面向限定领域问答系统的自然语言理解方法综述

王东升[1,2] 王卫民[1] 王 石[3] 符建辉[3] 诸 峰[1]

(江苏科技大学计算机科学与工程学院 镇江 212003)[1] (北京工业大学WIC研究院 北京 100022)[2]

(中国科学院计算技术研究所智能信息处理重点实验室 北京 100190)[3]

摘 要 领域无关的自然语言理解(NLU)技术在最近的十年中取得了长足的进步,然而由于基础研究与现实应用之间存在着强烈的实际需求与当前处理能力不足的矛盾,因此很多通用技术还不能在现实的问答系统中得到有效使用。针对现有的需求,开发面向领域的自然语言理解技术显得非常必要。首先对开放领域与限定领域问答系统进行了比较,并对一些典型的面向限定领域问答系统的自然语言理解技术进行了分析,然后介绍了面向限定领域问答系统的自然语言理解技术的评测标准,最后总结了目前限定领域问答系统研究存在的主要问题及未来发展方向。

关键词 限定领域,问答系统,自然语言理解,评测

中图法分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.08.001

Research on Domain-specific Question Answering System Oriented Natural Language Understanding: A Survey

WANG Dong-sheng[1,2] WANG Wei-min[1] WANG Shi[3] FU Jian-hui[3] ZHU Feng[1]

(School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China)[1]

(International WIC Institute, Beijing University of Technology, Beijing 100022, China)[2]

(Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China)[3]

Abstract Natural language understanding (NLU) in the last decade has made considerable progress in the line of domain-independent research. Such research is very important, but the gap between research results and real-world applications is large, and the strong realistic demand is in contradiction with the current processing capacity of the NLU community, which makes the development of domain-specific NLU technologies very necessary. We firstly compared open domain and domain specific question answering (QA) system. Then some typical natural language understanding technologies used in restricted domain question answering system were detailedly analyzed. The evaluation standard of the natural language understanding technology for domain-specific question answering system was introduced. At last, main problems and the future development of domain specific QA system were summarized.

Keywords Domain-specific, Question answering system, NLU, Evaluation

1 引言

自动问答通常被定义为这样一种任务:用户将以自然语言表述的问题提交到一个系统中,系统自动理解用户问题并产生答案。问答系统的典型应用场景为:用户想知道某事或某物的信息,但由于没有时间或者不愿意从大量的电子文档中查找相关的信息,则可以通过自然语言表述的信息诉求,直接从自动问答系统中精确、快速地得到问题的答案。

自问答系统提出以来,研究人员分别从人工智能(AI)和信息检索(IR)两个角度开展了大量的研究工作。在人工智能的早期发展阶段,问答系统主要是以各类机器可读的知识库作为信息源,来回答用户提出的自然语言问题。这类系统通常只能够回答答案已知的提问(即问题的答案作为一条知识存储于知识库中)。这类方法的优点是,对于所应用的领域,通常事先构造了领域模型,如领域本体或数据库模式等,从而可以使用各种复杂的定理证明、推理等技术来满足用户复杂的信息诉求。而自从1999年TREC举办了首届自动问答竞赛(QA TRACK)以来,越来越多的研究人员开始从信息检索(IR)角度开展问答系统的研究工作。参加TREC竞赛的各类系统的性能每年都有较大幅度的提升(这类评测竞赛还包括CLEF以及NTCIR等)。这类系统主要关注于怎样从海量的电子文档中找到包含问题答案的文本片断,并从文本片断

到稿日期:2016-11-22 返修日期:2017-02-06 本文受国家自然科学基金重点国际(地区)合作研究项目:海量语义数据处理关键技术研究(61420106005),国家自然科学基金面上项目:面向Web文本的因果知识获取方法研究(61173063),北京市博士后基金:面向语义大数据的交通路牌智能管理关键方法研究研究(2015ZZ-25),北京市朝阳区博士后基金:面向领域自然语言理解的带约束语义文法研究(2015ZZ-11),江苏科技大学博士启动基金:面向领域的自然语言理解研究资助。

王东升(1982-),男,博士,讲师,主要研究方向为问答系统、自然语言处理,E-mail:wds_ict@163.com;王卫民(1977-),男,博士,讲师,主要研究方向为自然语言处理,E-mail:wangweimin@gmail.com;王 石(1980-),男,博士,助理研究员,主要研究方向为知识工程、文本处理。

中抽取答案。参加各种评测竞赛的开放域问答系统一般比较容易评测,其使用的方法也与领域无关,通常只对问句做浅层次的分析。

这两类系统一直并行发展,但各有优缺点。其中,基于结构化知识的问答系统适合于知识易于被形式化并且用户的提问通常需要经过各种复杂处理后方能得到答案的领域,一般只能处理与限定领域相关的问题,而对于非领域相关的问题通常不能作答。这类系统包括基于自然语言接口的数据库查询系统、自然语言界面的专家系统或知识库查询系统、基于FAQ的问答系统等。而兴起于TREC,CLEF以及NTCIR的基于信息检索的问答技术的应用领域较广,适合于用户提问为简单的事实型问句的领域,这类系统包括各种基于Web的问答系统,如AskJeeves,START等。其中,IBM Watson系统[60]是近年来开放领域问答系统发展水平的典型代表,该系统综合利用了自然语言处理、信息检索、知识表示、自动推理、机器学习等开放式问答技术,其知识源包括非结构化的Web网页以及结构化的知识库,如DBpedia,YAGO等[3]。

面向限定领域的问答系统是问答系统的一个重要分支。其与开放域问答系统的主要不同点在于:限定领域的问答系统可以借助很多领域相关资源,极大地提高了系统的可用性及实用性。在目前开放域问答系统取得长足进步的同时,发展面向限定领域的问答系统是十分必要的,其原因有:1)限定领域的问答系统能为深层次的自然语言处理、人工智能等技术提供试验场。由于目前的开放域问答系统主要处理一些较简单的问题类型,主要使用信息检索、信息抽取等技术,而在限定领域中,研究人员可以关注处理复杂问题的一般理论和方法,如规划、推理等,随着这些理论的不断成熟,可逐步扩展到开放领域中。2)限定领域的问答系统可以为专业人员提供更好的人机接口。开放域的问答系统通常针对普通用户,其接受的查询语句多使用日常用语进行表达,系统反馈的答案也相对较简单,如关于某人或某物的介绍性信息等;而限定领域的问答系统则更多地面向熟悉此领域的用户,如领域专家,他们多使用领域专业术语来构造查询语句,且问题一般较复杂,其对系统反馈的答案的要求也较高,而这通常需要建立在深层次理解用户查询语句的基础上才能做到。但是,目前面向限定领域的问答技术还不是很成熟(虽然在某些很小的限定领域中问答系统已接近实用),还没有形成普适的理论和方法。

