

自然语言问答系统的发展与现状

冯志伟

(杭州师范大学 浙江 杭州 310018)

摘要: 本文简要地介绍自然语言问答系统的发展与现状。首先回顾自然语言问答系统的发展历程,然后介绍自然语言问答系统的类别和结构。

关键词: 自然语言; 问答系统; Turing 测试; 危险边缘

Question-Answer System of Natural Language: Past and Present

FENG Zhiwei

(Hangzhou Normal University, Hangzhou 310018, China)

Abstract: This paper briefly introduces the past and present of Question-Answer system (QA system) of natural language. Firstly the historical development of QA system is reviewed, then the classification and structure of QA system is introduced.

Key words: natural language; QA system; Turing test; Jeopardy

自然语言理解(Natural Language Understanding, 简称 NLU) 研究如何让计算机理解和运用人类的自然语言,使得计算机懂得自然语言的含义,并对人给计算机提出的问题,通过人机对话(man-machine dialogue)的方式,用自然语言进行回答。因此,自然语言理解系统又叫做自然语言问答系统(Question Answering System),简称问答系统(QA system),这样的问答系统可以用于知识工程、信息检索、自动问答、专家系统、自然语言人机接口等领域,有很大的实用价值。

本文首先介绍自然语言问答系统的发展情况,然后介绍自然语言问答系统的类别和结构。

1. 自然语言问答系统的发展

早在计算机出现以前,英国数学家 A. M. Turing 就预见到未来的计算机将会对自然语言研究提出新的问题。

他在《计算机与智能》一文中指出:

“我们可以期待,总有一天机器会同人在一切的智能领域里竞争起来。但是,以哪一点作为竞争的出发点呢?这是一个很难决定的问题。许多人以为可以把下棋之类的极为抽象的活动作为最好的出发点,不过,我更倾向于支持另一种主张,这种主张认为,最好的出发点是制造出一种具有智能的、可用钱买到的机器,然后,教这种机器理解英语并且说英语。这个过程可以仿效小孩子说话的那种办法来进行。”[4]

Turing 提出,检验计算机智能高低的最好办法是让计算机来讲英语和理解英语,他天才地预见到计算机和自然语言将会结下不解之缘,他设计了如下页图 1 所示的 Turing 测试(Turing test)。

在 Turing 测试中,Turing 采用“问”与“答”模式,即观察者通过控制打字机向两个测试对象通话,其中一个是人,另一个是机器。要求观察者不断提出各种问题,从而辨别回答者是人还是机器。

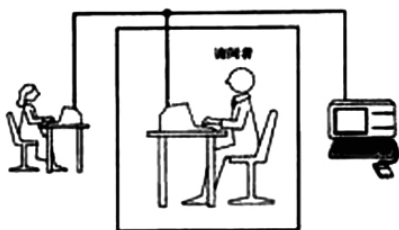


图 1 Turing 测试

Turing 还为这项测试亲自拟定了几个示范性问题:

问: 请给我写出有关“第四号桥”主题的十四行诗。

答: 不要问我这道题, 我从来不会写诗。

问: 34957 加 70764 等于多少?

答: (停 30 秒后) 105721。

问: 你会下国际象棋吗?

答: 是的。

问: 我在我的 K1 处有棋子 K; 你仅在 K6 处有棋子 K, 在 R1 处有棋子 R。现在轮到你走, 你应该下那步棋?

答: (停 15 秒钟后) 棋子 R 走到 R8 处, 将军!

Turing [4] 指出“如果机器在某些现实的条件下, 能够非常好地模仿人回答问题, 以至提问者在相当长时间里误认它不是机器, 那么机器就可以被认为是能够思维的。”

从表面上看, 要使机器回答按一定范围提出的问题似乎没有什么困难, 可以通过编制特殊的程序来实现。然而, 如果提问者并不遵循常规标准, 编制回答的程序是极其困难的事情。例如, 提问与回答呈现出下列状况:

问: 你会下国际象棋吗?

答: 是的。

问: 你会下国际象棋吗?

答: 是的。

问: 请再次回答, 你会下国际象棋吗?

答: 是的。

你多半会想到, 面前的这位回答者是一部笨机器。

如果提问与回答呈现出另一种状态:

问: 你会下国际象棋吗?

答: 是的。

问: 你会下国际象棋吗?

答: 是的, 我不是已经说过了吗?

问: 请再次回答, 你会下国际象棋吗?

答: 你烦不烦, 干嘛老提同样的问题。

那么, 你面前的这位大概是人而不是机器。

上述两种对话的区别在于, 第一种可明显地感到回答者是从知识库中提取简单的答案, 第二种则具有分析综合的能力, 回答者知道观察者在反复提出同样的问题。“图灵测试”没有规定问题的范围和提问的标准, 如果想要制造出能通过试验的机器, 以我们现在的技术水平, 必须在电脑中储存人类所有可以想到的问题, 储存对这些问题的所有合乎常理的回答, 并且还需要理智地作出选择。

Turing 预言, 大约到 20 世纪末叶, 具有 10GB 存储量的计算机经过 5 分钟的提问之后, 将会有 30% 的机会使提问者感到困惑, 辨不清回答问题的是人还是计算机。如果你也总是像计算机那样地去愚弄别人, 你也可能说不清楚究竟采用什么样的严格标准去愚弄, 才可以使别人感到困惑。然而, 对于图灵来说, 关键的问题是, 只要使用人的语言就足以作为测定计算机是否有智能的可操作标准了。

令人惊讶的是, 在 1966 年, 美国 MIT 的 J. Weizenbaum 就设计了一个叫做 ELIZA 的程序来验证 Turing 的思想 [5]。

ELIZA 是一个早期的自然语言处理系统, 这个系统能与用户进行有一定限制的对话。下面是 ELIZA 系统模拟心理医生的一段对话:

用户: You are **like** my father in some way.
(你在某些方面很像我的父亲。)

ELIZA: WHAT **RESEMBLANCE** DO YOU SEE? (你看是哪些方面像呢?)

用户: You are not very **aggressive** but I think you don't want me to notice that. (你不是一个非常坚持自见的人, 可是不愿意我提起这一点。)

ELIZA: WHAT MAKES YOU THINK I AM NOT **AGGRESSIVE**? (你凭什么说我不是一

个非常坚持自见的人?)

用户: **You don't argue with me.** (因为你不想和我争论。)

ELIZA: **WHY DO YOU THINK I DON'T ARGUE WITH YOU?** (为什么你认为我不想和你争论呢?)

用户: **You are afraid of me.** (因为你怕我。)

ELIZA: **DOES IT PLEASE YOU TO BELIEVE I AM AFRAID OF YOU?** (凭什么使得你认为我怕你?)

