句法分析

杨沐昀
教育部-微软语言语音重点实验室
MOE-MS Joint Key Lab of NLP and Speech (HIT)

内容提要

- PCFG概述
- PCFG算法实现：CYK and Beam Search
- PCFG的构建：树库

CFG缺陷

- 对于一个中等长度的输入句子来说，要利用大覆盖度的语法规则分析出所有可能的句子结构是非常困难的，分析过程的复杂度往往使程序无法实现
- 即使能分析出句子所有可能的结构，也难以在巨大的句法分析结果集中实现有效的消歧，并选择出最有可能的分析结果
- 手工编写的规则一般带有一定的主观性，对于实际应用系统来说，往往难以覆盖大领域的所有复杂语言
- 写规则本身是一件大工作量的复杂劳动，而且编写的规则对特定的领域有密切的相关性，不利于句法分析系统向其他领域移植

统计句法分析

- 鉴于基于规则的句法分析存在诸多局限，20世纪80年代中期研究者们开始探索统计句法分析方法
- 目前研究较多的统计句法分析方法是语法驱动的(grammar-driven)。其基本思想是由生成语法(generative grammar)定义被分析的语言及其分析出的类别，在训练数据中观察到的各种语言现象的分布以统计数据的方式与语法规则一起编码
- 在句法分析的过程中，当遇到歧义情况时，统计数据用于对多种分析结果的排序或选择
- 基于概率上下文无关文法(probabilistic context-free grammar,PCFG) 可以说是目前最成功的统计句法分析方法

概率上下文无关文法(PCFG)

- 基于概率上下文无关文法的句法分析既有规则方法的特点，又运用了概率信息，因此可以认为是规则方法和统计方法的紧密结合
- 概率上下文无关文法就是一个为规则增添了概率的简单CFG，指明了不同重写规则的可能性大小
- PCFG的规则表示形式为：$A \to \alpha \quad p$，其中A为非终结符，p为A推导出α的概率，即$p = P(A \to \alpha)$
- 该概率分布必须满足如下条件：

$$\sum_{\alpha} P(A \to \alpha) = 1$$

概率上下文无关文法(PCFG)

- 一个PCFG包括：
  - 一个终结符集合Σ
  - 一个非终结符集合N
  - 一个指定的初始符$S \in N$
  - 一个规则集合R，$R = \{A \to \alpha\}$，其中A为非终结符
  - 一个对应的规则概率之和如下：

$$\sum_{\alpha} P(A \to \alpha) = 1$$

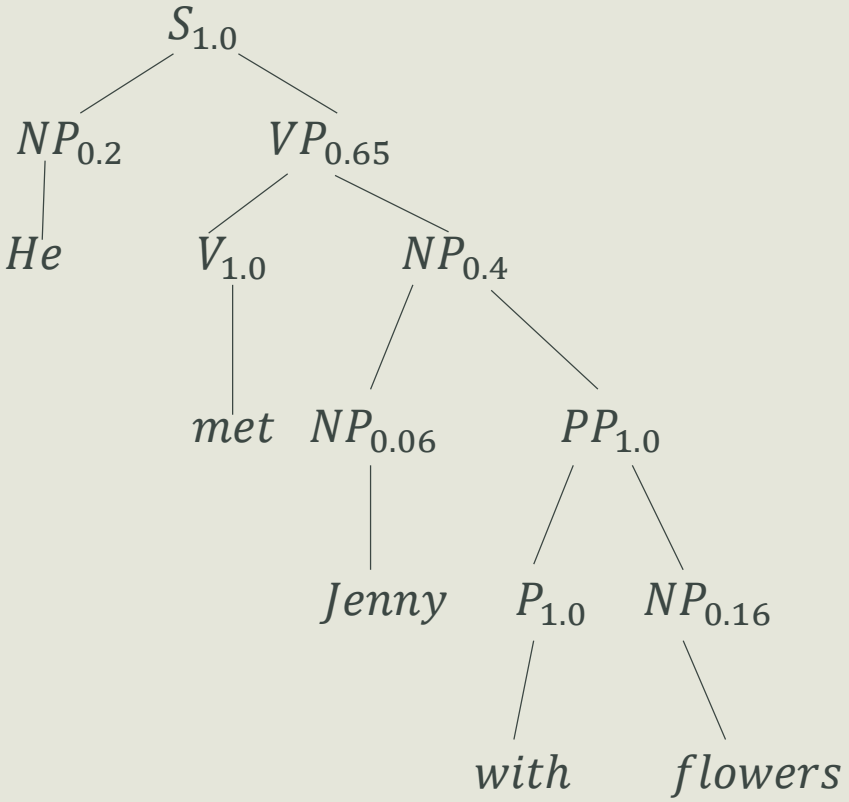
概率上下文无关文法(PCFG)

- 给定如下文法G(S):
- $S \rightarrow NP\ VP$ 1.0
- $PP \rightarrow P\ NP$ 1.0
- $VP \rightarrow V\ NP$ 0.65
- $VP \rightarrow VP\ PP$ 0.35
- $P \rightarrow with$ 1.0
- $V \rightarrow met$ 1.0
- $NP \rightarrow NP\ PP$ 0.4
- $NP \rightarrow He$ 0.2
- $NP \rightarrow Jenny$ 0.06
- $NP \rightarrow flowers$ 0.16
- $NP \rightarrow books$ 0.18

- 根据上述文法，句子He met Jenny with flowers有两个可能的句法结构

概率上下文无关文法(PCFG)

