

汉语块扩句的自动分析

韦向峰 张全

(中国科学院声学研究所 北京 100190)

摘要 块扩句是一类概念预期知识十分明确的句子。基于概念层次网络理论介绍了块扩句对应的块扩句类,总结了能够激活块扩句类的典型块扩动词。根据块扩动词的概念知识得到句子的特征语义块及块扩句类,依据块扩句类的知识对句子进行检验后可给出句子句类的分析结果。在已有的句类分析系统的基础上对真实语料中的块扩句进行了自动分析,实验表明正确率达到了71.29%,错误主要来自特征语义块动词辨识、动词多句类代码等。正确分析块扩句将有助于解决汉语句子的多动词处理难点。

关键词 概念层次网络理论,句类,块扩,句类分析

中图分类号 TP391 **文献标识码** A

Automatic Analysis of Chunk-extension Sentences

WEI Xiang-feng ZHANG Quan

(Institute of Acoustics, Chinese Academic of Sciences, Beijing 100190, China)

Abstract The chunk-extension sentence is a kind of sentence, which is defined with specific conceptual knowledge in the theory of hierarchical network of concepts. This paper introduced some semantic categories of chunk-extension sentences and typical head verbs in them. Through these typical verbs, the eigen semantic chunk and semantic category of a chunk-extension sentence can be activated. After testing with the conceptual knowledge of chunk-extension semantic categories, the semantic structure of a sentence can be revealed. An experiment to automatically analyze some chunk-extension sentences from corpus was done based on a system which is designed for analyzing semantic categories of sentences. The precision of experimental result is 71.29%. The errors are mainly caused by recognizing head verbs and multi-semantic-category of a verb. To analyze chunk-extension sentences well will be help to tackle the puzzle of multi-verb in processing Chinese sentences.

Keywords Theory of hierarchical network of concepts, Semantic category, Chunk-extension, Analyzing sentence categories

现代汉语的句子中常出现多个动词,这些动词在句子中扮演着不同的语义角色。除动词本身存在的词性兼类模糊外,动词之间的语义角色模糊也是计算机自动分析句子时的主要难题之一,这就是汉语处理的多动词难点。

针对这一难题,语言学研究者提出了连动句、兼语句以及主谓词组作宾语等概念^[1]。这些概念立足于语法特征,为理解和分析多动词语句提供了理论依据,要把这些理论知识应用到句子自动分析中,除了需要准确的句法分析结果外,还需要把供人理解用的知识形式化为计算机可用的知识。

在概念层次网络理论^[2,3]中,句类就是句子的概念联想脉络类型。语义块则是构成句类的语义单位,是句子中有结构的概念团块,在形式上可表现为一个词、一个短语或者一个句子。表征句子基本语义信息的语义块称为特征语义块,特征语义块类似于句子的谓语,主要由动词组成。而非特征语义块可以包含句子,该句子中的特征语义块又有动词。这样,

不同层级和地位的特征语义块之间的语义关系也反映到动词之间的关系上。

根据特征语义块之间的关系,两个动词之间的关系分为块扩、句蜕、并列。并列关系是指两个特征语义块分属同一层级的两个句类;块扩关系和句蜕关系都是指一个动词属于某个句类的特征语义块,另一个动词则属于该句类的非特征语义块中嵌套句子的特征语义块。块扩关系与句蜕关系的本质区别在于块扩句类有明确的预期知识:前一动词构成特征语义块,能够激活后一动词构成句子的联想,并要求构成的语句充当它的语义块。块扩句与兼语句类似,但它更强调特定动词对句中主语语义块扩展成句子这一情况的预期。下面是关于块扩、句蜕和并列关系的例句:

例1 大家扛着锄头跑了。

例2 我希望大家来。

例3 水门事件导致尼克松下台。

到稿日期:2009-05-08 返修日期:2009-07-17 本文受国家973项目课题(2004CB318104),国家科技支撑计划课题(2007BAH05B02-05),中科院声学所知识创新工程项目(O654091431),中国科学院青年人才领域前沿项目(O754021432)资助。

