

# Chart 分析算法

詹卫东

<http://ccl.pku.edu.cn/doubtfire>

# 线图分析算法

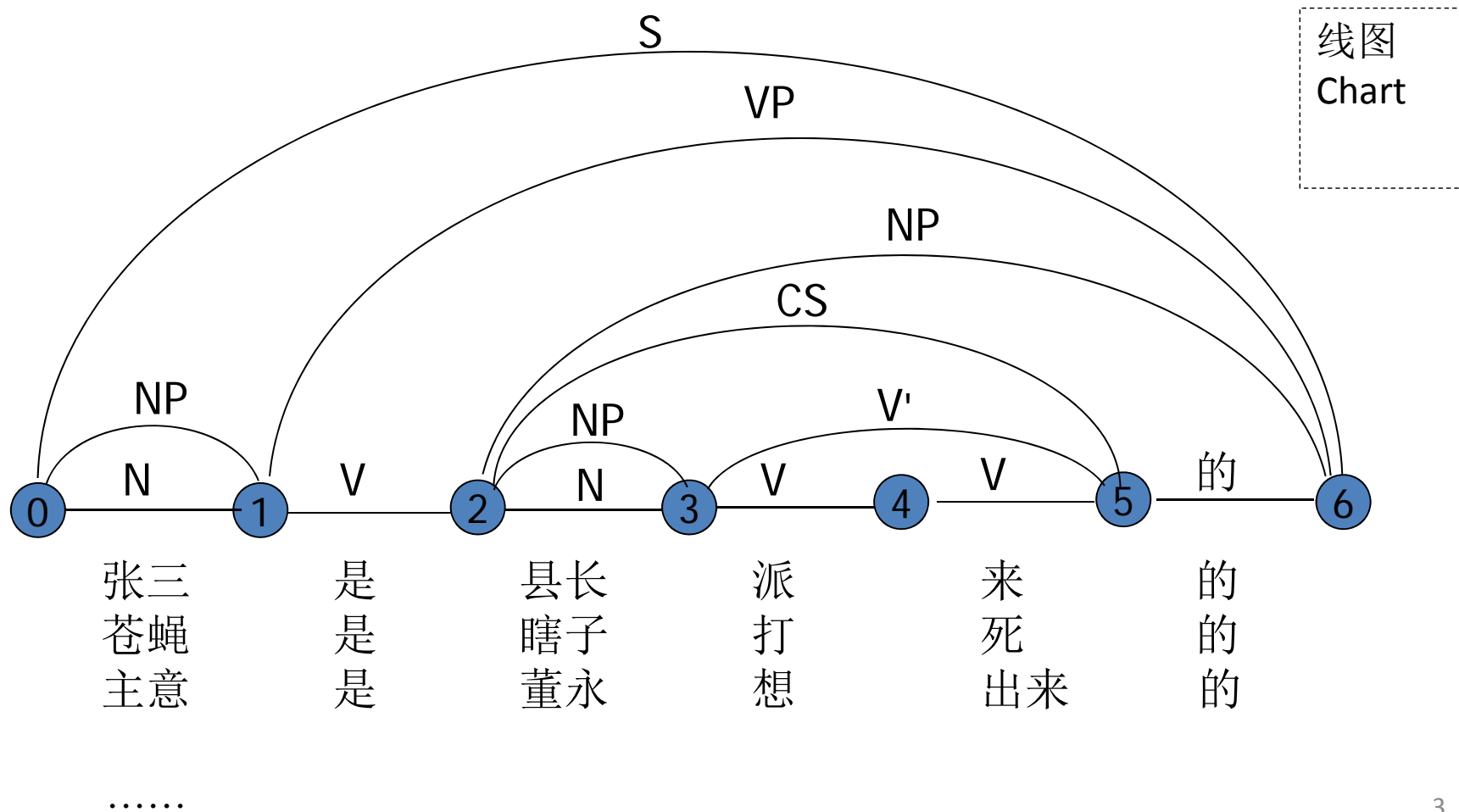
张三是县长派来的  
苍蝇是瞎子打死的  
主意是董永想出来的

.....