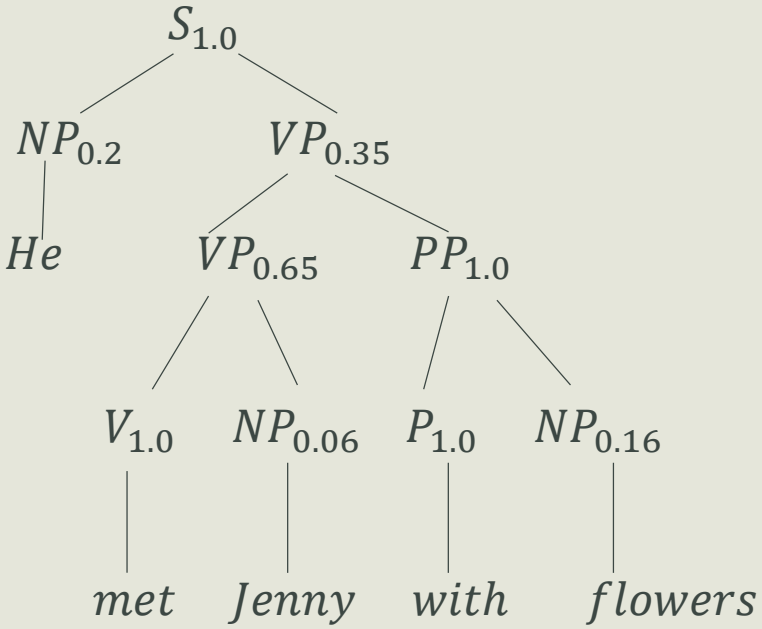
t1:



$$P(t_1) = 1.0 \times 0.2 \times 0.65 \times 1.0 \times 0.4 \times 0.06 \times 1.0 \times 1.0 \times 0.16 = 0.0004992$$

概率上下文无关文法(PCFG)

t2:



$$P(t_2) = 1.0 \times 0.2 \times 0.35 \times 0.65 \times 1.0 \times 0.06 \times 1.0 \times 1.0 \times 0.16 = 0.0004368$$

概率上下文无关文法(PCFG)

- 在基于PCFG的句法分析模型中，假设满足以下三个条件：
  - 上下文无关性(context-free)：子树的概率不依赖于子树控制范围以外的单词
  - 祖先无关性(ancestor-free)：子树的概率不依赖于推导出子树的祖先节点
  - 位置不变性(place invariance)：子树的概率不依赖于该子树所管辖的单词在句子中的位置

内容提要

- PCFG概述
- PCFG高效算法：CYK

PCFG句法分析模型：CYK算法

- 给定一个句子$W = w_1 w_2 \ldots w_n$和文法G，如何选择该句子的最佳结构？即选择句法结构树t使其具有最大概率:$argmax_t P(t|W, G)$？
  - CFG 句法分析算法可以直接用于PCFG
  - 既然有了概率，能否找到更高效的句法分析算法？
  - 能否用有向图求最优的方式解决？

概率CYK算法

- CYK算法的伪代码如下：

```
# 输入：待分析的词序列words和PCFG规则集R
# 输出：最有可能的句法结构书及其概率

function Probabilistic-CYK(words,R){
    for j <- from 1 to LENTGH(words) do
        for all { A | A -> words[j] ∈ R} //初始化三角阵
            table[j-1, j, A] <- P(A -> words[j])
        for i <- from j-2 downto 0 do //计算Viterbi变量，记录树结构的推导路径
            for k <- i+1 to j-1 do
                for all {A | A->BC ∈ R, and table[i,k,B] > 0 and table[k,j,C] > 0}{
                    if(table[i,j,A] < P(A->BC) X table[i,k,B] X table[k,j,C]) then
                        table[i,j,A] <- P(A->BC) X table[i,k,B] X table[k,j,C]) //存放最大的概率
                        back_trace[i,j,A] <- {k,B,C}
                }
    return BUILD_TREE(back_trace[1,LENGTH(words),S]),[1,LENGTH(words),S]// 返回生成的树结构和概率
}
```

- 其中，函数LENGTH(words)用于计算词序列words的长度，table[i,j,X]用于存放三角阵中以X为根节点，跨度范围从i到j的片段的概率；back_trace[i,j,A]用于存放以A为根节点，跨度范围从i到j的片段的子树结构；函数BUILD_TREE用于构造整个分析序列的句法结构树，根节点为S

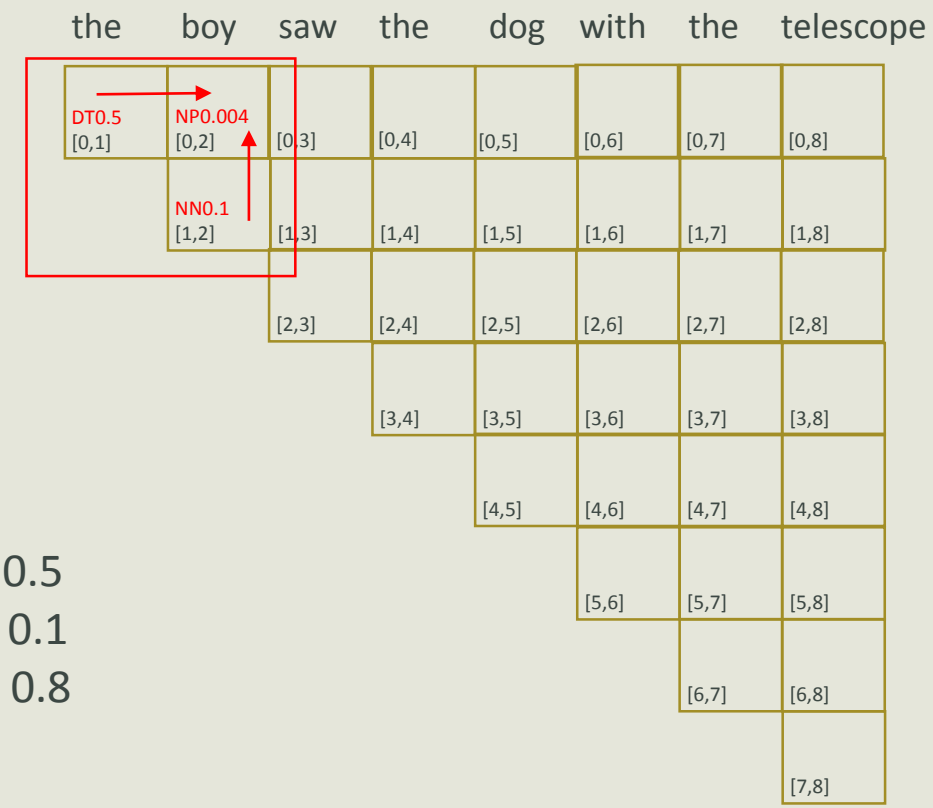
概率CYK算法：实例

- 给定如下PCFG G(S)：
- 非终结符集：N = {S,NP,VP,PP,DT,Vi,Vt,NN,IN}
- 终结符集合：$\Sigma = \{sleeps, saw, man, woman, dog, telescope, the, with, in\}$
- 规则集：