韦向峰(1976-),男,博士,助理研究员,CCF会员,主要研究方向为自然语言处理等,E-mail:wxf@mail.ioa.ac.cn;张全(1968-),男,博士,研究员,博士生导师,主要研究方向为自然语言处理、机器翻译等。

例1中,“扛着”和“跑了”是并列关系,即该句可以分解为两个句子“大家扛着锄头”、“(大家)跑了”。两个句子属于同一层级,分别构成两个不同的句类。

例2中,“希望”和“来”是块扩关系,两个句子“我希望”和“大家来”不在同一层级上,“大家来”这个句子充当了“希望”所在句类的一个语义块。

例3中,“导致”和“下台”是句蜕关系。虽然“尼克松下台”也充当了“导致”所在句类的一个语义块,但该句类的句类知识没有明确预期这个语义块必须是一个句子。

文献[4]在理论研究、实际词汇和语料分析的基础上,对“块扩”和“句蜕”进行了较为细致的研究,初步理清了块扩的范围。文献[5]系统地研究了 HNC 基本句类的句类知识,在文献[4]的基础上,进一步明确了块扩的内涵,将块扩句类细分为先验块扩和条件块扩两类,确定了 8 组先验块扩和两组条件块扩的具体句类和块扩条件,并将块扩作为一项知识融入到句类知识中,为计算机自动分析处理这一语言现象提供了先验知识。

本文将以此为基础,具体说明各块扩句类及其特性,结合具体语料给出能激活块扩句类的词语,在词语概念知识库和句类分析方法的基础上实现汉语块扩句的自动分析。通过实验观察块扩句的处理情况,分析错误原因和改进方向。块扩关系是处理多动词难点的一个有效突破口,块扩句鲜明的句类知识为分析块扩句的语义结构提供了良好的知识基础。

1 块扩句类与块扩动词

句子意义不是词汇意义的简单堆砌。要理解句子意义还需要建立句子层面的概念联想脉络,将组成句子的词语形成有机的结构体。句子层面的概念联想脉络就是句类。概念层次网络理论按照作用效应链^[2],划分出 57 组基本句类。基本句类可以进一步组合形成混合句类。

$$SC = GBK1 + EK + GBK_m (m = 2 - 3) \quad (1)$$

$$SCR = SC + fK_m \quad (2)$$

句类的一般形式用式(1)和式(2)表示。其中,SC 是句类描述,SCR 是增加辅语义块的描述。辅语义块 fK_m 根据句子表述的需要添加,数量不定;主语义块一般不超过 4 个,包括 GBK(广义对象语义块)和 EK(特征语义块),具体到 57 组基本句类中有更具体的符号表示^[5]。EK 一般直接采用其句类代码作为表示。GBK 尽管在不同句类中有其各自的表示符号,却都来自 3 类共同的基元,分别用符号表示为 A, B, C, 其中 A 表示作用的制造者,或者施事; B 表示作用的承受者,或者受事; C 表示作用的内容。

块扩句类是指句类表示式中 GBK2 或 GBK3 以一个句子的形式出现的句类,简单地说就是某个句子的宾语位置是另一个句子。这一知识是作为句类知识先验地赋予特定句类的。换句话说,块扩是某些句类的特性,是可预期的。

在 HNC 理论的 57 组基本句类中,有 8 组属于先验块扩句类,还有 2 组属于条件块扩句类。所谓先验块扩,就是句类知识约定需要块扩;而条件块扩则是指激活该句类的动词对应的 HNC 基元映射符号需要取特定的值,只有在这种情况下,句类才具有块扩的特性。先验块扩的句类,根据句类就可以判定其具有块扩的特征;而条件块扩句类,除了根据句类外,还需要根据 EK 核心的 HNC 概念基元符号进行判断。下

面给出这 10 组句类中一些典型句类的描述。

(1) 块扩作用句

句类表示式为 $X03J = X03A + X03 + X03BC$, $X03BC$ 块扩。激活这一句类的动词常常具有使令、强迫的意义。这种情况实际上就是语言学研究中的部分兼语句。