显而易见,ELIZA只是使用“模式匹配”(pattern matching)的方法来处理输入并且把它转换成适当的输出,这是一个非常简单的系统。事实上,ELIZA并没有必要懂得如何去模拟心理医生,它只是使用简单的模式匹配就取得了成功。正如系统的设计人Weizenbaum所说的,在ELIZA系统中,听话者的所作所为就好像他们对于周围的世界一无所知。

ELIZA与图灵思想的深刻联系在于,很多与ELIZA进行过交互的人都相信,ELIZA确实理解了他们所说的话以及他们所提出的问题。Weizenbaum在1976年指出,甚至在他把程序的操作过程向人们作了解释之后,仍然有不少的人继续相信ELIZA的能力。因此,ELIZA软件一时在美国市场上成为畅销货。

哲学家和人工智能研究者对于Turing测试究竟是否适合用来测试智能的争论已经持续很多年了,但是,上述比赛的结果,并没有平息这样的争论。

不过,这样的比赛结果与计算机究竟能否思维,或者计算机究竟能否理解自然语言的问题是风马牛不相及的。更为重要的是,在社会科学中的有关研究证实了Turing在同一篇文章中的预见[4]:

然而,我相信,在本世纪的末叶,词语的使用和教育的舆论将大为改观,使我们有可能谈论机器思维而不致遭到别人的反驳。

现在已经清楚,不管人们相信什么,不管人们是否已经知道了计算机的内部工作情况,他们都在谈论计算机,并且都在与计算机进行着交互,把计算机当作一个社会实体。人们把计算机当作人一样地对待,他们要对它讲礼貌,他们把它当作团队中的成员,并且期望计算机能够理解人们的需求,能够非常自然地与人们进行交互。给出这样的一些预设,使用自然语言处理系统就能够给众多的用户在很多应用方面提供更加自然的交互界面。

1966年美国公布了否定机器翻译的ALPAC报告之后,处于草创时期的机器翻译研究转入低潮,于是,同自然语言的计算机处理有关的研究,逐渐转向了自然语言问答系统方面。学者们采用了各种精巧的方法,尝试着建立计算机系统,让计算机理解自然语言,而根据图灵的意见,判断计算机是否理解了自然语言的最直观的方法,就是人同计算机对话。根据计算机对于人们用自然语言所提的问题的回答,就可以看出计算机是否理解了自然语言。这一方面的研究不久便取得了令人鼓舞的进展。因此,当20世纪60年代末期机器翻译困难重重、一筹莫展的时候,自然语言问答系统的研究却左右逢源、后来居上,而当机器翻译东山再起、重振旗鼓而进入复苏期的时候,自然语言问答系统却已获得了累累的硕果[1]。

下面,我们简要地介绍自然语言问答系统研究的发展情况。

自然语言问答系统的发展可以分为第一代系统和第二代系统两个阶段。第一代系统建立在对词类和词序分析的基础之上,分析中经常使用统计方法;第二代系统则开始引进语义甚至语用和语境的因素,几乎完全抛开了统计技术[6;7;8]。

第一代自然语言问答系统又可分为四种类型:

(1) 特殊格式系统:早期的自然语言问答系统大多数是特殊格式系统,根据人机对话内容的特点,采用特定的格式来进行人机对话。

a. SAD-SAM系统:

1963 年 ,R. Lindsay 在美国卡内基技术学院用 IPL - V 表处理语言设计了 SAD - SAM 系统 ,就采用了特定格式来进行关于亲属关系方面的人机对话 ,系统内建立了一个关于亲属关系的数据库 ,可接收关于亲属关系方面的问题的英语句子提问 ,用英语作出回答。这个系统分为两个模块: SAD 模块和 SAM 模块。SAD 模块的任务是作句法分析 ,它接收输入的英语句子 ,从左到右进行分析 ,建立起这个英语句子的推导树 ,然后 ,把这个能表示该英语句子结构的推导树传给 SAM。SAM 模块的任务是作语义分析并作出回答。首先 ,它从语义的角度抽取有关亲属关系的信息 ,建立起亲属关系树 ,然后根据数据库中存储的信息 ,找出问题的答案。SAD 模块处理英语句法结构的能力较强 ,除一般简单句外 ,还能处理一些结构复杂的句子。SAM 模块只能处理亲属关系方面的语义信息 ,不能处理其它方面的语义问题。SAM 在建立亲属关系树时并不考虑输入信息的顺序。如果先输入的信息可说明 B 和 C 是 X 的后代 ,D 和 E 是 Y 的后代 ,那么 ,就建立起两个家庭单元;而如果根据别的信息还可以说明 E 和 C 有兄弟姐妹关系 ,那么 ,就可以把这两个家庭单元合并为一个家庭单元。但是 ,SAM 不能处理某些歧义问题。例如 ,在句子 “Joe plays in his aunt Jane’s yard” 中 ,Jane 或者是 Joe 的姑妈 ,或者是 Joe 的姨妈 ,SAM 对此不能作出判断。

b. STUDENT 系统:

1968 年 ,D. Bobrow 在美国麻省理工学院 (MIT) 设计了 STUDENT 系统。这个系统能读懂用英语写的高中代数应用题 ,列出方程求解并给出答案。例如 ,STUDENT 系统能解决如下的用英语写的应用题:

If the number of customers Tom gets is twice the square of 20 per cent of the number of advertisements he runs , and the number of advertisements he runs 45 , what is the number of customers Tom gets? (如果汤姆争取得到的顾客数是他所出的广告数的百分之二十的平方的两倍 ,已知他出的广告数是 45 ,那么 ,汤姆争取

得到的顾客数是多少呢?)

STUDENT 系统中能识别的英语句子可以从如下的基本模式推出来:

(what are * and *)
 (what is *)
 (How many * 1 is *)
 (How many * do * have)
 (How many * does * have)
 (find *)
 (find * and *)
 (* is multiplied by *)
 (* is divided by *)
 (* is *)
 (* (* 1/verb) * 1 *)
 (* (* 1/verb) * as many * as * (* 1/verb) *)

其中 ,* 表示任意长度的词串 ,* 1 表示一个单独的词 ,(* 1/verb) 表示必须用词典来识别的一个动词。

当计算机解应用题时 ,首先要分析英语句子 ,理解这个应用题的意思 ,然后根据意思列出方程 ,最后 ,利用一个叫做 SOLVE 的求解模块来求解。如果 SOLVE 模块求解失败 ,STUDENT 系统还可利用探索法进一步辨识题意 ,或者利用一个叫做 REMEMBER 的模块来补充有关事实 ,以便进一步理解题意。

例如 ,REMEMBER 模块中可存储如下信息:

feet is the plural of foot(feet 是 foot 的复数)
 one half always means 0.5(一半总是意味着 0.5)

Successful candidates sometimes means students who passed the admissions test(成功的投考者有时是指那些通过了入学考试的学生)

distance equals speed times time(距离等于速度乘时间)

one foot equals 12 inches(一英尺等于 12 英寸)