本文主要回顾面向限定领域的问答系统的研究历史,总结其中所使用到的自然语言理解方法的研究现状,而面向开放域的基于信息检索技术的问答系统及其相关技术的最新进展可参考TREC的相关论文集。

2 早期工作

LUNAR和BASEBALL是两个较早的领域相关问答系统,其中,LUNAR可回答关于阿波罗发射过程中岩石样本分析的一些问题[1],BASEBALL可回答某一个赛季中棒球比赛的一些相关问题[2]。这两个系统都是基于自然语言接口的数据库查询系统,即关于特定主题的答案都存放在一个数据库中。系统首先将用户的问题转化成数据库的查询语句,再将查询结果作为答案返回给用户。

面向特定领域的人机口语对话系统是基于自然语言接口的数据库查询系统的另一个重要分支,这类系统面临的主要挑战是口语的不规范性。自20世纪80年代中后期以来,随着语音识别技术和自然语言处理技术的快速发展,世界各国的学术界和企业界都对人机对话系统给予了极大的关注。这些系统及项目包括美国的DARPA项目、欧洲的SUNDIAL项目以及SUNSTAR计划[3-4]。还有一大批在很多领域中的实际应用也推动了面向领域的人机对话系统的发展,包括飞机旅行查询、火车信息查询、餐馆导航、渡船时间查询、天气查询、电子汽车分类广告等[5]。

在20世纪70和80年代,很多研究人员热衷于计算语言学的研究,这也推动了问答系统在一些更为复杂领域的应用,问答系统作为一个应用框架来验证计算语言学的研究成果。Berkeley大学的Unix Consultant(UC)系统可以回答一些关于Unix操作系统的问题,其通过将自动规划、推理、自然语言处理以及知识表示等理论相结合,对用户问题进行分析,并生成问句的形式化的意义表示,然后通过查询用户模型,并用目标分析技术猜测用户的信息查询需求[6]。国内也研制出了很多相关系统,如RCHIQL,NCHIQL,NLCQI等[7],它们所使用的是类似于语法和模板的技术,由于查询的对象是数据库,因此大部分系统都充分利用了ER模型。

面向领域的基于自然语言界面的专家系统或知识库系统则多从传统的人工智能角度来构建问答系统。早期的专家系统包括Feigenbaum等人于1968年在斯坦福大学建成的DENDRAL、MIT大学开发的数学符号运算专家系统MACSYMA以及肺功能测试专家系统PUFF等。由于专家系统的知识库通常是形式化的,这类系统通常会先将用户的查询语句转化为某种逻辑表示,再利用"合一"等逻辑推理技术从知识库中推理得到答案。国内相关系统包括陆汝钤院士主持开发的"Pangu"人机对话系统以及由曹存根研究员建立于NKI(国家知识基础设施)海量知识库基础上的NKI问答系统等。

3 开放领域与限定领域问答系统的比较

面向限定领域的问答系统与面向开放领域的问答系统既有联系又有区别,很多因素决定了现有的比较成熟的开放域的问答系统技术对一些限定领域问答系统会失效。而两者所使用的技术既有重叠,又各有特殊之处,具体将在第5节进行阐述。总体而言,面向限定领域的NLU通常依赖领域知识库、领域语境等领域资源来提高问答系统的整体性能。下面将从数据规模、领域语境、可用语义资源以及可用语料4个方面对两者进行比较。

3.1 数据规模

在开放领域的自动问答系统中,通常会充分利用Web等数据资源的数据冗余的特点,并利用一些统计方法来为问题查找正确的答案,而不需要使用很复杂的语言模型[8]。但这种方法通常不适用于限定领域的问答系统,特别是对于语料资源比较少的领域,限定领域的问答通常需要使用相对复杂的语言处理技术,包括句子语意理解、推理等。

3.2 领域语境

面向限定领域的自动问答系统通常提供了非常具体的领域语境，在这个语境中，每个词所对应的义项通常是其所有义项的一个子集，尽管有些词在某个领域中存在多义的现象，但语义消歧在面向限定领域的 QA 中的作用有限。另外，面向限定领域的问答系统所接受的问句通常与开放域的问答系统的输入有很大的不同。限定领域的问答系统的用户，特别是一些领域专家，在向系统提问时，通常会使用一些非常具体的领域术语，而且这些问句通常会比较复杂。而在开放域的问答系统中，用户的问题会相对比较简单（如事实型问题），用户的用词也是一些平常用语。这些特点决定了限定域问答系统通常需要较复杂的问句处理技术，而开放域的问答系统则只需要采用较浅层次的语言理解方法。

3.3 任务导向

面向限定领域的问答系统通常以具体的任务或目标为导向，不同的任务或目标决定了系统的领域知识库、领域词典等，从而决定了系统所能回答的问题的范围，对于超出范围的问题一般无法给出答案；而开放领域问答系统通常基于信息检索等技术从 Web 中检索出相关答案列表，并通过相关度等对答案进行排序，对所能回答的问题范围一般没有限制。

3.4 可用语义资源

面向限定领域与开放域的问答系统的最大不同点在于：前者通常有很多领域相关资源可以使用，而这些资源对于理解用户的提问具有重要价值。对于某个特定领域来说，可用的领域资源通常与领域的复杂性以及领域用户的需求密切相关，包括简单的领域实体列表、形式化的领域本体等。

而开放域的问答系统通常没有这样的具体领域信息。其可以利用的资源包括一些通用本体资源，如 WordNet、HowNet、常识本体 CYC，以及近些年发展较快的基于语义网（Semantic Web）技术的 linked data，如 FreeBase，Dbpedia[61-62] 等。而对于某个具体领域来说，这些通用的本体资源中的信息通常是失衡的。比如，对于一些比较特殊的领域来说，有些术语在通用本体中有可能根本找不到相关义项；而另一些部分的知识描述粒度又过细，比如，有些词通用本体中有多个义项，但在某个具体领域里有些义项根本不会出现，这些词的使用通常是没有歧义的，当使用通用本体资源来解析问句时，会增加很多不必要的歧义。最坏的情况是，通用本体中的某些信息对于具体领域来说是不合适的[9]。

3.5 可用语料

开放域问答直接面对的是 Web 等资源，可以通过网络爬虫、搜索日志等渠道抓取大量的语料进行训练和测试，机器学习等技术通常可以取得较好的效果；而大部分的限定领域问答系统则因为领域较窄、系统需求各异、用户量较小等原因，可用于限定领域问答系统的语料资源显得尤为宝贵，而将领域知识与统计相结合的语言理解方法通常可以在一定程度上克服这一困难[26]。

