[19]中华人民共和国国家知识产权局

[51]Int.Cl[7]
G06F 17/27

[12]发明专利申请公开说明书

[21]申请号 99119539.6

[43]公开日 2000年5月10日

[11]公开号 CN 1252575A

[22]申请日 1999.9.2 [21]申请号 99119539.6
[30]优先权
[32]1998.10.26 [33]JP [31]303775/1998
[71]申请人 松下电器产业株式会社
地址 日本大阪府
[72]发明人 郭俊桔

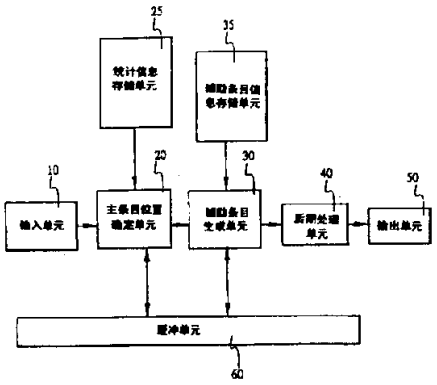
[74]专利代理机构 中科专利商标代理有限责任公司
代理人 朱进桂

权利要求书1页 说明书9页 附图页数13页

[54]发明名称 用于机器翻译的中文生成装置

[57]摘要

用于机器翻译的中文生成装置使用中文生成的统计数据取代大量的与语义、语法相关的规则，统计数据是从加标记的中文语言资料库中搜索的。为克服中文生成需要大量语义或语法规则的问题，中文句子的构成被分成基本句型条目和其他条目。每一中文动词可能生成的基本句型的概率和其他条目的位置概率是从中文语言资料库中搜索并存储在缓冲单元中。使用分支界限法，能找出最佳基本句型和其他条目的位置并生成最适当的中文句子。



ISSN1008-4274

权 利 要 求 书

1、一种用于机器翻译的中文生成装置，其将中文句子的从属结构变换成中文句子，其中所述装置包括:

统计信息存储单元，用于存储中文句子的从属结构的变元条目、可能的句型、每一槽的可能的格位标记排列和相应的概率值;

辅助条目信息存储单元，用于存储格位标记、源语言表面格位标记、变元语义的编码、修饰成分的语义码和相应的句首表面格位标记和句尾表面格位标记;

主条目位置确定单元，用于从输入中文句子的从属结构中搜索子结构，用于以子结构的主要变元作为检索键从统计信息存储单元中搜索在每一槽中的相应的可能的句型和相应的可能的格位标记排列，以及对应的概率值，并且按照评估函数顺序地生成中文句结构;

辅助条目生成单元，用于从中文句结构中搜索页节点条目的格位标记、源语言表面格位标记、变元语义的编码和对应节点条目语义码作为检索键，用于按照检索键从辅助条目信息存储单元搜索句首表面格位标记和句尾表面格位标记，并顺序地为中文句结构生成介词结构；以及

后期处理单元，用于从中文句结构中搜索每一从句结构，用于生成疑问句或“把”字句或否定句或被动句或祈使句和相应的时态标记和标点符号，并用于通过使用线形化方法将中文句结构变换成中文句子。

说　　明　　书

用于机器翻译的中文生成装置

本发明设计用于机器翻译的中文生成装置，其使用统计数据代替大量的语义的和语法规则。

在二十世纪，人们必须不断学习以避免与社会隔绝。然而，大量新的知识是来自外国，在提高读外国文献的效率方面，翻译是非常重要的。为了改善文件翻译的质量和效率，近来的趋势是采用计算机代替人的翻译工作。这样的设备通常称作机器翻译设备。在这样的机器翻译设备中，将被翻译的语言作为源语言，而已经从输入语言翻译的输出语言称作目标语言。例如，日文至中文的机器翻译设备的源语言是日文，目标语言是中文。此外，在机器翻译设备中使用翻译格式可以是直接形式、中间变换形式或核心语言形式，这取决于将要翻译的语言的特征。通常，中间转换形式是常用的一种形式。

参照图 8，采用中间变换形式的常规机器翻译设备包括一源语言语法分析单元 1、一中间结构转换单元 2、一目标语言生成单元 3 和一词典单元 4。然而机器翻译的质量取决于输入句子是否在源语言语法分析单元中被正确地分析，取决于在中间结构变换单元 2 中是否将源语言和目标语言之间的差别消除（例如，解决在文法或意思上的差别，或者是词汇条目翻译的选择），并取决于在目标语言生成单元 3 中是否根据目标语言的语法规则正确地生成目标语言。

然而，中文句子随着在句子中词汇位置的变化将具有不同的意思。例如，在这些句子中，[他*在桌子上*跳]和[他跳*在桌子上*]，由于前面的句子中“在桌子上”的位置不同于后面的句子中“在桌子上”的位置，所以两个句子具有不同的意思。因此，在中文句子中一些词汇的排列具有给定的顺序，除非那样排列，否则将生成不正确的中文句子。下面将是一个例子，其中时间词汇必须被放在地点词汇之前。

（正确的中文句子）他*昨天**在学校*吃饭。

（错误的中文句子）他*在学校**昨天*吃饭。

在另一方面，在中文句子中一些词汇的顺序是不受限制的。下面将是一个例子，其中时间词汇可以放置在主语之前或之后。

（时间词汇放在主语之前）昨天他去学校。

（时间词汇放在主语之后）他昨天去学校。

因此，如果机器翻译设备的目标语言是中文，那么将要解决的最重要的问题是如何正确地确定在中文句子中词汇的排列顺序。参照图 9，在中国台湾的专利公报 3 2 4 8 0 4 中揭示了一种机器翻译使用的中文生成装置。

在图 9 中中文生成装置的预处理单元 200 使用虚节点对省略了如图 10A 所示输入的从属结构(一种中间结构)中的主语的子结构恢复主语节点。接着，通过使用每一子结构的主条目（动词或形容词）的动词分类码作为检索键，基本条目展开单元 300 根据基本句型产生包括基本条目的如图 10B 所示的一基本句子结构，其中的基本句型是存储在基本句型存储器单元 350 中的。

不受限制条目展开单元 400 通过使用在从属结构中的每个不受限制条目的格位标记、源语言的表面格位标记、语义支配码和自身的语义码作为检索键，按照句子条目信息存储单元 450 检索句首的表面格位标记、句尾的表面格位标记和句子条目槽，并按照在句子结构中句子条目槽位置的对应位置生成图 10C 的每一不受限制条目的句子结构。

一特定句型生成单元 500 根据每一动词或形容词的特定句型属性，产生图 10D 的特殊句型句子结构。一条目位置调整单元 600 从句子格式条目顺序存储单元 650 中顺序地检索在句子结构中每一句子条目槽中的条目排列顺序限制，并调整在句子结构中每一句子条目槽中条目的排列顺序，如图 10E 所示。然后，一后期处理单元 700 进行在句子结构上其他辅助条目和标点符号的生成，并且排列该句子结构。一输出单元 800 输出翻译结果“我把这本书放在车子里”。一缓冲器单元 900 被用于临时地存储来自基本条目展开单元 300，不受限制条目展开单元 400 和条目位置调整单元 600 的输出。

前面所述的用于机器翻译的常规中文生成装置存在的缺点如下:

1. 中文的动词或形容词有可能生成的多个中文基本句型，例如，动词“送”可以生成如下的基本句型。(其中 S 代表主语，V 代表动词，O 代表直接宾语或间接宾语，C 代表补语)。

SVOO: 我送他书。

SVOOC: 我送他书当作纪念。

SVOC: 我送他回家。

SVO: 他会送命。

因此，用于机器翻译的常规中文生成装置不能用动词分类表编码解决基本句型中的差别问题。这个问题必须由启发式的方法解决，因此，不能确保翻译的质量。

2、由于不受限制的条目的位置是按照句子条目信息存储单元的内容而不是按照相关条目状态指定的，所以翻译的质量不能得到改善。例如，如果时间词汇“今天”的位置被指定为 2，常规中文生成装置只能生成句子“我今天毕业”，而不能生成强调“今天”的句子“今天我毕业”。

3、由于在同一槽中在不受限制条目中相关位置的调整是与句子条目顺序存储单元的内容相关的，所以当句子条目顺序存储单元的内容不完整时，可能会产生奇怪或不正确的中文句子。