如果查了 REMEMBER 模块还失败 ,STUDENT 系统还可以向用户提问 ,了解更多的

信息 继续利用探索法求解,每当探索成功,就可以把得到的新信息存入 SOLVE 模块中,从而增强 SOLVE 模块的能力。最后,如果求解成功,STUDENT 系统就把求得的解用英语打印出来,如果解不出来,则回答它不能解决这个应用题。例如,上面的那个应用题求解成功后,STUDENT 系统用英语打印出如下的解:

“The number of customers Tom gets is 162”
(汤姆争取得到的顾客数是 162)

STUDENT 系统解决高中代数应用题的能力很强,算题速度也很快。有一次在麻省理工学院(MIT)试验时,它解题的速度甚至比一个研究生还要快。

c. BASEBALL 系统:

20 世纪 60 年代初期,B. Green 在美国林肯实验室建立了 BASEBALL 系统,也使用 IPL-V 表处理语言,系统的数据库中存贮了关于美国 1959 年联邦棒球赛得分记录的数据,可回答有关棒球赛的一些问题。该系统的句法分析能力较差,输入句子十分简单,没有连接词,也没有比较级形式的形容词和副词,主要靠一部机器词典来进行单词的识别,使用了 14 个词类范畴,所有的问题都采用一种特殊的规范表达式来回答。BASEBALL 系统句法分析能力较差,输入句子十分简单,没有连接词(如 and,or,not),也没有比较级(如 higher,longer),主要是靠一部大词典来进行单词的识别,使用十四个词类范畴。工作时,BASEBALL 系统从右到左扫描输入的英语句子,把该句子转换为功能短语,找出关键词,再把该功能短语改写成一份说明表。这种说明表实质上是代表所提的问题的意义的规范表达式。例如:

“How many games did the Yankees play in July?”(七月间 Yankees 队进行了几次比赛?)

这个问题经过 BASEBALL 处理后,变为如下的规范表达式:

TEAM = YANKEES

MONTH = JULY

GAMES(数目) = ?

其中,TEAM 表示队名,分析出队名为

YANKEES,MONTH 表示月份,分析出月份为 JULY(七月),GAMES 数表示比赛次数,是需要回答的问题,用问号“?”表示。根据这样的问题,BASEBALL 在数据库中进行搜索,查出数据库中与该问题相匹配的数据条目,然后,输出这些数据,作出回答。由于 BASEBALL 系统的词典容量较大,可用试探法解决某些歧义问题(例如,score 可为动词“记分”,亦可为名词“记录”,Boston 可为地名“波士顿市”,亦可为球队名“波士顿队”),BASEBALL 可作出判断。

BASEBALL 的程序不能修改数据库中的数据,因此,这个系统没有演绎推理的能力。

(2) 以文本为基础的系统:某些研究者不满意在特殊格式系统中的种种格式限制,因为就一个专门领域来说,最方便的还是使用不受特殊格式结构限制的来进行人机对话,这就出现了以文本为基础的系统。

1966 年 R. F. Simmons、J. F. Burger 和 R. E. Long 设计的 PROSYNTHESIS - I 系统,就是以文本信息的存贮和检索方式工作的。

(3) 有限逻辑系统:有限逻辑系统进一步改进了以文本为基础的系统。在这种系统中,自然语言的句子以某种更加形式化的记号来替代,这些记号组成一个有限逻辑系统,可以进行某些推理。

a. SIR 系统:

1968 年 B. Raphael 在美国麻省理工学院用 LISP 语言建立了 SIR 系统。SIR 中针对英语提出了 24 个匹配模式,把输入的英语句子与这些模式相匹配,从而识别输入句子的结构。SIR 还能处理人们对话中常用的一些概念,如集合的包含关系、空间关系等等,这样,它就能作初步的逻辑推理,使这个系统有一点“智能”。

例如,SIR 系统中提出了如下的模式:

* is *

* is part of *

there are * on *

Is * * ?

How many * does * have?

what is the * of * ?

其中,*表示名词,这样的名词前可用限定词 a ,the ,every ,each 或数词来修饰,当输入句子与这些模式相匹配时,SIR 系统会触发出一定相应的动作,从而进行人机对话。

下面是 SIR 系统进行的人机对话的一个片段:

人: Every boy is a person. (每个男孩都是人。)

机: I understand. (我明白了。)

人: A finger is part of a hand. (指头是手的一部分。)

机: I understand. (我明白了。)

人: There are two hands on each person. (每个人有两只手。)

机: I understand. (我明白了。)

人: How many fingers does John have? (约翰有几个指头?)

机: The above sentence is ambiguous but I assume (has) means (has as parts) .

I don't know whether finger is part of John.

(上句话是有歧义的,但我假定“有”的意思是“作为各个部分的有”。我不知道指头是不是约翰的一部分。)

人: John is a boy. (约翰是一个男孩。)

机: I understand. (我明白了。)

人: How many fingers does John have? (约翰有几个指头?)

机: The above sentence is ambiguous but I assume (has) means (has as parts) . How many fingers per hand? (上句话是有歧义的,但我假定“有”的意思是“作为各个部分的有”。每只手有几个指头?)

人: Every hand has 5 fingers. (每只手有五个指头。)

机: The above sentence is ambiguous but I assume (has) means (has as parts) . The answer is 10. (上句话是有歧义的,但我假定“有”的意思是“作为各个部分的“有”。答案是 10。)

人: Who is President of the United States? (谁是美国的总统?)

机: Statement form not recognized. (该句子的型式我不能辨识。)

在上述对话中,机器分析与 24 个模式相匹配的句子,对于这些句子,机器能利用系统中已存储的一些概念进行简单的推理。例如,在前面几个问答中,机器知道了如下知识:

① 男孩是人;

② 指头是手的一部分;

③ 每个人有两只手。

当人问“约翰有几个指头?”的时候,由于这时机器不知道约翰是不是人,因而就无法判断指头是不是约翰身体上的一部分,在这种情况下,机器急切地想知道约翰是不是人。当人提示机器“约翰是一个男孩”的时候,机器利用已经知道的信息,马上可以推出“约翰是一个人”,而人是有两只手的,指头是手的一部分,因而指头也必然是约翰身体上的一部分。确定了指头是约翰身体上的一部分之后,由于机器知道约翰是人,人有两只手,故机器可推出约翰有两只手。这时,为了算出约翰有几个指头,必须知道每只手有几个指头,于是,机器便问:“每只手有几个指头?”人回答后,机器知道了每只手有五个指头,因此,机器便可作出判断,作出回答“答案是 10”,即约翰有 10 个指头。

我们可以看到,在这个人机对话中,机器一方面要识别句子的结构,另一方面也得进行一些简单的推理,自己在对话中进行学习,并记住已学到的知识,从事一些初步的智能活动。对于 24 个匹配模式之外的句型,机器是不能识别的。当人问“Who is President of the United States?”时,由于机器没有分析这种句型的能力,因此它回答“该句子的型式我不能辨识”。

b. DEDUCOM 系统:

