

基于领域的语用研究和实现

赵礼彬, 赵克, 许威, 林丽姝

(西安电子科技大学机电工程学院, 陕西 西安 710071)

摘要:要准确理解自然语言,不仅需要与自然语言进行语义理解,更需要把语义理解和具体的语境结合,即要进行语用方面的研究。在教学智能辅助系统中,我们对语用的实现进行了一些尝试,把自然语言理解的语义理解结果和教学中的某个领域(这里是行程领域)结合研究,通过语用推理得到深层隐含信息,能够比较完整地理解自然语言的本质。

关键词:语用学;语用推理;语境假设;隐含信息

中图分类号:TP18

文献标识码:A

文章编号:1673-629X(2007)06-0049-04

Research and Realization of Pragmatics Based on Given Domain

ZHAO Li-bin, ZHAO Ke, XU Wei, LIN Li-shu

(Mechatronics College of Xidian University, Xi'an 710071, China)

Abstract: To understand natural language accurately, it is not enough to research natural language semantic understanding just, furthermore, it demands to combine the semantic with the language environment, namely the research of the aspect of pragmatics. In the intelligent teaching assistant systems, have made some attempts to achieve pragmatics. Combine the result of semantic understanding with the special teaching domain, and get the concealed knowledge through pragmatics consequence, so people can understand the essence of natural language in extenso.

Key words: pragmatics; consequence; circumstance-hypothesis; connotative information

0 引言

人们在交往过程中,往往用语言来表达自己的思想,也通过语言来理解别人的思想。但是单单理解语言的语义对语言理解来说是不够准确的,许多隐含的信息语义处理不能够挖掘出来。只有把语义和具体的语境结合起来理解才能够正确地反映语言理解的实质。这种把语境和语言理解结合起来进行研究属于语言学中的新领域——语用学^[1]。语用学是对语言与使用环境之间关系的研究^[1]。语用学在语言学中占有非常重要的位置,目前在自然语言理解领域,语用学是一个比较薄弱的环节,如何把语义和语境有机结合推理出文字背后的隐含信息是自然语言理解研究中的一个难点^[2]。我们做的教学智能辅助系统就对这方面进行了一些探索性的尝试。把语义和行程领域(一个具体的语境)结合起来,在语义理解的基础上,挖掘在行程这个领域内隐含的信息,准确地表达出语言的信息。

同时,这个系统具有一定的通用性,适合在多个领域中应用。

1 语用学

语用学是对语言与使用环境之间的关系(某些部分)的研究。使用环境包括像人和物这样的本体,也包括话语的上下文等等^[3]。在该文中,主要研究的是语言与应用领域之间的关系。

语用学分成两大流派^[4]:社交语用学和认知语用学。这里主要讨论与教学智能系统有关系的认知语用学。认知语用学是以 Sperber & Wilson^[5]为代表从认知角度出发,提出关联理论,探讨人们如何对某一语言行为产生理解,人们大脑中发生了什么。关联理论不以规则为基础,而是以准则为基础,话语的内容、语境和各种暗含,都会使听话人产生不同的理解,但听话人不一定在任何场合下对话语的所有含义理解,他只用一个单一的标准去理解话语,这个标准足以使听话人认定一种唯一的可行理解,这个标准就是所谓的关联。关联^[4]的含义是:当且仅当新的信息在某一特定语境下取得语境效果,该信息才算与该语境有关联。认知理论^[4]认为:语言交际行为是一种有目的、有意图的活

收稿日期:2006-09-08

基金项目:科技部科技型中小企业创新基金(01c26226111002)

作者简介:赵礼彬(1982-),男,浙江东阳人,硕士研究生,研究方向为人工智能、知识工程、创新技术;赵克,教授,硕士生导师,研究方向为人工智能、知识工程和创新设计。