因此，本发明的主要目的是提供一种能够克服上述已有技术所具有的缺点的用于机器翻译的中文生成装置。

按照本发明，一种用于机器翻译的中文生成装置，其使用统计信息代替大量的语义的、语法的和句结构规则并将输入中文句子的从属结构变换成中文句子，该装置包括:

统计信息存储单元，用于存储中文句子的从属结构的变元条目、可能的句型、每一槽的可能的格位标记排列和相应的概率值;

辅助条目信息存储单元，用于存储格位标记、源语言表面格位标记、变元语义的编码、修饰成分的语义码和相应的句首表面格位标记和句尾表面格位标记;

主条目位置确定单元，用于从输入中文句子的从属结构中搜索子结构，用于以子结构的主要变元作为检索键从统计信息存储单元中搜索在每

一槽中的相应的可能的句型和相应的可能的格位标记排列，以及对应的概率值，并且按照评估函数顺序地生成中文句结构；

辅助条目生成单元，用于从中文句结构中搜索页节点条目的格位标记、源语言表面格位标记、变元语义的编码和对应节点条目语义码作为检索键，用于按照检索键从辅助条目信息存储单元搜索句首表面格位标记和句尾表面格位标记，并顺序地为中文句结构生成介词结构；以及

后期处理单元，用于从中文句结构中搜索每一从句结构，用于生成疑问句或“把”字句或否定句或被动句或祈使句和相应的时态标记和标点符号，并用于通过使用线形化方法将中文句结构变换成中文句子。

根据本发明的用于机器翻译的中文生成装置，主条目位置确定单元从输入从属结构中搜索子结构；以子结构的主要变元作为检索键从统计信息存储单元中搜索在每一槽中的相应的可能的句型和相应的可能的格位标记排列，以及对应的概率值；并且按照评估函数顺序地生成中文句结构；以及在缓冲单元存储中文句结构。然后，辅助条目生成单元从中文句结构中搜索页节点的格位标记、源语言表面格位标记、变元语义的编码和对应节点条目语义码作为检索键，按照该检索键从辅助条目信息存储单元中搜索句首表面格位标记和句尾表面格位标记，并在中文句结构的相应位置生成介词结构。

然后，后期处理单元从中文句结构中搜索每一从句结构，按照中文语法结构执行疑问句、“把”字句、否定句、被动句、祈使句和相应的时态标记和标点符号的生成，并最终通过使用线形化方法搜索生成的中文句子并将生成的中文句子输出到输出单元。

通过下面结合附图对本发明实施例的详细描述，本发明的其他特征和优点将更为清楚明显。

图 1 是本发明的一个实施例的用于机器翻译的中文生成装置的系统方块图；

图 2 是本发明的实施例的主条目位置确定单元的处理流程图；

图 3 是本发明的实施例的辅助条目生成单元的处理流程图；

图 4 是本发明的实施例的后期处理单元的处理流程图；

图 5 是本发明的实施例的统计信息存储单元的示意结构原理图；

图 6 是本发明的实施例的辅助条目信息存储单元的示意结构原理图；

图 7A 至 7D 是用于说明本发明的处理过程的图；

图 8 是常规机器翻译装置的系统方块图；

图 9 是用于机器翻译的常规中文生成装置的系统方块图；

图 10A 至 10E 是用于说明已有技术的处理过程的示意图。

为了降低在机器翻译中中文生成所需的语义的、语法的和特殊的句法规则的总数。由于在国内市场上标记的中文平衡语句资料库的出现，我们可以简化一些软件工具以从语句资料库位置提取所需要的信息以代替在机器翻译中使用的规则，例如，从动词或形容词析出的基本句型和各种条目在基本句型中出现的概率。此外，通过简单运算方式的使用，例如条件概率的运算，可以方便地产生其他的相关概率信息，例如，可以从单个单词的概率信息和两个互相连接的单词的概率信息中得出三个或四个相互连接的单词的概率信息。

图 1 是本发明的用于机器翻译的中文生成装置的示意性系统方块图。10 表示用于输入中文从属结构的输入单元。以日译中机器翻译作为一个例子，通过日文语法分析处理和中间结构变换从一日文句子中获得中文从属结构。例如，如图 7A 所示，一动词 V 作为一变元，而如“我”、“今天”这样的页节点是上述变元的修饰成分。标号 25 表示用于存储作为检索键的中文句子从属结构的变元条目的一统计信息存储单元，以及相应的可能的句型及在每一句型中的每一槽的可能情况的排列和相应概率值（例如，在图中的槽 1 至槽 6），存储单元 25 的基本结构原理图如图 5 所示。

标号 20 表示主条目位置确定单元，用于以每一子结构的变元作为检索键分别从统计信息存储单元 25 中搜索相关信息，以及用于根据最佳路径搜索逼近，如分支界限法，确定最佳基本句型和其他条目的相关位置并生成中文句结构，其处理流程图如图 2 所示。标号 35 表示辅助条目信息存储单元，用于存储格位标记、日文标记、变元语义码和语义码作为检索键和相应的句首和句尾标记，其如图 6 所示。

标号 30 表示一辅助条目生成单元，其用于通过用页节点的格位标记、日文标记（和日文词汇条目）、变元语义码和其语义码作为检索键搜索每一页节点，从辅助条目信息存储单元 35 中搜索在中文句结构的相应位置

的相应句首和句尾标记生成中文介词结构，此处理流程图如图 3 所示。标号 40 表示后期处理单元，用于根据线形化方法从中文句结构生成中文句子，并在进行否定、疑问、祈使、“把”字句、被动、时间标记和标点符号处理之后输出中文句子，该处理流程图如图 4 所示。

标号 50 表示如由监视器构成的一输出单元，标号 60 表示用于暂时存储中间结果的一缓冲单元。

图 2 是主条目位置确定单元 20 的处理流程图。在步骤 S201 中在从输入单元 10 发送的中文从属结构中搜索主变元之后，在步骤 S205 中确定主变元是否存在。如果主变元不存在，即，没有从句存在，处理过程进行到步骤 S270 以进行该特定句子顺序调整，例如，复合句“进入禁止”被调整为“禁止进入”。在完成步骤 S270 之后，该处理过程结束。如果在步骤 S205 中确定存在主变元，则执行步骤 S210 以确定是否存在修饰成分。如果不存在修饰成分，执行步骤 S265 以确定在其他修饰成分中是否存在未处理的从句变元。如果在其他修饰成分中不存在未处理的从句变元，则执行步骤 S270 且处理过程结束。否则，在步骤 S215 中以变元条目作为检索键从统计信息存储单元 25 中搜索相应可能的句型和相应槽的可能的格位标记排列和相应概率值。

然后，在步骤 S220 将可修饰成分 i 的初始初始值（可能的句型数）设置为 1。然后，在步骤 S225 将可修饰成分 j 的初始值（槽数）设置为 1。在步骤 S225 之后，处理过程进入到步骤 S230 以确定在槽（SLOT）ij 中是否存在可能生成的格位标记排列。如果在槽 ij 中存在格位标记排列，则通过使用未确定的修饰成分格位标记和相应的统计信息计算每一排列的评估函数值，并在步骤 S235 中将最高评估值的格位标记排列用作槽 ij 的排列。

在使 j 加 1 的步骤 S240 之后，在步骤 S250 中确定 j 值是否大于槽的最大数（在本实施例中槽的最大数是 6）。如果 j 值不大于槽的最大数，处理过程回到步骤 S230，否则，在步骤 S255 确定 i 值是否大于可能的句型数（来自步骤 S215 的结果）。如果 i 值不大于可能的句型数，在其中 i 被加 1 的步骤 S255 之后处理过程返回到步骤 S225，否则，在步骤 S260 中按照评估函数值搜索最佳生成排列。然后，在步骤 S265，确定在未处理的修

饰成分中是否存在未处理的变元。如果在未处理的修饰成分中有未处理的变元，处理过程返回到步骤S210，否则，在其中调整特定句子顺序的步骤S270之后，处理过程结束。

图3是辅助条目生成单元30的处理流程图。在步骤S301中，从主条目位置确定单元20发送中文句结构。在步骤S305，从上向下和从左向右搜索未处理的从句结构。在步骤S310中，如果确定未处理的从句结构的搜索失败，则处理过程结束，否则，在步骤S315中，以修饰成分的格位标记、源语言词汇条目标记和语义码（变元和其自己的）作为检索键，按照辅助条目信息存储单元35从未处理的从句结构中搜索相应句首标记和相应句尾标记。然后，在步骤S320中将具有句首和句尾标记的修饰成分生成为在句结构的相应位置中的介词（PP）结构。在步骤S325，在步骤S320生成的介词结构代替存储在缓冲单元60中的中文句结构的相应从句结构。然后处理过程返回到步骤S305。