4 限定领域 QA 分类

根据答案源的不同形式以及输入自然语言的表现形式，本文对面向限定领域的问答系统进行分类，如图 1 所示。

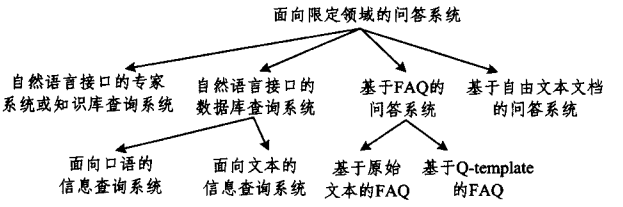

图 1　限定领域问答系统的分类

在不同的领域应用环境中，需要设计不同类型的问答系统。其中，基于自然语言接口的专家系统或知识库查询系统的信息源（答案源）为机器可读的领域知识或专家知识库，在用户输入自然语言查询时，系统首先将自然语言表述转化成特定的知识表示形式（取决于知识库的表示形式），并利用推理得到答案。这类系统的优点是：系统具有较强的推理能力，所使用方法的理论基础较好，包括知识表示、逻辑推理、自然语言的逻辑转换等；其缺点是：专家知识的形式化表示通常需要耗费大量的人力，知识库的完整性很难得到保证。基于结构化数据库的问答系统是从一个预先建立的结构化的数据库中查找提问的答案。系统需要将用户的自然语言表述转换成数据库的查询语言，如 SQL。从自然语言的表现形式来看，这类系统可分为口语和文本输入的信息查询系统。由于语音识别准确率低且口语本身具有语法不规范等特性，要求基于口语的信息查询系统必须具备较强的鲁棒性。基于常用提问集的问答系统是在已有的“问题-答案”对集合中找到与用户提问相匹配的提问，并将其对应的答案返回给用户。按照 FAQ 库中“问题”的表现形式，可分为基于原始文本的 FAQ 和基于 Q-template 的 FAQ，其中，基于 Q-template 的 FAQ 是在对用户的常问问题进行分析整理后形成的一种结构化表示形式，一个 template 可表示一簇表达相同或相似意思的问题。这类系统的优点是经过整理后的 FAQ 库的规模变小，但 FAQ 库的覆盖度变大。基于 FAQ 的问答系统的关键在于计算用户查询和 FAQ 知识库中问题或 Q-template 的相似度。计算相似度的方法有很多，大致可分为两种：1）不考虑语义信息，而是直接利用模式匹配技术、关键字（词）匹配技术、基于向量空间模型的 TF/IDF 方法等来计算相似度；2）考虑语义信息，利用 WordNet、HowNet、同义词词林等语义知识资源来计算用户查询与 FAQ 知识库中所有问题的语义相似度。另外，基于自由文本文档的问答系统与开放域的问答系统相似，即从预先建立的文本语料库中查找答案，这类问答系统的缺点是无法涵盖用户所有类型提问的答案，但它能够提供一个优良的算法评测平台，适用于对不同问答技术的比较研究。这类系统在限定领域的应用较少，因为在限定领域中，语料资源较少，预先收集文本语料库通常比较困难。

5 典型的面向限定领域的 NLU 技术分析

自 20 世纪 80 年代末以来，面向限定领域的 NLU 技术作为问答系统的核心模块，一直是一个较活跃的研究领域。总体来说，面向限定领域的 NLU 方法可大致分为基于自然语言处理的方法、基于语义文法的语义理解方法、基于统计的语言理解方法以及将后两者相结合的方法。其中，面向限定领域的基于自然语言处理的方法、基于统计的语言理解方法等

在开放领域中也有相似技术，与开放领域问答系统中所使用的技术相比，面向限定领域的相关技术需要依赖领域知识库以及领域语义规则等，并在语义分析过程中利用领域资源以提高系统性能。

5.1 基于传统自然语言处理的 NLU

这类系统的共同特征是将自然语言输入转换为某种形式化的意义表示，如逻辑（一阶谓词）、语义网络（semantic networks）、概念依赖图（conceptual dependency diagrams）、框架（frame）等。这类方法可应用于面向数据库的自然语言接口、专家系统或知识库查询系统、故事理解系统（story comprehension system）等。Moldovan[10]提出 Logic Form，即把问句和文本同时转化成统一的 Logic Form（QLF 和 ALF），通过对 QLF 和 ALF 的运算来抽取答案。Logic Form 最大的特色是它结合了词汇链，从而可以表达语义知识，实现推理功能。典型的基于自然语言处理的 QA 系统的架构如图 2 所示[11]，图中标出了语言分析模块以及领域相关知识模块。

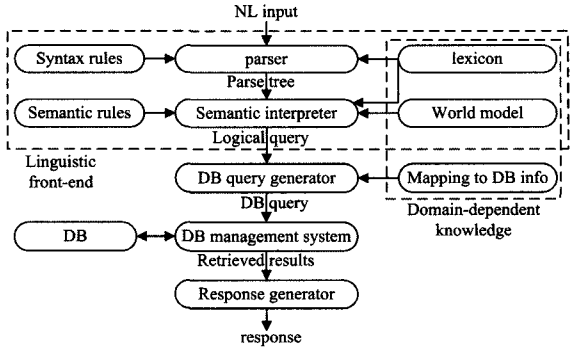

图 2 基于自然语言处理的 QA 典型系统结构[11]

语言分析模块首先对自然语言查询进行分析，然后将其翻译成中间逻辑表示，最终逻辑表示被转换成了特定的数据库查询语言。而领域相关知识则包含了与领域应用密切相关的信息，包括领域词典和领域模型。领域词典包含了领域相关术语，而领域模型则刻画了领域知识结构，比如领域类及对象的层次结构、对象的属性及其之间的关系。通常，领域模型使用领域本体进行描述。

基于自然语言处理的 QA 系统还可以利用机器学习方法来完善其中涉及到的句法规则[12]、词典[13]、语义规则[14]以及领域模型[15]等。对面向数据库的自然语言查询接口的相关综述可见文献[11]。典型的系统有 Symantec's Q&A，IBM's Language Access[16]以及 Bim's Loqui[17]。

Molla 等人[18]指出 NLP 方法比较适用于范围较窄且较复杂的领域，而信息检索技术则适用于数据冗余较多的领域。因为 NLP 方法需要依赖于知识库，增加了系统移植的难度。从图 2 可以看到，当问答系统所处理的语种改变时，一般需要修改语言分析模块中的 6 个子模块，而当变换所应用的领域时，则需要改变领域知识所对应的 3 个子模块。这些改变均需要经验丰富的编程人员、知识工程师以及数据库管理员的参与。