动,它要传递的是说话人的意图,也就是说,交际对于说话人来说有一个明示行为。起意图涉及两种:一是信息意图,只提供交际内容的意图,听话人可以从中获得语境假设;二是交际意图,它让听话人明白说话人有一个传递信息的意图,即从说话人话语行为推断出其意图。所以关联理论认为,明示和推理是交际中不可缺少的两部分。

2 行程问题的分析

文中所说的行程问题是指研究物体运动的速度、时间、距离三者之间关系的应用题。这里的行程问题只处理一维空间中的行程运动。在行程问题中,每个行程运动主体都有运动的时间、运动时所出的具体位置以及运动的速度信息。按照行程中的运动主体的个数分类,可以把行程问题分为:单主体行程问题和多主体行程问题,如图 1 所示。

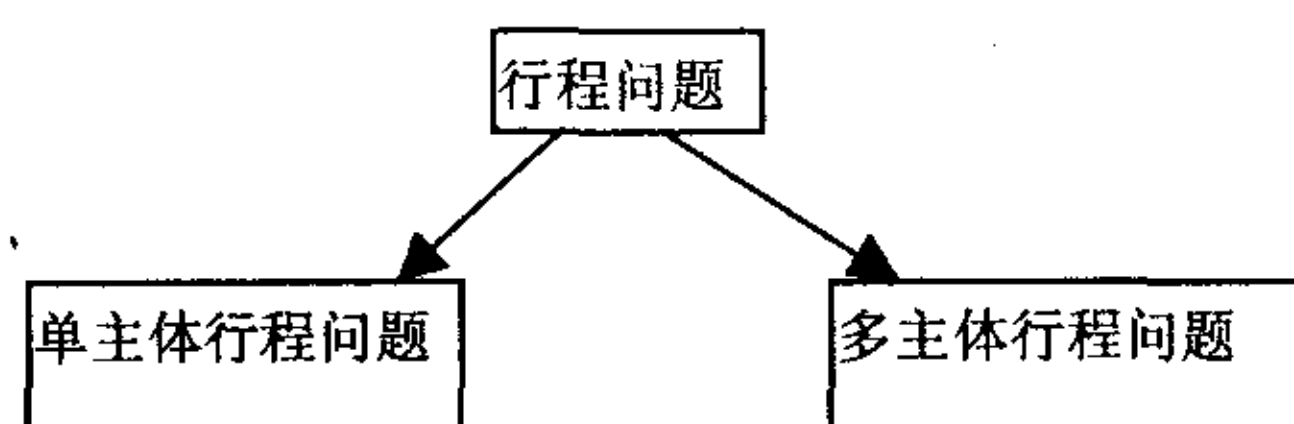


图 1 行程问题分类

单主体行程问题就是只有一个运动主体的行程,它根据运动物体的运动时间、速度和距离的变化情况,又可以把一个行程问题“分而治之”,切分成一个或者两个以上的阶段进行处理。而多主体行程问题就是由两个或者两个以上的运动主体组成的行程,它按照运动主体之间的运动关系又可以分成:追击问题、相遇问题和相背问题。追击问题就是两个行程主体“一前一后”运动,后面的主体对前面的主体追击,两个主体的速度方向是同向的;相遇问题是两个行程主体反向行驶,并且在一个定点遇上;相背问题是两个主体从同一个点出发,但是行驶的方向相反,是背对背行驶,两者越走越远,如图 2 所示。