图4是后期处理单元40的处理流程图。在步骤S401中，首先由上向下和由左向右地从缓冲单元60中搜索中文句结构。然后在步骤S405中分别由上向下和由左向右地搜索未处理的从句结构。在步骤S410中，如果确定未处理的从句结构的搜索失败，则在按线形化句结构搜索中文句子的步骤S465之后结束处理过程，否则，在步骤S415中，确定未处理的从句结构是否是疑问结构。如果未处理的从句结构是一疑问句，则在步骤S420中执行表示疑问的“吗”和“呢”的标记的生成处理过程，而且程序进行到步骤S425，否则，在步骤S415之后，程序直接进入到步骤S425以确定是否是“把”字句。如果是“把”字句，则执行步骤S430以生成一“把”字句而且过程进入步骤S435，否则，过程直接进行到步骤S435以确定是否是否定句。如果是否定句，则执行步骤S440以生成一否定句而且过程进入步骤S445，否则，过程直接进行到步骤S445以确定是否是被动句或是使役句。如果是被动句或使役句，则执行步骤S450生成被动句或是使役句子而且过程进入步骤S455，否则，过程直接进行到步骤S455以进行时态标记的生成。然后，在步骤S460执行标点符号的生成。然后，生成的从句结构代替在中文句结构中的相应从句结构，而且程序返回到步骤S405。

下面通过一个例子进一步地说明本发明的工作过程。由输入单元 10 输入如图 7A 所示的中文从属结构。然后，主条目位置确定单元 20 根据图 2 的处理流程图工作。由于变元是“送”，按照图 5 的统计信息存储单元 25 可以搜索到下面的信息:

可能句型和它的概率值:

SVO 0.41, SVOO 0.30, SVOC 0.18, SVOOC 0.11

(1) SVO 句型的每一槽（SLOT）的可能格位标记排列和它的概率值:

SLOT1: time 0.2, purpose 0.39

SLOT2: time 0.39, location 0.99, time_at 0.21, time time_at 0.17, time time_at 0.07

SLOT4: location_to 0.25

(2) SVOO 句型的每一槽的可能格位标记排列和它的概率值:

SLOT2: time 0.16

(3) SVOC 句型的每一槽的可能格位标记排列和它的概率值:

SLOT2: time 0.24

(4) SVOOC 句型的每一槽的可能格位标记排列和它的概率值:

SLOT2: time 0.11

最大 SLOT 数: 6

可能的句型数: 4。

在中文从属结构中修饰成分的格位标记: subject, time, time_at, object, loc_to 。

通过使用上述的信息和分支界限算法可以获得图 7 所示的中文句结构，换句话说，根据该计算结果，可知在 SLOT2 中的格位标记时间的排列好于在 SLOT1 中的。然后，辅助条目生成单元 30 根据在缓冲单元 60 中的中文句结构分别搜索格位标记、日文标记、每一页节点的语义码作为检索键，根据图 6 的辅助条目信息存储单元 35 搜索相应的句首标记和句尾标记，例如，time_at 的首标记是“在”，并生成中文介词结构，例如，“九点”将被生成为“在九点”的介词句子。这时生成的中文句结构被显示在图 7C 中。

然后，后期处理单元 40 确定这个句型是一“把”字句，所以进行“把”

字句处理。在这时产生的中文句子如图 7D 所示。使用线形化方法，可以生成中文句子“我今天在九点把书送到学校”。然后，输出单元 50 可以输出这个中文句子到一输出装置，如显示器或打印机。

从前面的描述可以看出，本发明的装置可以克服已有技术存在的问题，即，本发明的优点在于：

（1）由于使用了统计数据信息，用于机器翻译中文生成所需的规则数可以减少 1/3，从而大大地增加了机器翻译的工作效率（速度）。

（2）由于可以克服可能生成的句子中的差异性别并且在句子中每一条目的排列顺序可以同时确定，所以可以大大地改善中文生成的质量。

（3）由于规则减少，系统维护变得容易。此外，由于规则之间的竞争减少，所以翻译质量更稳定。

在结合最佳实施例描述了本发明的同时，应认识到本发明是不限于所揭示的实施例，而是趋于覆盖在本发明的范围内的各种变化的结构，以致于包括所有的这些变化形式和等同结构。例如，源语言不限于日语，或输入的语言结构可以直接是句结构，而不是从属结构。

说 明 书 附 图

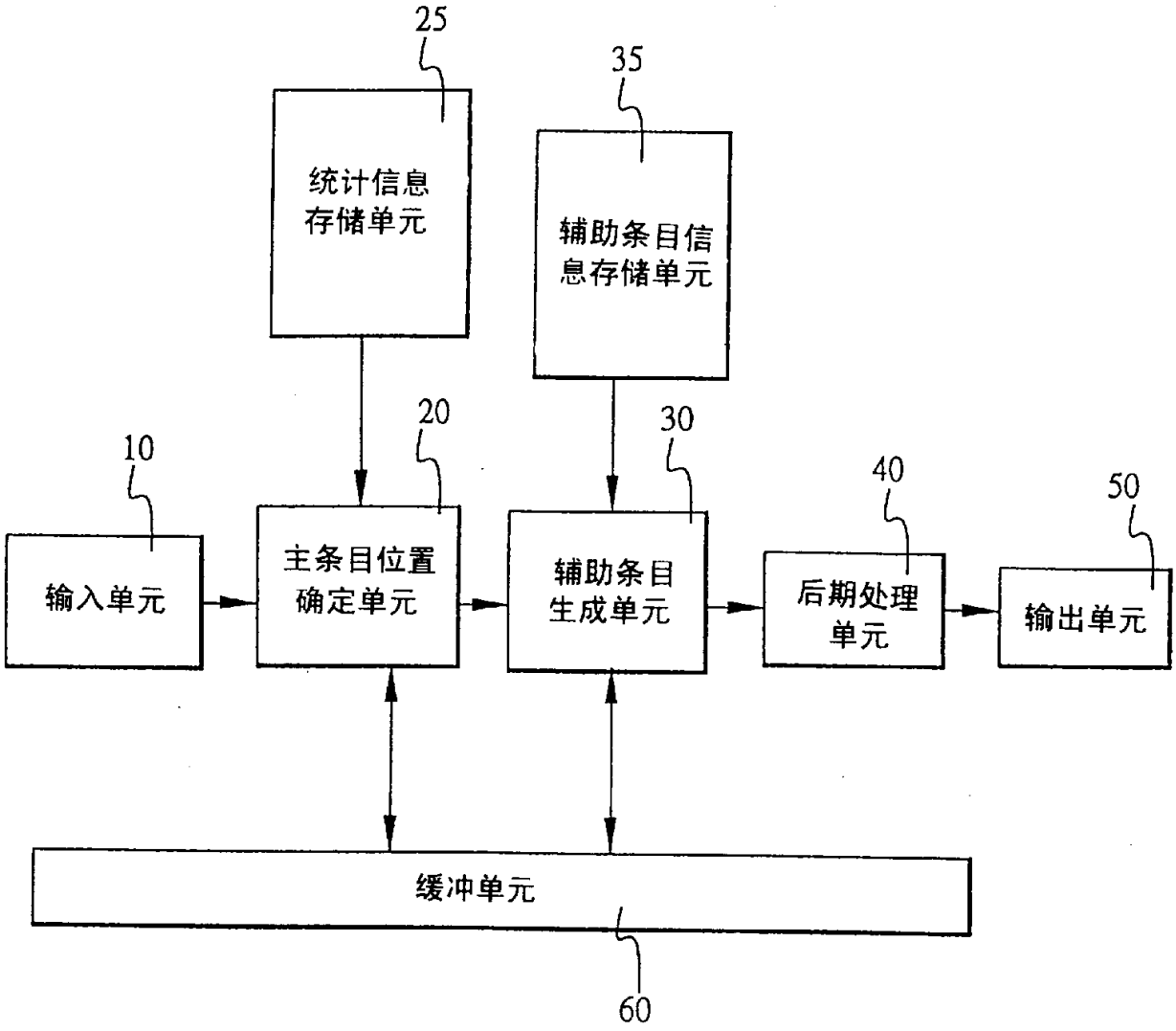

25
35
统计信息
存储单元
辅助条目信
息存储单元
20
30
10
40
50
输入单元
主条目位置
确定单元
辅助条目
生成单元
后期处理
单元
输出单元
缓冲单元
60