图 2 的语言分析模块是按照传统的自然语言理解方法来进行语言分析的，即先对句子按照句法规则进行句法分析，再依据句法与领域语义的对应关系（即图中的语义规则，Semantic rules）对句法分析结果进行语义解释，这一过程对于面向限定领域的问答系统特别是口语对话系统来说，通常是低效的；并且，问答系统的自然语言输入通常是含有噪音且不合语法规范的（ill-formed，病构的），从而导致上述语言分析过程的第一步就不能产生正确的结果，进而影响后续的语义处理；同时，短语结构语法存在局限性，即其生成能力过强，常常会产生一些不合语法规律的句子和歧义的句子，因此其不是一个特别适用于描写自然语言的语法，但是为了继承已有的研究成果，不宜完全抛弃短语结构语法，而应该在其基础上提出有效的手段来限制其过强的生成能力[20]。

5.2 基于语义文法的 NLU

Burton 首先提出了语义文法（semantic grammar）的概念，他认为语义文法在形式上与 Chomsky 类似，但语义文法可以在同一个模型中集成语法和语义信息，其中的语义类都是针对特定的应用领域。在限定的领域应用中，基本可以无歧义地建立词与语义类的对应关系[21]。语义文法 G 可以用一个 4 元组 $G=(V_T,V_{NT},S,R)$ 表示，其中 V_T 表示语义文法中的终结符（Terminals），V_{NT} 表示语义文法中的非终结符（Non-terminals）。注意，语义文法中的非终结符可以包括两类不交的非终结符集合，即 $V_{NT}=V_N+V_W$ 且 $V_N \wedge V_W=\varnothing$，其中 V_N 表示句法层次的非终结符集合，如 NP，VP 等；V_S 表示带有领域语义的非终结符，如航班名、道路名等。S 表示语义文法的开始符集合，$S=\{S_1,\cdots,S_n\}$，不同的开始符对应不同的领域语意，其通常与用户的查询意图相对应。R 表示语义文法中的确定型产生式集合。

PLANES[22]，LADDER[23]，REL[24]，EUFID[25]等系统均以语义文法为基础建立了面向数据库的自然语言问答系统。语义文法是一个宽泛的概念，一般来说，语义文法是指文法中带有语义单元，可以根据实际系统的需要采用不同的表示形式，包括 Finite State Machine（FSM），Augmented Transition Networks（ATN）[30]，Semantic Frame[19]，Context-Free Grammar（CFG）[26]，Question Template[43]等。

语义文法与句法文法最大的不同点在于：语义文法中的开始符及一些非终结符被赋予了领域语义，从解析结果中可以直接获取句子的语意信息；而传统的基于句法文法的解析结果只能得到句子的句法信息，须经过进一步的语义处理方能得到句子的语意信息。语义文法中可以包括句法层次和语义层次上的非终结符，也可以只包括语义层次上的非终结符。采用前一种设计模式的语义文法由于包含了领域无关的句法层次上的规则和领域相关的语义规则，因此在应用于新的领域时可以重用领域无关的那部分规则，具有较好的可移植性，但文法相对较复杂，文法规则数目较大；而后一种设计模式由于只包含了语义层次上的规则，因此文法数目相对较少，较容易设计，但当将系统应用于一个新的领域时，已有语义文法往往很难被重用。下面以航班查询问答系统为例来说明语义文法的作用，此处的语义文法只包括语义层次上的非终结符。语义文法的部分规则如下所示：

QUERY_FOR_FLIGHTS→DATE_TIME ROUTE FLIGHT QUERY_WHAT

DATE_TIME→ _WEEKLY_DATE

ROUTE→DEPARTURE_INFO ARRIVAL_INFO
DEPARTURE_INFO→从 _CITY
ARRIVAL_INFO→到 _CITY
_WEEKLY_DATE→周一 | 周二 | 周三 | 周四 | 周五 | 周六 | 周日
_CITY→北京 | 上海 | 深圳 | …
FLIGHT→航班 | 班次 | …
QUERY_WHAT→有哪些 | 有什么 | …

假设用户查询语句为：周五从北京到深圳航班有哪些？使用上述语义文法对该句子的解析结果如图 3 所示。

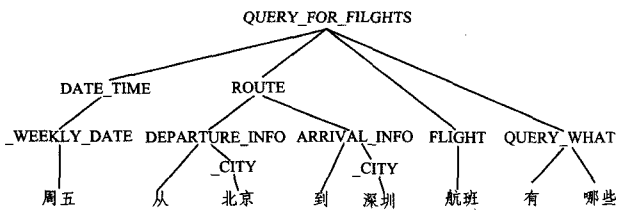

图 3　语义文法分析结果示例

图 3 中除叶子节点外的每个节点都是一个语义单元：根节点(QUERY_FOR_FLIGHTS)表示一个查询航班的意图，子节点 DATE_TIME 和 ROUTE 分别表示时间和航线两个查询条件，其中航线又分为起飞城市(DEPARTURE_INFO)和到达城市(ARRIVAL_INFO)信息。根据图 3 所示的分析树可以很容易地了解用户意图以及句子的深层语义信息，抽取的句子语义可以用意义表示形式进行表示。这里以框架形式为例，给出上述句子的意义表示，如图 4 所示。

图 4　语意表示框架示例

上述框架由属性-值对组成。抽取解析树中含有特殊标记的非终结符(本例中以“_”开头的非终结符)作为属性，将非终结符所覆盖的终结符作为属性的值。属性的值可以是一个嵌套结构，即属性的值还可以是一个属性-值对的集合。

使用语义文法描述语言的优点是可以给出句子丰富的语义结构。与句法文法相比，语义文法中的非终结符更多，文法规则的数量也更多，因此为了使设计的语义文法具有较好的领域覆盖度和准确度，通常需要文法设计人员经过多轮调优(包括增加、删除、更新等操作)，这一过程相当费时、费力。

