

多媒体支持的智能信息查询系统¹

陈 芳 袁保宗

(北方交通大学信息科学研究所 北京 100044)

摘 要 本文介绍多媒体技术支持下的智能信息查询系统,该系统可以使用图像、语音、文本等多种媒体对知识库进行查询,对同一语义可以有不同的询问方式。系统产生的回答内容丰富、形式多样,而且系统知识库可以方便地扩充。

关键词 多媒体, 知识表示, 语法, 语义, 自然语言生成

中图分类号 TP391

1 系统结构

系统首先确定查询中使用的媒体,也就是确定选择图像、语音、文本或键盘方式来提问。之后我们需要对媒体进行预处理,即对语音信息进行语音识别和理解,对图象信息进行图象识别和理解,对文本信息进行文字识别和理解。由媒体预处理得到查询语义后,可以通过多媒体知识表示库把用户的问题转换成一般的数据库风格的查询语句。然后对知识库进行查询,并可以通过学习方便地向该知识库中添加新知识。从知识库中查询到相应内容后,形成基本的语义。利用自然语言生成技术,把语义表示转换成自然语言形式的输出。最后通过媒体转换以图象、文字或语音的形式给出所需的查询结果。系统结构如图 1 所示。

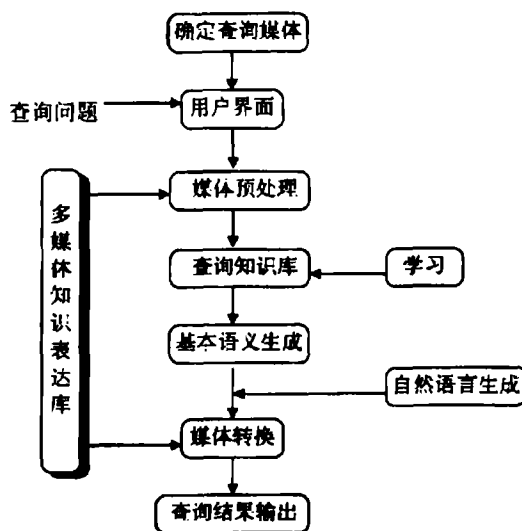


图 1 智能信息查询系统结构

¹ 1996-02-14 收到, 1996-10-30 定稿
国家自然科学基金资助课题

2 系统实现的关键问题

整个系统由 PC-486 主机配以摄象机、扫描仪、声卡、打印机、扬声器等设备组成。其中扫描仪可以对图象资料、文本信息进行扫描,声卡完成语音的输入输出。PC 机为主控设备,它实现对各个部分的控制,并对多媒体信息进行语义分析和综合处理。

2.1 多种媒体的知识表示^[1-3] 我们采用基于面向对象方法的框架表示与语义网表示相结合的混合型知识表达方法。这种方法具有可以起规划作用的框架结构,框架的具体槽面为对象,即用对象来代替框架的细节。添加规则就是添加对象,这不会影响整个知识体系结构。规则查找方便,添加也很容易。语义网很灵活,用它可以简洁地建立对象和对象内涵中的某一部分内容与具体实例之间的对应关系,易于学习。这种方法不仅在框架的选择上体现出了灵活性,而且框架和语义网的加入弥补了面向对象方法中逻辑功能和推理能力的不足。框架和面向对象方法的模块化规划提高了语义网的搜索效率。同时这几种方法都具有继承性,保证了推理机制的一致性。这种混合型知识表达方法表达力强、模块性好,易于增加新的知识,是一种较好的知识表达方法。

2.2 多种媒体的预处理及媒体转换^[4,5] 完整的多媒体信息查询系统是应该具有完整的媒体处理功能的,它的整体功能应包括语音的识别和理解、文字的识别和理解、图象的识别和理解(媒体的预处理)以及语音、文字、图象的输出(媒体转换)。

