **33**(6): 30–37(38)
- 33 Kwok C, Etzioni O, Weld D S. Scaling question answering to the web. *ACM Transactions on Information Systems*, 2001, **19**(3): 242–262
- 34 Dumais S, Banko M, Brill E, Lin J, Ng A. Web question answering: is more always better? In: *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Tampere, Finland: ACM, 2002. 291–298
- 35 Khalid M A, Jijkoun V, Rijke M D. The impact of named entity normalization on information retrieval for question answering. In: *Proceedings of the 30th European conference on Advances in information retrieval*. Berlin, Heidelberg: Springer-Verlag, 2008: 705–710
- 36 Liu Ya-Jun, Xu Yi. Automatic question answering system based on weighted semantic similarity model. *Journal of Southeast University (Natural Science Edition)*, 2004, **34**(5): 609–612
(刘亚军, 徐易. 一种基于加权语义相似度模型的自动问答系统. 东南大学学报(自然科学版), 2004, **34**(5): 609–612)
- 37 Zhou Fa-Guo, Yang Bing-Ru. New method for sentence similarity computing and its application in question answering system. *Computer Engineering and Applications*, 2008, **44**(1): 165–167
(周法国, 杨炳儒. 句子相似度计算新方法及其在问答系统中的应用. 计算机工程与应用, 2008, **44**(1): 165–167)
- 38 Soubbotin M M. Patterns of potential answer expressions as clues to the right answers. In: *Proceedings of the 10th Text Retrieval Conference*. Gaithersburg, USA: NIST, 2001. 293–302
- 39 Lin D K, Pantel P. Discovery of inference rules for question-answering. *Natural Language Engineering*, 2001, **7**(4): 343–360
- 40 Mollá D. Learning of graph-based question answering rules. In: *Proceedings of the 1st Workshop on Graph Based Methods for Natural Language Processing*. New York, USA: ACL, 2006. 37–44
- 41 Moldovan D, Clark C, Harabagiu S M, Maiorano S J. Cogex: a logic prover for question answering. In: *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology*. Edmonton, Canada: ACL, 2003. 87–93
- 42 Tang L R, Mooney R J. Using multiple clause constructors in inductive logic programming for semantic parsing. In: *Proceedings of the European Conference on Machine Learning*. Freiburg, Germany: Springer, 2001. 466–477
- 43 Zadeh L A. Fuzzy logic = computing with words. *IEEE Transactions on Fuzzy Systems*, 1996, **4**(2): 103–111
- 44 Clark P, Thompson J, Porter B. A knowledge-based approach to question-answering. In: *Proceedings of the 6th National Conference on Artificial Intelligence*. Orlando, USA: AAAI, 1999. 43–51
- 45 Barker K, Chaudhri V K, Chaw S Y, Clark P E, FAN J, Israel D, Mishra S, Porter B, Romero P, Tecuci D, Yeh P. A question-answering system for AP chemistry. In: *Proceedings of the 9th International Conference on Knowledge Representation and Reasoning*. Whistler, Canada: AAAI, 2004. 488–497
- 46 Liu Kai-Ying. Research on Chinese FrameNet construction and application technologies. *Journal of Chinese Information Processing*, 2011, **25**(6): 46–53
(刘开英. 汉语框架语义网构建及其应用技术研究. 中文信息学报, 2011, **25**(6): 46–53)
- 47 Wang Zhi-Qiang, Li Ru, Liang Ji-Ye, Zhang Xu-Hua, Wu Juan, Su Na. Research on question answering for reading comprehension based on Chinese discourse frame semantic parsing. *Chinese Journal of Computers*, 2016, **38**(4): 795–807
(王智强, 李茹, 梁吉业, 张旭华, 武娟, 苏娜. 基于汉语篇章框架语义分析的阅读理解问答研究. 计算机学报, 2016, **38**(4): 795–807)
- 48 Bollacker K, Evans C, Paritosh P, Sturge T, Taylor J. Freebase: a collaboratively created graph database for structuring human knowledge. In: *Proceedings of the ACM SIGMOD International Conference on Management of Data*. Vancouver, Canada: ACM, 2008. 1247–1250