(2) 信息转移句

句类表示式为 $T3J = TA + T3 + TB + T3C$, 表示 TA 将信息 T3C 传递给 TB, TA 是信息的传出者, TB 是接收者, T3C 块扩。

(3) 块扩判断句

句类表示式为 $DJ = DA + D + DBC$, DA 是做出判断者, DBC 是判断出的内容, DBC 块扩。

(4) 后继反应句

句类表示式为 $X200J = X2B + X2 + X2C$, $X2B$ 是反应者, $X2C$ 扩展成句子, 表示反应者产生反应以后的表现。

(5) 扩展主从关系句

句类表示式为 $R100J = RB1 + R + RB2 + RC$, RC 发生块扩。RC 是 RB1 与 RB2 共同完成的内容。

(6) 扩展双向关系句

句类表示式为 $R0J = RB + R + RC$, RC 发生块扩, 与扩展主从关系句不同的是关系对象的双方同时出现在特征语义块的一侧, RB 实际上包括 RB1 和 RB2 两部分, 两者的关系对等。

(7) 扩展替代句

句类表示式为 $T4a10J = T4B1 + T4a + T4B2 + T4C$, 表示 $T4B1$ 代替 $T4B2$ (或反之, $T4B2$ 代替 $T4B1$) 完成 $T4C$, $T4C$ 块扩。 $T4C$ 是 $T4B1$ (或 $T4B2$) 单独去做的事, 不是双方共同去做的。

(8) 扩展双向替代句

句类表示式为 $T4a0J = T4B + T4a + T4C$, $T4C$ 块扩。这里的 $T4C$ 不是 $T4B$ 同时都去做的事, 而是 $T4B$ 中的不同对象分别做的。

(9) 一般反应句

$X20J = X2B + X20 + XBC$ 。当特征语义块 $X20$ 的核心动词具有“期望(待实现的愿望)”(概念节点 7121)的内涵时, XBC 扩展成句子。 $X2B$ 是愿望的产生者, XBC 是其愿望。

(10) 一般效应句

$Y30J = YB + Y + YC$ 。当充当特征语义块 Y 的核心动词具有“信息的显现”(概念节点 331)的内涵时, YC 扩展成句子。这个句类和前面提到的信息转移句在内容上很接近, 主要区别在于信息转移句是由 TA 给出的信息, 而这里是 YB 显现出来的信息。

表 1 给出了这 10 类块扩句类的分布情况, 句子来自国家语委现代汉语语料库(2001 版)。从表 1 可知, 前 3 类(块扩作用句、信息转移句、块扩判断句)合起来占到了块扩句的 89.9%, 因此在块扩句的分析中需要对这 3 类进行重点分析。

表 1 块扩句类及其分布

| 句类 | 句类代码 | 句子数 | 百分比 |
|---------|-------|-----|--------|
| 块扩作用句 | X03J | 209 | 40.58% |
| 信息转移句 | T3J | 148 | 28.74% |
| 块扩判断句 | DJ | 106 | 20.58% |
| 后继反应句 | X200J | 9 | 1.75% |
| 扩展主从关系句 | R100J | 6 | 1.17% |

| | | | |
|---------|--------|-----|-------|
| 扩展双向关系句 | R0J | 5 | 0.97% |
| 扩展替代句 | T4a10J | 5 | 0.97% |
| 扩展双向替代句 | T4a0J | 7 | 1.36% |
| 一般反应句 | X20J | 15 | 2.91% |
| 一般效应句 | Y0J | 5 | 0.97% |
| 合计 | 10类 | 515 | 100% |

句类的核心语义块是特征语义块,而特征语义块的核心一般是动词。因此,本文根据已经建好的词语知识库(包含4万多个义项,每个义项包括词语的词形、音调、概念类别、HNC概念符号、句类代码、句类知识等),从中整理得到能够产生块扩的动词148个,如表2所列。这些能激活块扩句类的词语在句子中通常构成句子第一层级的全局特征语义块,而之后的动词一般会构成块扩句类中语义块扩展形成的句子的特征语义块。