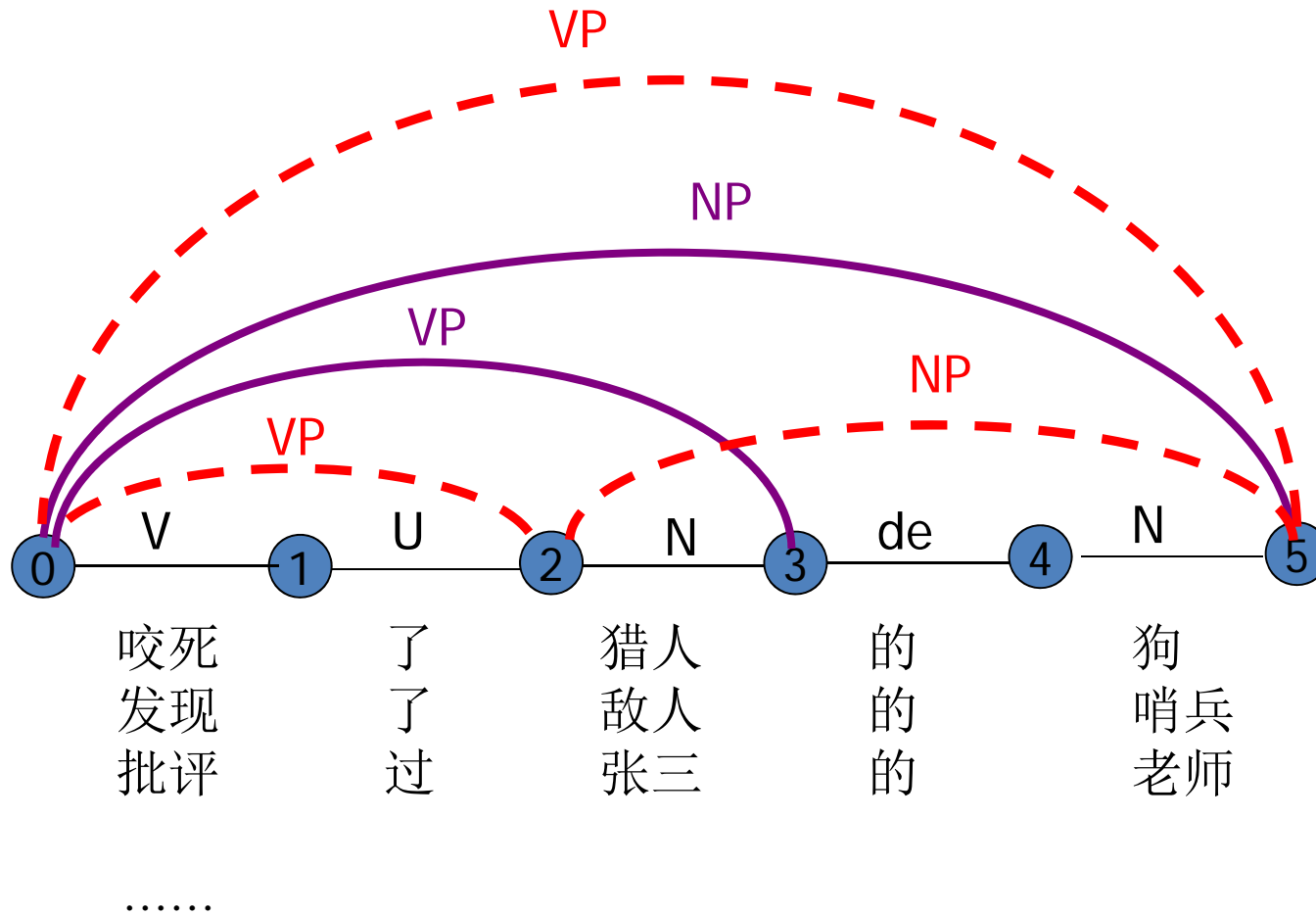
N V N V V 的

- (1)  $S \rightarrow NP VP$
- (2)  $NP \rightarrow N$
- (3)  $NP \rightarrow CS$  的
- (4)  $VP \rightarrow V NP$
- (5)  $CS \rightarrow NP V'$
- (6)  $V' \rightarrow V V$

# 线图结构图示-1



# 线图结构图示-2



# 基本概念：线图/chart

线图是一组节点(node)和边(edge)的集合

**节点**：对应着输入字符串中的字符间隔

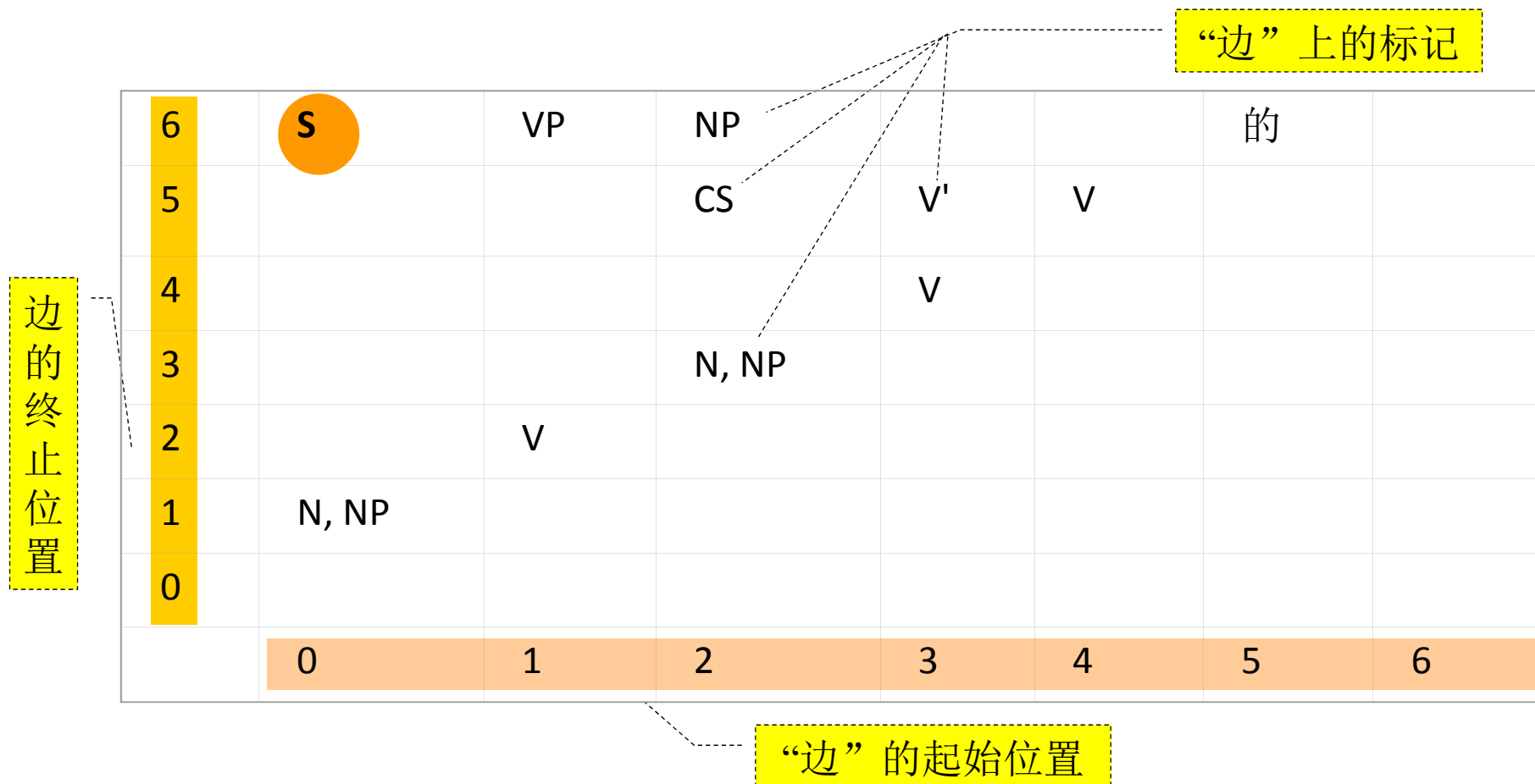
**边**：<起点, 终点, 标记>

其中标记为非终结符或终结符

问题：

如何从输入串开始，一步步形成chart，使得存在一条边可以覆盖全部节点，并且边上标记为S？

# Chart的另一种表示形式



# Chart算法的基本数据结构

1) chart (线图)

$\{edge_{[i]}\} \quad i=1,2,\dots \quad edge := \langle P1,P2, Label \rangle$

2) agenda (议程表)

存放等待加入到chart中的边 (edge)

可以栈 (stack) 结构, 或队列 (queue) 结构实现

3) active arc (活动弧)