- 49 Suchanek F M, Kasneci G, Weikum G. Yago: a core of semantic knowledge. In: Proceedings of the 16th International Conference on World Wide Web. Banff, Canada: WWW, 2007. 697–706
- 50 Hoffart J, Suchanek F M, Berberich K, Weikum G. Yago2: a spatially and temporally enhanced knowledge base from wikipedia. *Artificial Intelligence*, 2013, **194**: 28–61
- 51 Yao X C, Van Durme B. Information extraction over structured data: question answering with freebase. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore, USA: ACL, 2014. 956–966
- 52 Yih W T, Chang M W, He X D, Gao J F. Semantic parsing via staged query graph generation: question answering with knowledge base. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the AFNLP. Beijing, China: ACL, 2015. 1321–1331
- 53 West R, Gabrilovich E, Murphy K, Sun S, Gupta R, Lin D K. Knowledge base completion via search-based question answering. In: Proceedings of the 23rd International Conference on Worldwide Web. Seoul, Korea: ACM, 2014. 515–526
- 54 Unger C, Cimiano P. Pythia: compositional meaning construction for ontology-based question answering on the semantic web. In: Proceedings of the Natural Language Processing and Information Systems-International Conference on Applications of Natural Language to Information Systems. Alicante, Spain: Springer, 2011. 153–160
- 55 Zhou Yong-Mei, Tao Hong, Chen Jiao-Jiao, Zhang Zai-Yue. Study on sentence similarity approach of Automatic Ask & Answer System. *Computer Technology and Development*, 2012, **22**(5): 75–78
(周永梅, 陶红, 陈姣姣, 张再跃. 自动问答系统中的句子相似度算法的研究. 计算机技术与发展, 2012, **22**(5): 75–78)
- 56 Du Wen-Hua. Comparative study of ontology construction methods. *Journal of Information*, 2005, **24**(10): 24–25
(杜文华. 本体构建方法比较研究. 情报杂志, 2005, **24**(10): 24–25)
- 57 Wei Shun-Ping, He Ke-Kang. Semi-automatic building approach of domain ontology based on text mining — a case study of building instructional design domain ontology. *Open Education Research*, 2008, **14**(5): 95–101
(魏顺平, 何克抗. 基于文本挖掘的领域本体半自动构建方法研究 — 以教学设计学科领域本体建设为例. 开放教育研究, 2008, **14**(5): 95–101)
- 58 Iyyer M, Boyd-Graber J, Claudino L, Socher R, Iii H D. A neural network for factoid question answering over paragraphs. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing. Doha, Qatar: ACL, 2014. 633–644
- 59 Yih W T, He X D, Meek C. Semantic parsing for single-relation question answering. In: Proceedings of the Meeting of the Association for Computational Linguistics. Baltimore, USA: ACL, 2014. 643–648
- 60 Zhang Y Z, Liu K, He S Z, Ji G L, Liu Z Y, Wu H, Zhao J. Question answering over knowledge base with neural attention combining global knowledge information. arXiv: 1606.00979, 2016.
- 61 Werbos P J. Beyond Regression: new Tools for Prediction and Analysis in the Behavioral Science [Ph. D. dissertation], Harvard University, USA, 1974
- 62 Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, **18**(7): 1527–1554
- 63 Silver D, Huang A, Maddison C J, Guez A, Sifre L, Van Den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D. Mastering the game of go with deep neural networks and tree search. *Nature*, 2016, **529**(7587): 484–489
- 64 Liu Kang, Zhang Yuan-Zhe, Ji Guo-Liang, Lai Si-Wei, Zhao Jun. Representation learning for question answering over knowledge base: an overview. *Acta Automatica Sinica*, 2016, **42**(6): 807–818
(刘康, 张元哲, 纪国良, 来斯惟, 赵军. 基于表示学习的知识库问答研究进展与展望. 自动化学报, 2016, **42**(6): 807–818)
- 65 Zheng Z P. AnswerBus question answering system. In: Proceedings of the 2nd International Conference on Human Language Technology Research. San Diego, USA: ACM, 2002. 399–404
- 66 Tunstall-Pedoe W. True knowledge: open-domain question answering using structured knowledge and inference. *AI Magazine*, 2010, **31**(3): 80–92
- 67 Hajishirzi H, Mueller E T. Question answering in natural language narratives using symbolic probabilistic reasoning. In: Proceedings of the 25th International Florida Artificial Intelligence Research Society Conference. Marco Island, USA: AAAI, 2012. 38–43
- 68 Lally A, Prager J M, McCord M C, Boguraev B K, Patwardhan S, Fan J, FODOR P, Chu-Ca J. Question analysis: how watson reads a clue. *IBM Journal of Research and Development*, 2012, **56**(3–4): 2:1–2:14
- 69 Kalyanpur A, Patwardhan S, Boguraev B K, Lally A, Chu-Carroll J. Fact-based question decomposition in DeepQA. *IBM Journal of Research and Development*, 2012, **56**(3): 388–389
- 70 Gondek D C, Lally A, Kalyanpur A, Murdock J W, Duboue P A, Zhang L, Pan Y, Qiu Z M. A framework for merging and ranking of answers in DeepQA. *IBM Journal of Research and Development*, 2012, **56**(3–4): 14:1–14:12
- 71 Dang H T, Kelly D, Lin J J. Overview of the TREC 2007 question answering track. In: Proceedings of the 16th Text Retrieval Conference. Gaithersburg, USA: NIST, 2007. 115–123