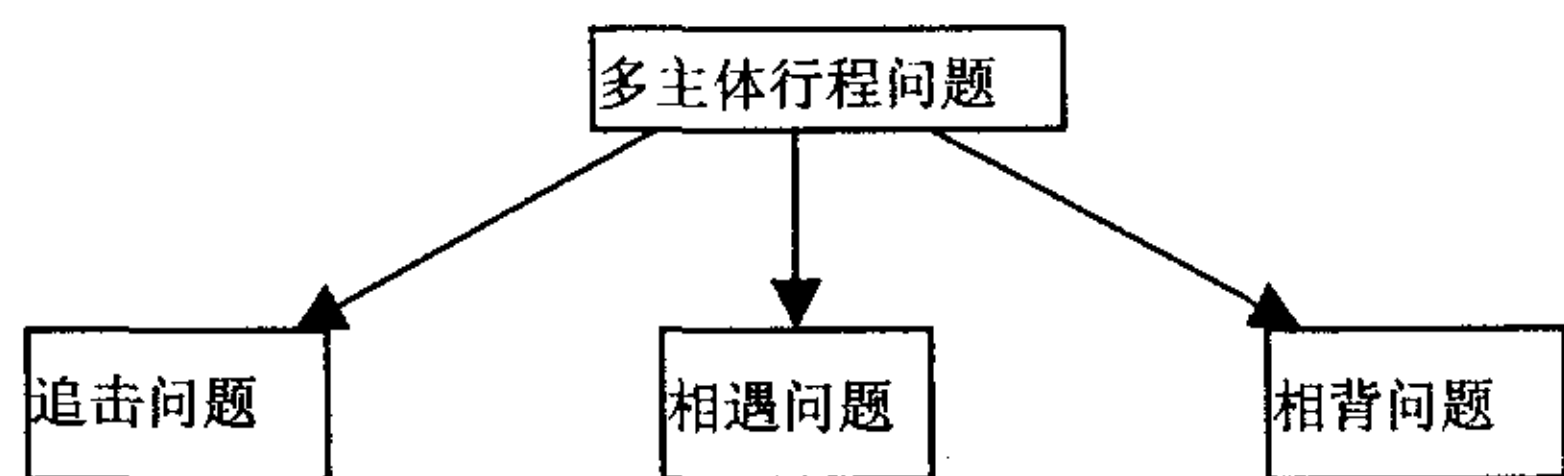


图 2 多行程问题分类

由于行程问题具有多主体和多阶段性的特征,要充分提取问题所表达的显式和隐式信息,就必须首先按主体和阶段分别分析清楚,最后再将主体和阶段综合起来分析。

3 行程领域中的应用模板

在本系统中,以模板的形式来进行知识表示,下面

介绍一些需要的模板。

语境假设就是我们通常所说的语境^[6],只是语境假设是存在于人的大脑中的已有信息。认知语用学将认知语境看成一个心理结构体,即存在于听话人脑中的一系列假设,即在语言使用过程中,语言使用者在处理信息获取含义的过程中,“通过经验或者思维已经将具体语境内在化、认识化了,这个语用因素内在化、认识化的结果就是大脑中的认知语境”。在行程领域中,本系统建立一个静态的模板来描述对行程领域的认识。

行程问题的知识表达要能够充分反映行程问题的特征,如距离、速度、时间等。领域知识的表示是为了把人们对语境的内在认识形式化,为后面进行的语用推理提供依据,我们建立了几个模板,如下:

(1)静态问题模板:用来描述领域中问题的所有必要要素,是体现行程领域语境特征的模板,即行程领域的语境假设。在系统中定义了下面的几个槽:对象、要素、问题辅助和事件(见表 1)。

表 1 静态问题模板中各个槽的定义

槽名	注释
对象	行程问题中涉及的所有行程主体,需要建立行程主体概念树
要素	行程问题中的要素,如时间、速度、路程等
问题辅助	行程问题中具体的辅助要素,如副词同时等
事件	行程中的瞬时事件和阶段性事件,如相遇事件等

(2)动态问题模板:是具体的话语语义结果与行程领域的语境结合产生的语境效果的模板,是行程问题中信息挖掘的主要原材料。动态问题模板中定义了几个槽:对象、要素、问题辅助和事件(见表 2)。

表 2 动态问题模板中各个槽的定义

槽名	注释
对象	具体行程问题中涉及的所有行程主体
要素	具体行程问题中的要素,如时间、速度、路程等
问题辅助	具体行程问题中具体的辅助要素,如副词同时等
事件	具体行程中的瞬时事件和阶段性事件,如相遇事件等