1. S -> NP VP	1.0	8.Vi -> sleeps	1.0
2.VP -> Vi	0.3	9.Vt -> saw	1.0
3.VP -> Vt NP	0.4	10.NN -> boy	0.1
4.VP -> VP PP	0.3	11.NN -> girl	0.1
5.NP -> DT NN	0.8	12.NN -> telescope	0.3
6.NP -> NP PP	0.2	13.NN -> dog	0.5
7.PP -> IN NP	1.0	14.DT -> the	0.5
		15.DT -> a	0.5
		16.IN -> with	0.6
		17.IN -> in	0.4

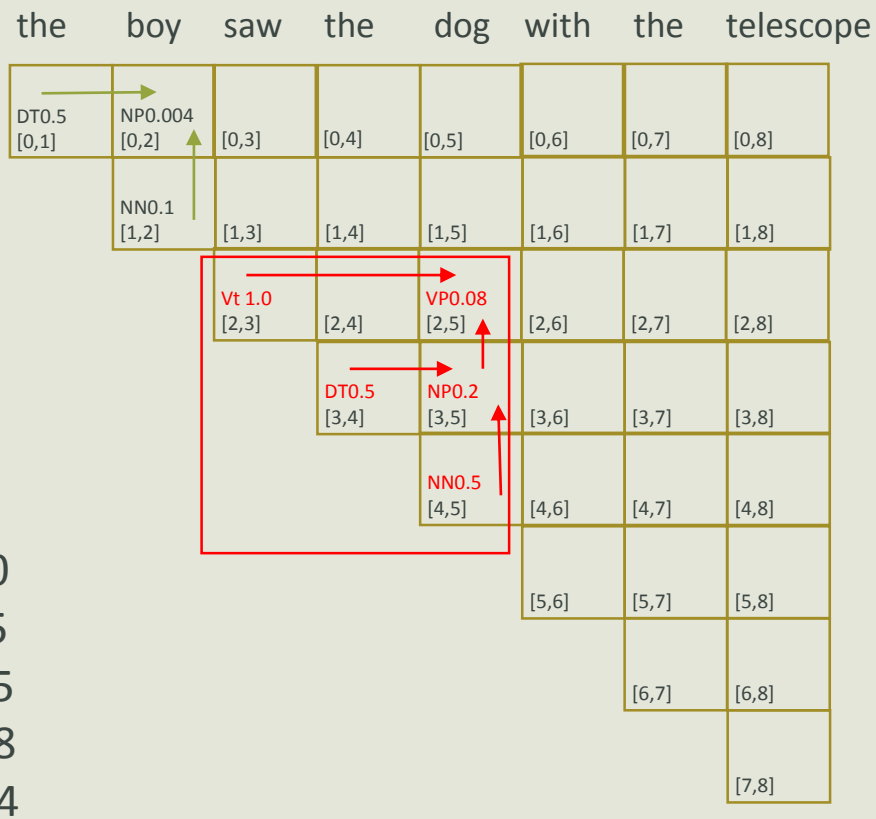
- 输入句子：the boy saw the dog with a telescope

概率CYK算法：实例



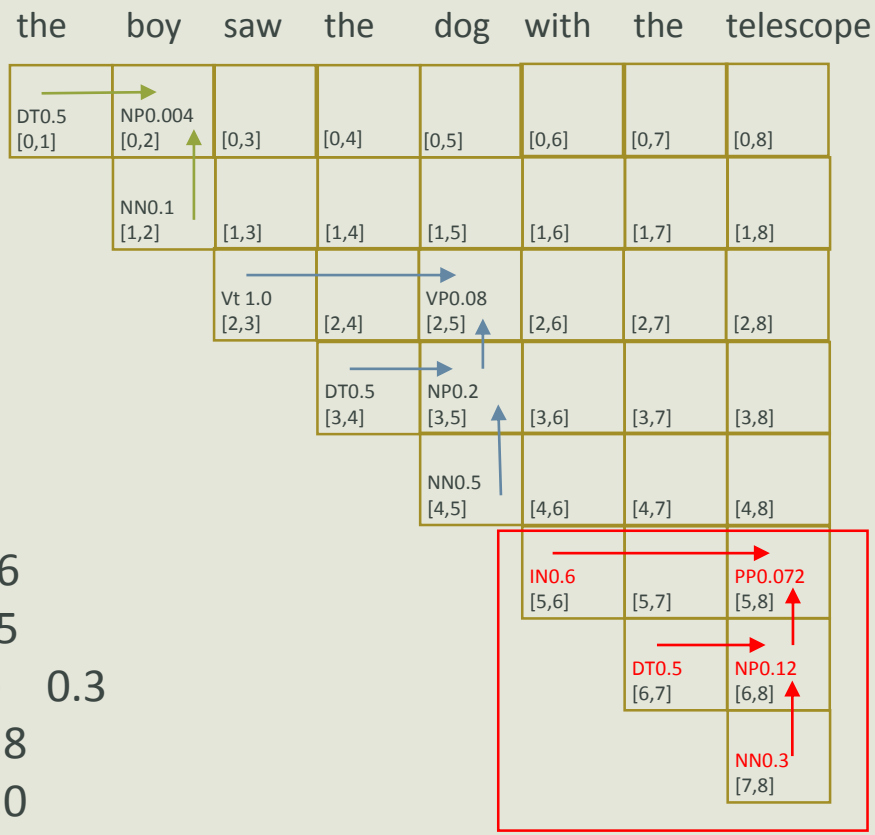
DT -> the 0.5
NN -> boy 0.1
NP -> DT NN 0.8