图象的识别和理解是利用神经网络建立了二维地图识别系统完成的。该系统可以通过扫描仪或摄象机采集地图,根据地图识别某个国家,并形成对该国家进行查询的语义。语音识别的任务是识别通过语音形式输入的自然语句查询。根据时序矢量量化(TSVQ, Time Sequence Vector Quantization)方法建立了一个可以识别中等词汇量(200 个左右)词组的实时汉语词组识别系统。该系统由 PC 机和一块 TMS320C25 高速信号处理板构成,以 12 阶 LPC 倒谱系数作为语音信号的特征参数,采用 TSVQ 方法进行模式匹配。文字识别和理解系统是利用文字识别技术把扫描仪输入的文本转换成计算机内部表示。

通过语音识别和文字识别后得到的结果都要通过语句理解系统转换成计算机可以理解的语义。我们设计的语句理解系统有语法分析和语义分析两个模块组成,其中语法分析模块又由词法分析和句法分析两个子模块构成。

(1) 语法分析模块

这一模块的功能是对输入的自然语言语句进行语法分析,得出它的语法结构。

(A) 词法分析

(a) 自动分词:采用修正的最大匹配法。

(b) 断句:根据句子结束的标志,如标点符号判定句子结束与否。

(B) 句法分析

(a) 归一化:凡是输入的句子首先根据系统中已存储的句型改写成规范化的无歧义语句。

如果超出了系统存储范围,就要进行学习。

(b) 词属性的分析:根据汉语系统特征网络,给出各个词的具体功能和特征的分析。

(c) 判断句子的语气类型和结构类型:根据系统特征网络判断句子的语气类型是属于疑问句、陈述句、感叹句或祈使句,判断句子的结构类型是属于主谓句、非主谓句等。

(d) 阶判定:进一步判定该句是短语、单句或复句。

(2) 语义分析模块

对句子作了形式上的分析后,需要通过语义分析来提取用户真正需要查询的知识库内容。查询语句主要是两类语句:陈述句和疑问句。这两种语句有不同的语义重点。

(A) 陈述句 陈述句的一般结构为:查询语句 = 附加结构 + 动词短语 + 特殊字结构。

特殊字结构指的是由“的”、“里”、“上”等字的前后词组成的字符串,它充当句子中的某一成分。在这类结构中,特殊字结构为查询的核心。只要寻找特殊字结构中的词所对应的内容、这些词与别的词之间的语义联系等,就可以得到需要表达的原始内容。

(B) 疑问句 疑问句分为两种,一种是含有特殊字结构的,这类的询问语句同陈述句一样处理。另一种是不含特殊字结构的,其一般形式为:查询语句 = 附加结构 + 动词短语 + 疑问结构。这类问句的查询核心为动词宾语和疑问结构中不含疑问词的部分。

得到了自然语句的文本之后,可以通过媒体转换变成容易听懂的声音输出。我们的汉语合成系统基元为:基本四声音节、轻声音节和常用词。它们的发音构成基本音库。同时我们还建立了规则库,其中包括:多音字、句子的发音模式规则库和语义表达发音模式规则库。以基本音库为基础,以规则库为控制手段,完成文本到发音描述的转换。语言学处理部分完成词的分割、词法分析、句法分析和语义分析,然后根据相应的规则库完成文本到发音描述的转换;声学处理部分根据发音描述完成语音数据的读取、调整 and 连接工作,利用声卡和扬声器输出语音。语音合成方法采用了波形编辑方式。查询结果同样也可以转换成相应的图象媒体输出。

2.3 系统的智能化 特别需要提一下的是因为汉语的句子组织十分自由,很难定义一个完备的语义集合,包括所有的提问表达方式,因而我们为系统设计了学习功能。当某种表达不能为机器接受时,我们可根据句法库中标准句式提取出机器不理解的部分进行学习。同时也可以进行词汇学习和具体词汇的特征学习。通过不断的学习,机器可以提高自身的理解能力。