为了使设计的文法具有较好的可移植性并减轻文法设计人员的工作量，通常可以在语义文法中添加一些领域无关的内容。

Seneff[27]提出了一种降低语义文法设计过程难度的方法，即重复利用文法中与领域无关的部分，从而在将系统应用于新领域时，只需要重新设计领域相关的那部分文法，而不需要重新设计全部的文法。该方法基于以下假设：句法结构在不同领域之间通常改变不大，因此高层次的上下文无关句法文法可以被不同领域所共享。通过使用领域相关的语义标签替换低层次的句法标签来引入领域知识。比如，可以使用领域相关的概念标签“宾馆名称(HOTEL_NAME)”对名词短语标签(NP)进行替换。有研究者使用类似的思想在“合一文法(unification grammar)”中引入领域相关的语义特征。这种方法的缺点是对文法设计人员的要求较高，其必须同时对句法知识以及领域知识有深刻的理解[28]。Ward[29]设计的 Phoenix 系统则使用了一种完全不同的方法来刻画语义信息。其缺点是限制了文法在多个领域中的共享，但文法设计人员可以在不需要了解任何句法规则的情况下对语义文法进行调优。为了解决系统开发初期训练数据不足的问题，文献[26]将领域知识融入到统计模型中来辅助对语料的语义标注，并将之用于语义文法的训练。该方法的优点是可以减少训练语料的数量。文献[48]通过在网络上检索与领域相关的文档，利用手工构造的种子文法来辅助创建搜索引擎的 query，从而检索出候选文档，再利用困惑度(perplexity)从文档中抽取领域最相关的部分用于语义文法的归纳学习。文献[49-50]将语义理解方法应用于语音识别控制领域，首先利用框架槽对语料进行概念标注，利用非负矩阵分解方法识别意图与用户用词之间的关联(个性化用词识别)，然后将用语与动作框架进行关联从而进行文法归纳学习。该方法的优点是可以对用户的个性化用语进行学习和建模。为了解决对话训练语料不足的问题，文献[51]提出利用众包(crowsourced)形式来收集文法学习所需语料，通过设置不同任务如答案补全、对话补全、问题转述(paraphrasing)、问题补全等来收集语料，并利用困惑度模型从语料中抽取质量较高的部分用于语义文法中的语义类学习。文献[52]在学习 non-ternimal 型规则时融入了词法、语法、语义特征，使用词汇化的本体资源来提高所学习规则的准确性，并在文法学习的不同阶段融入数据驱动的分析方法来提高文法学习的覆盖度。文献[53]首先计算未识别文本段与已有文法规则的相似度，包括字符串相似度和语境相似度，然后将相似度最高的文本段与规则相匹配，从而学得新的文法规则。文献[54]采用弱监督方法来学习语义文法，并基于语料给得到的规则赋予权重，将其同时用于语音识别和语义理解。该方法的优点是统一考虑语音识别和语义理解这两个步骤，从而避免错误传递。

除了上述语义文法学习之外，还有很多研究工作集中于对语义类(semantic class)的学习和抽取，语义类对应于语义文法的 terminal 型的规则是语意识别的重要基础[55-56]。

即使可以通过上述提到的一些技巧来改进手工设计文法的枯燥过程，但手工工作效率低、文法覆盖度难以保证等缺点还是无法避免。在人机交互过程中，输入的自然语言(口语或文本)通常是不规范的(重复、省略、错误输入、词序混乱和病句等现象)，这一特点对 QA 系统中的自然语言理解模块提出了巨大的挑战，要求其能够鲁棒地处理用户的自然语言输入，容忍用户输入中的“小缺点”，并能够准确理解用户的意图。

为了处理自然语言的非规范输入，面向语义文法的鲁棒式解析器在解析过程中可以跳过一些无意义的单元，对输入自然语言语句的片段产生部分解析结果，这些由句子片段产生的解析树可以分别被转化成语义框架，最后再使用一些启发式规则将这些片段对应的语义框架组合成整个句子的语义框架表示[58]。

5.3 基于机器学习的语言理解方法

数据驱动的、基于机器学习方法的 NLU 系统可以将设计人员从繁重的文法设计过程中解脱出来，从而获得了很多研究人员的青睐。

文献[31-32]将 HMMs 和 n 元语法模型相结合来刻画自然语言。模型使用标注语料进行训练。前者使用一种相对扁平的结构来表示语义，而后者则可以处理带嵌套的语义结构。文献[33]基于统计机器翻译的研究成果提出了一种模型，在这种模型中，语义表示中的一个概念可以覆盖句子中多个不连续的词。文献[34]研究了 HMMs 在 NLU 中的使用，发现两个状态之间的转移概率可以分解为栈操作的概率，其中栈操作可以将一个状态转化成另一个状态。

Meng H M[35] 提出了一种基于置信网络(Belief Networks)的语言理解方法。依据标注语义信息的语料，采用特征选择方法为每个信息查询目标选择语义特征，并为每一个信息查询目标分别建立语义特征与信息查询目标的贝叶斯网，最后基于标注语料库对贝叶斯网进行参数训练。Wei-Lin Wu[36]将口语理解看作两个分类问题，提出了一种基于弱指导学习的面向领域的口语理解方法。该方法使用主题分类器识别输入句子的主题，再使用训练好的槽分类器从句子中提取出槽-值序对，最终生成句子的语义表示。Minker W 等人[37]提出了一种基于一阶隐马尔科夫模型的语义分析方法，基于标注语料训练一阶隐马尔科夫模型的参数，其中句子成分为观察序列，语义标签为状态序列，在解析新句子时，基于模板将状态序列转换成语义框架表示。

这类系统的优点是鲁棒性及可移植性较强，但其通常采用机器学习方法如贝叶斯分类器、VSM 模型构造分类器，产生的是一个无层次的分类，并不能产生一个带有嵌入变元的结构，如分析树，这类方法对那些需要详细分析句子结构的应用并不适用。这类方法常应用于电话分类(call routing)等只需要获取句子浅层语义的应用[45]。另外，这类系统通常需要海量的训练数据，而在很多实际应用中通常很难得到大量的训练数据，数据稀疏问题是亟需解决的问题。虽然有些系统可以在预处理过程中用领域相关概念替换字符串(如在 ATIS 领域中，可以使用“城市名”替换语料中的实际出现的城市名称)来部分解决数据稀疏问题，但对于自底向上、纯数据驱动的文法归纳算法来说，在稀疏语料上学习到的文法的质量难以得到保证，需要语言专家人工校验学习到的文法，并且要手工为自动学习到的语言结构赋予确当的语义标签。

5.4 基于规则与统计相结合的方法

Pieraccini[38] 分析了设计 NLU 系统不同阶段可能面临的一些困难：1)在系统设计初期，通常只有很少的数据可用于训练，因此系统设计几乎是从零训练数据开始的，这时唯一的选择就是手工设计文法；2)在系统部署运行一段时间后，通常会积累大量的数据，而通过人工来分析这些数据以发现系统存在的问题是极其困难和不现实的，这时自动学习功能可以十分有效地提高系统的性能。