存放分析过程的中间状态,

状态由三部分组成  $\langle P1,P2, \text{点规则} \rangle$

# Chart算法的过程描述

- 1) 将待分析字符串 $w$ 置入输入缓冲区，agenda清为空栈；
- 2) 循环，反复执行下面步骤，直至输入缓冲区和agenda均为空
  - a) 若agenda为空，则从输入缓冲区取一个字符，并把该字符及其起止位置 $(P1, P2)$ 推入agenda栈；
  - b) 若agenda不为空，则从agenda中弹出栈顶的边，该边的起止位置为 $(P1, P2)$ ，边上标记为 $L$ ；
  - c) 检查规则集中的规则，对所有形如 $A \rightarrow L \beta$ 这样的规则，在active arc集合中增加一条起止位置为 $P1, P2$ ，弧上为 $A \rightarrow L \cdot \beta$ 这样的点规则；
  - d) 把从agenda中弹出的标记为 $L$ 的边，加入到chart中的 $P1, P2$ 之间；
  - 扫描吃进 → e) 检查所有active arc，如果存在起止位置为 $P0, P1$ ，且弧上点规则为 $A \rightarrow \alpha \cdot L \beta$ 的active arc，就增加一条新的active arc，起止位置为 $P0, P2$ ，弧上点规则为 $A \rightarrow \alpha L \cdot \beta$
  - 归约 → f) 如果一条active arc（起止位置为 $P0, P2$ ）上点规则形如 $A \rightarrow \alpha L \cdot$ （点号在规则最右端），就将起止位置为 $P0, P2$ ，边上标记为 $A$ 的边压入agenda栈。
  - 预测 →

agenda中的这条边将为下面的分析指出方向：吃进所有右部以 $A$ 开头的规则（步骤b, c）

# Chart算法分析示例-1

- (1)  $S \rightarrow NP VP$
- (2)  $NP \rightarrow N$
- (3)  $NP \rightarrow CS$  的
- (4)  $VP \rightarrow V NP$
- (5)  $CS \rightarrow NP V'$
- (6)  $V' \rightarrow V V$

agenda

“议程表”  
用于中  
转当前  
正在处  
理的边

c  
h  
a  
r  
t  
—  
—  
a  
c  
t  
i  
v  
e  
a  
r  
c

“线图”用于存放已经确定的分析结果



“活动弧”用于存放分析的中间结果  
(包括尚未确定的, 和已经确定的分析结果)

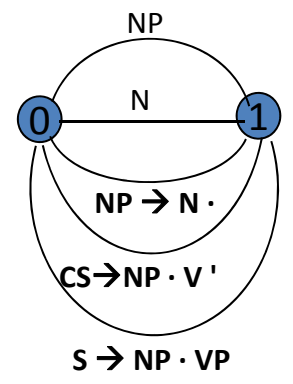
N	V	N	V	V	的
---	---	---	---	---	---

输入缓冲区

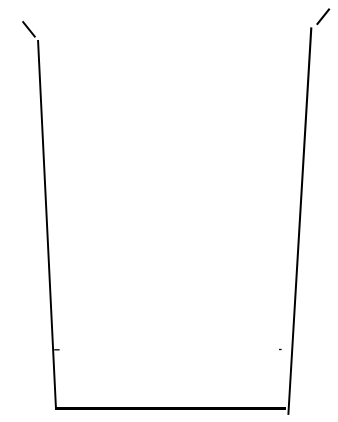
# Chart算法分析示例-2

- (1)  $S \rightarrow NP VP$
- (2)  $NP \rightarrow N$
- (3)  $NP \rightarrow CS$  的
- (4)  $VP \rightarrow V NP$
- (5)  $CS \rightarrow NP V'$
- (6)  $V' \rightarrow V V$

c  
h  
a  
r  
t  
  
a  
c  
t  
i  
v  
e  
a  
r  
c



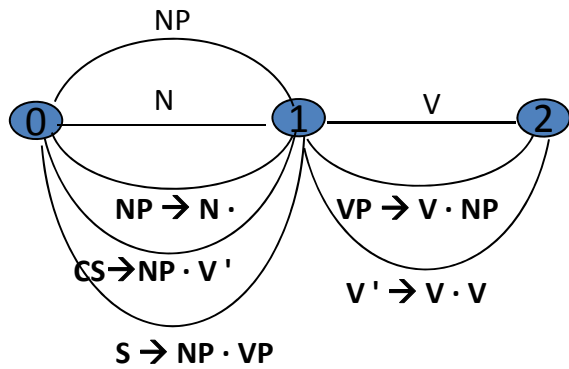
agenda



输入缓冲区

# Chart算法分析示例-3

- (1)  $S \rightarrow NP VP$
- (2)  $NP \rightarrow N$
- (3)  $NP \rightarrow CS$  的
- (4)  $VP \rightarrow V NP$
- (5)  $CS \rightarrow NP V'$
- (6)  $V' \rightarrow V V$



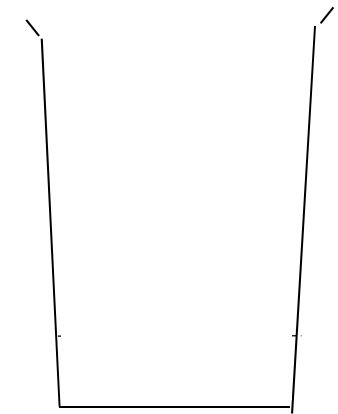
3

4

5

6

agenda



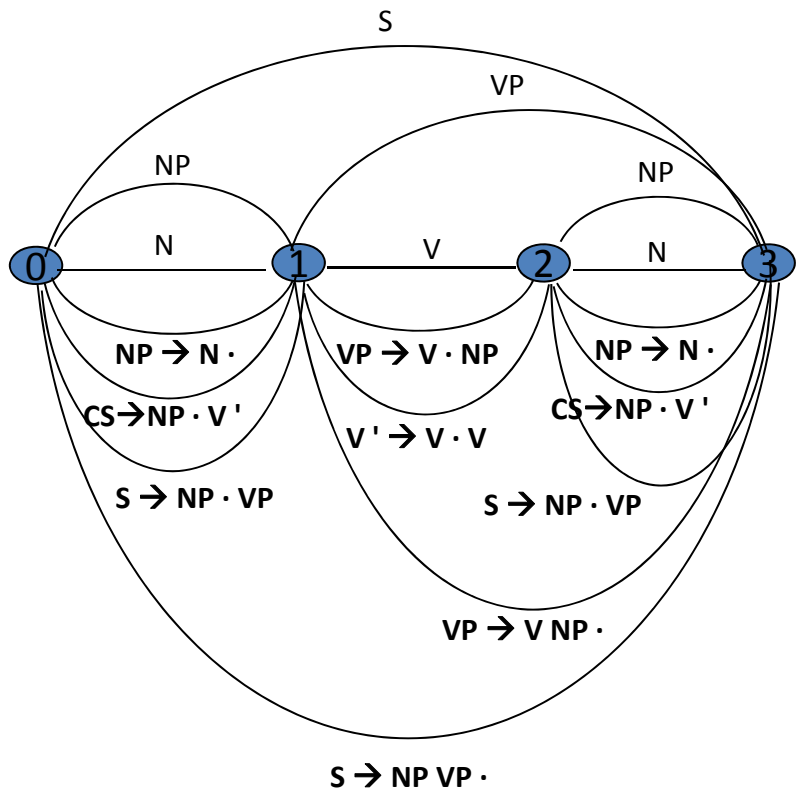
N	V	V	的
---	---	---	---

输入缓冲区



# Chart算法分析示例-5

- (1)  $S \rightarrow NP VP$
- (2)  $NP \rightarrow N$
- (3)  $NP \rightarrow CS$  的
- (4)  $VP \rightarrow V NP$
- (5)  $CS \rightarrow NP V'$
- (6)  $V' \rightarrow V V$