在行程问题中有一些模板用来记录行程主体的信息,记录行程主体在某个阶段的运动速度、阶段持续时间、主体运动的地点,以及行程主体在这个阶段的起始时间、阶段的终止时间、阶段的起始地点以及阶段的终止地点,见表 3。把这些信息记录在主体模板当中,它记录的就是每个主体划分阶段之后的信息,其中包含部分隐含信息。

行程中的主体不是相互孤立的,它们之间包含一定的关系,为了能够挖掘行程中的深层含义,不仅要比

较主体之间的关系,还要比较行程主体在划分阶段后的不同主体的不同阶段之间的关系。文中采用了一个行程关系模板,记录不同行程间的关系。行程主体之间的关系包括行程主体之间的速度关系、行程主体之间的时间关系和行程主体之间的路程关系,见表4。

表3 主体模板中各个槽的定义

槽名	注 释
主体号	一个代表主体的数字
阶段号	一个数字,表示主体所处的阶段
起始时刻	主体在该阶段的起始时间
终止时刻	主体在该阶段的终止时间
起始地点	主体在该阶段的起始地点
终止地点	主体在该阶段的终止地点

表4 行程关系模板中各个槽的定义

槽名	注 释
主体	记录相互比较的两个主体
阶段	记录相互比较两个主体阶段
速度	记录主体比较后的速度差值,没有关系就记录为“空”
时间	记录主体比较后的时间差值,没有关系就记录为“空”
路程	记录主体比较后的路程差值,没有关系就记录为“空”

4 隐含信息语用推理

所谓语用推理^[4]就是根据语用准则,结合一定的语境,从语句的字面意义推导出其话语意义的推理。在行程领域中,隐含的信息需要一个推理过程才能发现。

隐含的信息可以以多种形式隐藏在话语中,如隐藏在副词中,也可以隐藏在行程事件当中。以语句“汽车甲和汽车乙同时从 A、B 两地相向开出,在途中 C 地相遇”为例子,事件“相遇”隐含汽车甲和汽车乙在发生这个事件时处在相同的位置,如果这句话抛开语境不管,那么“相遇”事件关注的侧重点就会不一样,在行程领域关注的是相遇的时刻、相遇的地点和相遇的主体,并且会需要了解相遇时的主体运动速度等信息;然而在天气预报的应用范畴上,当两片云相遇时,关注点就可能是这两片云之间是否存在温度差,是否携带了足够的水分,是否能够引起降雨等。所以即使是相同的事件也可能产生不同的信息,这种情况在通用知识语义处理中是无法解决的。因此,运用语用推理知识挖掘隐含信息,推理过程必须要和具体的语境结合,这样才能准确地描述出话语中的所有信息。

5 行程领域中的语用实现

文中,行程领域的语用实现的思想是:首先将自然语言理解的语义结果和行程领域的语境假设(静态问题模板知识)有机结合,产生一个具体的语境效果(动态问题模板知识),然后根据话语中的副词和行程事件等信息挖掘出隐含信息,最后通过建模得到行程表达式,以数学形式给出行程问题的信息。