表2 典型块扩动词

| 句类 | 典型块扩动词 | 动词数 |
|---------|----------------------------------------------------------------------------------------|-----|
| 块扩作用句 | 逼迫、鞭策、差遣、促使、聘请、迫使、强逼、强使、驱使、强迫、强制、唆使、煽动、推荐、责成、指派、指使、胁迫、诱使、教唆、动员、恫吓、督促、威逼、委派、举荐 | 26 |
| 信息转移句 | 暗示、表示、倡议、重申、请示、声明、申请、提议、透露、通知、责问、宣称、反映、声言、妄称、指斥、表白、炫耀 | 18 |
| 块扩判断句 | 猜想、筹划、揣测、揣度、揣摩、猜测、确认、认定、认为、推测、推断、预谋、预想、预期、想象、以为、预定、预计、预料、酝酿、相信、假想、觉得、打算、估计、怀疑、立志、断言、运筹 | 29 |
| 后继反应句 | 懊悔、懊恼、悲伤、气愤、恐惧、苦恼、难堪、恼恨、恼怒、内疚、感动、难过、难受、委屈 | 14 |
| 扩展主从关系句 | 帮助、陪同、配合、团结、统帅、统率、指引、协助、依靠、引导、引诱、援助、支持、指挥、扶持、带领、跟随、勾结、联合、领导 | 20 |
| 扩展双向关系句 | 团结、协作、勾结、合作、联合、缔交 | 6 |
| 扩展替代句 | 托付、代表、代替、委托 | 4 |
| 扩展双向替代句 | 交替、轮流 | 2 |
| 一般反应句 | 期求、企求、情愿、期待、期望、向往、渴望、要求、愿意、自愿、祝愿、祝愿、希望、渴念 | 14 |
| 一般效应句 | 表明、表示、表现、显露、显现、反映、描写、预示、赞扬、展示、展现、显示、发表、意味着 | 14 |

高,纯v概念动词比兼类概念动词高;动词在句子中的位置越靠前,则优先级越高。

假设的广义对象语义块主要依靠l类逻辑概念和特征语义块,即句子中的广义对象语义块以l类逻辑概念和特征语义块为边界。l类逻辑概念是指“把、被、向、对、由”等这些词语,它们通常后接一个语义块。

从假设的特征语义块中的v概念动词可以获得对应的句类代码,一个v概念动词在知识库中可以有多个句类代码,这些句类代码就是假设的句类代码。对这些句类代码也应该做一个优先级的排序,然后根据句类进行检验,以便更快、更准地找到正确的句类代码。

根据假设的句类代码可以得到句类的具体语义块和句类知识,依据句类知识和句子现场的语义块的要素词语信息,可进行句类知识的检验。这些检验包括语义块个数的检验、语义块相对位置的检验、语义块要素词语概念符号的检验、语义块内部构成方式的检验等。

如果一个v概念动词有多个句类代码,那么对该动词的所有句类代码都依据句类知识进行检验。若通过则得到句子的句类、语义块及其内部构成等结果。如果动词的所有句类代码都没有通过,那么继续对排在下一位的动词的句类代码做检验。对于所有通过检验的结果,取最优的一个句类作为句子的句类分析结果。

整个句类分析的简要过程如图1所示。关于句类分析更为详细的论述可以参阅文献[6-8]。

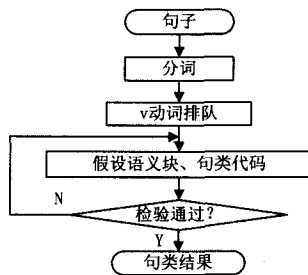


图1 句类分析过程

2 块扩句类的自动分析

2.1 句类分析基础

对一个句子进行句类分析,其首要任务是分词,即把句子切分成词语。然后根据词语的概念符号知识获得假设的特征语义块和广义对象语义块,通过特征语义块获得假设的句类代码和句类表示式,根据句类知识和句子中的语义块要素进行检验。检验通过,则获得句子的句类及其语义块;检验通不过,则假设的句子句类不正确。