从上述分析可以发现，对于一些领域应用来说，单纯地使用基于文法规则的方法或基于统计的方法的 NLU 都不能满足现实需求，同时，基于文法规则的 NLU 方法能够获得句子的丰富的结构信息，而基于统计的方法又能够容忍自然语言的随意性。因此，在实际应用中，通常将两类方法以不同的方式进行结合。微软研究院的 Yeyi Wang 等人[39] 提出了一种将统计模型与语义文法相结合的口语理解系统，该系统首先采用 SVM 和 Naïve Bayes 等分类器识别用户的查询意图，再根据识别出的查询意图选择与其相关的文法对句子进行解析，并最终依据解析结果生成句子的语义表示。此方法充分结合了统计模型的鲁棒性以及语义文法能够生成复杂结构的特性。Schapire 等人[40] 将先验知识引入统计模型，以弥补在构建鲁棒分类器时遇到的训练数据稀疏的缺陷；为了克服某些领域训练数据较少的问题，其提出使用已有的领域知识来补充训练数据，通过修改 AdaBoost 中的损失函数(Loss function)，在训练数据的适应度量以及领域知识的适应度量间取得均衡。Rayner 和 Hockey[41] 提出了一种基于决策列表的口语理解系统，该系统根据训练数据是否可用以及训练数据是否足够，动态决定是否使用基于规则的方法或数据驱动的方法以及将两者相结合的方法。在早期训练数据缺乏时，系统以基于规则的语言理解方法为主；当训练数据积累得越来越多时，系统逐渐演变到以数据驱动方法为主的语言理解方法。为了解决系统移植性问题，文献[42]提出将领域相关的 CFGs 融合到领域无关的 N-Gram 模型中，融合后的模型可以提高 CFG 的通用性以及 N-Gram 模型的领域相关性。与单纯的 N-Gram 语言模型相比，融合后的语言模型可以大大降低其在测试集上的困惑度。文献[46-47]通过将领域知识(领域专家定义的语义框架)与条件随机模型结合进行语义理解，来克服领域训练语料不足的问题。

6 面向限定领域的问答系统评测

面向限定领域的问答系统的评测与开放领域的问答系统评测相比，既有相同部分，也有其特殊之处。开放领域的问答系统由于所处理的问题类型较少，且问题较为简单，如事实型、列表型等，在评测时可分别对不同的问题类型制定特定的评测方法。目前，对开放域问答系统进行评测的国际会议有英语问答评测平台 TREC QA Track、日语问答评测平台 NICIR 和多语言问答评测 CLEF 等。其所采用的评测指标主要包括平均排序倒数(mean reciprocal rank，MRR)、准确率(accuracy)、CWS(confidence weighted score)等。其中：

$$MRR=\sum_{i=1}^{N}\frac{1}{\text{标准答案在系统给出的排序结果中的位置}} \quad (1)$$

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。

$$CWS=\frac{1}{N}\sum_{i=1}^{N}\frac{\text{前 } i \text{ 个体文中被正确回答的提问数}}{i} \quad (2)$$

其中，N 表示测试集中提问的个数。

上述指标是将问答系统看作一个黑盒，即只对系统输出的答案质量进行评价，这些指标对于开放领域是可行的，因为对于用户的输入，几乎都能在存在大量冗余的资源中查找到答案；而对于限定领域的问答系统而言，即使对用户输入的提问分析正确，由于答案源中没有包括该条知识，也不能反馈答

案，如果此时还是按照开放域的评测方法对系统进行评价，则不太合理。

对于限定领域的问答系统，要分别对几个主要部件的性能进行评价，主要包括 NLU 模块的评价以及答案源的评价。对于不同的 NLU 方法来说，具体的评价方法不尽相同。比如，对于使用语义文法的 QA 系统而言，NLU 的评价指标包括[43]：文法 G 对测试语料的句子意图识别准确度（Intent identification Accuracy，IAccu）、正确理解结果的平均排序倒数（IMRR）、语义槽识别准确度（Slot identification Accuracy，SAccu）以及语义槽识别召回率（Slot identification Recall，SRecall）等。其中，句子意图是一种较浅层次的语意理解，是对说话者言语行为（speech action）的总体概括，对于限定领域的某个应用来说，言语行为的数量是有限的。而语义槽是指 NLU 模块从句子中抽取出的特定概念[46]。定义如下：

$$IAccu=\frac{\#\text{意图识别正确的句子数}}{\#\text{测试语料中的总句子数}} \tag{3}$$

$$IMRR=\frac{1}{N}\cdot\sum_{i}^{N}\frac{1}{\text{正确的解析结果在所有候选解析结果中的位置}} \tag{4}$$

$$SAccu=\frac{\#\text{正确识别的概念槽数}}{\#\text{语料中被识别的概念槽数}} \tag{5}$$

$$SRecall=\frac{\#\text{正确识别的概念槽数}}{\#\text{语料中所有应该识别的概念槽数}} \tag{6}$$

答案源的评价公式如下：

$$Coverage=\frac{\#\text{所有理解正确并得到正确答案的问题数}}{\#\text{所有理解正确的问题数}} \tag{7}$$

其中，分母表示所有正确理解的问句，分子表示正确理解且能够检索到答案的问句。

Anne 等[59]以航天工程领域问答为例，提出以用户为中心，基于任务的限定领域问答系统评测方法以及基于用户的评测指标集，包括系统响应性能评价（速度、可靠性等）、答案评价（完整性、准确性、相关性、有用性等）、数据库内容评价（数据可靠性、领域覆盖度、大小、时效性）、用户交互评测等。Benamara[63]提出从领域问答系统的可移植性以及答案展示的语言和认知可靠性进行评估，其中前者包括评价词典、领域本体等的可移植性。对具有自然语言生成模块的系统来说，语言和认知可靠性评估包括自动生成的答案的可推理评估（justification）、答案可读性评估、相关性评估、流畅性评估、答案模板的可移植性等。

结束语　通过以上分析可以看出，在面向限定领域的问答系统（尤其是汉语问答系统）的发展过程中，应该注意以下一些问题：1）与开放域的问答系统一样，需要重视大规模的公开评测技术，以评测推动问答技术的发展。面向英语的限定领域 QA 已经出现了类似的评测，如 ATIS 等，现阶段对于限定领域的汉语问答技术的研究，我们迫切需要一个公开、公认、合理的评测平台。2）重视领域资源对限定领域 QA 中 NLU 的作用，建立一套普适的面向领域 QA 的 NLU 技术框架，并逐步建立自动获取这些领域资源的理论和方法。3）将在面向限定领域 QA 中形成的成熟的理论和方法推广到开放域的问答系统中，以促进问答系统技术的整体提高。4）目前的技术主要支持一问一答式的场景，文献[44]提出了一种建立于普通 QA 系统之上的面向领域的交互式问答技术，但须进一步探寻用户模型以及领域模型在限定领域 QA 系统中的作用，用户模型的引入可以使系统显得更加友好与智能。

参 考 文 献

[1] WOODS W A. Conceptual indexing: A better way to organize knowledge: Technical Report SMLI TR-97-61[R]. Sun Microsystems, Inc., 1997.

[2] GREEN J, BERT F, et al. Baseball: an automatic question-answerer[C]// Proceedings Western Computing Conference, 1961. ACM, 1961: 219-224.

[3] FERRUCCI D, BROWN E, CHU-CARROLL J, et al. Building Watson: An overview of the Deep QA project[J]. AI Magazine, 2010, 31(3): 59-79.