图3给出语用推理的总体流程框架。

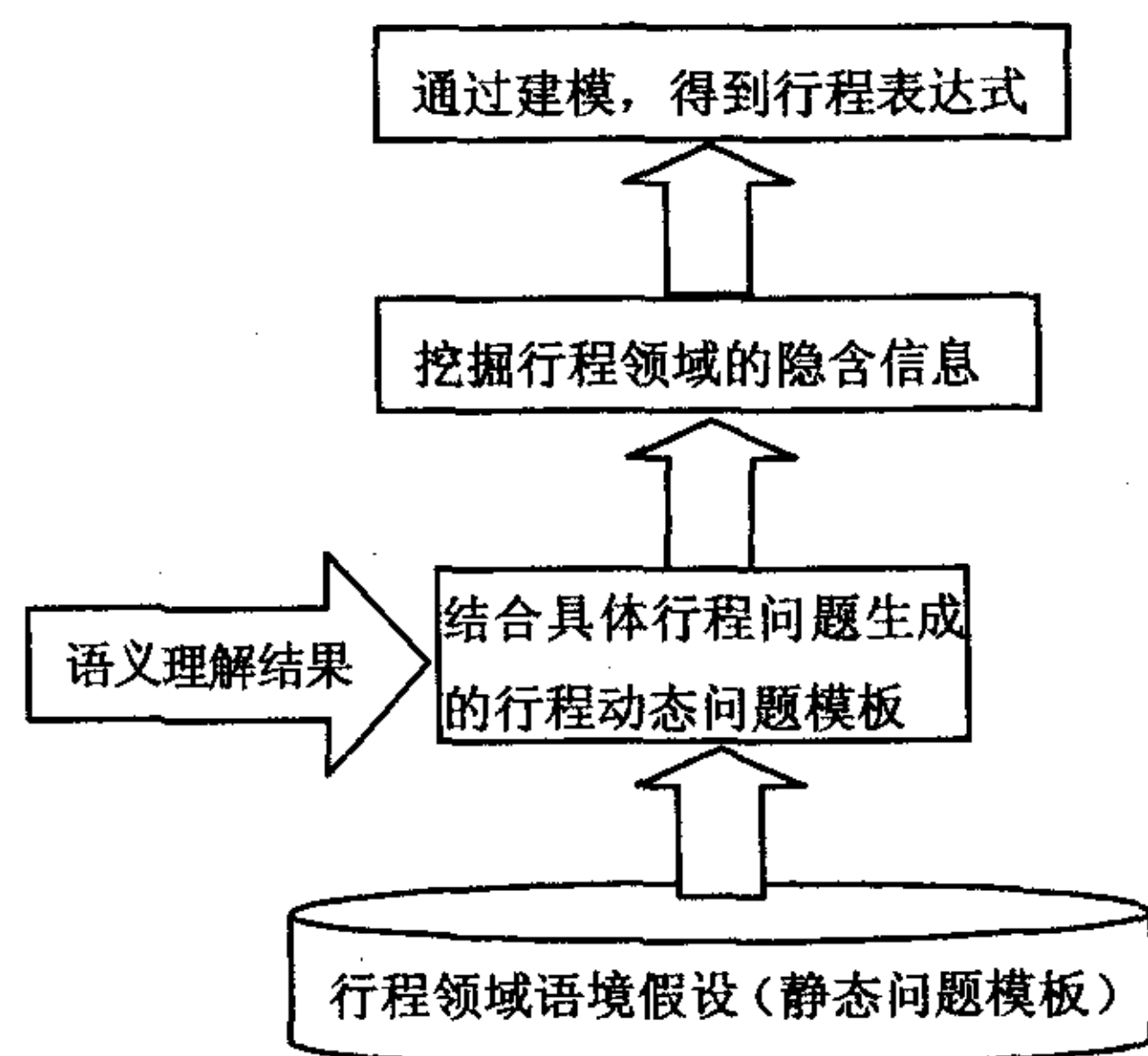


图3 语用推理总体流程

语用理解就是对语义理解的加深理解,处理语义理解没有办法和基础处理的问题,通过和语境结合找到话语中的隐含信息,通过语用推理来挖掘隐含信息。语用推理不同于逻辑推理,在推理过程中需要结合语境效果进行推理。在文中系统中把动态问题模板认为就是行程领域的语境效果。图4给出了在行程中的语用推理流程图,即隐含信息挖掘流程图。