图 1

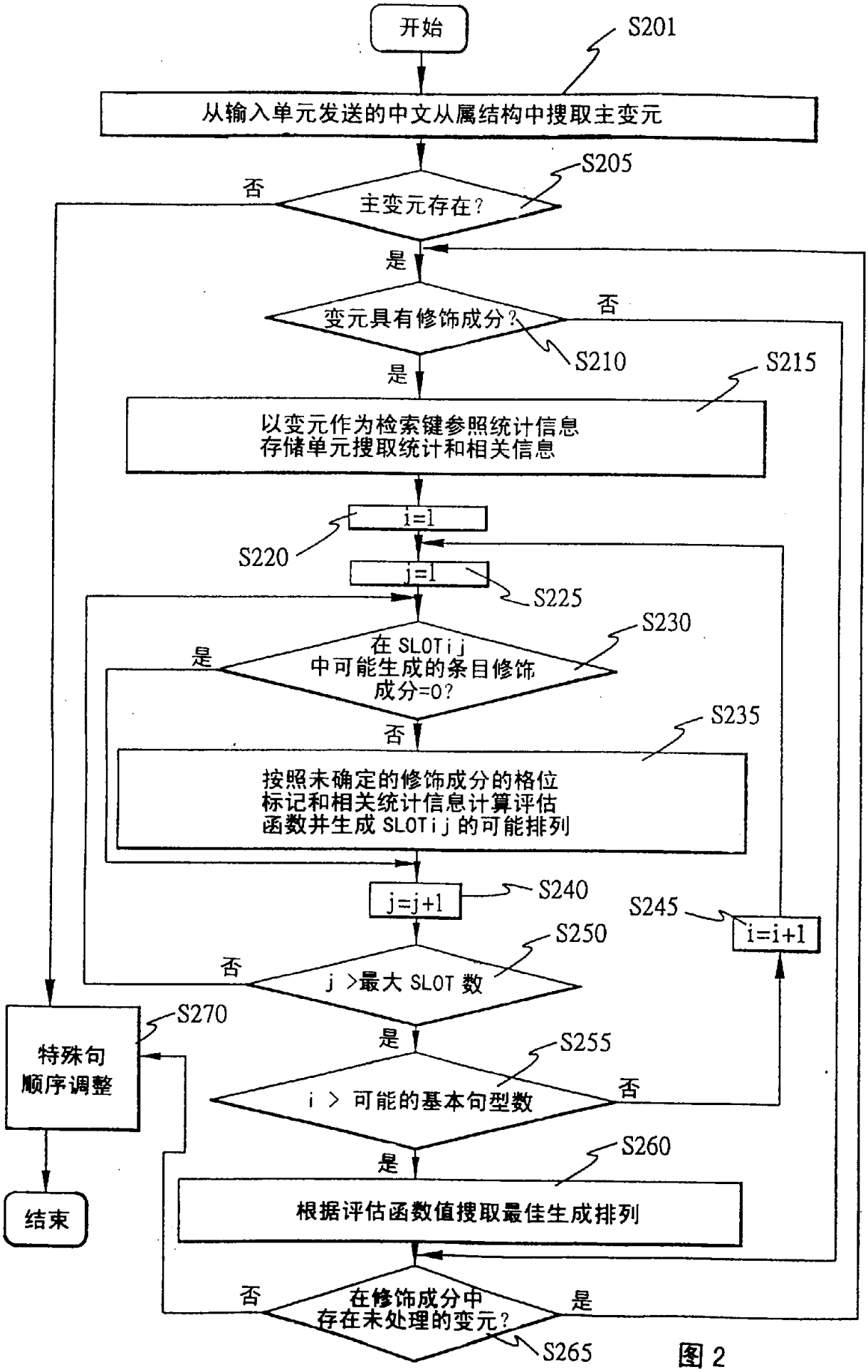

开始
S201
从输入单元发送的中文从属结构中搜取主变元
S205
否
主变元存在?
是
变元具有修饰成分?
否
S210
是
S215
以变元作为检索键参照统计信息
存储单元搜取统计和相关信息
i=1
S220
j=1
S225
S230
是
在 SLOT i j
中可能生成的条目修饰
成分=0?
否
S235
按照未确定的修饰成分的格位
标记和相关统计信息计算评估
函数并生成 SLOT i j 的可能排列
j=j+1
S240
S245
i=i+1
S250
否
j >最大 SLOT 数
是
S270
特殊句
顺序调整
S255
i > 可能的基本句型数
否
是
S260
结束
根据评估函数值搜取最佳生成排列
否
在修饰成分中
存在未处理的变元?
是
S265

图 2

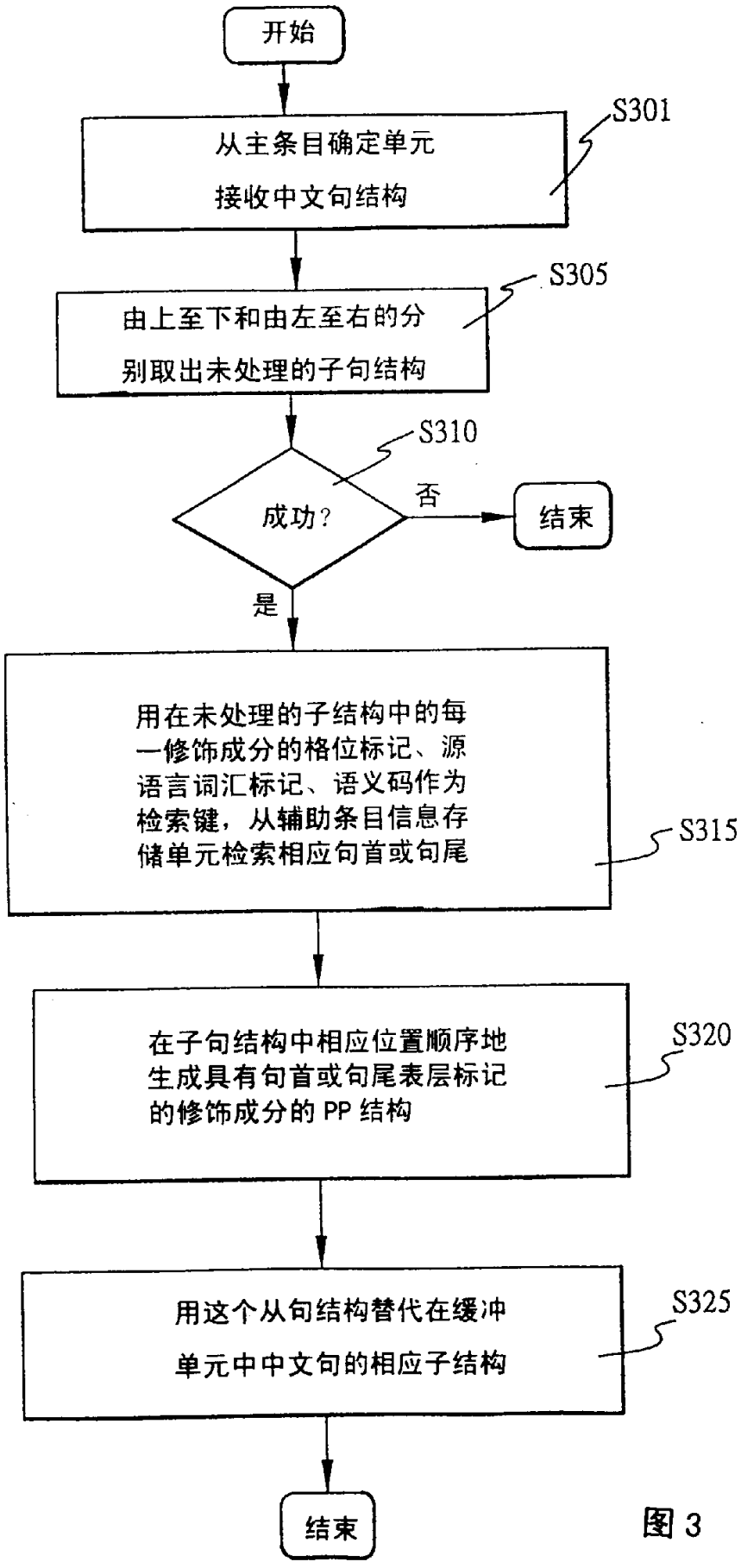

开始
从主条目确定单元
接收中文句结构
S301
由上至下和由左至右的分
别取出未处理的子句结构
S305
成功？
S310
否
结束
是
用在未处理的子结构中的每一修饰成分的格位标记、源语言词汇标记、语义码作为检索键，从辅助条目信息存储单元检索相应句首或句尾
S315
在子句结构中相应位置顺序地生成具有句首或句尾表层标记的修饰成分的 PP 结构
S320
用这个从句结构替代在缓冲
单元中中文句的相应子结构
S325
结束

