

基于规则的汉语句法分析方法研究¹

王鹏, 戴新宇, 陈家骏, 王启祥

(南京大学计算机科学与技术系 南京大学计算机软件新技术国家重点实验室 南京 210093)

摘要: 本文从汉语的句法结构特点出发对基于规则的汉语句法分析方法进行讨论, 从中总结出这类方法的一般特征, 提出实用化的基于规则的汉语句法分析器必须引入其它辅助分析手段才能提高性能和可靠性。

关键词: 句法分析, 上下文无关文法, 汉语, 语料库, 知识库

中图分类号: TP391.2

Researches on Rule-based Chinese Parsing Techniques

WANG Peng, DAI Xin-yu, CHEN Jia-jun, WANG Qi-xiang

(Department of Computer Science and Technology, Nanjing University, State Key Laboratory for Novel Software Technology at Nanjing University Nanjing 210093)

Abstract: In this paper, we discuss the Rule-based Chinese Syntactic Parsing techniques based on the characteristics of Chinese sentence structure. The common features of these techniques are summarized. Based on the analysis of these techniques, we put forward the viewpoint that in order to improve the performance and reliability of the practical Rule-based Chinese Syntactic analyzer, other auxiliary analytic measures must be introduced.

Key Words: syntactic parsing, context-free grammar, Chinese, corpus, knowledge base

一、引言

句法分析在自然语言处理领域中具有十分重要的地位, 同时它也是公认的一个研究难题。以传统的基于转换规则的机器翻译系统为例, 前端的句法分析结果的好坏直接影响到后续处理的正确性。汉语作为自然语言的一种, 具有自然语言所共有的最本质的东西, 因此汉语句法分析有必要借鉴和吸收成熟的语言学理论。同时, 由于汉语的特殊性, 决定了汉语句法分析不可能照搬现成的方法, 必须从汉语自身特点出发, 以先进的语言学理论为指导, 研究和开发适合汉语的句法分析技术。

目前国内很多有代表性的汉外机译系统的汉语句法分析模块都是通过改进已有的经典句法分析算法设计的, 不少已在实践中取得了很好的效果。在这中间, 传统的基于规则的方法又占了主导地位。这是因为基于规则的方法从汉语句子最本质的特征出发, 例如构词法、词组构造法、造句法等, 从宏观上总结出句法规则。无论汉语的句式结构多么自由, 其中蕴涵的最基本的规则是相对稳定的。规则易于表达汉语句子成分的构成规律。其次, 经过长时间的发展和不断改进, 这类方法已较为成熟, 应用在一些实用化的系统中具有较高的效率。

进入 20 世纪 90 年代以来, 基于语料库的统计方法获得了较快的发展, 但是单独用它作为汉语句法分析的例子却不多见, 统计方法不依赖汉语语言学理论, 仅仅依靠从语料库中统计得到的数据进行分析, 本质上是一种经验主义方法。统计方法在处理汉语词的切分歧义、兼类词消解、句法结构排歧等方面有较强的优势, 可以为基于规则的句法分析过程提供有用的统计数据, 从而提高系统的健壮性和适应性, 两者有着融合的趋势。

本文首先阐述汉语句法结构的特点及其对句法分析造成的困难。接着, 通过分析一些基

¹本课题得到国家 863 计划资助, 项目编号为 2001AA114102。

于规则的汉语句法分析器的运行过程得出这类方法的一般特征,然后结合几个句法分析算法来逐个说明这些特征,最后对未来汉语句法分析器的发展趋势作出展望。

二、汉语句法结构的特点

与其它自然语言相比,汉语句法有以下两个主要特征:

(一) 同一词类可担任多种句法成分且无形态变化

汉语既不同于以英语为代表的屈折语,也有别于以日语为代表的粘着语。无论是屈折语还是粘着语,词类与句法成分之间有相对简单的对应关系,而“汉语的词类和句法成分之间的关系是错综复杂的。”^[1]

汉语的根本特点是同一词类可担任多种句法成分且无形态变化。在汉语中,名词除了主要担任主、宾语外,也可以直接担任定语,在一定的条件下还可担任谓语,例如“明天星期天”,“星期天”就作谓语;作定语虽然是形容词的主要功能,但形容词和名词的主要区别在于形容词经常用作谓语和补语,例如:“苹果熟了”,其中的“熟”就作谓语,“小偷跑得真快”,“真快”可作为补语。名词不能作补语,也很少作谓语。动词的主要功能是作谓语,但汉语动词可直接用作定语,例如:“毕业的日子不远了”,动词“毕业”就是定语。此外,动词和形容词直接用作主宾语都不是罕见的现象。总而言之,汉语的词类具有多种句法功能,单纯依靠词类来确定句法成分是行不通的。汉语句子要表达一个完整的意思,更多取决于词或短语间的语义约束,句法成分之间的语法约束反而弱化了,词类的多功能就反映了这一特点。人们只要求“看懂”一个汉语句子,而不必对这句话的句法结构有深入的理解。