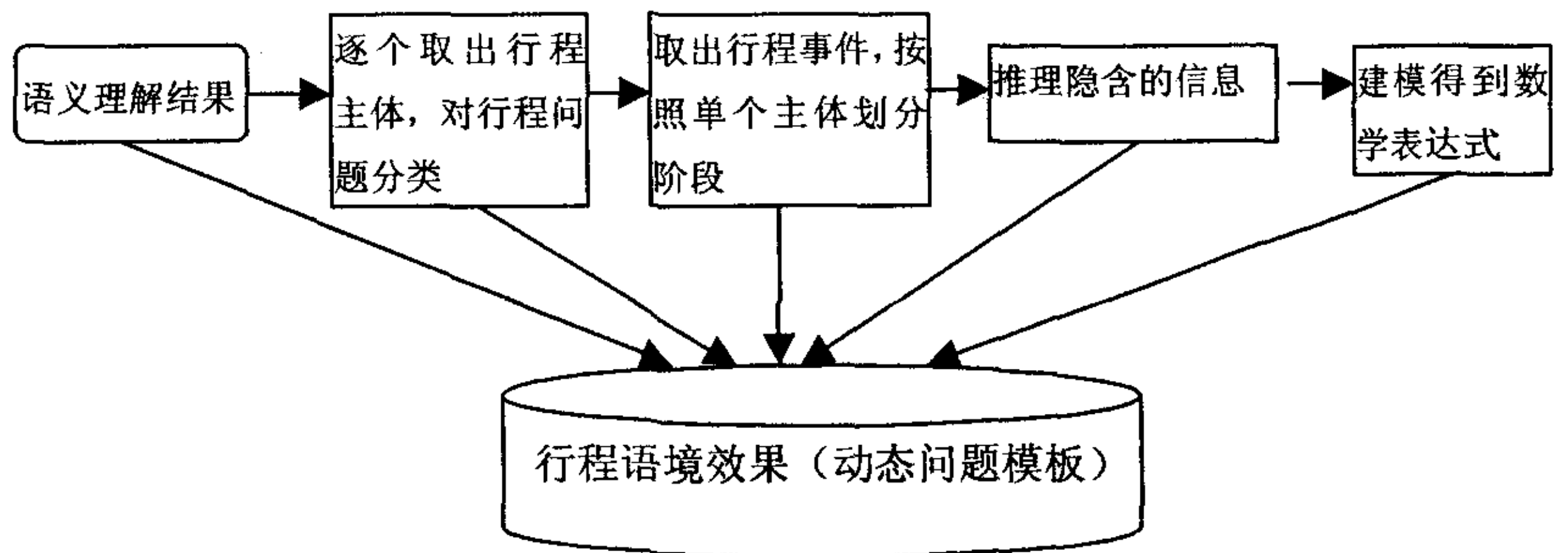


图4 隐含信息挖掘流程图

下面以句子“汽车甲和汽车乙同时从 A、B 两地相向开出,在途中 C 地相遇”为例来说明行程领域中语用的实现过程。经过自然语言的语义分析,并且根据行程领域的语境假设(静态问题模板),可以得到一个语境效果(动态问题模板):

(动态问题(主体 汽车甲 * 汽车乙) (要素)
(问题辅助 同时) (事件 出发 * 相遇))

然后根据动态问题模板中的信息来进行语用推理,挖掘出语义结合语境后更深层次的信息。先按照行程中的主体来进行扫描,取出汽车甲(主体1),汽车

甲在整个行程中经历了从出发到和汽车乙在途中相遇的过程,把它划分为一个阶段,出发表示一个阶段的开始,相遇表示一个阶段的结束。同理,汽车乙也可以分成一个阶段。由于行程中的隐含信息可以由副词产生,也可以由事件产生,通过对主体模板来挖掘隐含的信息。

根据出发和相遇的两个事件,可以描述出汽车甲的主体模板:

(主体(主体号 1)(阶段号 1)(起始时刻 StartTime11)(终止时刻 EndTime11)(起始地点 A)(终止地点 C))

同理,可得到汽车乙的主体模板:(主体(主体号 2)(阶段号 1)(起始时刻 StartTime11)(终止时刻 EndTime11)(起始地点 B)(终止地点 C))

上面的两个主体模板就把隐含信息挖掘后的结果给出,由于两个主体的出发事件是“同时”发生的,所以两个主体的阶段起始时刻槽设置成相同;由于“相遇”事件的发生,两个主体的阶段终止时刻槽和阶段终止地点槽相同。