为此,建立了词语概念知识库。知识库中包含了词语的词形、概念类别、概念符号、句类代码、句类知识等。根据知识库可以获得词语的概念类别,动词一般用“v类概念”表示。

假设的特征语义块主要来自v类概念动词,而一个句子中往往有多个v动词,因此需要对多个v按优先级进行排序,称为“v动词排队”。排队的主要依据是:动词的前后是否有否定概念,如果有则动词的优先级被降低;单字词的优先级比双字词或多字词的优先级低;有先验块扩句类代码的动词比没有的动词高,前后有特征语义块修饰概念的动词比没有的

2.2 块扩句类的处理

如果一个句子出现了两个动词,而且第一个动词是块扩动词,即能形成块扩句类EpJ,那么第二个动词必然要构成句类ErJ,并且ErJ充当句类EpJ的一个语义块。这一推论是由块扩动词的概念联想脉络和块扩句类预期知识得到的。它也可以推广到多个动词的情况。也就是说,如果一个动词是块扩动词,其后有其它动词,那么该动词优先假设为块扩句类并进行句类知识的检验。

具体到图1所示的句类分析过程中,在“v动词排队”阶段就考虑动词的句类代码。如果含有块扩句类,那么该动词的优先级提高,即排在其它动词前面;在“假设语义块、句类代码”阶段,块扩句类也应该排在其它句类代码的前面,优先进行检验。

此外,在基本格式的块扩句类中,块扩语义块都在特征语义块Ep的后面,即只有GBK2或GBK3可以块扩,特征语义块之前的GBK1不能块扩。因此,在基本格式的条件下,两个动词中的第二个动词不能假设为块扩句类。如果句子中只有一个动词并且是块扩句类,那么很可能是句类分析系统没有识别出其后的动词,出现了动词辨识错误。

块扩句类和块扩动词的相关知识是句类分析中十分鲜明的先验知识,它可以指导句类分析的语义块感知、句类假设甚至检验阶段,为解决多动词难题提供了一些有效的依据;还可以排除一些不恰当的句类假设和不必要的句类检验,使句类分析系统更加高效,结果也更加精确、可靠。

3 实验与分析

为了测试句类分析系统对真实语料中的块扩句子的处理状况,从人工标注的语料库中选取了 303 个出现块扩的句子。这些句子都是以句号、问号、叹号以及它们和反引号连用为结束标志的,因此句子一般都较为复杂。有的句子还形成了块扩的嵌套,即块扩句的块扩语义块中又出现了块扩句类,句子结构非常复杂。首先用句类分析系统对句子进行全自动的标注,然后将标注结果与人工标注结果进行块扩句类标注的比较,得到的实验数据如表 3 所列。

表 3 块扩句的分析结果

| 正确 | 错误 | 句子总数 | 跨越号 | 正确率 |
|-----|----|------|-----|--------|
| 216 | 87 | 303 | 109 | 71.29% |

在表 3 中,句类分析系统能给出正确块扩句类的句子有 216 句,其中块扩语义块跨越逗号的占 109 句,超过了句子总数的一半。在跨越逗号的块扩句类中,发生块扩的语义块的边界未能很好处理,例如“说、表示、认为、指出、强调”等块扩动词后面常常会跟很长的内容,作为块扩动词激活的块扩句类中的块扩语义块。此外,计算机标注结果中虽给出块扩句类,但没给出块扩语义块内部句类的有 30 多句。从这一情况看,对于块扩语义块的边界以及内部构成的处理还需要进一步改进和完善,使得句类分析的结果更加接近人的标注。