[4] FRASER N, DALSGAARD P. Spoken Dialogue System: A European Perspective[C]// Proc. Int'l Symp. Spoken Dialogue. 1996: 25-36.

[5] 宗成庆. 统计自然语言处理[M]. 北京：清华大学出版社，2008.

[6] WILENSKY R, CHIN D N, LURIA M, et al. The Berkeley Unix Consultant project[J]. Computational Linguistics, 1994, 14(4): 35-84.

[7] CARPENTER B, CHU-CARROLL J. Natural language call routing: A robust, self-organized approach[C]// Proc. Int. Conf. Spoken Language Proc.. 1998: 2059-2062.

[8] BRILL E, LIN J J, BANKO M, et al. Data-Intensive Question Answering[C]// TREC. 2001: 90.

[9] MOLLÁ D, VICEDO J L. Question answering in restricted domains: An overview [J]. Computational Linguistics, 2007, 33(1): 41-61.

[10] MOLDOVAN D, RUS V. Logic Form Transformation of WordNet and its Applicability to Question Answering[C]// Proceedings of 37th Meeting of Association of Computational Linguistics (ACL'2001). 2001.

[11] ADROUTSOPOULOS I, RITCHIE G D, THANISCH P. Natural Language Interfaces to Databases: An Introduction[J]. Journal of Natural Language Engineering, 1995, 1(1): 29-81.

[12] UNOLD O, DABROWSKI G. Use of learning classifier system for inferring natural language grammar[M]// Learning Classifier Systems. Springer Berlin Heidelberg, 2007: 17-24.

[13] KOZAREVA Z, HOVY E. Tailoring the automated construction of large-scale taxonomies using the Web[J]. Language Resources and Evaluation, 2013, 47(3): 859-890.

[14] ZETTLEMOYER L S, COLLINS M. Learning to map sentences to logical form: Structured classification with probabilistic categorial grammars[C]// Proceedings of the 21st Conference on Uncertainty in AI. 2005: 658-666.

[15] ANDRENUCCI A, SNEIDERS E. Automated Question Answering: Review of the Main Approaches[C]// ICITA (1). 2005: 514-519.

[16] BINOT J L, et al. Natural language interfaces: a new philosophy [J]. Sun Expert Magazine, 1991.

[17] OTT N. Aspects of the Automatic Generation of SQL statements in a Natural Language Query Interface[J]. Information Systems, 1992, 17(2): 147-159.

[18] MOLLA D, et al. NLP for Answer Extraction in Technical Domains[C]// Proc. of EACL 2003. Morgan Kaufmann, USA, 2003.

[19] WANG Y Y, DENG L, ACERO A. Semantic Frame-Based Spoken Language Understanding [J]. Spoken Language Understanding: Systems for Extracting Semantic Information from Speech, 2011, 48(3): 41-91.

[20] FENG Z W. Complex Features in Description of Chinese Language[J]. Journal of Chinese Information Processing, 1990, 4(3): 20-29. (in Chinese)
冯志伟. 汉语句子描述中的复杂特征[J]. 中文信息学报, 1990, 4(3): 20-29.

[21] BURTON R R. Semantic Grammar: An Engineering Technique for Constructing Natural Language Understanding Systems[J]. ACM Sigart Bulletin, 1976(61): 26.

[22] WALTZ D L. An english language question answering system for a large relational database[J]. Communications of the ACM, 1978, 21(7): 526-539.

[23] HENDRIX G, SACERDOTI E, SAGALOWICZ D, et al. Developing a natural language interface to complex data [J]. ACM Transactions on Database System, 1978, 3(2): 105-147.

[24] THOMPSON F B, THOMPSON B H. Practial natural language processing: the REL system prototype[J]. Advances in Computers, 1975, 13: 109-168.

[25] TEMPLETON M, BURGER J. Problems in Natural language interface to DBMS with examples from EUFID[C]// Proceedings of the 1st Conference on Applied Natural Language Processing. Santa Monica, California, 1983: 3-16.

[26] WANG Y Y, ACERO A. Rapid development of spoken language understanding grammars [J]. Speech Communication, 2006, 48(3): 390-416.

[27] SENEFF S. TINA: A Natural Language System for Spoken Language Applications [J]. Computational Linguistics, 1992, 18(1): 61-86.

[28] FRANCEZ N, WINTNER S. Unification Grammars[M]. New York: Cambridge University Press, 2012.

[29] WARD W, ISSAR S. Recent improvements in the CMU spoken language understanding system[C]// Proc. ARPA Human Language Technology. 1996: 213-216.

[30] SHAPIRO S C. Generalized augmented transition network grammars for generation from semantic networks[J]. Computational Linguistics, 1982, 8(1): 12-25.

[31] PIERACCINI R, LUBENSKY D. Spoken language communication with machines: the long and winding road from research to business[C]// International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. Springer Berlin Heidelberg, 2005: 6-15.

[32] PSAROLOGOU A. A Stochastic Petri Net based NLU Scheme for Technical Documents Understanding[D]. Wright State University, 2016.

[33] WONG Y W, MOONEY R J. Learning for semantic parsing with statistical machine translation [C]// Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics. Association for Computational Linguistics, 2006: 439-446.

[34] HE Y, YOUNG S. Semantic processing using the hidden vector state model[J]. Computer Speeech and Language, 2005, 19(1): 85-106.

[35] MENG H M, WAI C, PIERACCINI R. The use of belief networks for mixed-initiative dialog modeling[J]. IEEE Transactions on Speech and Audio Processing, 2003, 11(6): 757-773.

[36] WU W L, LU R Z, DUAN J Y, et al. Spoken language understanding using weakly supervised learning[J]. Computer Speech & Language, 2010, 24(2): 358-382.

[37] MINKER W, WAIBEL A, MARIANI J. Stochastically-based semantic analysis [J]. Springer International, 1999, 13(2): 177-194.

[38] PIERACCINI R. Spoken language understanding, the research/industry chasm[C]// HLT/NAACL Workshop on Spoken Language Understanding for Conversational Systems. Boston, 2004.

[39] WANG Y Y, ACERO A. SGStudio: rapid semantic grammar development for spoken language understanding [C]// INTERSPEECH. 2005: 1869-1872.

[40] SCHAPIRE R E, ROCHERY M, RAHIM M, et al. Boosting with prior knowledge for call classification[J]. IEEE Transactions on Speech and Audio Processing, 2005, 13(2): 174-181.

[41] RAYNER M, HOCKEY B A. Transparent combination of rule-based and data-driven approaches in a speech understanding architecture[C]// Proceedings of the Tenth Conference on European Chapter of the Association for Computational Linguistics. Budapest, Hungary, 2003: 12-17.

[42] WANG Y, MAHAJAN M, HUANG X. A unified context-free grammar and n-gram model for spoken language processing[C]// International Conference of Acoustics, Speech, and Signal Processing. Istanbul, Turkey, 2000: 1639-1642.