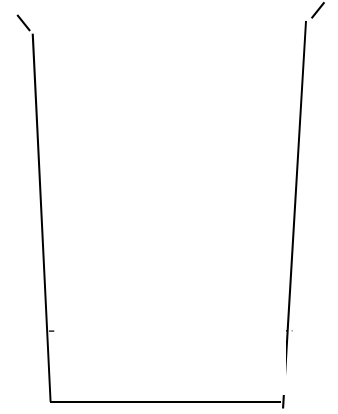


4

5

6

agenda

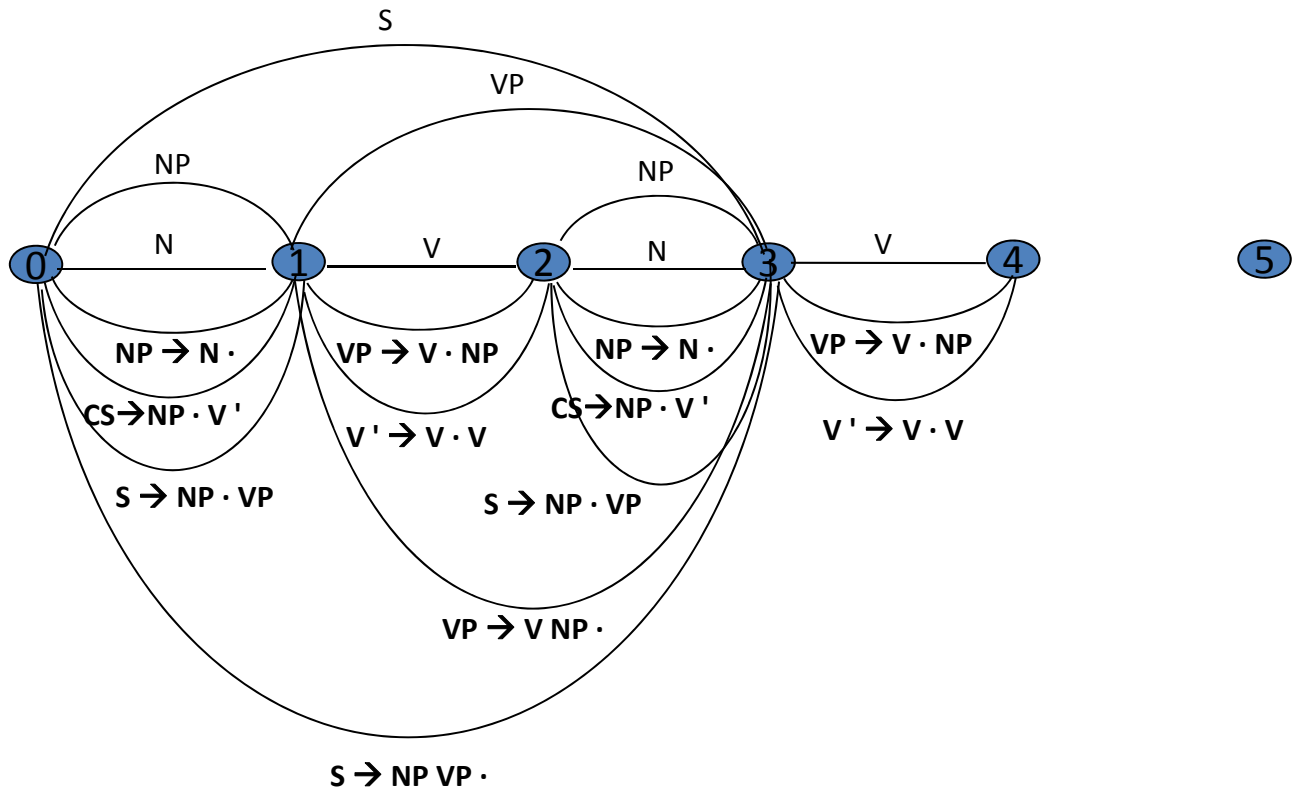


V	V	的
---	---	---

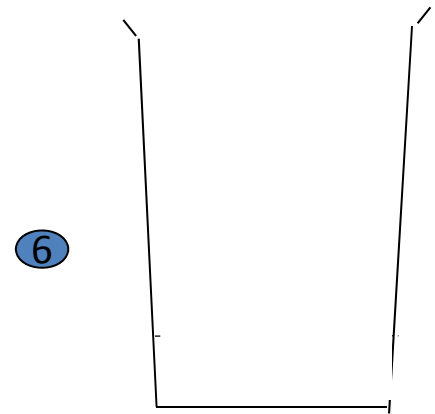
输入缓冲区

# Chart算法分析示例-6

- (1)  $S \rightarrow NP VP$
- (2)  $NP \rightarrow N$
- (3)  $NP \rightarrow CS$  的
- (4)  $VP \rightarrow V NP$
- (5)  $CS \rightarrow NP V'$
- (6)  $V' \rightarrow V V$