图 3

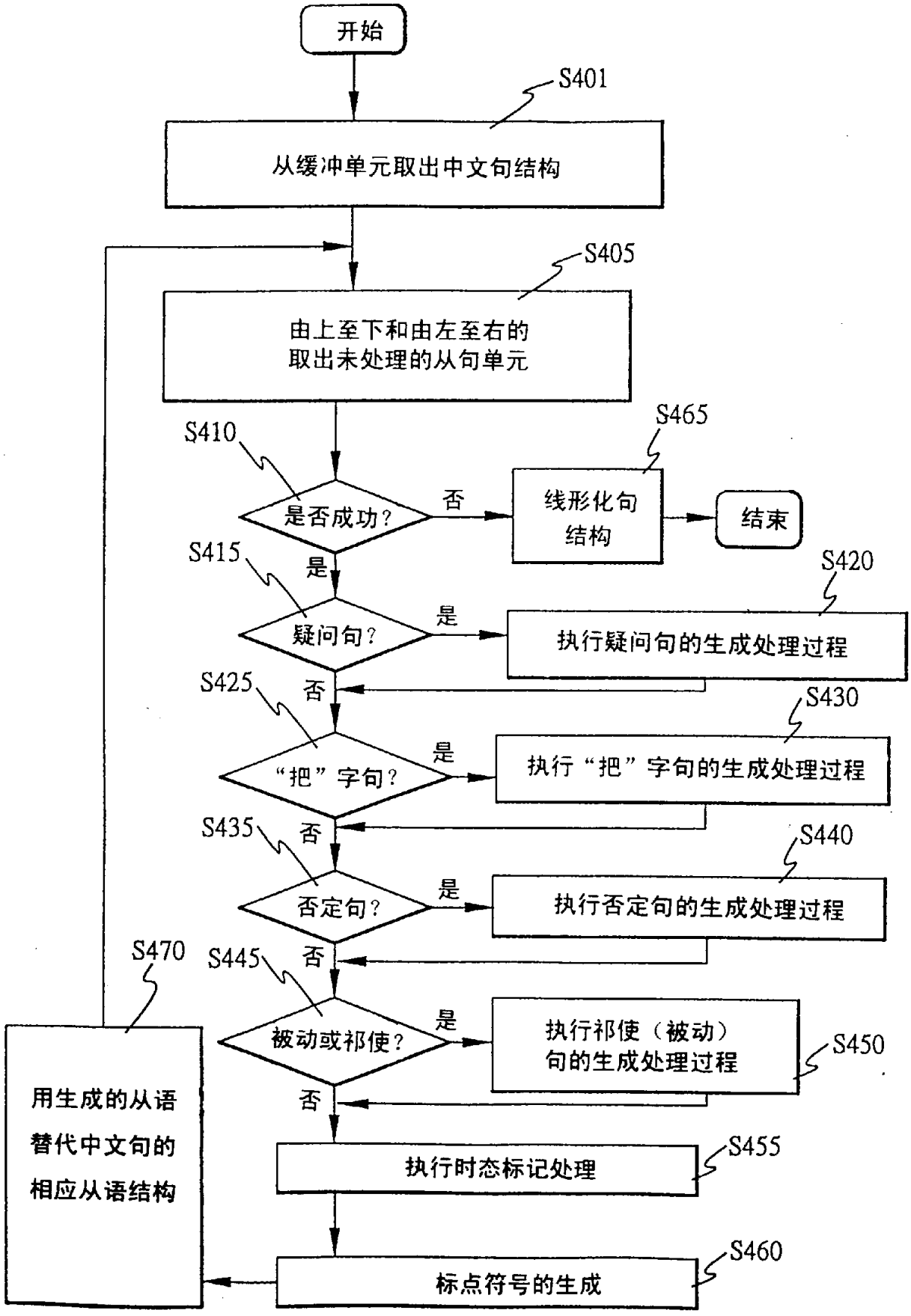

开始
S401
从缓冲单元取出中文句结构
S405
由上至下和由左至右的
取出未处理的从句单元
S410
是否成功？
否
S465
线形化句
结构
结束
是
S415
疑问句？
是
S420
执行疑问句的生成处理过程
否
S425
“把”字句？
是
S430
执行“把”字句的生成处理过程
否
S435
否定句？
是
S440
执行否定句的生成处理过程
否
S445
被动或祁使？
是
执行祁使（被动）
句的生成处理过程
S450
否
执行时态标记处理
S455
标点符号的生成
S460
S470
用生成的从语
替代中文句的
相应从语结构

图 4

变元条目	可能的句型	SLOT1		SLOT2		SLOT3		SLOT4		SLOT5		SLOT6	
		格位排列	概率	格位排列	概率	格位排列	概率	格位排列	概率	格位排列	概率	格位排列	概率
送	SVO (0.41)	time	0.2	time	0.39			loc-to	0.25				
		Purp	0.06	loc	0.09								
				Time -at	0.21								
				time, time -at	0.17								
				time, time -at loc	0.07								
	SVOO (0.3)			time	0.16								
	SVOC (0.18)			time	0.24								
	SVOOC (0.11)			time	0.11								
:	:	:	:	:	:	:	:	:	:	:	:	:	:

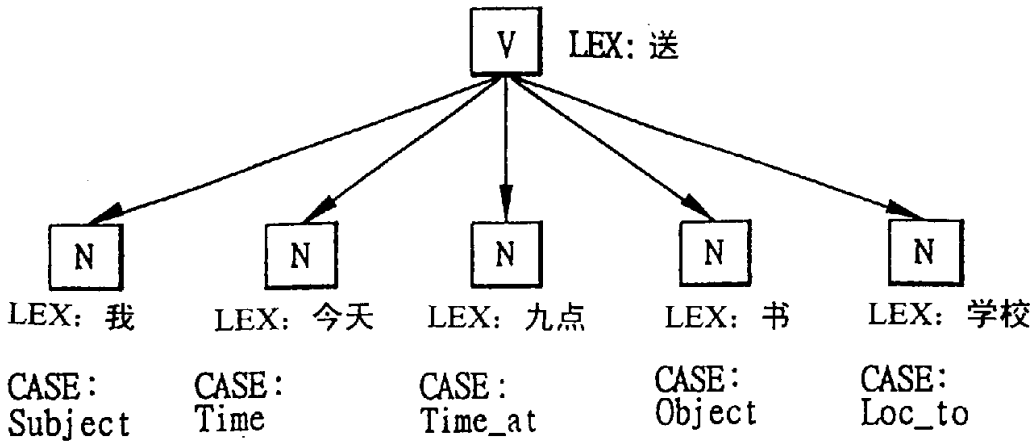
Loc: location(位置格位) purp: purpose(目的格位)

图 5

格位标记	日文标记	变元语义码	语义码	句首标记	句尾标记
LOCATION		S383 S232		在	
LOCATION				在	
LOC_FROM				从	
LOC_TO				到	
STATE_FROM				从	
CAUSE		S2989		依	不同而
CAUSE				由于	
DURATION				在	之中
INSTRUMENT				借着	
INSTRUMENT				用	
CONDITION				在	的情况下
PURPOSE				为了	
:	:	:	:	:	:
ACCOMPANY				和	
TIME-FROM				从	
TIME-TO				到	