1965 年 J. R. Slagle 建立了 DEDUCOM 系统,可在信息检索中进行演绎推理。

c. DEACON 系统:

1966 年 F. B. Thompson 建立了 DEACON 系统,通过英语来管理一个虚构的军用数据库,设计中使用了环结构和近似英语的概念来进行推理。

d. CONVERSE 系统:

1968 年 ,C. Kellog 在 IBM 360/67 计算机上 ,建立了 CONVERSE 系统 ,该系统能根据关于美国 120 个城市的 1000 个事实的文件来进行推理。

(4) 一般演绎系统: 一般演绎系统使用某些标准数学符号(如谓词演算符号) 来表达信息。例如:

Some girls are pretty (有些女孩是漂亮的)

这个英语句子可表示为

$x(\text{Girl}(x) \ \& \ \text{Pretty}(x)) ,$

Every girl is pretty (所有的女孩都漂亮)

这个英语句子可以表示为

$x(\text{Girl}(x) \rightarrow \text{Pretty}(x)) .$

其中 , \exists 是存在量词 , x 表示存在某个 x ,
 \forall 是全称量词 , x 表示对于一切的 x , $\&$ 是合
取符号 , \rightarrow 是蕴涵符号 ,表示“如果……,则
……”。

这样一来 ,逻辑学家们在定理证明工作上取得的全部成就 ,就可以用来作为建立有效的演绎系统的根据 ,从而能够把任何一个问题用定理证明的方式表达出来 ,并实际地演绎出所需要的信息 ,用自然语言作出回答。一般演绎系统可以表达那些在有限逻辑系统中不容易表达出来的复杂信息 ,从而进一步提高了自然语言问答系统的能力。

1968 - 1969 年 ,Green 和 Raphael 建立的 QA2 ,QA3 系统 ,采用谓词演算的方式和格式化的数据(formatted data) 来进行演绎推理 ,解答问题 ,并用英语作出回答 ,这是一般演绎系统的典型代表。

以上介绍的各种系统都属于第一代自然语言问答系统。

1970 年以来 ,出现了一定数量的第二代自然语言问答系统 ,这些系统绝大多数是程序演绎系统 ,大量地进行语义、语境以至语用的分析。其中比较有名的系统是 LUNAR 系统、SHRDLU 系统、MARGIE 系统、SAM 系统、PAM 系统。

a. LUNAR 系统:

LUNAR 系统是 W. Woods 于 1972 年设计的一个自然语言情报检索系统 ,其目的在于帮助地质学家们比较和评价从阿波罗 - 11 火箭得到的关于月球岩石和土壤的组成成分的化学分析数据 ,这个系统采用形式提问语言(formal query language) 来表示所提问的语义 ,从而对提问的句子作出语义解释 ,最后把形式提问语言执行于数据库 ,产生出对问题的回答。这个系统有一定的实用性 ,显示了自然语言问答系统对科学和生产的积极作用 ,因而大大地推动了这方面的研究工作。

LUNAR 系统的工作可分为三个阶段:

第一阶段: 句法分析

采用 ATN(扩充转移网络) 及语义探索方法产生出所提问题的推导树。LUNAR 系统能处理大部分英语的提问句型 ,词典容量是 3 500 词 ,可以解决时态、语式、代词所指、比较级、关系从句以及某些嵌入成分结构等较为困难的问题。不过 ,在分析连接词以及解决修饰词的某些歧义问题时 ,还常常会出现麻烦。该系统已足以处理地质学家们经常用来提问的那些英语句型了。

下面是 LUNAR 系统能够理解的一些英语句子:

1. What is the average concentration of aluminium in high alkali rocks? (高碱性岩石中铝的平均密集度是多少?)

2. What samples contain P205? (哪一些样本中含有 P205?)

3. Give me the modal analyses of P205 in those samples. (给我作出这些样本中 P205 的常规分析。)

第二阶段: 语义解释

用形式提问语言(formal query language) 来表示所提问题的语义 ,从而对提问的句子作出语义解释。形式提问语言由三部分组成: i . 标志符: 它标志在数据库中所存储事物的类别; ii . 语句: 它由谓词及论元组成 ,而论元就是标志符; iii . 指令: 它可启动一个动作。例如: (TEST(CONTAIN S10046 OLIV)) 是形式提问

语言的一个表达式。其中 ,S10046 是某种样本的标志符 ,OLIV 是橄榄石这种矿物的标志符 ,CONTAIN 是谓词 ,TEST 是真值检查指令。这个表达式的意思是:检查在样本 S10046 中是不是含有橄榄石这种矿物。

形式提问语言有一种带有量词函数 FOR 的表达式 ,形式如下:

(FOR QUANT X/CLASS: PX; QX)

其中 ,QUANT 是如 each ,every 数字等这样的逻辑量词 ,X 是要用这样的量词来说明的变量 ,CLASS 确定量词所涉及的事物的范围 ,PX 表示对这个范围加的限制 ,QX 是要用量词来说明的语句或指令。

例如 ,(FOR EVERY X1/(SEQ TYPECS): T; (PRINTOUT X1)) 就是一个这样的形式提问语句。其中 ,SEQ 表示枚举 ,PRINTOUT 表示打印论元的标志符 ,由于对量词的范围没有限制 ,所以 ,PX = T。这个形式提问语句的意思是:“枚举出所有类型为 C 的样本的样本数并打印出来。”

第三阶段: 回答问题

把形式提问语言表达式执行于数据库 ,产生出对问题的回答。

LUNAR 系统的一个完整的操作例子如下:

提问:

Do any samples have greater than 13 percent aluminium (举出任意的含铝量大于百分之十三的样本)

经过分析后得出的形式提问语言为

(TEST (FOR SOME X1/(SEQ SAMPLES): T; (CONTAIN X1 (NPR * X2/AL203) (GREATERTHAN 13 PCT))))

回答: YES

然后 ,LUNAR 系统可枚举出一些含铝量大于百分之十三的样本。

LUNAR 系统的专业范围有严格的限制 ,在语言处理中尽量解决那些常见的语法现象 ,不花过多的精力去解决那些目前水平还不能解决的复杂问题 ,因而能很快地投入实用 ,为地质学家

们提供了一个有用的自然语言信息检索系统。

b. SHRDLU 系统:

SHRDLU 系统是 T. Winograd 于 1972 年在美国麻省理工学院建立的一个用自然语言指挥机器人动作的系统。该系统把句法分析、语义分析、逻辑推理结合起来 ,大大地增强了系统在语言分析方面的功能。该系统对话的对象是一个具有简单的“手”和“眼”的玩具机器人 ,它可以操作放在桌子上的具有不同颜色、尺寸和形状的玩具积木 ,如立方体、棱锥体、盒子等 ,机器人能够根据操作人员的命令把这些积木捡起来 ,移动它们去搭成新的积木结构 ,在人机对话过程中 ,操作人员能获得他发给机器人的各种视觉反馈 ,实时地观察机器人理解语言、执行命令的情况。在电视屏幕上还可以显示出这个机器人的模拟形象以及它同一个真正的活人在电传机上自由地用英语对话的生动情景。