概率CYK算法：实例



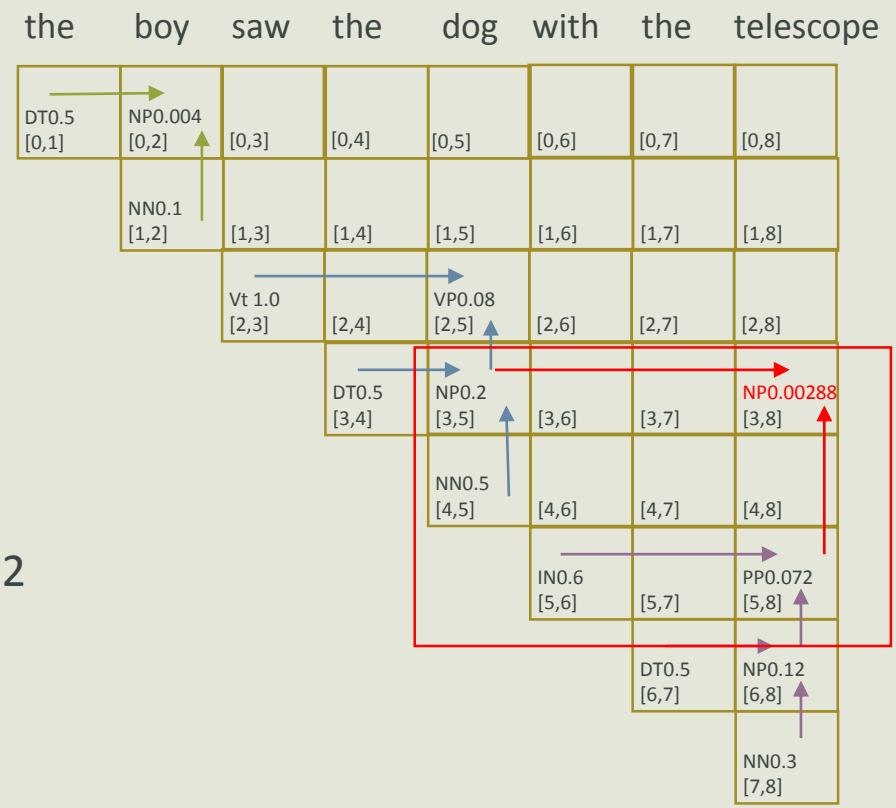
Vt -> saw 1.0
DT -> the 0.5
NN -> dog 0.5
NP -> DT NN 0.8
VP -> Vt NP 0.4

概率CYK算法：实例



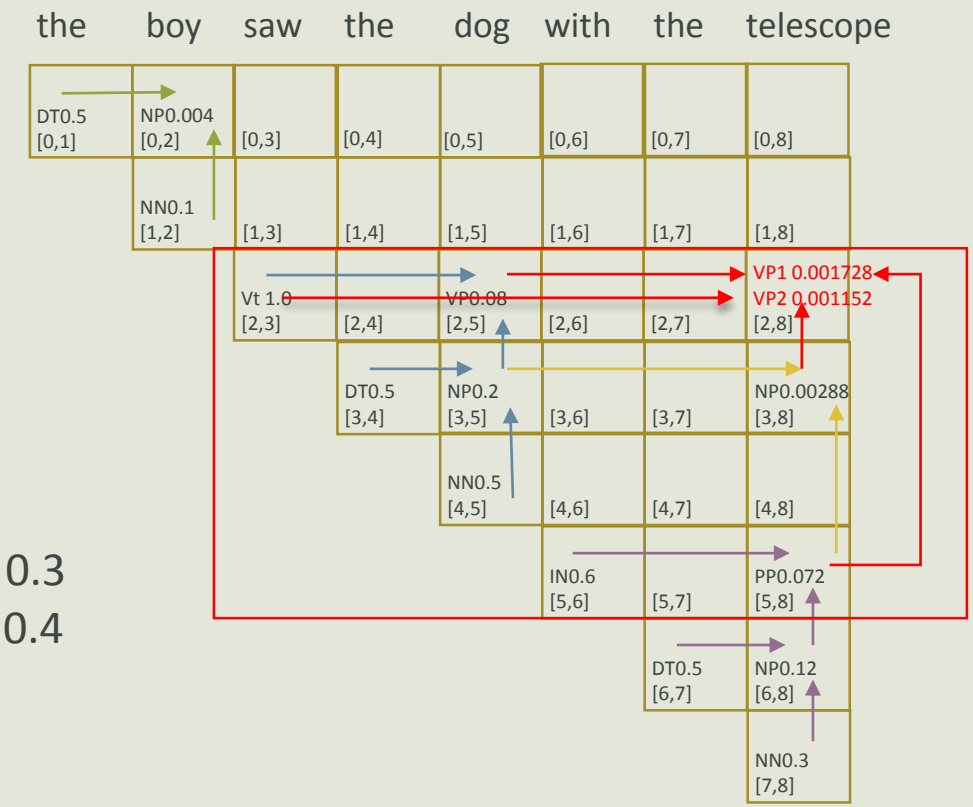
IN -> with 0.6
DT -> the 0.5
NN -> telescope 0.3
NP -> DT NN 0.8
PP -> IN NP 1.0