与词类的多功能相联系,还有一个兼类词的问题。汉语的词类虽有多功能,但不能无限制地扩充某个词类的功能。例如,不能将作状语也扩充为名词的功能,副词基本上只能作状语。汉语中有个别的词,如“重点、决心”等,它们既有名词的功能,可以作主宾语;又有副词的功能,可以作状语。需要将它们分别归入名词和副词。“重点”和“决心”这些词就成了兼类词。汉语中的兼类词是同形兼类词。无论在不在句子中,形态都是一样的。这点跟英语有很大的不同。英语中的同形兼类词当以不同的词类作为句法成分时,其形态差别还是会显现出来的。而汉语就不一样,如“这篇文章的重点是第3段”中的重点是名词,“这篇文章重点讲述机器翻译问题”中的“重点”是副词,两者在形态上没有任何差异。这使得在词的兼类与词类的多功能之间划一条界限变得困难了,这也是汉语句法分析的一大难点。

(二) 汉语句子的构造原则与短语的构造原则基本一致

汉语的短语在汉语语法体系中有及其重要的地位,这是因为汉语短语的构造原则与句子的构造原则基本上是一致的^[1]。在汉语中,短语是一种静态的抽象句法结构,它不与具体的语句相联系。任何短语的只要能单独站得住,带上句调后就能表示相对完整的意思,而成为汉语的句子。在汉语中,从短语到句子是一种“实现”关系。只要把各类短语的结构和功能都描述清楚了,句子的结构实际上也就描述清楚了,因为句子不过是独立的短语而已。这里“实现”关系是相对于“组成”关系而言的。

在汉语中,由词构成短语是“组成”关系。按照组成方式的不同,可以把汉语短语结构划分成主谓、述宾、述补、定中、状中、联合等各种类型。假设超过短语这个层次,由词直接构成句子,其组成方式或者说句子结构基本上也就是这几种。因此可以说汉语句子的构造原则与短语的构造原则基本上是一致的。在“词组为本”的汉语语法体系中,除只有一个调整的句子外,通常不认为汉语的句子由词直接组成,而是由词先组成短语,再由短语实现为句子。汉语句法的这个重要特点是和其他语言相比较显现出来的。英语句子的结构模式是“主语部分+谓语部分”。汉语句子并不限于这种唯一的模式。各种类型的自由短语都可实现为句子。汉语句式结构的多样性也增加了汉语句法分析的难度。

汉语短语结构的另一个重要特点是各类短语的组成成分又可以是各种类型的短语。这些句法成分包括主谓短语中的主语与谓语，述宾、述补短语，偏正短语中的定语、状语与中心语。这表现出了汉语句法成分特有的套叠现象^[2]。当然，各种自然语言在句法结构上都具有递归性。汉语表现这种普遍递归性的特殊之处在于短语担任不同句法成分时形态不发生任何变化。汉语中的主谓结构不仅可以作更大短语中的宾语、主语和定语（相当于英语中的宾语从句、主语从句和定语从句），而且可以作谓语。由于英语中每个句子必须有一个限定形式的谓语动词，而且也只能有一个限定形式的谓语动词（假设不考虑并列情况），因此在英语中没有“谓语从句”这个概念。但在汉语中，就有主谓谓语句。因此，在结构上，汉语的短语可能相当于英语的短语，也可能相当于句子——简单句或者主从复合句。不仅如此，比短语低一个层次的汉语合成词的组成方式也主要是这些。这正体现了汉语多级语法单位的构造的一致性和简明性。因此，研究短语的结构对理解汉语多级语法单位的构造极富启发性。当然，汉语短语的概念并不能完全覆盖英语复合句的概念。在汉语中，也有“单句”和“复句”之分。复句是比单句高一个层次的句法单位^[3]，复句是由分句组成的。无论是单句还是复句，在汉语中都叫句子。