agenda

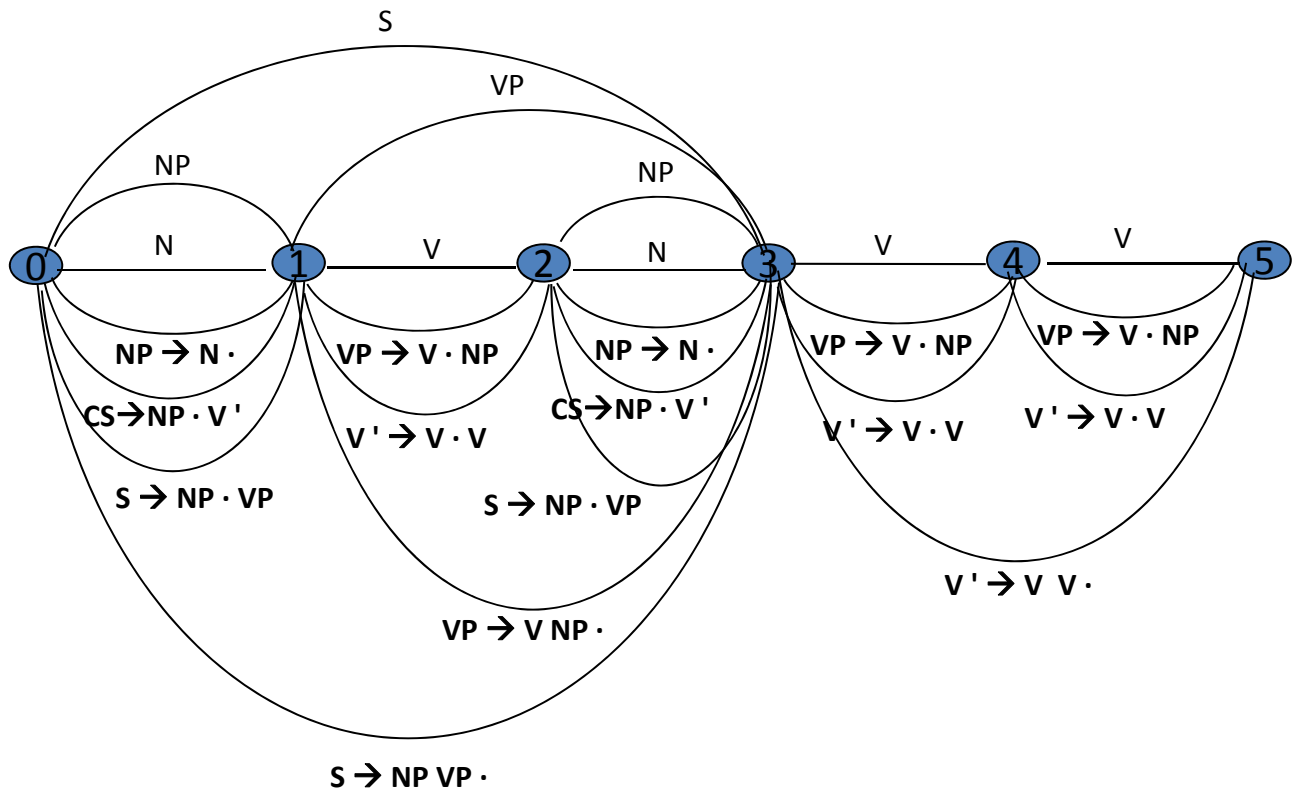


V	的
---	---

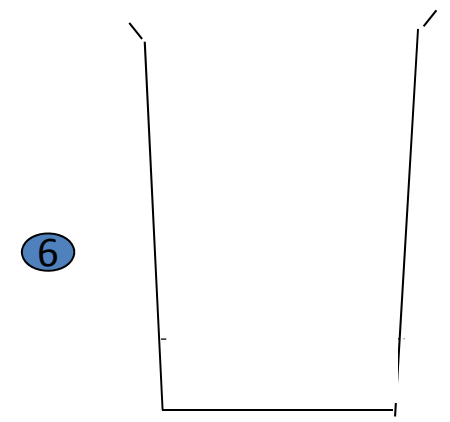
输入缓冲区

# Chart算法分析示例-7

- (1)  $S \rightarrow NP VP$
- (2)  $NP \rightarrow N$
- (3)  $NP \rightarrow CS$  的
- (4)  $VP \rightarrow V NP$
- (5)  $CS \rightarrow NP V'$
- (6)  $V' \rightarrow V V$