概率CYK算法：实例



NP -> NP PP 0.2

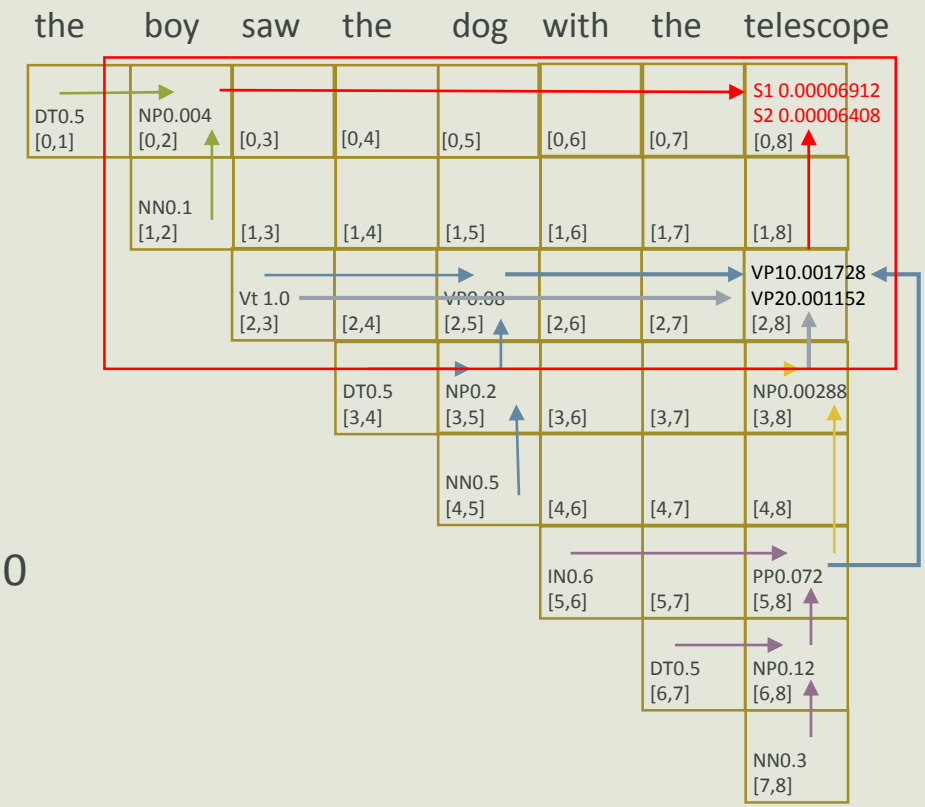
概率CYK算法：实例



VP1 -> VP PP 0.3

VP2 -> Vt NP 0.4

概率CYK算法：实例



S -> NP VP 1.0

剪枝策略：beam search

- 动态规划算法本身复杂度可以是指数级
  - 中等规模的CFG下，一个自然语言句子很容易出现百万parse
- Beam search（集束搜索）是一种启发式图搜索算法，通常用在图的解空间比较大的情况下
- 为了减少搜索占用的时间和空间，在每一步深度扩展的时候，减掉一些质量比较差的节点，保留质量较高的一些节点
- 优点是减少空间消耗，提高时间效率
- 缺点是有可能存在潜在的最佳方案被丢弃，beam search算法是不完全的

剪枝策略：beam search

- beam search使用广度优先策略建立搜索树，在树的每一层，按照启发代价对节点进行排序，然后仅留下预先确定的个数(Beam Width 集束宽度)的节点，仅这些节点在下一层继续扩展，其他节点就被剪掉了。如果集束宽度无穷大，那该搜索就是广度优先搜索
- 概念：
  - 为了达到搜索的目的，beam search引入了启发函数的概念(h)来估计从从当前节点到目标节点的损失
  - beam width B为每一层广度优先搜索算法保存的节点数目
  - BEAM用来保存下一轮扩展的节点
  - set 保存BEAM中，作为启发函数的输入空间
  - hash table 用于保存所有已经访问过的节点

剪枝策略：beam search

- 算法流程：

1. 将开始节点增加到BEAM和hash table
2. 循环遍历BEAM的所有后续节点并增加到set中，然后清空BEAM
3. 从set中选择B个启发函数值最优解的节点增加到BEAM及hash table中（已经存在hash table中的节点不能增加）
4. 以上过程循环持续进行直到找到目标节点或hash table已满或主循环结束后BEAM为空

内容提要

- PCFG概述
- PCFG算法实现：CYK and Beam Search
- PCFG的构建：树库与工具

PCFG来源：宾州树库

- 宾州大学语料库(Upenn Tree Bank)在1980年代末开始发起
- 由该校计算机系M.Marcus教授主持
- 1993年，完成了对近300万英语词的句子语法结构标注
- 2000年发布中文树库(第一版)，10万词，4185个句子，325 data files（新华社语料）
- 2004年发布中文树库4.0版，404156词，664633汉字，15162个句子，838data files（大陆，香港，台湾语料）

宾州树库

- 中文树库示例：
- 原始句子：他还提出一系列具体措施的政策要点。
- 词性标注：他/PN 还/AD 提出/VV 一/CD 系列/M 具体/JJ 措施/NN 和/CC 政策/NN 要点/NN 。/PU

```
例子：( IP ( NP-SBJ (PN 他 ))
            ( VP (ADVP ( AD 还 ))
                  ( VP ( VV 提出 ))
                        ( NP-OBJ ( QP ( CD 一)
                                          ( CLP ( M 系列 )))
                                    ( NP ( NP ( ADJP ( JJ 具体)
                                                      ( NP (NN 措施)))
                                          ( CC 和)
                                          ( NP ( NN 政策 )
                                                ( NN 要点 ))))))
            ( PU 。 ))
```

宾州树库

- Penn Treebank中包含了一个数据集the Wall Street Journal(WSJ)
- 该数据集被parsing community分为了一个标准的训练集(sections 02-21),一个开发集(section 24),一个测试集(section 23)，还有一些剩余的部分

句法分析器性能评测

- 目前比较广泛的句法分析器评价指标是PARSEVAL测度，三个基本的评测指标：
- 精度(precision):句法分析结果中正确的短语个数所占的比例，即分析结果中与标准分析树中的短语相匹配的个数占分析结果中所有短语个数的比例

$$P = \frac{\text{分析得到的正确的短语个数}}{\text{分析得到的所有的短语个数}} \times 100\%$$

- 召回率(recall):句法分析结果中正确的短语个数占标准分析树中全部短语个数的比例

$$P = \frac{\text{分析得到的正确的短语个数}}{\text{标准树库中的短语个数}} \times 100\%$$

- F-measure:$F = \frac{(\beta^2+1)\times P\times R}{\beta^2\times P+R} \times 100\%$

一般的，$\beta = 1$，称作F1测度

宾州树库推动了句法分析研究

- 关于parsing,近年来有相当多的工作，国外的著名代表人物是Collins和Charniak，近年来也有一些新人，如Bikel的多语言分析器的工作(http://www.cis.upenn.edu/~dbikel/home.html)
- 在宾州树库准确率接近90%