朱德熙先生认为上述两个特点是关系到对汉语语法的全局认识的。造成这两个特点的根源都在于汉语词类没有屈折语那样的形态变化^[1]。分析汉语句子不能象分析其它自然语言句子那样，仅依靠一些语法公式就能进行。很多情况下需要对经验知识进行收集、整理和抽象。而这些知识正是计算机所缺乏的，它不可能象人类那样能够通过“意合法”绕过句法分析直接“理解”句子的深层语义。

三、基于规则的汉语句法分析方法的主要特征

自然语言是由无限多的句子构成的，计算机能够存储的只能是构成无限多句子的有限的词和有限的语法规则^[4]。计算机所能处理自然语言句法规则的是严格的形式化规则，这些规则必须经过严格的数学定义。用来描述这些规则的语法被称作计算语言学语法。一般的基于规则的自然语言句法分析系统都必须建立在某个形式语法基础上。这些语法理论有短语结构语法、广义短语结构文法、中心语驱动的短语结构语法、词汇功能语法和树邻接语法等等。其中短语结构语法（PSG）是自然语言句法分析中常用的语法理论。它是在 20 世纪 50 年代后半期，由美国语言学家乔姆斯基（Chomsky）提出来的。他把短语结构文法按其生成能力的不同划分成 4 类：

- （1） 无约束短语结构文法，即 PSG，也称 0 型文法；
- （2） 上下文有关文法（Context-Sensitive Grammar，简称 CSG），也称 1 型文法；
- （3） 上下文无关文法（Context-Free Grammar，简称 CFG），也称 2 型文法；
- （4） 正则文法（Regular Grammar），也称 3 型文法。

在这中间，上下文无关文法（CFG）被广泛地应用于自然语言的句法分析。汉语当然也不例外。CFG 可以用来描述任何一种递归可枚举的语言，而且 CFG 有高效的分析算法。但 CFG 在处理自然语言时有生成能力过强的缺陷，使用它进行句法分析容易产生多种有歧义的分析结果，这些结果在句法层面上可能合法，但却不符合语义约束。如何消解句法结构歧义是句法分析的一大难点。这个问题在汉语句法分析中表现得尤为突出，这是由于汉语句子的构造原则与短语的构造原则基本一致，短语结构的多样化导致了句法结构的多样化，从而导致分析结果的多样化，不仅仅是类似英语 NP+VP 的结构。这里所提及的歧义一般指的是句子本身没有结构歧义，但经过句法分析器的处理产生了多种分析结果。所以基于规则的汉语句法分析器如何正确高效地进行句法分析是人们关注的焦点。

关于汉语句法分析的研究有很多^{[5][6][7][8][9]}，其中，论文^[5]给出的“一种生成复杂特征集句法树的汉语句法分析方法”是一个比较典型的基于规则的汉语句法分析方法，它用一棵复

杂特征集多叉树来描述汉语句子的句法结构。多叉多标记树模型（MMT）是我国学者 80 年代初提出的，多叉易于描述汉语句子的复杂结构，而多标记体现了汉语特征的复杂性，每个树节点携带有相应的特征和特征值。

论文^[5]的句法分析器输入的是经过分词、词性标注和兼类处理的词串，输出是一棵句法树。规则库中存放的是上下文无关文法的产生式规则，每条规则就是一棵子树的生成规则。规则的条件部分包括被扫描节点的特征值和约束条件（可选）两项，动作部分包括新根节点生成和特征集合的合并两项。通过限制被扫描节点特征的取值和节点扫描范围来实现句子的确定分析，该分析算法是一种确定性算法。该算法采用人机交互的方式消除句法结构歧义。在碰到歧义时，采用一种对话模式，由系统给出几种可能的结构以及上下文，用户选择一种正确结构。由于该算法是在 90 年代初提出来的，当时大规模的汉语语料库尚不完备，缺乏可以从语料库中学习得到的排歧知识，因此使用人机交互不失为一种可行的方法。这个算法的运行效率较高。由于在分析过程中用于消歧的启发式知识不多，因而当分析一个较为复杂的句子时，会产生大量分析结果。通过完善词典和规则可以缓解这个问题。

分析算法实际上是树的合并过程。分析成功体现为森林能够合并成一棵树，如果森林最终不能合并为一棵树，则分析失败。分析开始，输入句子的每个词都作为一棵树的根结点，每一遍从左到右扫描，当扫描到词节点时根据读入的词查词典获得词的静态特征信息，同时进行规则匹配，根据规则进行节点的合并或上升操作。对于不同的词类和词组进行规则分类，以加快匹配速度。匹配成功，则生成新树，新生成的根节点作为被匹配的那部分原根节点的父节点，它替换了根节点队列中被合并的原根节点。根节点队列中森林子树的数目越来越少，最后合并为一棵树，这棵树的根结点或者为单句，或者为复句^[5]。