在计算机自动标注出错的 87 个句子中,大部分是因为没有把块扩动词作为句子的全局特征语义块,发生这一类错误的句子有 65 句。如果词语知识库没有给出块扩动词的块扩句类,也会使句类分析系统无法得到正确的块扩句类,发生这类错误的句子有 17 句。还有一些句子的标注结果中,虽然把块扩动词作为全局特征语义块,但是由于动词有多个句类代码,系统给出的句类与人标注的块扩句类不一致,发生这类错误的句子只有 5 句。

例 4 概括起来,主要表现在两个方面:一是盲目崇拜 || 西方的新闻自由, + 总以为 || [# 西方 | 是 | 新闻自由的乐园 #]。

UNKNOWN + X20J + DJ[# DBC #] = jDJ

例 4 是句类分析系统对句子的自动标注结果,“+”表示与前面的句子共享 GBK1,“||”表示主语义块之间的分隔,“|”表示嵌套句子(句蜕、块扩)句类中主语义块的分隔,“[# ”和“#]”分别表示块扩语义块的左、右边界。句子下面一行的符号表示各个语句对应的句类代码符号,如“DJ”代表块扩判断句,对应“总以为西方是新闻自由的乐园”这个句子。“UNKNOWN”对应“概括起来”,表示系统不知道它是什么句类。

例 5 当时 ||, 市场 || 严重失灵, ++ { 使 | 世界资本主义经济 } || 陷于 || 崩溃。

CnYS * 11J ++ SOJ # SB = { ! 31X03 }

例 5 中,符号“{”和“}”分别表示原型句蜕的左、右边界。句子的正确标注应该是“使 || [# 世界资本主义经济 | 陷于 | 崩溃 #]”,但是由于“使、陷于、崩溃”都是动词,单字词“使”的优

先级比其它两个双字词低,因此最后得到的句子全局特征语义块是“陷于”,没有给出正确的块扩句类。避免这类错误的措施在于对动词排队时不仅要考虑单字词的优劣先级,还要考虑其作为块扩句类时的高优劣先级,才能得到准确的全局特征语义块和句类标注结果。

例 6 尼克松总统就曾经坦言, ++ 现代总统 || 必须学好 || \ { 驾驭 | 舆论 } 的艺术 /。

UNKNOWN ++ Y0J # YC = \ { ! 31XJ } H /

例 6 中的“坦言”没有作为词语收入词语知识库,因此系统无法对其分析得到块扩句类。有的词语虽然收入到词语知识库中,但没有给出块扩句类代码,句子中这些词语充当特征语义块时也会使系统给出错误的分析结果。

例 7 但 { 油价 | 大幅攀升 } || 的确威胁到 || [# 石油安全并 | 影响 | 经济的稳定增长 #]。

T3X03 * 21J # TA = { T2bP * 22J } [# X03BC #] = XY0 * 21J

在例 7 中,动词“威胁”有“T3X03 * 21”和“XY20 * 21”两个句类代码,但人标注的是“XY20 * 21”,并给出了语义块的正确边界是“……的确威胁到 || 石油安全 + 并影响 || < 经济的 | 稳定增长 >”。这里出现了兼类的块扩动词在句子中不充当块扩动词的情况,还出现了两个动词属于并联关系的情形。

总之,块扩句的分析是句类分析中的一个有机组成部分,它涉及到动词的识别与排序、动词句类的假设与检验以及语义块内部的句类构成等。同时,它又是多动词处理难点中的一部分,块扩句的处理也不能与动词之间的句蜕、并联关系脱离开来。

结束语 本文从概念层次网络理论的角度阐述了类似于兼语句的块扩句的分析方法。基于先验预期知识的块扩句类可以从词语知识库中的块扩动词激活,使用句类分析的假设检验方法可以得到句子的句类、语义块和语义块的内部构成结果。实验表明,其对于块扩句类的分析可以达到 70% 左右的正确率。

由于受块扩动词兼类、句子中多个动词之间复杂层次结构关系的影响,加上句类分析系统中词语知识、动词排序、句类代码排序以及检验知识等的不完善,分析结果中还有不少错误。这些错误一方面可通过改进排序方法,增加知识库中的相关知识来克服,另一方面也需要多动词处理中的其它处理环节如句蜕处理、并联句的处理等的帮助。