agenda

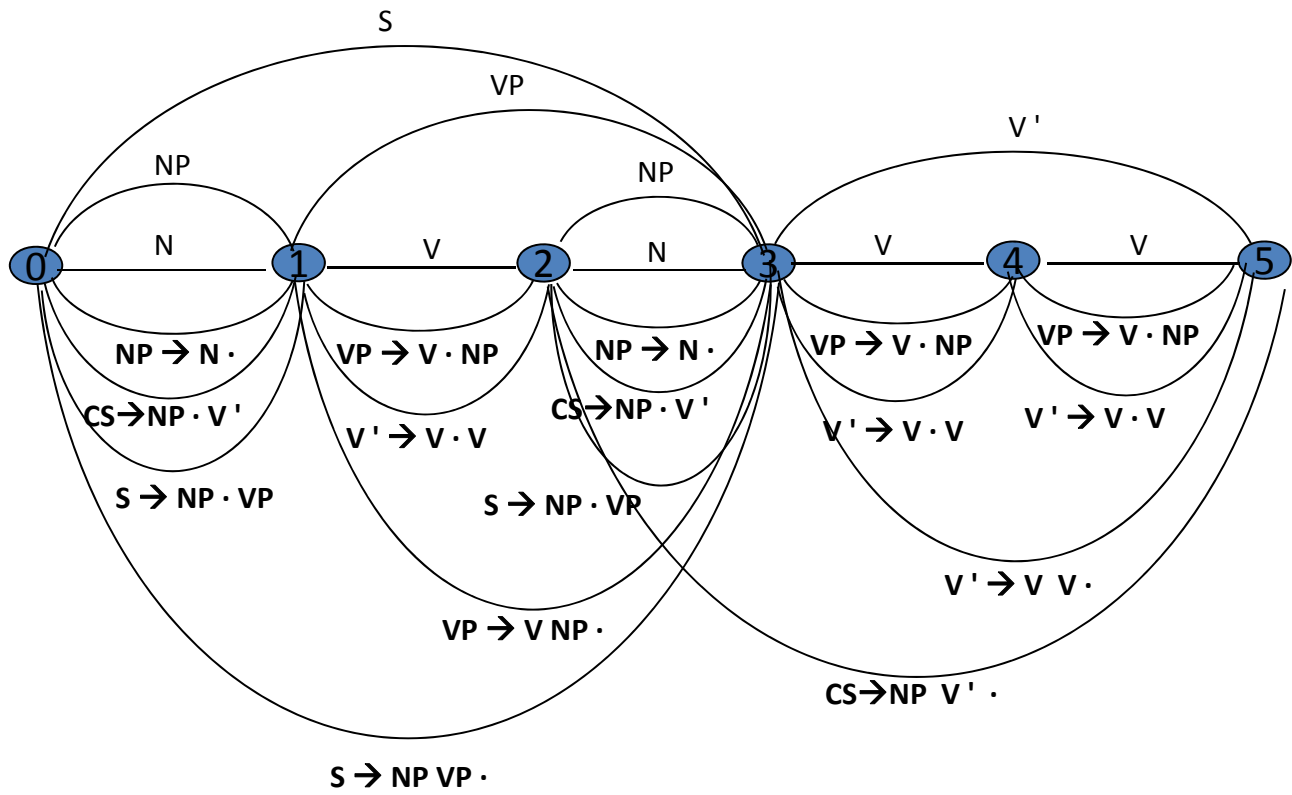


的

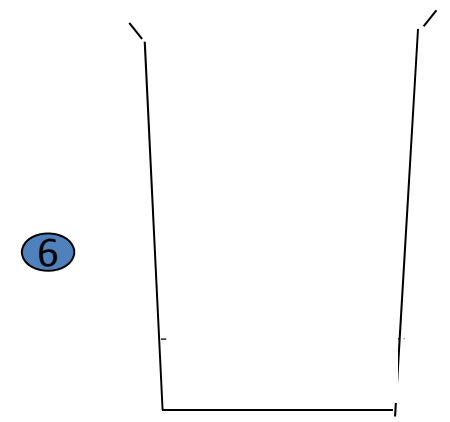
输入缓冲区

# Chart算法分析示例-8

- (1)  $S \rightarrow NP VP$
- (2)  $NP \rightarrow N$
- (3)  $NP \rightarrow CS$  的
- (4)  $VP \rightarrow V NP$
- (5)  $CS \rightarrow NP V'$
- (6)  $V' \rightarrow V V$



agenda

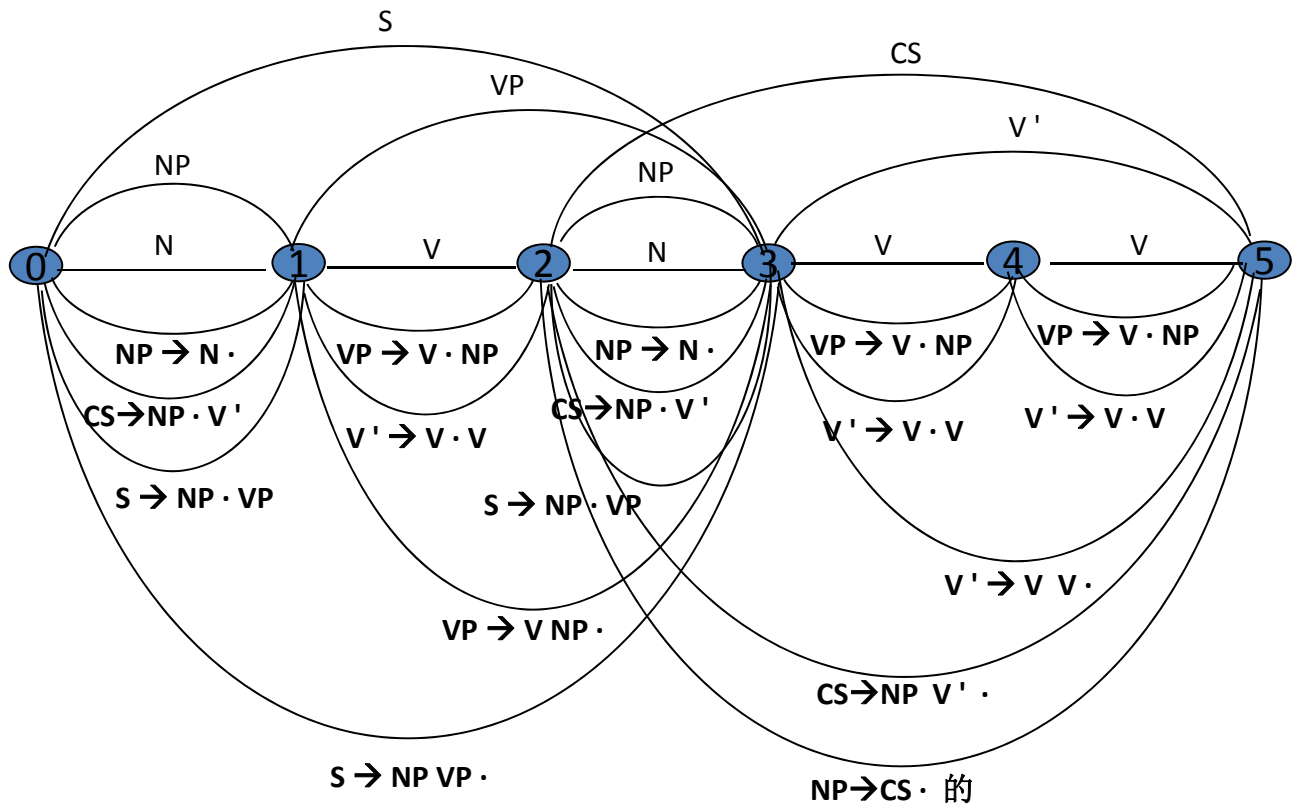


的

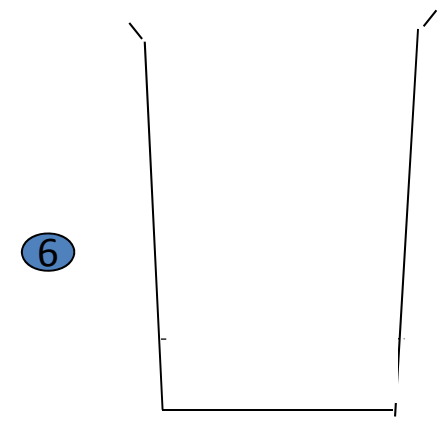
输入缓冲区

# Chart算法分析示例-9

- (1)  $S \rightarrow NP VP$
- (2)  $NP \rightarrow N$
- (3)  $NP \rightarrow CS$  的
- (4)  $VP \rightarrow V NP$
- (5)  $CS \rightarrow NP V'$
- (6)  $V' \rightarrow V V$