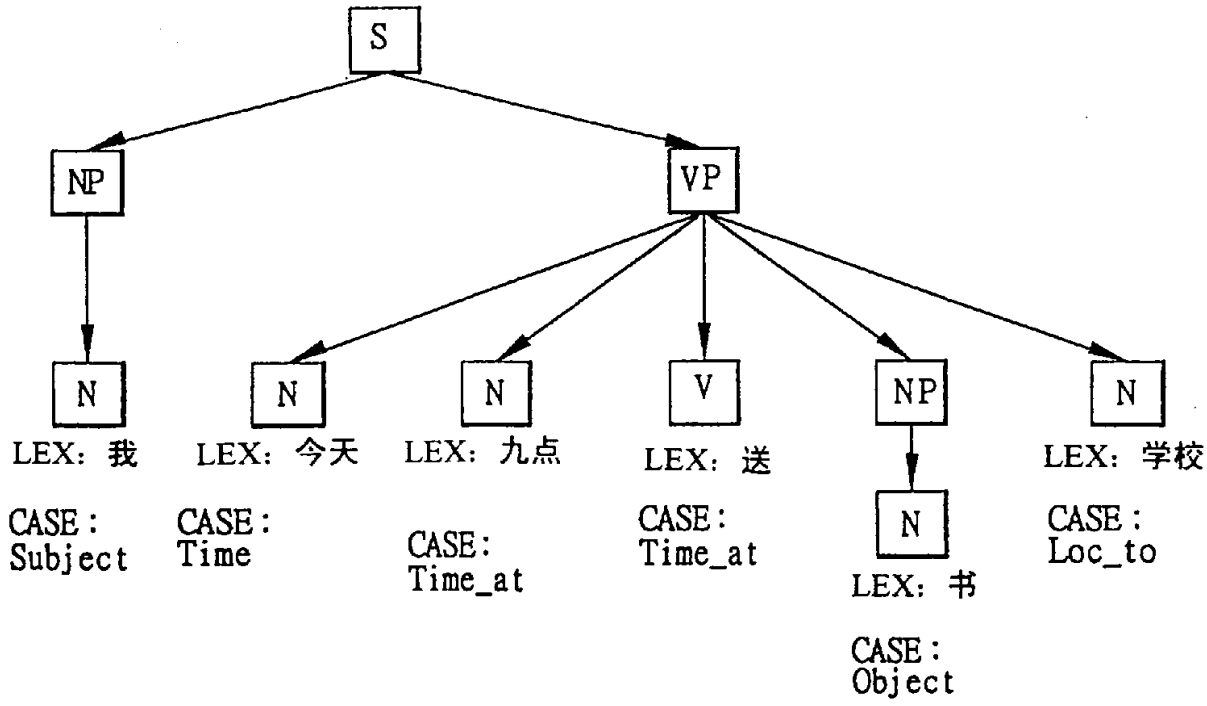
空白格：任何数据均可

图 6



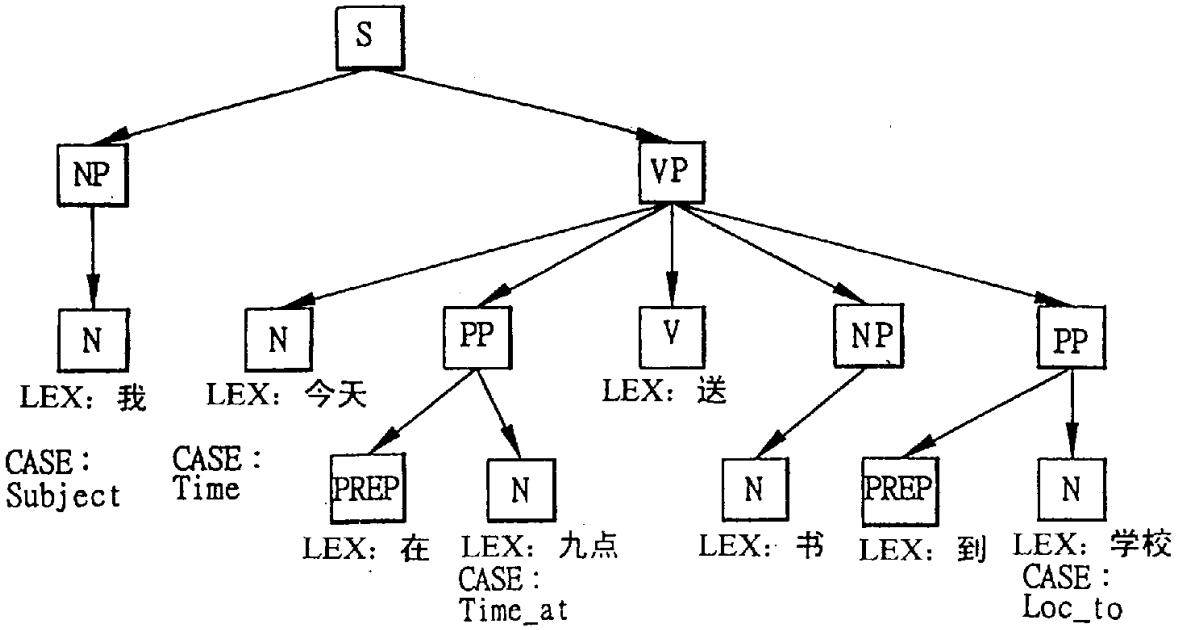
中文从属结构

图 7A



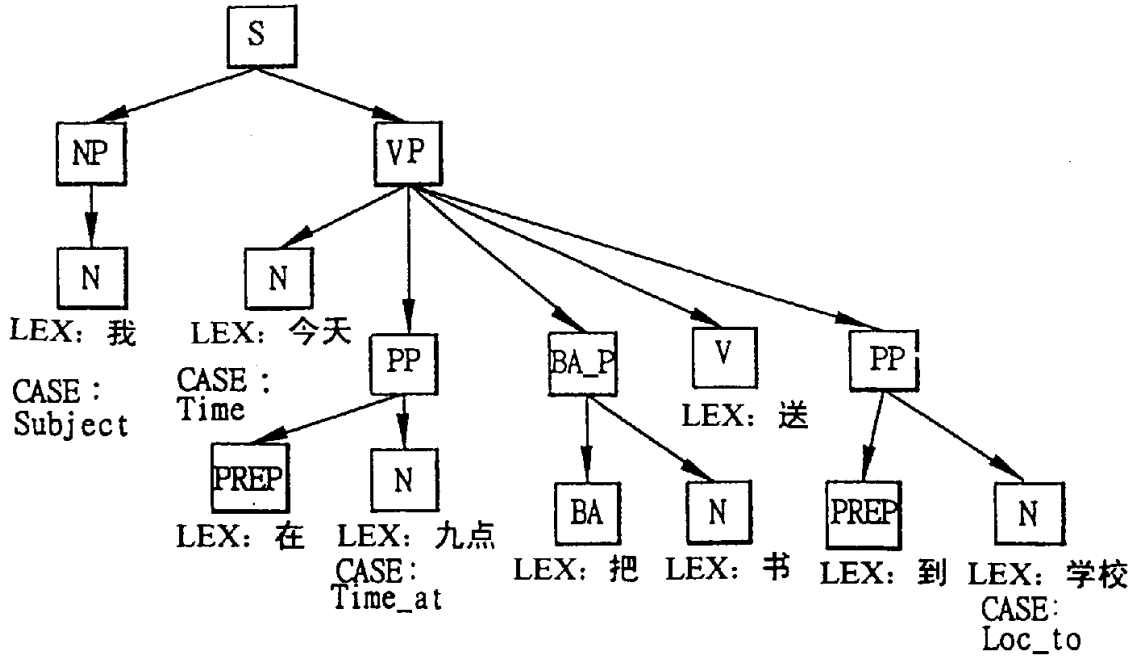
主条目展开单元展开之后的句结构

图 7B



辅助条目生成单元的展开句结构

图 7C



后期处理单元的展开句结构

图 7D

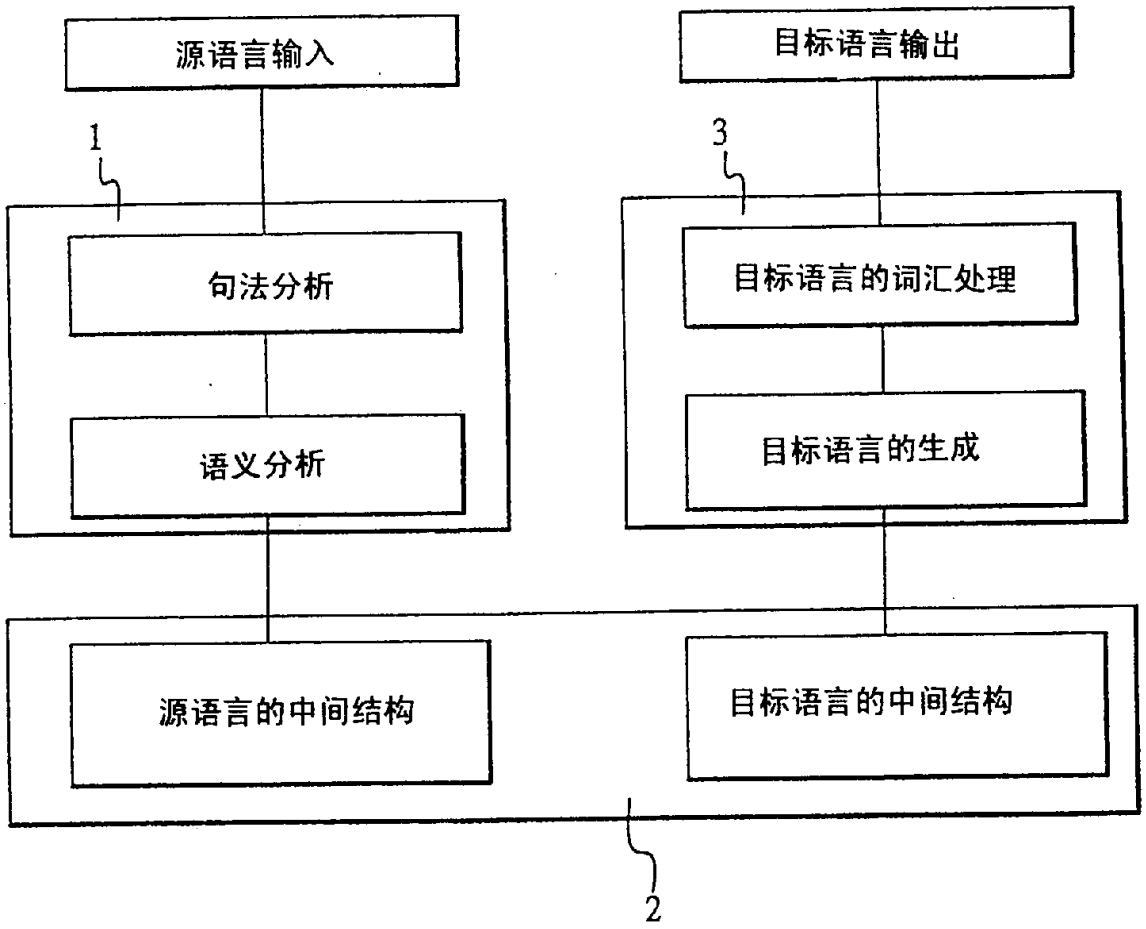

源语言输入
目标语言输出
1
3