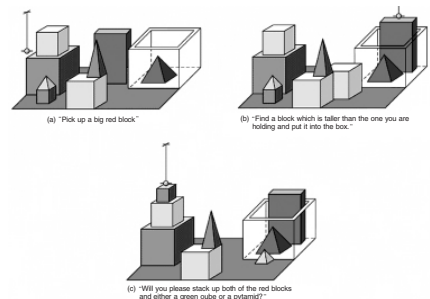


图 2 SHRDLU 系统示意图

下面是 SHRDLU 系统中人和计算机之间的一段对话:

人: PICK UP A BIG RED BLOCK. (抓起大的红色积木块)

机: OK.

人: GRASP THE PYRAMID.

机: I DON ' T UNDERSTAND WHICH PYRAMID YOU MEAN.

人: FIND A BLOCK WHICH IS TALLER THAN THE ONE YOU ARE HOLDING AND PUT IT INTO THE BOX. (抓起比计算机拿着的积木块更高的积木块 ,并把它放入箱子中)

机: BY "IT" , I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I

AM HOLDING.

机: OK.

人: WHAT DOES THE BOX CONTAIN?

机: THE BLUE PYRAMID AND THE BLUE BLOCK.

人: WHAT IS THE PYRAMID SUPPORTED BY?

机: THE BOX.

人: WILL YOU PLEASE STACK UP BOTH OF THE RED BLOCKS AND EITHER A GREEN CUBE OR A PYRAMIDE. (你可以把两个红色的积木块以及一个绿色的方块或一个金字塔积木块堆起来垛在一起吗?)

机: OK.

读者可以参照图 2 中的示意图来体会这段人机对话的内容。

c. MARGIE 系统:

MARGIE 系统是 R. Schank 于 1975 年在美国斯坦福人工智能实验室研制出来的。该系统的目的在于提供一个自然语言理解的直观模型。系统首先把英语句子转换为概念依存表达式,然后根据系统中有关信息进行推理,从概念依存表达式中推演出大量的事实。由于人们在理解句子时,总要牵涉到比句子的外部表达多得多的内容,因此,该系统的推理有 16 种类型,如原因、效应、说明、功能等等,最后,把推理的结果转换成英语输出。

d. SAM 系统:

SAM 系统是 R. Abelson 于 1975 年在美国耶鲁大学建立的。这个系统采用“脚本”(script)的办法来理解自然语言写的故事。所谓脚本,就是用来描述人们活动(如上饭馆、看病)的一种标准化的事件系列。Schank 和 Abelson 假定,每个人在他自己的生活实践中,会自然而然地意识到这样的脚本,在理解故事时,这些脚本可以用来构建时间发生的语境,因而也就可以用来预料它所代表的事件的情况,并以这些脚本为背景来理解自然语言,对故事中的人物、地点、事件进行推理,在推理过程中,给它们补充新的信息,最后采用“同义互训”(paraphrase)的方法,根据计算机理解的结果,

由计算机复述原来的故事。复述时,由于在推理过程中补充了许多新的信息,因而所复述的故事的内容会比原来的故事要丰富得多。计算机似乎像一个有理智的活人,把在推理过程中所推出的新信息加到故事中,添油加醋地把原来的故事说得更加精彩。例如,输入这样的简单的故事“约翰走进了一家饭馆。他坐了下来。他生气了。他走了。”SAM 系统的输出为:“约翰饿了。他决定到饭馆去。他走进了一家饭馆。服务员没理他。于是约翰生气了。他决定离开这个饭馆。”计算机推论出,约翰离开饭馆的原因是坐下来之后没有得到服务。这是因为在关于饭馆的“脚本”中,有“服务员送菜单”的项目,而输入句子中没有这样的内容,却有约翰生气的句子,因此,SAM 系统作出了这样的推论。

e. PAM 系统:

PAM 系统是 R. Wilensky 于 1978 年在美国耶鲁大学建立的另一个理解故事的系统。PAM 系统也能解释故事情节,回答问题,进行推论,作出摘要。它除了“脚本”中的事件序列之外,还提出了“计划”(plan)作为理解故事的基础。所谓“计划”,就是故事中的人物为实现其目的所要采取的手段。如果要通过“计划”来理解故事,就要找出人物的目的以及为完成这个目的所采取的行动。系统中设有一个“计划库”(plan box),存贮着有关各种目的的信息以及各种手段的信息。这样,在理解故事时,只要求出故事中有关情节与计划库中存贮的信息相重合的部分,就可以理解这个故事的目的是什么。当把一个一个的故事情节与脚本匹配出现障碍时,由于“计划库”中可提供关于一般目的的信息,就不致造成故事理解的失败。例如,营救一个被暴徒抢走的人,在“营救”这个总目的项下列举出若干个子目的,包括到达暴徒的巢穴以及杀死暴徒的各种方法,就可以预期下一步的行为。同时能根据主题来推论目的。例如,输入故事“约翰爱玛丽。玛丽被暴徒抢走了。”PAM 系统即可预期约翰要采取行动营救

玛丽。故事中虽然没有这样的内容,但是,根据计划库中的“爱情主题”,可以推出“约翰要采取行动营救玛丽”的情节。

Schank 等学者还进一步研究语言理解和记忆的关系,概括各种具体知识结构为一般经验,综合句法、语义、知识、推理为一体,建成 FRUMP 和 IPP 两个快速阅读系统。这两个系统存贮 2000 多个英语单词,对输入故事无需逐字逐句地分析,而是跳过某些无关的词语提取故事中的主要信息。这样的系统可以对报刊上一些新闻故事自动地做出摘要。

上述的系统都是书面的自然语言问答系统,输入输出都是用书面文字。口头的自然语言问答系统,还牵涉到语音识别、语音合成等复杂的技术,显然是更加困难的课题,口头自然语言问答系统的研究近年来也有很大的发展[9]。

2. 自然语言问答系统的类别

自然语言问答系统给某个提问提供简单而精确的回答,与信息检索任务和信息提取任务极为不同。目前的信息检索系统能让我们对与提问切题的相关文献进行定位,把从文本的等级列表中抽取答案的任务留给用户。在信息检索中,相关文本的识别是使用将提问与文献集匹配的方法来实现的,信息检索系统并不负责回答用户的问题。

信息抽取与信息检索不同,信息抽取系统抽取的东西是用户感兴趣的信息,抽取的条件是信息已经存在于预先规定的被称为模板的目标表现形式中。从总体上,信息抽取系统在一个与提取任务相关的文献集合上操作。信息抽取系统在完成抽取的任务时,可以成功地组拼模板。