句法树的生成包括节点生成操作和特征传递操作。后者是一个特征合一运算的过程。因此，随着合并操作的进行，根节点将携带越来越多的特征信息。这些特征值是从下层节点经过合一运算后传递上来的。当分析森林中只剩一棵树时，根节点（代表整个句子）携带的全局信息也就得到了。

通过现有一些汉语句法分析算法的分析，大致可以得出基于规则的汉语句法分析器有如下几个特点：

- （1） 基于规则的分析算法，确定句子成分之间的搭配关系，把输入的词序列最终规约为一棵句法树或者其它等价的表示形式。
- （2） 必须包括歧义消解机制用来排除不正确的句法结构。
- （3） 有一个知识库，包括规则库和面向句法分析的电子词典。

以上任何一方面都关系到一个句法分析器实际性能。分析速度和正确率是衡量分析器性能的重要指标。两者是此消彼涨的关系。要提高分析的正确率就必须完善歧义消解机制，而这往往会降低分析的速度，反之亦然。不同的系统对这两个指标有不同的要求，关键要在两者之间找到合理的平衡。

四、基于规则的汉语句法分析方法的深入考虑

国内很多研究人员结合汉语特点通过改进国外的一些成熟高效的句法分析算法开发了很多实用化的汉语句法分析器，其中包括：LR 分析算法、线图分析算法和移进规约分析算法等等。例如由长沙工学院周会平等人提出的用于分析汉语句子中词语搭配关系的扩充 LR 分析算法^[6]。它对 LR 算法作了改进，并且结合上下文无关文法，句法成分间的语法关系，最终得到的句法树包含了句子词语之间的语法关系。该算法对上下文无关文法做了扩充，加入了语法关系集合和布尔函数集合，相应地增加了 2 张表格：规则语法属性表和约束函数表。

前者描述了推理机在使用某一条规则进行规约时,被规约的各元素之间的语法关系;后者以函数的形式描述了每个规约动作的前提条件,系统设置这些函数可以调整推理机的动作,阻止对错误分析树的进一步分析,从而减少了歧义句法结构的数目。

汉语在词类这个层次上存在很多的歧义结构,这给汉语的句法分析带来了很大的障碍。仅仅依靠词类作为句法分析依据的直接后果是造成大量歧义的分析结果。对某些复杂句分析更会发生组合爆炸,导致分析失败。在句法分析中加入语义信息,限制句法规则过强的生成能力能够明显提高汉语句法分析的效率和可用性。采用这种思路的根本原因在于汉语缺乏形态特征,各类词用法灵活,句子成分间语法约束弱而语义约束强。

北师大的杨开城提出了一种基于句法语义特征的汉语句法分析器^[7]。该分析器的设计思想是从词类中提取句法语义特征,并把句法语义特征根据其在句法分析时的作用分为三个层次:中心词在句法规则中对自身的约束、应用于规则局部上下文(即一条规则内部)中单元间的约束、应用于全局上下文的约束。不同词类具有不同层次的语义特征,如动词具有三个层次的语义特征,而名词、代词等体词只具备第二层次的语义特征。相应的,在词典中也要加入每个词条的句法语义特征以及贡献特征。句法语义特征包括兼类消解规则、状态特征和约束特征。其中兼类消解规则规定了当这个词处在不同的上下文中时应选的词类;状态特征是指词汇自身具有的并且作为句法分析单元中心词可以传递的句法语义特征;约束特征是指词汇在用于生成更大语段时对自己或其它组合单元句法语义特征的要求。贡献特征是指当词条参与语段构造后为组合成的语段贡献的句法语义特征。所采用的句法规则属于增强型上下文无关文法,规则左部包括了执行这条规则的上下文约束以及规则执行后生成的句法语义特征,规则右部除了基本的组合单元以外,还包括了规则单元的制约条件、贡献句法特征、角色、关键字和可选单元。其中,制约条件指的是在符号匹配成功后要进一步检查的特征集,只有制约条件通过该单元才匹配成功;贡献句法特征是句法单元为规则头即规约后的句法成分提供的句法语义特征;角色指的是句法单元是否为中心词、是否为规则头贡献句法语义特征等信息;关键字用于标识一条规则。