开源的短语句法分析器

- Collins Parser
- http://people.csail.mit.edu/mcollins/code.html
- Bikel Parser
- http://www.cis.upenn.edu/~dbikel/software.html#stat-parser
- Charniak Parser
- http://www.cs.brown.edu/people/ec/#software
- Oboe Parser
- http://www.openpr.org.cn/index.php/NLP-Toolkit-for-Natural-Language-Processing/

开源的短语句法分析器

- Berkeley Parser
- http://nlp.cs.berkeley.edu/Main.html#Parsing
  - 目前最好的句法分析器？
- Stanford Parser
- http://nlp.stanford.edu/downloads/lex-parser.shtml
  - SMT常用

内容提要

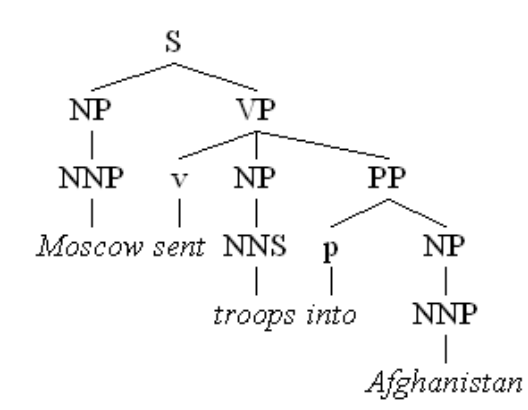
- PCFG概述
- PCFG算法实现：CYK and Beam Search
- PCFG的构建：树库
- 小结

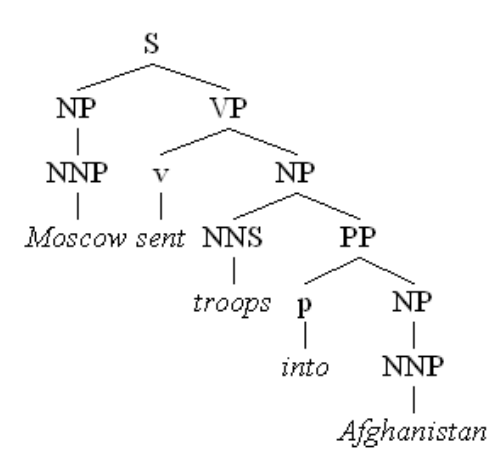
PCFG的优点

- 可利用概率减少分析过程的搜索空间
- 可以利用概率对概率较小的子树剪枝，加快分析效率
- 可以定量地比较两个语法的性能

PCFG的缺陷

- PCFG无法描述的两种现象：
- 结构相关性：
  - 同一个非终结符在不同的位置的推导概率不一样
  - 不同产生式的推导不是独立的。比如在双宾语结构中，第一个宾语一般是简单宾语。
- 词汇相关性
  - *Moscow sent troops into Afghanistan*