尽管在信息检索系统的输出和信息抽取系统的输入之间有重合现象,但是把信息检索技术和信息抽取技术简单地组合起来,直接应用到开放领域的自然语言问答系统中是行不通的。其原因在于:这种解决办法需要建立适用于所有可能领域的信息抽取规则;这种解决办

法会把可能问及的问题的类型仅仅局限在信息抽取模板信息的形式范围之内。

不过,自然语言问答系统可以使用信息检索的方法来识别那些可能包含问题答案的文献,同时使用信息抽取技术来进行命名实体的辨识。

不管怎么说,成功的自然语言问答系统要对复杂的自然语言处理技术进行编码,捕获提问的语义,并对提问和候选答案进行词汇语义的合成。由于自然语言问答系统集中地使用了大量的句法、语义和语用的处理方法,因此,对自然语言问答系统技术的关注势必促进自然语言处理技术的发展,将自然语言理解推到研究与系统开发的前沿。

自然语言问答系统技术一定会在今后的数年内,在信息技术中发挥重要的作用。自然语言问答系统的用户可能是随意的提问者,他们只是问一问简单的具体问题;也可能是寻找具体产品特性和价格的顾客;也可能是正在收集市场、财经或商业信息的调研分析人员;还可能是查询非常具体、需要大量专门技术的信息的专业信息分析人员。所以,对回答提问的需求是很广泛的,正是由于自然语言问答系统有广泛的用户群,它的研究便具有广阔的应用前景。

根据处理提问与答案的形式,自然语言问答系统可以大致分为定型的自然语言问答系统和开放领域的自然语言问答系统两种[3]。

a. 定型的自然语言问答系统:

在定型的自然语言问答系统中,系统需要回答的问题或者是关于特定事实的,或者是具有专业性的。定型的自然语言问答系统对一个新提问首先进行最佳匹配,匹配对象是已知答案的预置问题的一个集合。若有合适匹配,就提供正确答案。定型的自然语言问答系统的客户群众多,客户们迫切希望依靠定型的自然语言问答系统,对自己特定的问题找出正确的答案。定型的自然语言问答系统在受限领域内表现较好,因在这些领域中,比较容易预测问题的答案。

定型的自然语言问答系统中的问题大致可

以分为两类:一类是关于具体事实的问题,一类是专业问题。

下面是关于具体事实的问题的例子:

—Who was the first American in space? (第一个进入太空的美国人是谁?)

—Where is capital airport? (首都机场在哪里?)

—When did the Neanderthal man live? (尼安德特人生活在什么时候?)

系统只要查询到有关的事实,就可以轻而易举地回答这一类的问题。

例如,最后一个问题的答案是:65万-30万年以前。

下面是关于专业问题的例子:

—What will the US response be if Iran closes the Strait of Hormuz? (如果伊朗关闭霍尔木兹海峡^①,美国将如何回应?)

—What effects on the price of oil on the international market are likely to result from the terrorist attacks on Saudi facilities? (恐怖分子袭击沙特阿拉伯的设施,对于国际市场上的油价会产生什么影响?)

回答这一类的问题,需要根据专业知识来进行推理,需要从各种专业文献中收集证据碎片,然后将这些证据碎片合并,才能形成最后的答案。

b. 开放领域的自然语言问答系统:

开放领域的自然语言问答系统要对来自任何领域的提问都能够提供答案,为了达到这个目的,需要运用句法、语义、语用等自然语言处理手段,从大量联机文献集合中搜寻并发现对于提问的答案。设计这种开放领域的自然语言问答系统的难点在于系统需要处理的提问的宽泛性。提问可能是问具体的信息的,例如,在文本检索会议(TREC)评估时所提的问题;提问也有可能问及复杂事件、事实或情况。

鉴于开放领域的自然语言问答系统具有提问的宽泛性,仅对提问类型分类是不够的,因为对同一问题,由于所查询文献的情况不同,或者由于文本中有关答案的遣词造句的方法不同,

答案的提取有难有易。因此,我们不对问题处理技术或答案提取技术进行分类,而是对整个自然语言问答系统进行分类,把开放领域的自然语言问答系统进一步细分为如下五类:

第一类:能够处理事实问题的自然语言问答系统。这类系统从一个或几个文献的集合中抽取文本片段作为回答。在通常的情况下,系统只需要逐字逐句地进行搜索,在文献中直接找出问题的答案。例如:

人: Who is the author of the book 'The Iron Lady: A Biography of Margaret Thatcher'? (谁是《铁娘子:撒切尔夫人传》的作者?)

机: THE IRON LADY: A Biography of Margaret Thatcher by Hugo Young. (《铁娘子:撒切尔夫人传》的作者是雨果·杨。)

第二类:具有简单推理机制的自然语言问答系统。这类系统需要在不同的文本片段中找出答案,并且用简单的推理形式,找出问题与这些答案之间的关系,从而把它们关联起来。在这种形式下,答案的发现需要使用更加精细的本体概念知识或者更加精细的语用知识,而答案的抽取则需要在这些知识的基础上进行推理。由于简单释义的不足,这样的推理通常必须使用世界知识和普通的常识。例如,在下面的问答中,就使用了“喝有毒饮品是死亡的一个原因”这样的假设。

人: How did Socrates die?

(苏格拉底是怎么死的?)

机: Similarly, it was to refute the principle of retaliation that Socrates, who was sentenced to death for impiety and the corruption of the city's youth, chose to drink the poisonous **hemlock**, the state's method of inflicting **death**, rather than accepting the escape from prison that his friends had prepared. (类似地,这驳斥了那种报复的原则,认为苏格拉底被判处死刑是由于他的不敬

^① 霍尔木兹海峡(the Strait of Hormuz)位于波斯湾与阿曼湾之间,是出入波斯湾的重要海上通道,地理位置至关重要。

行为以及他腐蚀城市的青年,他喝了毒芹,这是国家执行死刑的一种方式,而不是他接受了他的朋友们策划的越狱计划。)

在词网(WordNet1.6)中,名词 poison(毒)的第一个意思解释为 any substance that causes injury or illness or **death** of a living organism(“能对生物体造成伤害、疾病或死亡的任何物质”),根据这样的因果链进行推理,就为 poisonous hemlock(毒芹)可能是苏格拉底死亡的原因提供了证据。

第三类:能够从不同文献中融合出答案的自然语言问答系统。这种系统的特征是,它们能够提取散落在不同的若干个文献中的局部的信息,然后形成一个融合的答案。这样的回答格式决定着这些自然语言问答系统的多层复杂性。例如:

人: Name three countries that banned beef imports from Britain in the 1990? (列举出1990年禁止从英国进口牛肉的3个国家的名字?)