agenda

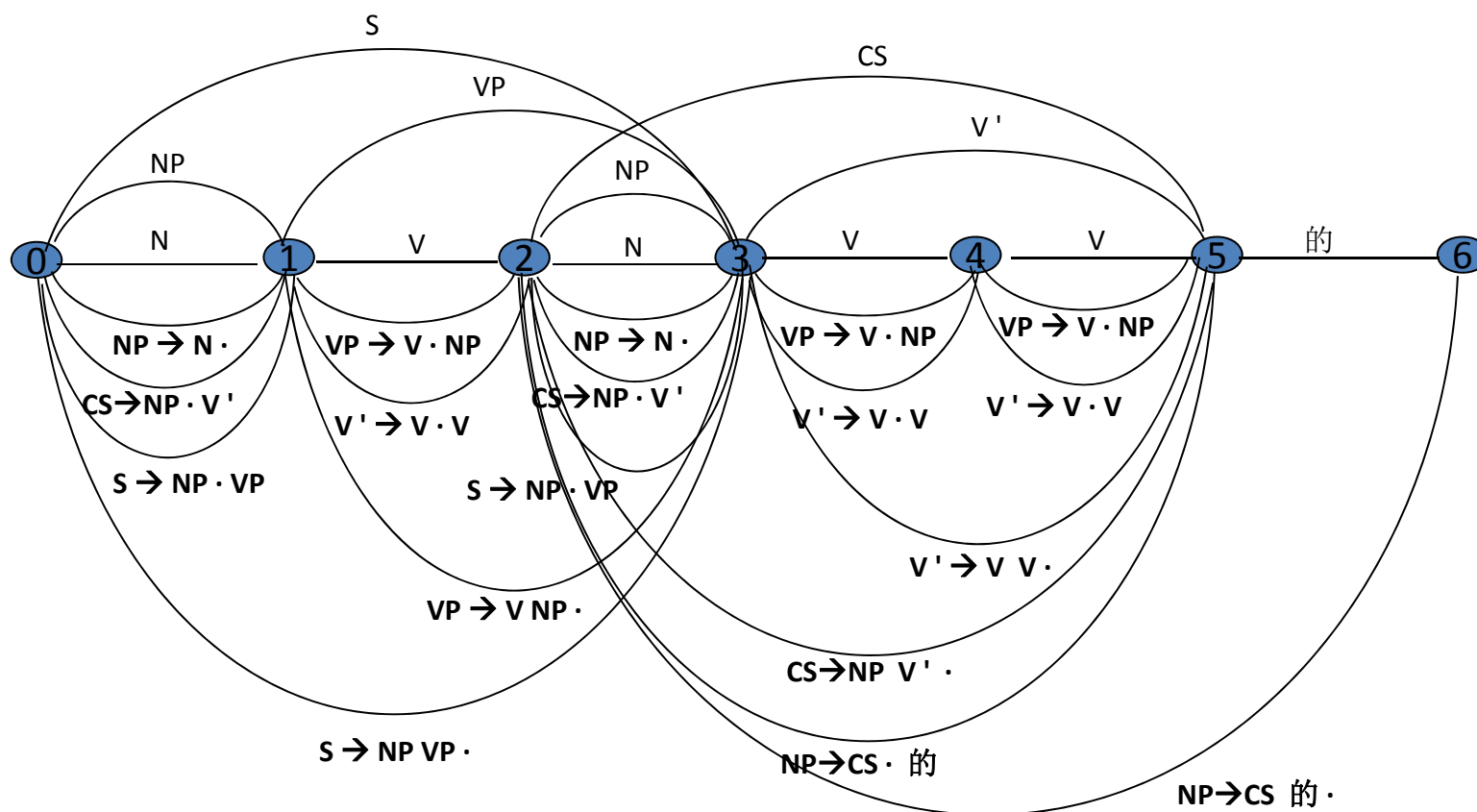


的

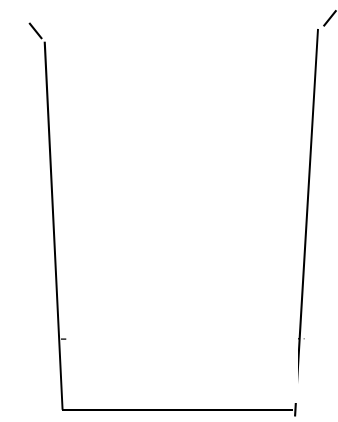
输入缓冲区

# Chart算法分析示例-10

- (1)  $S \rightarrow NP VP$
- (2)  $NP \rightarrow N$
- (3)  $NP \rightarrow CS$  的
- (4)  $VP \rightarrow V NP$
- (5)  $CS \rightarrow NP V'$
- (6)  $V' \rightarrow V V$