如果行程主体之间只进行孤立的处理,这样得到的结果是不全面的,必须处理不同行程主体之间的相互关系。通过对主体模板的比较分析,可以得到行程关系模板:

(行程关系(速度关系空)(时间关系 0)(路程关系 0)(阶段号 1- * 1)(主体号 1- * 2))

这个行程关系模板表示主体 1 的第 1 阶段和主体 2 的第 1 阶段进行比较,速度关系槽中的空表示速度上不能得到结果,时间关系和路程关系槽中的 0 表示两个主体阶段的差值为 0,也就是主体 1 第 1 阶段和主体 2 的第 1 阶段在速度上和在运行的路程上都是相等的。根据给出的行程关系,然后通过系统建模,就能得到这个行程问题的表达式, $T11 = T21, S11 = S21$ 。这样就挖掘出了行程中的隐含信息:两个运动主体的

运动时间和运动路程是相同的。通过这样的流程处理就比较完整地理解了行程问题中的信息。

6 结束语

文中介绍了自然语言研究中语用学的一些理论,着重提到了认知语用学的一些研究成果。并且结合教学智能辅助系统,讨论了在语用研究中很重要的两个方面的内容:行程领域语境假设和行程领域的语用推理。最后给出了语义处理结果如何与具体语境结合产生语境效果,再根据行程领域的语境特征运用语用学理论实现行程领域中的语用推理,推导出一段自然语言在行程领域中的隐含信息,然后以数学形式给出全部信息。本系统要实现语用理解必须要建立一个关于应用领域的描述,即给出语境假设^[2],所以一个领域的理解程度和该领域的描述状况对语用的理解、对隐含信息的挖掘具有相当重要的影响。建立合理完整的语境假设,才能够理解好自然语言。对于其他领域的语用实现,只需要建立一个属于该领域的完整的静态问题模板,并且和语义结果结合,就能推理出隐含信息。文中的系统留出了一些接口,便于领域扩展。

参考文献:

- [1] Allen J. 自然语言理解[M]. 第 2 版. 北京:电子工业出版社,2005.
- [2] 孙红梅. 认知语境和语言的理解[J]. 宜宾学院学报,2002, 4(13):43-45.
- [3] Jurafsky D, Martin J H. 自然语言处理综合[M]. 北京:电子工业出版社,2005.
- [4] 王 微. 论认知语用推理中的传统语境因素[J]. 天津外国语学院学报,2005,12(2):44-46.
- [5] Sperber D, Wilson D. Relevance: communication and Cognition[M]. 2nd edition, Oxford: blackWell, 1995.
- [6] 刘根辉,李德华. 计算语用研究的方法和途径[J]. 计算机工程应用,2005,25(1):4-8.

(上接第 48 页)

- tems,2003,E86-D(10):2114-2120.
- [7] Nomoto H, Katahira M, Fukatsu T, et al. Independent test [C]//Proceedings of the First IAASS Conference - Space Safety, a New Beginning. Nice, France: European Space Agency,2005:501-506.
- [8] Yamaura T. How to design practical test cases[J]. IEEE Software,1998,15(6):30-36.
- [9] Indue S, Yamada S. Testing - coverage dependent software reliability growth modeling[J]. International Journal of Reliability, Quality and Safety Engineering,2004,11(4):303-312.
- [10] Nakanishi T, Saeki M. Applying multiple program graphs to

modify specifications[J]. IEICE Transactions on Information and Systems,2000,E83-D(4):669-678.

- [11] Paul R, Ohara S, Tsunoda F, et al. A software test and evaluation environment based on longitudinal database[J]. International Journal of Software Engineering and Knowledge Engineering,2002,12(3):223-244.
- [12] Sawada K, Sandoh H. Software reliability demonstration testing with consideration of damage size of software failures[J]. Electronics & Communications in Japan, Part III: Fundamental Electronic Science,1999,82(5):10-21.