通过以上的分析可知,引入句法语义和兼类处理排除了分析过程中大量不合法的句法成分组合,使得句法分析的正确率得到了明显提高。尽管这个算法还有一些不完善的地方,例如缺乏对连词的语义特征描写,但它提出的把句法语义特征加入到汉语句法分析的思路是很有价值的,也是研究人员对汉语本质特征有了更深体会的结果。

汉语句法分析的最终目标是能够高效准确地分析大规模真实文本。基于规则的汉语句法分析方法虽然能抽象出汉语的句法规则,但却难以处理分析真实文本中出现的种种复杂现象。例如某些句法现象很难用规则来加以描述。为了弥补基于规则方法的缺陷,可以引入基于语料库的统计方法辅助句法分析。与前者相比较,统计方法处理语言不确定性方面的能力较强;能够获取微观的统计数据,在消除句法结构歧义方面有一定优势。

清华周强等人提出的基于局部优先的汉语句法分析方法^[8]就是一种把统计方法引入汉语句法分析的典型例子。它利用从语料库中自动获取的结构优先关系数据作为相邻句法成分之间结合的优先判断标准。算法基本思想是对扫描过程中识别出的每个待匹配项 A,根据规则如果发现有歧义结合结果的时候,利用当前待匹配成分 A 的左右语境,分别定义为交段前境和交段后境。从语料库中统计出 A 的前向组合概率和后向组合概率。这些信息可以保存在一张结构优先关系描述项(SPR)表中,每个 SPR 表项的结构是(交段前境,当前待匹配项,交段后境,前向组合概率,后向组合概率)。检索这张表,如果找到一个语境片段和当前分析的相同,则取出前向组合概率和后向组合概率。两者相减,把得到的值和预定义的阈值相比较,大于阈值的话,取左向结合,反之取右向结合。这个分析算法的实质是利用从局部语境中统计得到的概率值对句法结构排歧。

采用上述方法使分析器的整体效率提高了 30%,而且正确率较高。这个算法的缺陷在

于难以优化左右结合概率相当的情况，往往会排除一些事实上正确的匹配组合情况。尽管如此，我们不难发现把统计手段引入到传统的基于规则的句法分析方法后，分析器的效率和健壮性得到了很大提高。基于语料库的统计方法可以在句法分析遇到歧义的情况下提供消歧所需的统计数据。但是经验参数的好坏也会影响到分析器的性能，例如上述算法，如果预定义的阈值选得恰当的话，利用统计进行优化的效果将更明显。好的经验参数除了需要有一个相当规模的汉语语料库作为训练语料库外，成熟的统计模型也是必需的。

知识库是基于规则的汉语句法分析的一个不可或缺的资源，包括规则库和电子词典。规则库一般独立于句法分析模块，这样便于维护；规则库在系统运行的过程中也不是一成不变的，必须易于更新和扩充。汉语句法结构复杂多样，不可能在系统的设计之初就能给出完整的能够解释所有结构的规则库。规则库中的规则是有限的，在分析大规模真实文本时，经常会遇到新的句法现象，而规则库中却没有对应的规则可以解释它或者这种现象和现存的某条规则冲突，因此实用化的汉语句法分析器要求规则库具备可更新和可扩充性。更新用于修正已有规则，扩充用于向规则库中添加新规则，藉此不断完善规则库，提高句法分析器处理“小粒度”句法现象的能力。对规则库进行更新和扩充主要有两种方式：

(1) 从语料库中自动学习句法规则，例如清华苑春法等人提出的基于词性和语义知识的汉语句法规则学习^[9]，这种方法的规则学习和表示利用了词性，语义以及语境约束信息，能自动学习上下文无关的二元规则，并能自动发现词类搭配中的歧义结构。该方法需要一个相当规模的训练树库。采用基于错误驱动的机器学习机制是更新扩充规则库的另一有效途径，初始状态可以给出一系列基本的句法规则，通过学习可以自动获取新的规则。如果训练语料库的选材较广、规模较大的话，那么规则库中通过学习获得的规则将更具体，更具一般性。

(2) 通过人机交互界面录入新规则。这是一种简洁直观的方式，可以由普通用户或者语言学家手动完成。在这种方式下，要求规则在不影响它表达能力的前提下尽量简化书写格式，以便于人的学习。系统也可以提供一种规则的内部转化机制，把用户输入的句法规则转化为机器可读的句法规则。

规则库的不断扩充带来了规则一致性问题，特别是第二种方式，很难保证新加入的规则和规则库中已有规则不发生冲突，因此系统必须提供必要手段保证规则的一致性。为了提高规则查询的速度，可以根据某种分类标准把规则库分为若干模块，每个模块内部按规则的访问优先级从高到低排序。