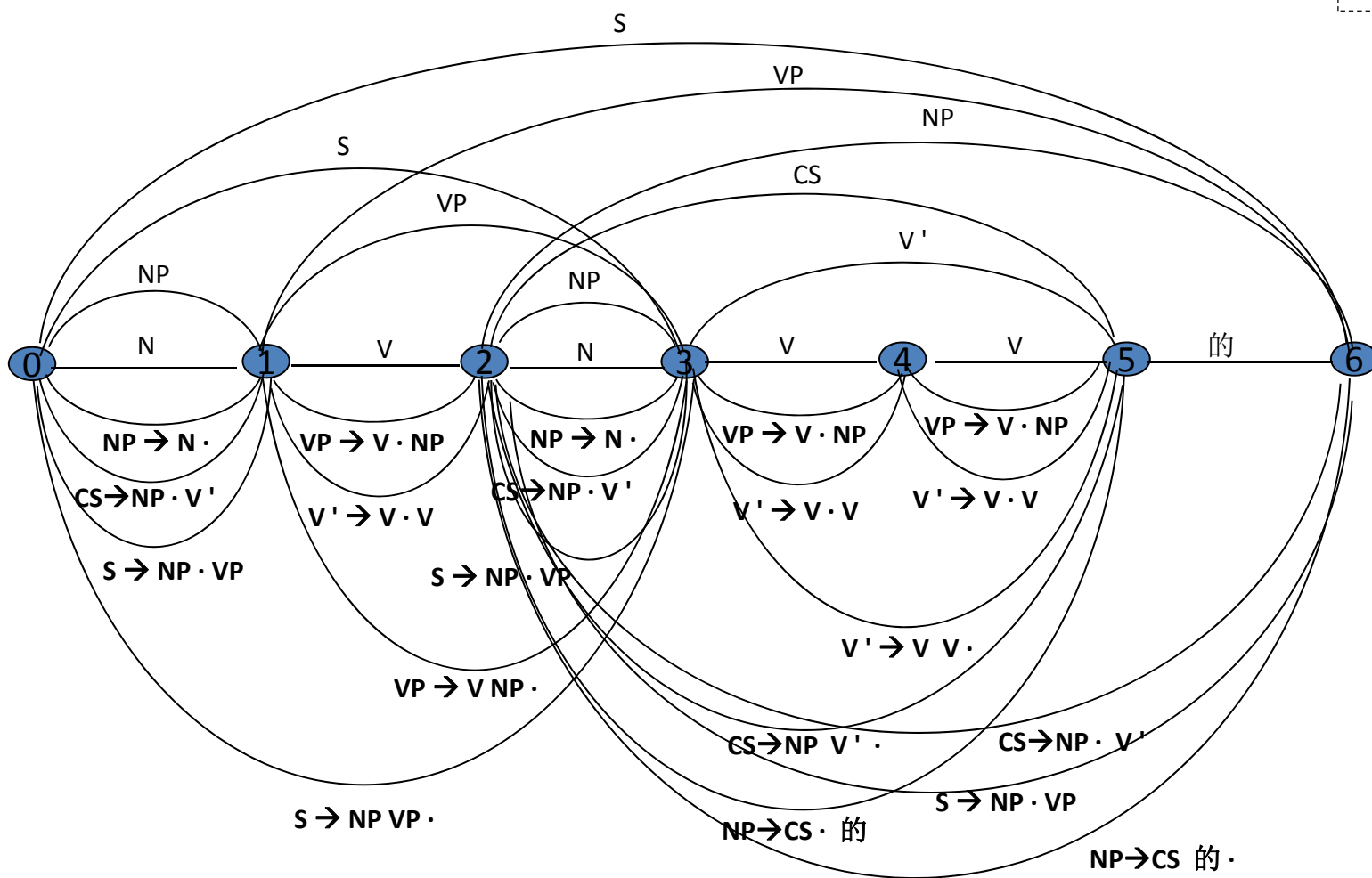
agenda



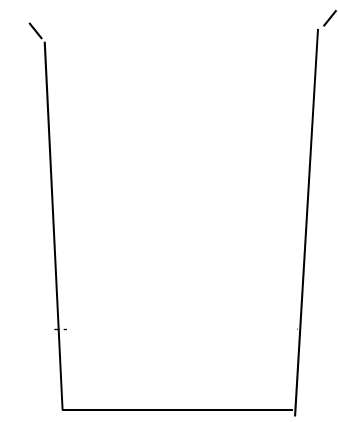
输入缓冲区

# Chart算法分析示例-11

- (1)  $S \rightarrow NP VP$
- (2)  $NP \rightarrow N$
- (3)  $NP \rightarrow CS$  的
- (4)  $VP \rightarrow V NP$
- (5)  $CS \rightarrow NP V'$
- (6)  $V' \rightarrow V V$



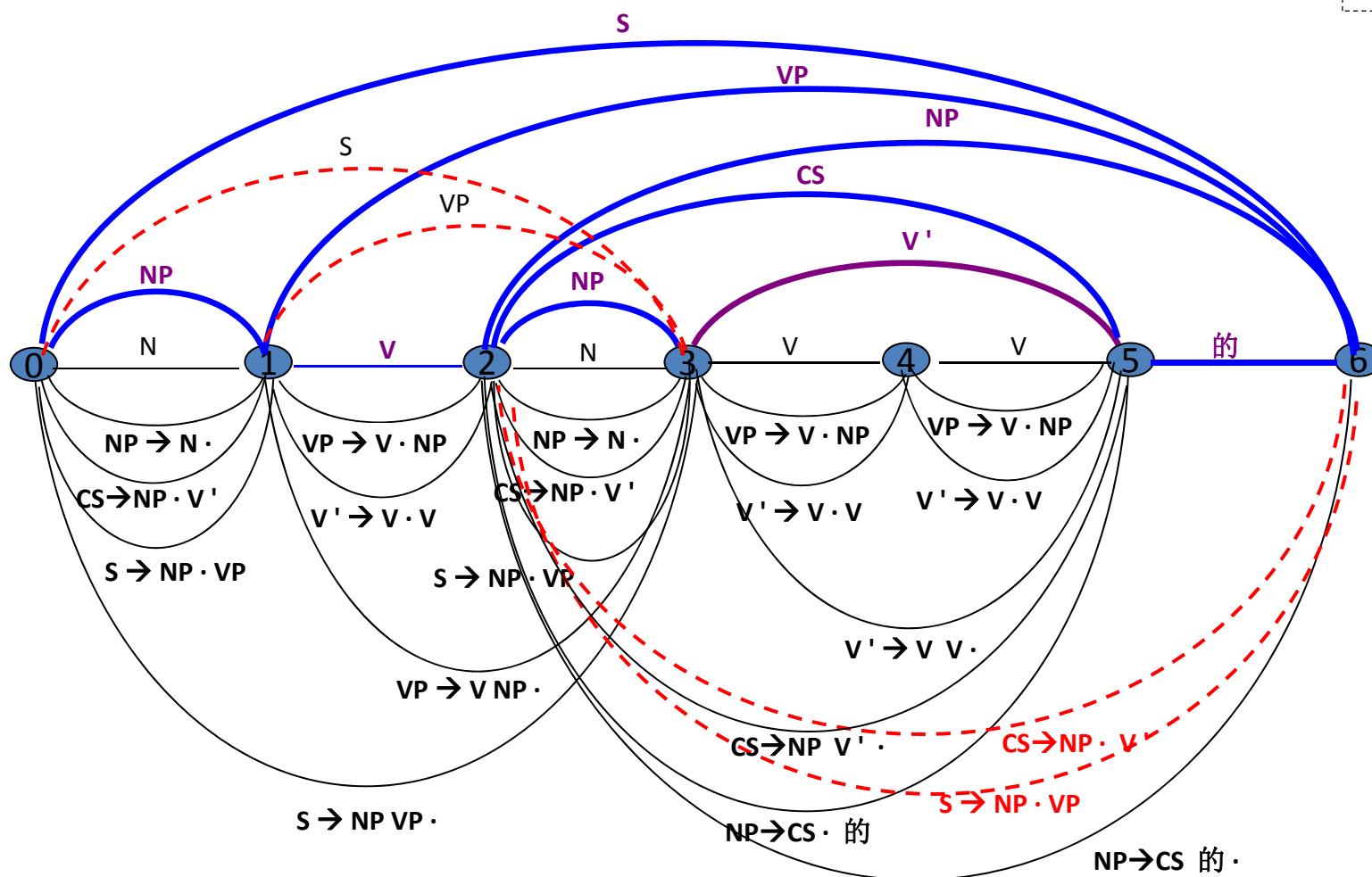
agenda



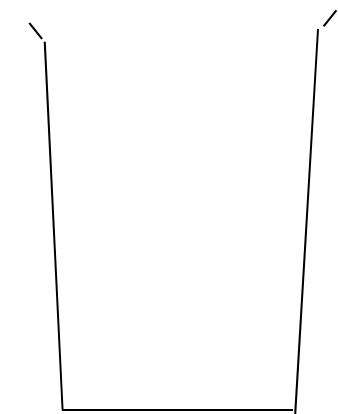
输入缓冲区

# Chart算法分析示例-12

- (1)  $S \rightarrow NP VP$
- (2)  $NP \rightarrow N$
- (3)  $NP \rightarrow CS$  的
- (4)  $VP \rightarrow V NP$
- (5)  $CS \rightarrow NP V'$
- (6)  $V' \rightarrow V V$



agenda



输入缓冲区