机: [France, West Germany, Luxembourg, Belgium]([法国,西德,卢森堡,比利时])

这种融合的开放领域自然语言问答系统需要具有更高级的语义处理能力和名称别名的识别能力。例如,在不同的若干个文献中,可能会使用 Britain 和 UK 等不同的名称来称呼“英国”,系统要能够识别出 Britain 和 UK 是同一个国家,才有可能在若干个不同的文献中进行知识的融合,把 Britain 和 UK 融合在一起。

第四类:可以进行类比推理的自然语言问答系统。这类自然语言问答系统的特征是,它们具有类比推理的能力。在这种自然语言问答系统中,问题的答案不会在任何文献中明确表述出来,而是需要将不同的答案进行类比推理,预测它们之间的相似点和不同点。在类比推理时,系统需要将问题分解成提取证据碎片的若干个小问题,然后使用类比的方式进行推理来构造对于问题的答案。例如:

问: Is the Fed going to raise interests at their next meeting? (Fed 打算在他们的下一次会议上提高利息吗?)

问: Is the US out of recession? (美国摆脱了经济萧条吗?)

问: Is the airline industry trouble? (航空工业出现了什么麻烦?)

要回答上述的问题需要从各种文本中提取证据的碎片,然后进行类比推理,构造出问题的答案。

第五类:交互式自然语言问答系统。这类自然语言问答系统的特征是能够在前期与用户互动形成的语境的基础上提问题,而不是孤立地提问,人与计算机之间可以交互。例如:

语境中的提问 1: Which museum in Florence was damaged by a major bomb? (佛罗伦萨的哪一个博物馆被炸弹破坏了?)

答: On June 20, the Uffizi gallery reopened its doors after the 1993 bombing. (1993 年爆炸之后,在 6 月 20 日,乌菲齐美术馆又重新开门了。)

语境中的提问 2: On what day did it happen? (爆炸是在那一天发生的?)

答: (Thursday) (May 27 1993) (星期四) (1993 年 5 月 27 日)

语境中的提问 3: Which galleries were involved? (包括哪一些画廊呢?)

答: One of the two main wings. (两个主要侧面画廊当中的一个。)

语境中的提问 4: How many people were killed? (死了多少人呢?)

答: Five people were killed in the explosion. (在爆炸中死了 5 个人。)

在回答这些问题的时候,计算机需要在前面已经回答的问题的基础上,检查提问前后的语境,才有可能做出回答。

3. 自然语言问答系统的结构

一个自然语言问答系统通常由三个模块组成:提问处理模块(Question – Processing)和文献处理模块(Document – Processing) 答案的提取和构造模块(Answer Extraction and Formulation)。

a. 提问处理模块:

在提问处理模块中包含着自然语言问答系统的很多技术,这些技术能够对提问加以进一步的说明,以便在所采集到的文献中找出对于有关问题的回答。

在自然语言问答系统中,自然语言的提问不能使用信息检索中的关键词和算子来表示,而是使用人类所能理解的、并且能够由自然语言问答系统处理的一套固有语义来表示。

这套固有语义也就是回答应当归属的语义类别。例如,当问 Who is best known for breaking the color line in baseball? (在打破职业棒球的肤色界限方面谁最有名?),预期的答案的语义类型是“人”(Person),以姓名的形式来表征,例如 Jackie Robinson。

开放领域问题的相关段落检索是建立在提问关键词的基础之上的。我们使用经验的方法来提取提问关键词,从提问的语义形式中提取实词,并优先考虑(a)引用表达;(b)命名实体;(c)复合名词。可能的关键词包含所有的名词和它们的形容词性修饰语,还有提问中的主要动词。

自然语言文本中的关键词会出现形态变化、同义表达、语义变换等变体形式,在自然语言问答系统中,有必要对这些关键词进行必要的变换。

我们可以从语言学的角度把关键词的变换分为如下三类:

(1) 形态变换。在自然语言问答系统的提问表达式中,可以列举出关键词有关的各种形态变化的形式。

例如,对于问题“who invented the paper clip?”(谁发明了回形针?)而言,预期的回答类型是“人”(Person),而且这个“人”是动词 invented (“发明了”)的主语,即词汇上名词化了的 inventor (“发明人”)。另外,由于在文献中搜索时不仅限于搜索关键词的词干形式,还要搜索该动词的所有屈折变化形式。这样,问题就可以使用如下的提问表达式来表示:

QUERY: [paper AND clip AND (invented OR inventor OR invent OR invents)]

其中,AND 表示“和”,OR 表示“或”,invented 的形态变换形式还有 inventor, invent, invents。

(2) 词汇变换。词网(WordNet)对于大量的很容易挖掘的语义信息都进行了编码,这样,我们就可以根据词网对关键词进行词汇变换,来检索关键词的同义词和其它语义相关项。这种词汇变换提高了答案的召回率。

例如,对于问题“Who killed Martin Luther King?”(“谁杀了马丁·路德·金?”),在搜索时除了 killer (“杀人者”)之外,还搜索 killer 的同义词 assassin (“刺客”),从而提高自然语言问答系统的召回率。

同样,对于问题“How far is the moon?”(“月亮离我们有多远?”),由于副词 far 在词网(WordNet)中被编码为 distance (距离)的属性特征,如果我们把 distance 这个名词添加到检索关键词中,也可以找到正确回答。

(3) 语义变换。词网(WordNet)中还记录了单词的上下位关系以及搭配,在自然语言问答系统之中,我们可以把单词的上下位关系或搭配定义为关键词的语义变换,这样,也可以提高自然语言问答系统的召回率。例如,对于问题“Where do lobsters like to live?”(龙虾喜欢生活在哪里?),由于在词网中,动词 like (喜欢)是动词 prefer 的上位词,它的定义是 like better (更加喜欢),所以,提问式可以写为:

QUERY: [lobsters AND (like OR prefer) AND live]

b. 文献处理模块:

在文献处理模块中,为了处理大范围的提问,开放领域的自然语言问答系统需要决定,它要寻找什么样的信息,或者要寻找什么样的预期的回答类型,并且还要决定,它到哪些文献中去搜寻这样的回答。

由于答案是靠文献中的文本碎片来呈现的,所以,这样的答案必定应当包含在能够被大多数提问概念辨识的文本碎片之中。因此,可能找到最终答案的文本碎片应当包含最具代表性的问题的概念,并且包含与预期的回答类别

相同的文本概念。

现有的检索技术还不能很好地模拟语义知识,因此,大多数自然语言问答系统只是将这样的搜索分解成基于问题关键词的检索以及文献的过滤机制两个部分,使得在文献中只保留那些预期的答案类型。

c. 答案的提取和构造模块:

在答案的提取和构造模块中,要确定答案的类型,答案的语义格式应该直接与提问的主干相关联,并且在提问的语义形式中具有最高的连通性。

答案类型的脱机分类可以依靠大型的词汇语义资源(例如,词网)来建立。词网(WordNet1.6)的数据库中包含100,000多个英语的名词、动词、形容词和副词,这些词使用“同义词集”(SYNSET)的方式组织起来。在对答案进行分类时,我们要设法建立起问答系统中的答案类型与词网中的同义词集之间的关联。