  - *correct* *incorrect*
  - 句法规则的推导依赖于具体单词（再比如rely on等搭配）

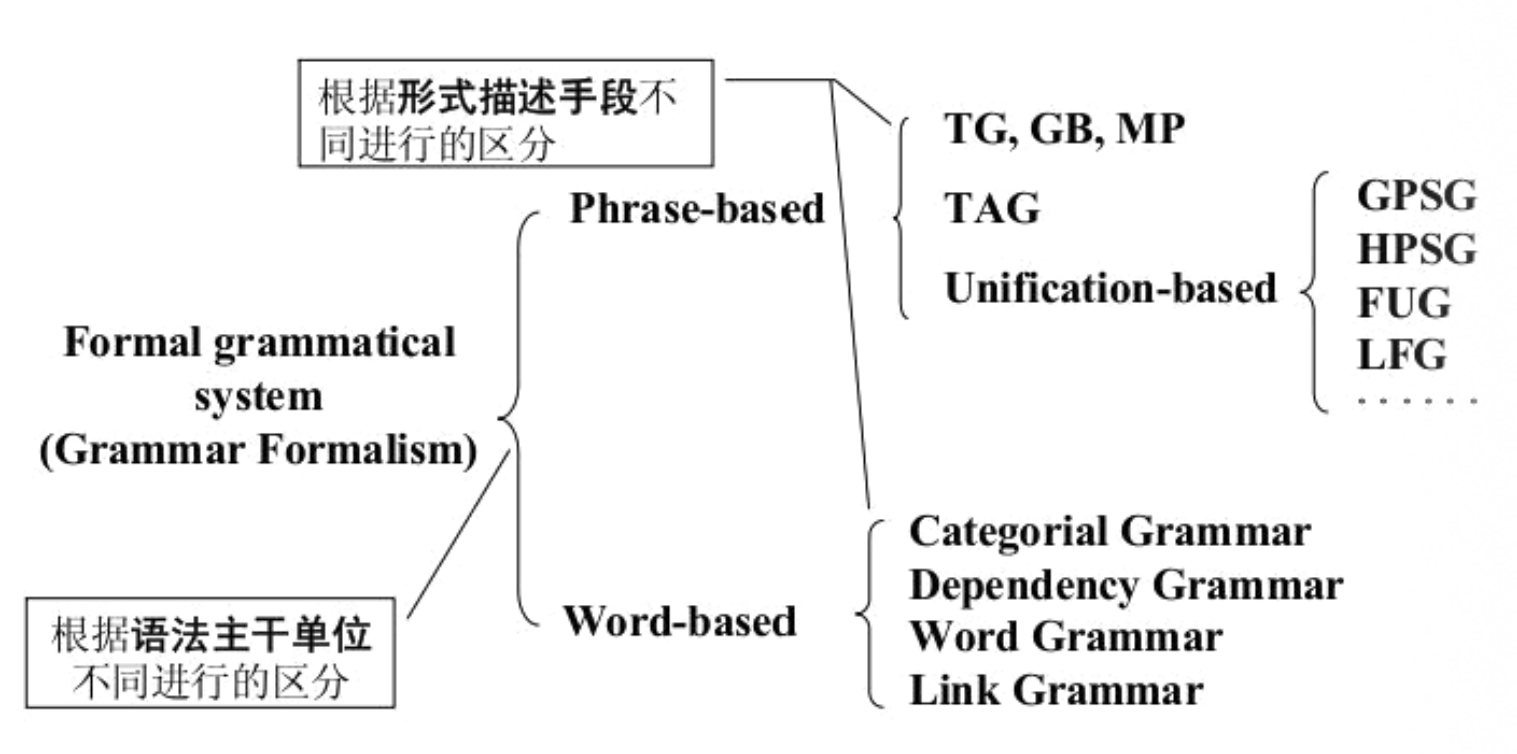
PCFG总结

- PCFG的句法规则：来自标记树库
  - 树库规模VS规则集的质量（完备性、可靠性）
- 自然语言是上下文相关的。
  - PCFG只是一个非常粗糙的概率模型，不一定适合描述自然语言.

附：当代形式语法理论体系分类

- 从目标来看：
  - 理论型语法系统
  - 应用型语法系统
- 从手段来看：
  - 产生式规则+特征结构
  - 树结构
  - 语法知识的词汇化
  - 自动机

分类层级


根据形式描述手段不同进行的区分
Formal grammatical system (Grammar Formalism)
Phrase-based
TG, GB, MP
TAG
Unification-based
GPSG
HPSG
FUG
LFG
. . . . . .
根据语法主干单位不同进行的区分
Word-based
Categorial Grammar
Dependency Grammar
Word Grammar
Link Grammar

词汇功能语法

- 词汇功能语法(Lexical Functional Grammar, LFG)于上个世纪70年代末由R.Kaplan 和 J.Bresnan在美国MIT提出
- 基本思想：依托短语结构语法已有的树结构，通过自底向上(bottom-up)层层传递的方式把词汇所负载的各种信息传播、汇集到上层节点中去，最终形成关于一个句子的完整的结构信息和功能信息描述
- 基本观点：句子由两个相对独立的层次来描述：

1. 成分结构层次：描述句子成分的结构关系
2. 功能结构层次：描述句子主语、谓语、宾语等之间的关系

词汇功能语法

- 要点：
- 突出词汇的作用，体现“语法结构”可以由某些词的意义预示出来
- LFG认为，动词、形容词和一部分名词在句子中的语法结构作用，相当于数理逻辑中的谓词逻辑(predicate),它们的词义决定它们拥有哪些论元(argument)。即如果知道谓词的意义，那么就可以知道以该谓词为中心的句子还会有哪类词出现，它们的语法语义功能是什么
- 两个基本作用：

1.可以准确的解释语言现象：谓词的管辖范围+谓词对论元的预示->确定语法结构和语义解释

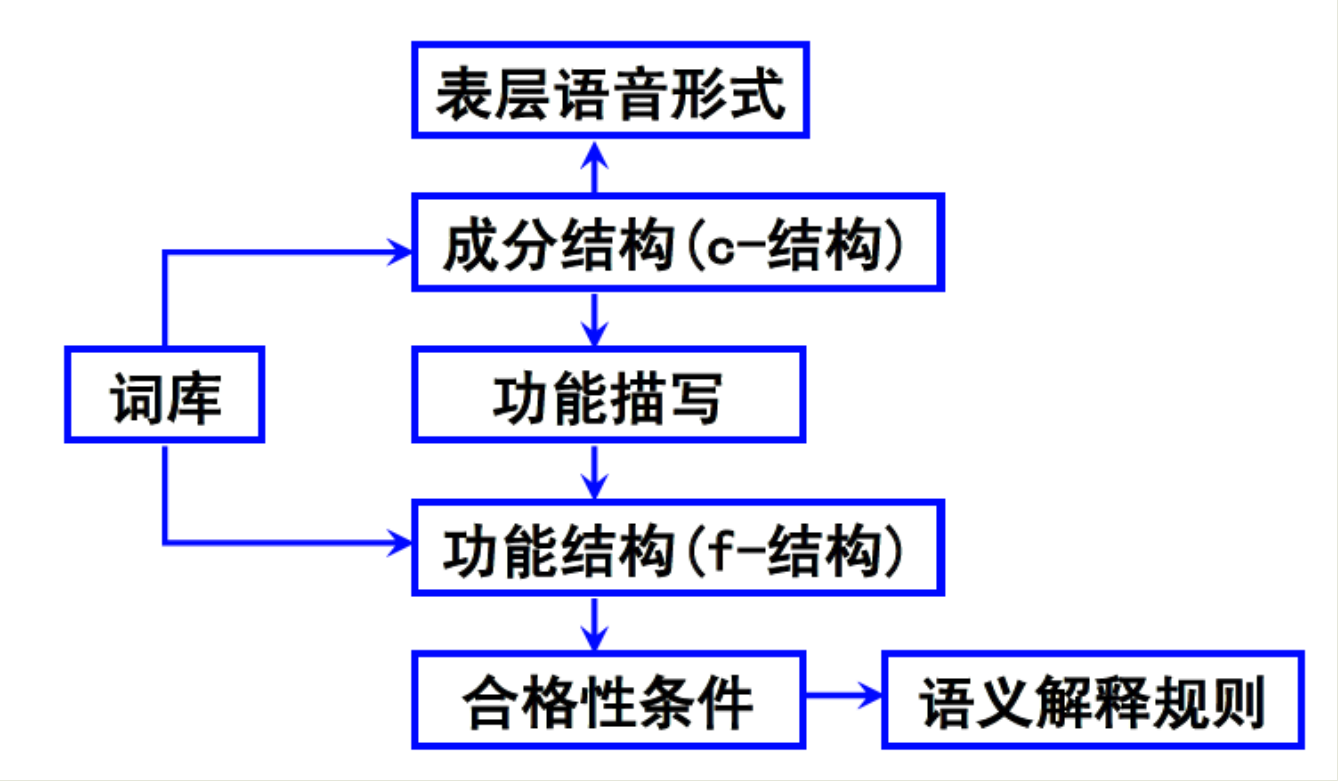
2.可以减轻语法规则的作用

词汇功能语法

- 要点：
- 把功能结构的描述作为语言描述中一个基本的独立层次
- LFG中的功能主要指语法功能，如主语、宾语、补语、修饰语，与传统的主语、宾语概念一致；时态、数、人称、格等语法特征；谓词功能
- LFG以功能为基础，定义句子的合格条件作为对成分结构的制约。有成分结构的句子不一定是合乎语法的句子，只有存在合法功能的句子，才是合乎语法的句子
- LFG本质上是一种以功能为基点的文法