句法分析
语义分析
目标语言的词汇处理
目标语言的生成
源语言的中间结构
目标语言的中间结构
2

图 8

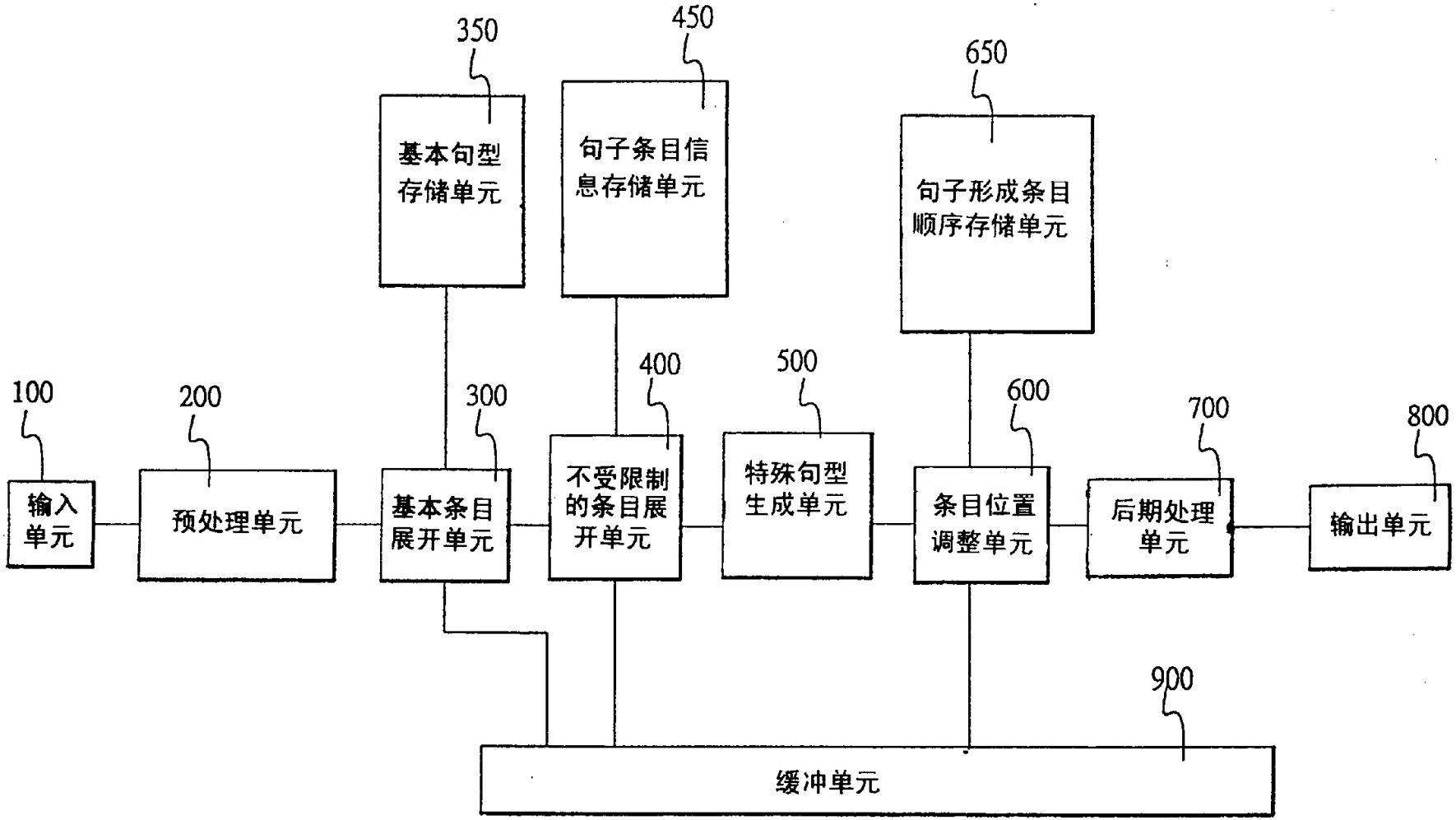
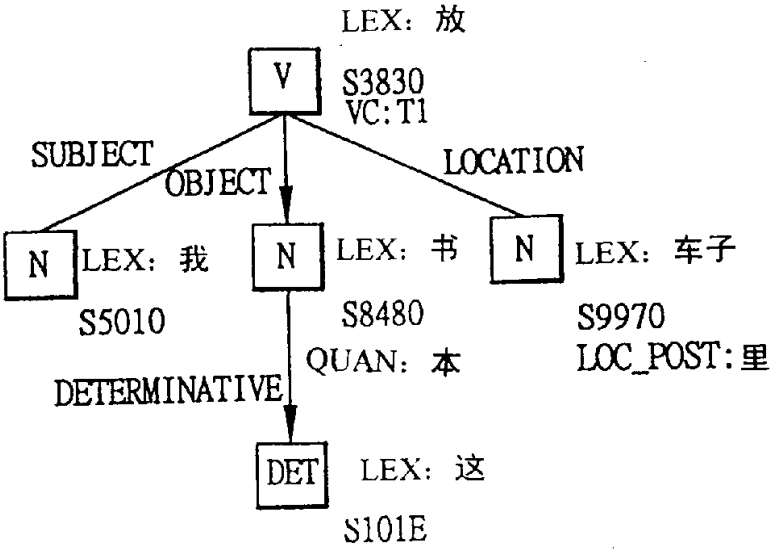

10
100
输入单元
200
预处理单元
300
基本条目展开单元
350
基本句型存储单元
400
不受限制的条目展开单元
450
句子条目信息存储单元
500
特殊句型生成单元
600
条目位置调整单元
650
句子形成条目顺序存储单元
700
后期处理单元
800
输出单元
900
缓冲单元