答案类型分类的过程可以分三步走:

第一步:对答案中的名词或动词的每一个语义类别,人工选择出它们最具代表性的概念结点,然后把这些概念结点加到答案类型分类(Answer Type Taxonomy)中。

第二步:由于预期的答案类型通常是命名实体,因此,我们需要在命名实体范畴和答案类型范畴之间建立多对多的映射。如图3所示:

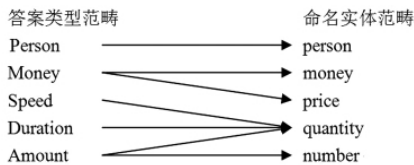


图3 答案类型范畴和命名实体范畴之间的多对多映射

从图3中可以看出,答案类型范畴的Speed, Duration和Amount三个范畴映射到命名实体范畴的quantity,形成3对1的映射;而答案类型范畴的一个范畴Money映射到命名实体范畴的money和price两个范畴,形成1对2的映射。可见,答案类型范畴和命名实体范畴之间的映射是多对多的。

第三步:把答案类型分类(Answer Type

Taxonomy)体系中顶端的每一个叶子结点,手工链接到词网的一个或者多个下属层次的结点上;这样,就可以把提问中的命名实体范畴与答案中的答案类型范畴联系起来,构造出相关的答案来。

4. 结语

自然语言问答系统的研究近年来取得突飞猛进的成绩。

2011年2月14日至16日,美国国际商用机器公司(IBM)研制的超级计算机Watson与两名人类智力竞赛高手Brad和Ken在美国著名的智力竞答电视节目《危险边缘!》(Jeopardy!)中进行竞答比赛。人类选手Brad和Ken曾经多次赢得《危险边缘!》的竞答冠军。然而,在这次竞答比赛中,超级计算机Watson却以绝对优势获得冠军,战胜了人类选手,这是计算机自然语言问答系统研究引起世界瞩目的重要成就。

尽管Watson存储了大量的百科全书和其他信息,但《危险边缘!》的问题十分复杂,并不会让Watson轻易地找到答案。自动问答比搜索引擎复杂得多。计算机的搜索引擎没法直接回答这些问题,搜索引擎只能给出符合搜索关键词的成千上万个似是而非的可能答案,而在自动回答问题时,Watson要通过各种不同的算法对所有的这些候选答案取得更多的证据支持,再根据各种证据的支持强度对每个候选答案计算出它们各自的置信度,最后根据置信度来判断是否向用户提供置信度最高的答案,并把这个答案当作是唯一正确的答案。显而易见,这样的搜索、计算和判断过程是极其复杂的。对于《危险边缘!》提出的任何问题,都需要动用几千个处理器的超级计算机来处理。Watson需要掌握大量的知识,并在相关的信息以及不相关的信息中反复权衡,发现线索。对计算机来说,这是一个巨大的挑战。因为人类可以在瞬间辨别出事物之间的联系,但是计算机却必须并行地考虑所有事情,从而得出结论。

Watson的胜利意味着IBM公司已经掌握

了对人类信息需求和问题给予更加准确而完善地处理的技术能力,并预见到了这个领域存在巨大商机。这项成果将被广泛应用于多个领域,例如,帮助医生更快、更准确地进行医疗诊断,帮助药物学家研究潜在的药物交互作用,帮助律师和法官寻找案例,帮助经济学家在金融领域实现“假设”的场景分析并遵从法规行事,帮助商业公司培养更加精明的销售人员,等等。

Watson 的胜利归根结底是人类智慧的胜利,因为 Watson 是由人类制造出来的,它的智慧是人类赋予的。Watson 的出现,改变了在此之前的简单的人机关系,并将带来一个崭新的人机合作时代[2]。

参考文献:

- [1] Allen, J. F. *Natural Language Understanding* [M]. Menlo Park, California: Benjamin/Cummings, 1995.
- [2] Androutsopoulos, I. & M. Aretoulaki. *Natural Language Interaction* [A]. R. Mitkov. *The Oxford Handbook of Computational Linguistics* (冯志伟导读) [C]. 北京: 外语教学与研究出版社 & 牛津大学出版社, 2009. 629 - 649.
- [3] Harabagiu, S. & D. Moldovan. *Question Answering* [A]. R. Mitkov. *The Oxford Handbook of Computational Linguistics* (冯志伟导读) [C]. 北京: 外语教学与研究出版社 & 牛津大学出版社, 2009. 560 - 582.
- [4] Turing, A. M. Computing machinery and intelligence [J]. *Mind*, 1950, 59: 433 - 460.
- [5] Weizenbaum, J. ELIZA-A computer program for the study of natural language communication between man and machine [J]. *Communications of the ACM*, 1966, 9(1): 36 - 45.
- [6] 冯志伟. 国外自然语言理解系统简介 [J]. 计算机科学, 1984, (2).
- [7] 冯志伟. 人机对话与语言研究 [J]. 语文建设, 1987, (6).
- [8] 冯志伟. 自然语言处理的形式模型 [M]. 北京: 中国科学技术出版社, 2009.
- [9] 宗成庆等. 基于对话知识的汉语口语理解 [A]. 1998 年中文信息处理国际会议论文集 [C]. 北京: 清华大学出版社, 1998.

收稿日期: 2012 - 07 - 02

作者简介: 冯志伟(1939 -), 昆明人, 计算语言学家, 国家教育部语言文字应用研究所研究员、博士生导师、学术委员会委员。研究方向: 语言学和计算机科学的跨学科研究。

“外教社博学文库”征稿启事

为及时反映我国英语学术研究的发展现状,展示各个研究领域的最新动态和成果,支持鼓励和扶植优秀的英语博士人才,上海外语教育出版社将继续遴选优秀英语语言文学博士论文,纳入“外教社博学文库”出版。

本文库征集书稿条件要求如下:

- 1) 书稿应为作者的博士毕业论文(中英文皆可),论文须有创新点;
- 2) 论文完成时间在 2000 年 1 月 1 日—2012 年 6 月 30 日期间;
- 3) 论文附三位或以上博导(校内一位、校外二位)的推荐信;
- 4) 作者在 2004 - 2012 年间在核心期刊上发表相关主题的论文不少于 2 篇;
- 5) 作者请从网上下载并填写投稿及推荐表格(<http://www.sflep.com>),同时提交论文影印本二份(论文中不出现作者姓名和工作单位等个人信息)。

本次征稿的截止日期为 2012 年 12 月 31 日。外教社将邀请专家匿名评审,遴选其中优秀论文出版。

投稿请寄: 上海市虹口区大连西路 558 号上海外语教育出版社 学术事业部 收 邮编: 200083