多动词处理的各个环节都不是孤立的,它们往往会交织在一起,这即使在预期知识非常明确的块扩句处理中也有反映。但是,块扩关系相对于句蜕关系和并联关系毕竟要容易处理得多。本文进一步的研究目标是深入和完善块扩句的分析,明确块扩语义块的边界确定和内部构成分析方法,同时考虑句蜕关系和并联关系对块扩句处理的影响。

参考文献

- [1] 胡裕树. 现代汉语(重订本)[M]. 上海:上海教育出版社,1995
- [2] 黄曾阳. HNC(概念层次网络)理论[M]. 北京:清华大学出版社,1998
- [3] 黄曾阳. 语言概念空间的基本定理和数学物理表示式[M]. 北京:海洋出版社,2004
- [4] 薛侃. 现代汉语的句蜕与块扩研究[D]. 北京:中国人民大学,1999

(下转第 196 页)

FA-SP 仅需要 28.17s。3) 框架公理占编码规模的 5% 以下, 对于不同类型的规划问题表现截然不同。与 PMA-SP 相比, 对于顺序的 Blocks World 问题域, 单纯的添加框架公理几乎没有作用, 求解效率反而会不同程度地下降; 对于并发的 Logistics 问题域, 求解效率有的提高了 3~4 倍, 最好的问题 Logistics11 的提高则超过了一个数量级, 也有部分问题求解效率有所下降。4) 值得注意的是, 对于 Logistics06 问题, SATPLAN2006 和 PMA-SP 均显示超时, 但 FA-SP 却能在 4.33s 内得到最优解, 这表明框架公理有效地对 SAT 判定分支进行了剪枝。

基于 FA 的编码方式理论上的有效性表明, 添加符合逻辑的相关公理可能形成新的‘冗余’的有效编码方式, 其目的是通过增加新的约束形式约简状态空间, 减少不必要的 SAT 求解分支, 代价是编码规模一定程度的增大。规划系统 FA-SP 的实验结果表明, 相对于著名的 SATPLAN2006, 增加框架公理不太适用于 Blocks World 等顺序规划问题域中, 而对于 Logistics 等并行规划问题域则表现出了明显的优势。现实世界问题中, 并发规划问题占了绝大多数, 所以基于 FA 的编码方式是一种有效的值得采用的编码方式。

结束语 基于可满足性的规划方法自提出以来, 由于其求解的高效, 得到了研究人员的高度重视。在具体的编码过程中, 为了保证有效性, 不得不添加若干个无法证明可删除的公理。本文设计并实现了基于 FA 的编码方式, 理论上证明了该编码方式的有效性, 实现了相应的规划系统 FA-SP, 通过在 Benchmark 问题上的测试, 证明了上述规划方法的性能较著名的 SATPLAN2006 要高。

基于上述几种编码方式的实验结果分析, 对于占统治地位的并行规划问题, 不仅需要动作互斥的部分放松, 还需要适当地添加框架公理以便用编码规模的小幅度增加换取求解效率的提高。上述对比实验证明了上述结论。有的改进效率有所下降, 这与求解问题本身的性质有关。添加框架公理导致编码必然存在一定的‘冗余’, 冗余的框架公理能否实现其约简 SAT 判定空间的目的, 对于不同的规划域问题也表现出了不同的性能。各种动态公理的适用性是目前相关研究者普遍关心的重要问题。不论编码策略如何实施, 各种编码方式的编码规模与求解效率都是判定一个编码方式(或规划系统)优劣的重要标准。公理的添加、约简和组合是编码过程中需要考虑的重要问题, 对于非经典逻辑下的各种编码方法更是如此^[18,19]。

参 考 文 献

[1] Laborie P. Algorithms for propagating resource constraints in AI planning and scheduling: Existing approaches and new results[J]. *Artificial Intelligence*, 2003, 143(2): 151-188