[43] WANG D S. A domain-specific question answering system based on ontology and question templates[C]// 2010 11th ACIS International Conference on Software Engineering Artificial Intelligence Networking and Parallel/Distributed Computing (SNPD). IEEE, 2010: 151-156.

[44] WANG D S, WANG S, WANG W M, et al. Interactive question answering based on Ontology and semantic Grammar[J]. Journal of Chinese Information Processing, 2016, 32(2): 142-152. (in Chinese)
王东升, 王石, 王卫民, 等. 基于本体和语义文法的上下文相关问答[J]. 中文信息学报, 2016, 32(2): 142-152.

[45] DUTA N. Natural Language Understanding and Prediction: from Formal Grammars to Large Scale Machine Learning[J]. Fundamenta Informaticae, 2014, 131(3): 425-440.

[46] WANG Y Y. Strategies for statistical spoken language understanding with small amount of data-an empirical study[C]// INTERSPEECH-2010. 2010: 2498-2501.

(下转第 41 页)

赵鹏,沈弼龙,王大伟,等. 面向异构 MPSoC 的多媒体应用程序任务分配方法[J]. 计算机辅助设计与图形学学报,2010,22(10):1671-1678.

[12] LI D,HOU Y B,HUANG Z Q,et al. A Fuzzy Dynamic Scheduling Algorithm for Multiple Characteristics of MPSoC System [J]. Journal of Computer-Aided Design & Computer Graphics, 2011,23(8):1447-1454. (in Chinese)
李达,侯义斌,黄樟钦,等. 面向 MPSoC 系统多特征的模糊动态调度算法[J]. 计算机辅助设计与图形学学报,2011,23(8):1447-1454.

[13] KARIM M,AMAROUCH M Y. An FPGA-based MPSoC for real-time ECG analysis[C]// Proceedings of 2015 Third World Conference on Complex Systems (WCCS). Los Alamitos: IEEE Computer Society Press,2015:1-4.

[14] ZHANG C,MA Y,LUK W. HW/SW Partitioning Algorithm Targeting MPSOC with Dynamic Partial Reconfigurable Fabric [C]// Proceedings of 2015 14th International Conference on Computer-Aided Design and Computer Graphics (CAD/Graphics). Los Alamitos:IEEE Computer Society Press,2015:240-241.

[15] WICAKSANA A,TANG C M,NG M S. A scalable and configurable Multiprocessor System-on-Chip (MPSoC) virtual platform for hardware and software co-design and co-verification[C]// Proceeding of 2015 3rd International Conference on New Media (CONMEDIA). Los Alamitos: IEEE Computer Society Press, 2015:1-7.

[16] Love R. Linux 内核设计与实现(第 3 版)[M]. 陈莉君,康华,张波,译. 北京:机械工业出版社,2011.

[17] 博韦,切萨蒂. 深入理解 LINUX 内核[M]. 陈莉君,张琼声,张宏伟,译. 北京:中国电力出版社,2007.

[18] AAS J. Understanding the Linux 2. 6. 8. 1 CPU scheduler[OL]. http://core. ac. uk/display/24603273.

[19] SAKAMURA K. μ ITRON4. 0 Specification: Ver. 4. 00. 00[R]. TRON Association,2002.

(上接第 8 页)

[47] WANG Y Y,LEE J,MAHAJAN M,et al. Combining statistical and knowledge-based spoken language understanding in conditional models[C]// Proceedings of the Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics,2006:882-889.

[48] KLASINAS I,POTAMIANOS A,IOSIF E,et al. Web data harvesting for speech understanding grammar induction[C]// INTERSPEECH. 2013:2733-2737.

[49] GEMMEKA J F,VAN DE LOO J,DE PAUW G,et al. A Self-Learning Assistive Vocal Interface Based on Vocabulary Learning and Grammar Induction[C]// Interspeech. 2012:831-834.

[50] VAN DE LOO J,DE PAUW G,GEMMEKE J F,et al. Towards shallow grammar induction for an adaptive assistive vocal interface: a concept tagging approach[C]// Natural Language Processing for Improving Textual Accessibility (NLP4ITA) Workshop Programme. 2012:27.

[51] PALOGIANNIDI E, KLASINAS I, POTAMIANOS A, et al. Spoken dialogue grammar induction from crowdsourced data[C]// 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE,2014:3211-3215.

[52] GEORGILADAKIS S,UNGER C,IOSIF E,et al. Fusion of knowledge-based and data-driven approaches to grammar induction[C]// Interspeech. 2014:288-292.

[53] CHORIANOPOULOU A,ATHANASOPOULOU G,IOSIF E, et al. tucSage: Grammar Rule Induction for Spoken Dialogue Systems via Probabilistic Candidate Selection [C] // SemEval 2014. 2014:668.

[54] GASPERS J,CIMIANO P,WREDE B. Semantic parsing of speech using grammars learned with weak supervision[C]// Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT). 2015.

[55] PARGELLIS A,FOSLER-LUSSIER E,LRR C H,et al. Auto-induced semantic classes [J]. Speech Communication, 2004, 43(3):183-203.

[56] PANGOS A,IOSIF E,POTAMIANOS A,et al. Combining statistical similarity measures for automatic induction of semantic classes[C]// IEEE Workshop on Auto-matic Speech Recognition and Understanding (2005). IEEE,2005:278-283.

[57] PALOGIANNIDI E. Using crowdsourcing for grammar induction with application to spoken dialogue systems[D]. Technical University of Crete,2013.

[58] KAISER E C,JOHNSTON M,HEEMAN P A. Profer: Predictive, robust finite-state parsing for spoken language[C]// IEEE International Conference on Acoustics, Speech, and Signal Processing,1999. IEEE,1999:629-632.

[59] DIEKERMA A R, YILMAZEL O, LIDDY E D. Evaluation of restricted domain question-answering systems[C]// Proceedings of the ACL Workshop on Question Aaswering in Restricted Domains. 2004.

[60] FERRUCCI D, LEVAS A, BAGCHI S, et al. Watson: beyond jeopardy![J]. Artificial Intelligence,2013,199:93-105.

[61] LOPEZ V,FERNÁNDEZ M,MOTTA E,et al. Poweraqua: Supporting users in querying and exploring the semantic Web [J]. Semantic Web,2012,3(3):249-265.

[62] LOPEZ V,UREN V,SABOU M,et al. Is question answering fit for the semantic web?: a survey[J]. Semantic Web,2011,2(2): 125-155.

[63] BENAMARA F. Cooperative question answering in restricted domains: the WEBCOOP experiment [C] // Proceedings of the Workshop Question Answering in Restricted Domains. ACL, 2004:21-26.