计算机自然语言生成是人工智能和语言学等学科交叉研究产生的新的研究领域。这是智能信息查询系统中体现智能的一个重要组成部分。一般的查询系统只能从知识库或数据库中找到一个现成的答案输出给用户。我们采用了自然语言生成技术,首先根据需回答内容进行文本规划和文本组织形成机器内部的生成结构。最后根据我们建立的汉语语法结构特征网络,完成机器的内部结构到自然语言的转换。利用这种自然语言语句生成技术后,系统不是把对问题的回答都固定地存好,系统每次都按进入条件进入语法网络寻求满意解,并随机输出其中任意一个。因而系统对同一问题的答案也不是唯一的。

3 系统的特点

虽然我们建立的是一个地理信息查询系统,但这个系统有别于一般的查询系统。

(1) 查询问题的形式不受限制 由于汉语表达方式很多,因此查询问题的形式可以多种多样,可以是简单问句,句子中直接体现出要查询的内容;也可以是复杂问句,这类问句中不直接体现要查询的内容,而由系统根据对问句内部涵义的理解给出相应回答;也可以是组合问句,系统要把知识库中与提问内容相关的知识都提取出来并加以组织后形成回答。

(2) 多种媒体查询和输出方式 目前的系统可以采用语音方式、扫描仪输入文本方式或键盘输入提问。有些问题来自对视觉信息识别的结果,如视觉系统可以从地图中识别出某一国家,而查询系统则选此作为查询主题。同时也可以选择图象、语音、文本等方式输出查询结果。

(3) 输出的答案不固定 由于问题的形式不固定,所以系统不是把对问题的回答都固定地存好,而是在分析查询的问题后找到有关的知识库节点后,利用自然语言生成技术生成回答。系统测试发现,对任何一种问题,系统都可以随机生成 2—15 种不同的输出。

(4) 具有学习能力 我们设计的系统具有初步的学习能力。系统处理提问时,遇到知识库没有的语义时,它可以自动需要学习。学习完这个语义后,再处理提问。同时对知识库现有的知识也可以随时通过学习来更新。知识库便于添加、易于维护,知识库的添加具有良好的人机交互方式,可以随意添加和删改而不影响知识库的结构。

4 结 论

多媒体技术可以很好地支持智能化的计算机信息查询系统。学习能力的加入和自然语言生成技术的应用进一步提高了查询系统的智能化水平。

参 考 文 献

- [1] Bes G. A Natural Language and Graphics Interface. Berlin, Springer-Verlag, 1992: 176-188.
- [2] Steier D. Combining multiple knowledge source in an integrated intelligent system. IEEE Expert, 1993, 7(6): 35-43.
- [3] 陈 芳, 袁保宗, 林碧琴. 多模式信息处理中的知识表达. 中国自动化学会第九届青年学术年会论文集, 湖南长沙, 1993: 230-232.
- [4] Young S. The use of syntax and multiple alternatives in the vovis: Voice operated database inquiry system. Computer Speech and Language, 1991, 5(1): 65-80.
- [5] 陈 芳. 自然语句生成系统的建模与实现: [博士论文]. 北方交通大学信息科学研究所, 1994: 98-128.

INTELLIGENT INFORMATION INQUIRING SYSTEM SUPPORTED
BY MULTIMEDIA TECHNOLOGY

Chen Fang Yuan Baozong

(Institute of Information Science, Northern Jiaotong University, Beijing 100044)

Abstract This paper introduces an intelligent information inquiring system. Users can inquire the knowledgebase by using the media of text, image and speech. The question may be in different forms of representation to the same semantic content. The system produces answers which have substantial content and form, also it can enlarge its knowledgebase conveniently through learning.

Key words Multimedia, Knowledge representation, Grammar, Semantic, Natural language generation

陈 芳: 女, 1967 年生, 讲师, 现从事多媒体信息处理和人工智能方面的科研和教学工作.

袁保宗: 男, 1932 年生, 教授, 现从事多媒体信息处理、通信等方面的科研和教学工作.