- 72 Olvera-Lobo M D, Gutiérrez-Artacho J. Question answering track evaluation in TREC, CLEF and NTCIR. *Advances in Intelligent Systems and Computing*, 2015, **353**: 13–22
- 73 Peñas A, Forner P, Sutcliffe R, Rodrigo Á, Forăscu C, Alegria I, Giampiccolo D, Moreau N, Osenova P. Overview of ResPubliQA 2009: question answering evaluation over European legislation. In: *Proceedings of the 10th Cross-Language Evaluation Forum Conference on Multilingual Information Access Evaluation: text Retrieval Experiments*. Corfu, Greece: Springer, 2010. 174–196
- 74 Agichtein E, Carmel D, Harman D, Pelleg D, Pinter Y. Overview of the TREC 2015 LiveQA track. In: *Proceedings of the 24th TextREtrieval Conference*. Gaithersburg, USA: NIST, 2015. 1–9



曾 帅 中国科学院自动化研究所复杂系统管理与控制国家重点实验室助理研究员。2011 年于北京邮电大学获得信号与信息处理专业博士学位。主要研究方向为社会计算与策略优化。本文通信作者。E-mail: shuai.zeng@ia.ac.cn

(**ZENG Shuai** Assistant professor at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. She received her Ph. D. degree in signal and information processing from Beijing University of Post & Telecommunication in 2011. Her research interest covers social computing and strategy optimaization. Corresponding author of this paper.)



王 帅 中国科学院大学博士研究生。主要研究方向为社会计算与平行管理。

E-mail: wangshuai2015@ia.ac.cn

(**WANG Shuai** Ph. D. candidate at the University of Chinese Academy of Sciences. His research interest covers social computing and parallel management.)



袁 勇 中国科学院自动化研究所复杂系统管理与控制国家重点实验室副研究员。2008 年于山东科技大学获得计算机软件与理论专业博士学位。主要研究方向为商务智能与计算广告学。

E-mail: yong.yuan@ia.ac.cn

(**YUAN Yong** Associate professor at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. He received his Ph. D. degree in computer software and theory from Shandong University of Science and Technology in 2008. His research interest covers business intelligence and computational advertising.)



倪晓春 中国科学院自动化研究所复杂系统管理与控制国家重点实验室工程师。2008 年于大连海事大学获得管理科学与工程专业硕士学位。主要研究方向为商务智能与知识自动化。

E-mail: xiaochun.ni@ia.ac.cn

(**NI Xiao-Chun** Engineer at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. He received his master degree in management science and engineering at Dalian Maritime University in 2008. His research interest covers business intelligence and knowledge automation.)



欧阳永基 解放军 61786 部队工程师。2015 年于解放军信息工程大学获得计算机软件与理论专业博士学位。主要研究方向为网络安全。

(**OUYANG Yong-Ji** Engineer at the Troops 61786 of People's Liberation Army. He received his Ph. D. degree in computer software and theory from PLA Information Engineering University in 2015. His main research interest is network security.)