电子词典作为句法分析知识库的一部分，它存放着每个词条及其相关的各种属性信息，比如词类信息、句法成分信息、语义信息以及语境信息等等。作为分析用的电子词典和规则库一样，必须具有可扩充性，包括新词条的录入和原有词条新属性信息的添加。词典中词条各项属性的定义的和规则中句法成分特征的定义必须一致。电子词典可以采用一种概念层次关系来组织。

基于规则的汉语句法分析器一般还需要一个语料库。从语料库中可以获得诸如词类搭配概率、规则使用概率等方面的统计数据，这些数据能辅助句法分析特别是句法结构排歧，提高分析的效率和正确率。此外，语料库也可作为规则学习的训练文本。这些功能在上面都进行了详细论述。当一个分析器开发完成，一般都要首先在封闭语料库中测试其性能。可见语料库对句法分析器的作用是不可忽视的。

一个完整的基于规则的汉语句法分析系统至少应该包括以下几个部分：基于规则的句法分析模块、规则库、电子词典、语料库以及知识库的维护工具。系统的整体性能除了和句法分析算法本身的效率有关，还和知识库的存储策略与访问策略，知识库维护工具是否健全等方面有关。实用化的汉语句法分析系统是以上几部分的有机组合。

五、启示与展望

本文就基于规则的汉语句法分析方法进行了讨论,给出了这类方法的一些共同特征,特别是当引入语义分析和语料库统计等辅助分析方法后带来的性能和可靠性方面的提高。通过分析,我们可以得到有关基于规则的汉语句法分析的以下几点启示:

(1) 分析不能仅仅依赖于语法知识,这是由汉语的本质特点决定的。扩充知识库的内容,引入词汇的语义、语境等信息,改进传统分析方法能提高分析器的效率。

(2) 采用规则和统计相结合的策略,随着大规模汉语树库的构建和统计模型的成熟,统计方法对句法分析的辅助作用将越来越明显。

(3) 知识库具备可扩充性,把人工智能和机器学习引入到汉语句法分析中来。

(4) 分析器的运行效率和正确率兼顾,视具体应用领域制定分析器的设计策略。

(5) 考察一个实用化的句法分析器的性能往往要把它整合到一个更大的系统中,例如双语机译系统。如果系统的整体性能不高,那么句法分析部分的效率再高也没意义。因此设计句法分析器的同时也要考虑系统的整体性能。

(6) 继续发展和完善汉语形式化理论体系,这是从根本上解决目前汉语句法分析所遇到困难的关键所在。

汉语句法分析是中文信息处理领域内的一个重要课题,它的发展能带动其它相关语言学应用的发展。未来基于规则的汉语句法分析将在规则为主体的前提下,从汉语的本质特征出发,融入多种辅助分析手段,提高分析的效率和整个系统的性能特别是在遇到复杂语言现象下系统的适应性。我们即将研制的基于转换规则的汉外机器翻译系统的句法分析部分就将采用规则为主体,结合其它辅助分析手段的分析策略。

我们坚信,随着汉语研究的不断深入和计算机软硬件技术的不断发展,汉语句法分析及其相关的研究和应用必将获得更大的突破,展示更美好的前景。

参考文献

- [1]朱德熙等,语法答问,北京商务印书馆,1985
- [2]陆俭明,陆俭明自选集,河南教育出版社,1993
- [3]朱德熙等,语法讲义,北京商务印书馆,1984
- [4]赵铁军等,机器翻译原理,哈尔滨工业大学出版社,2000
- [5]赵铁军等,一种生成复杂特征集句法树的汉语句法分析方法与系统实现,中文信息学报,1992,6(4),11~23
- [6]周会平等,用 LR 算法分析汉语的语法关系,软件学报,1999,10(9),967~973
- [7]杨开城,一种基于句法语义特征的汉语句法分析器,中文信息学报,1999,14(3),46~53
- [8]周强等,基于局部优先的汉语句法分析方法,软件学报,1999,10(1),1~6
- [9]苑春法等,基于词性和语义知识的汉语句法规则学习,中文信息学报,2000,15(3),1~8
- [10]蒋兴,汉语句法分析系统的设计和实现,南京大学硕士学位论文,1999
- [11]Masaru Tomita,An efficient augmented-context-free parsing algorithm,Computational Linguistics,1987,13(1~2):31~46
- [12]栾浩等,句法分析系统的知识表示和控制机制,中文信息 CIP 1994 年第 4 期