[2] Cross S E, Walker E. DART: Applying knowledge-based plan-

ning and scheduling to crisis action planning[M]. *Intelligent Scheduling*. San Mateo: Morgan-Kaufmann Publishers, 1994

[3] Muscettola N, Nayak P, Pell B, et al. Remote agent: To boldly go where no AI system has gone before[J]. *Artificial Intelligence*, 1998, 103(1/2): 5-47

[4] Kautz H, Selman B. Unifying SAT-based and graph-based planning[C]//*Proc. of the 16th Int'l Joint Conf. on Artificial Intelligence*. Stockholm: Morgan Kaufmann Publishers, 1999: 318-325

[5] Rintanen J, Heljanko K, Niemelä I. Planning as satisfiability: Parallel plans and algorithms for plan search[J]. *Artificial Intelligence*, 2006, 170(12/13): 1031-1080

[6] Kautz H, Selman B, Hoffmann J. SatPlan: Planning as satisfiability[C]//*The 5th International Planning Competition Booklet*. 2006

[7] 姜云飞, 杨强, 凌应标, 等译. 自动规划: 理论和实践[M]. 北京: 清华大学出版社, 2008

[8] 吕帅, 刘磊, 石莲, 等. 基于自动推理技术的智能规划方法[J]. *软件学报*, 2009, 20(5): 1226-1240

[9] 吴康恒, 姜云飞. 基于模型检测的领域约束规划[J]. *软件学报*, 2004, 15(11): 1629-1640

[10] Bonet B, Geffner H. Planning as heuristic search[J]. *Artificial Intelligence*, 2001, 129(1): 5-33

[11] Kautz H, Selman B. Planning as satisfiability[C]//*Proc. of the 10th European Conf. on Artificial Intelligence*. Vienna: John Wiley& Sons, 1992: 359-363

[12] Kautz H, Selman B. Pushing the envelope: Planning, propositional logic, and stochastic search[C]//*Proc. of the 13th National Conf. on Artificial Intelligence*. Portland: AAAI Press, 1996: 1194-1201

[13] Kautz H, McAllester D, Selman B. Encoding plans in propositional logic[C]//*Proc. of the 5th Int'l Conf. on Principles of Knowledge Representation and Reasoning*. Cambridge: Morgan Kaufmann Publishers, 1996: 374-384

[14] Blum A L, Furst M L. Fast planning through planning graph analysis[J]. *Artificial Intelligence*, 1997, 90(1/2): 281-300

[15] Lin H, Sun J G, Zhang Y M. Theorem proving based on extension rule[J]. *Journal of Automated Reasoning*, 2003, 31(1): 11-21

[16] 赖永, 欧阳丹彤, 蔡敦波, 等. 基于扩展规则的模型计数与智能规划方法[J]. *计算机研究与发展*, 2009, 46(3): 459-469

[17] 孙吉贵, 李莹, 朱兴军, 等. 一种新的基于扩展规则的定理证明算法[J]. *计算机研究与发展*, 2009, 46(1): 9-14

[18] 吕帅, 刘磊, 李莹, 等. 基于模态逻辑 D 公理系统的 Conformant 规划方法[J]. *计算机研究与发展*, 2009, 46(7): 1160-1168

[19] Shi L, Sun J G, Lü S, et al. Flexible planning using fuzzy description logics: Theory and application[J]. *Journal of Applied Soft Computing*, 2009, 9(1): 142-148

(上接第 186 页)

[5] 苗传江. HNC 句类知识研究[D]. 北京: 中科院声学所, 2001

[6] 晋耀红. 基于 HNC 理论的句类分析系统的设计与实现[D]. 北京: 中国科学院声学研究所, 1998

[7] 晋耀红. 汉语理解处理中多动词难点的研究与实现[D]. 北京: 中国科学院声学研究所, 2003

[8] 韦向峰. 基于 HNC 理论的扩展句类分析平台研究[D]. 北京: 中国科学院声学研究所, 2005