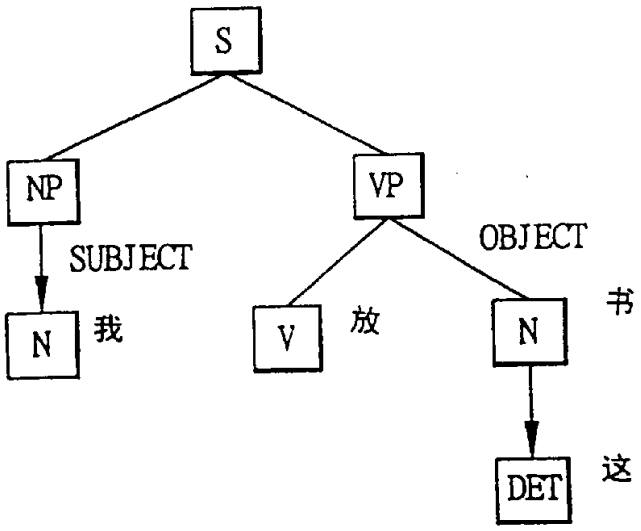
图 9



QUAN：数量词　LOC_POST：位置后缀　SXXX：语义码　LEX：词汇

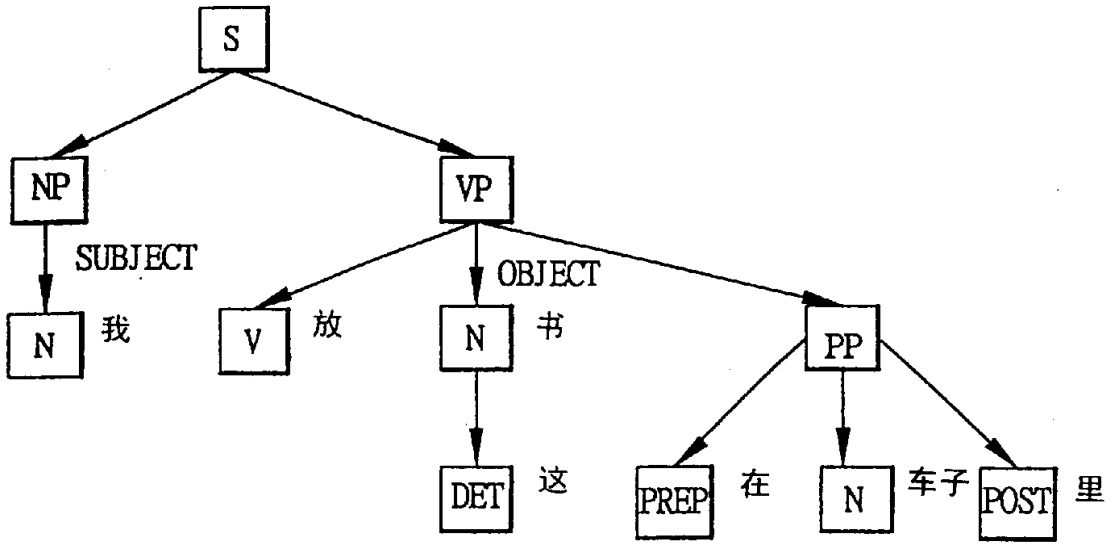
中文从属结构

图 10A



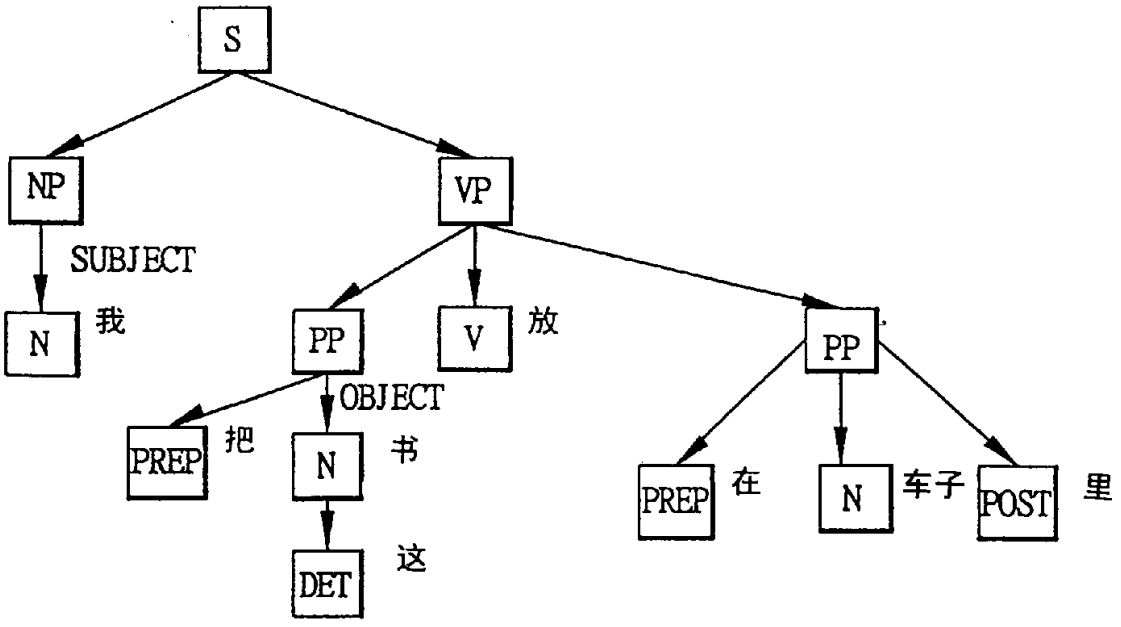
展开之后的基本条目句结构

图 10B



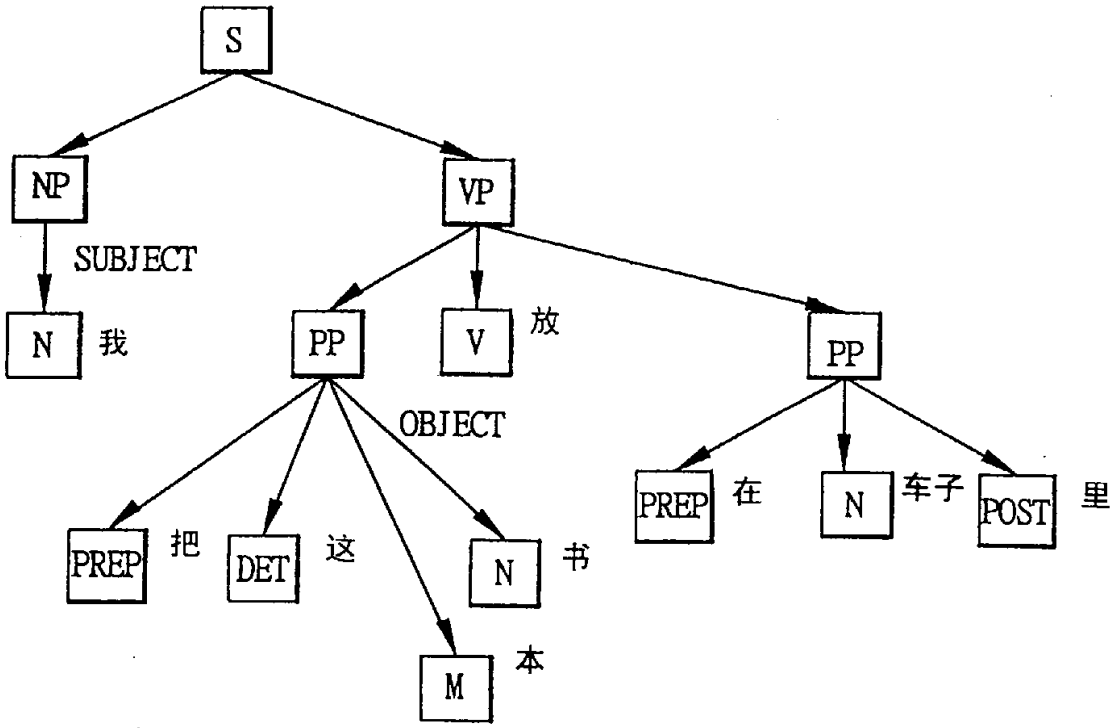
展开之后的不受限制条目句结构

图 10C



生成的特定句型的句结构

图 10D



生成的中文句结构

图 10E