词汇功能语法

- LFG理论的语言理解模式



词汇功能语法

- LFG的两个语法层次结构
- 成分结构(Constitute structure, c-结构)

用上下无关文法表示；树上的节点带有句子中词或短语所预示的功能信息，这些信息由语法规则右部的符号所带的功能注释表示

LFG的句法规则：

（1）

$$S \rightarrow \underset{(\uparrow SUBJ)=\downarrow}{NP} \quad \underset{\uparrow=\downarrow}{VP}$$

“↑”和“↓”称为直接支配元变量(immediate domination meta-variable)。“↑”表示规则的左部符号；↑$SUBJ$表示S的主语；“↓”表示带有该注释的符号本身

- 该例的含义是：句子S由NP和VP组成，其中Np所带的全部功能就是S的主语功能信息；Vp所带的全部功能信息就是S的谓词功能信息

词汇功能语法

LFG的句法规则：

(2) $NP \rightarrow DET \quad N$

表示 NP 可由限定词和名词组成。

(3) $VP \rightarrow \quad V \quad (NP) \quad (NP)$

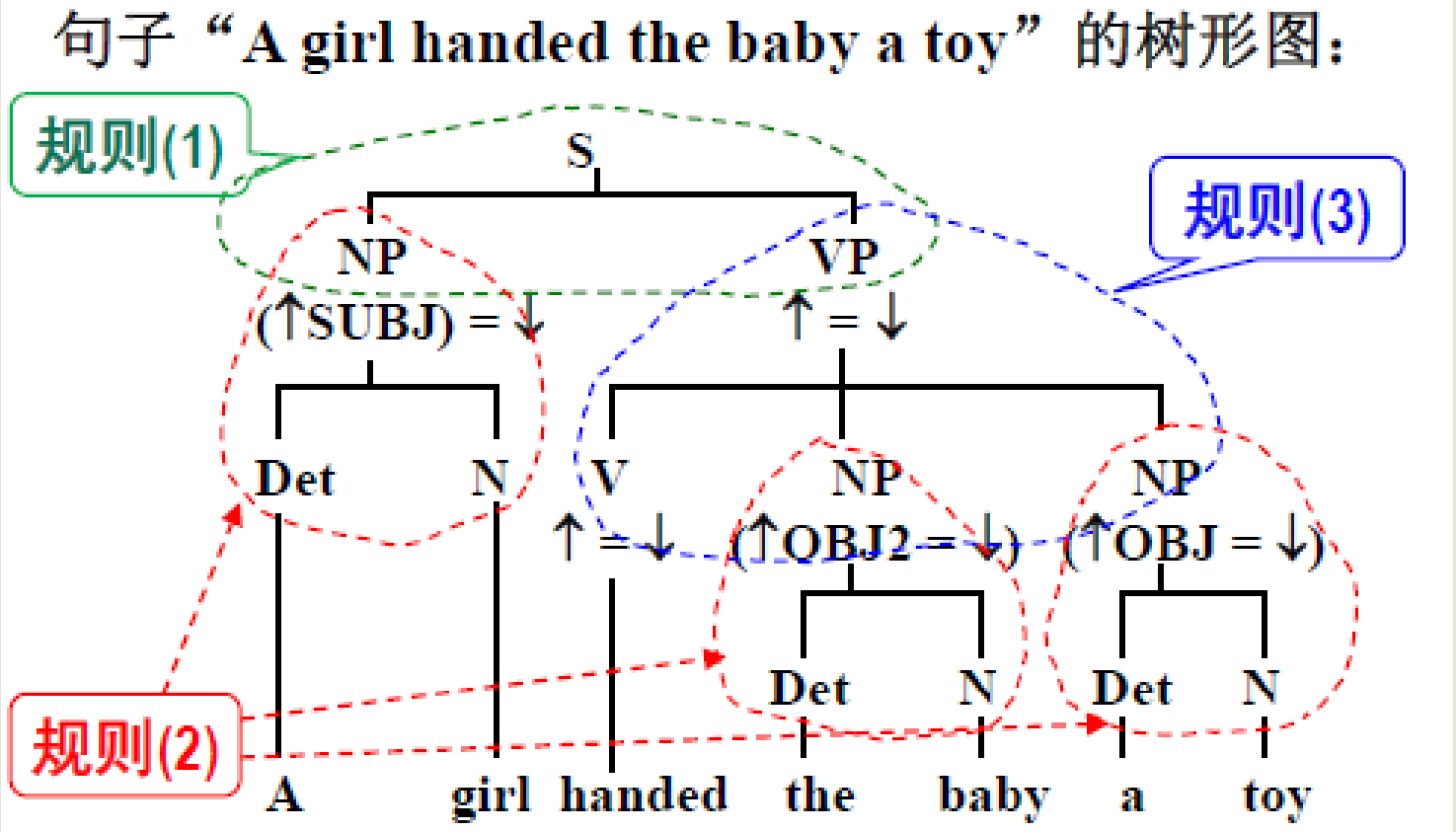
$\uparrow = \downarrow \quad (\uparrow OBJ2) = \downarrow \quad (\uparrow OBJ) = \downarrow$

NP 外的括号表示 NP 是可选的。

该规则表示动词所带的全部功能信息就是 VP 的功能信息；有如下3种可能: (*a*)VP 可由一个动词(不及物动词)组成；(*b*)由一个动词和一个 NP(及物动词带单宾语)，该 NP 的全部功能信息是 VP 的宾语的功能信息；(*c*)有另外一个 NP 参加(及物动词带双宾语)，该 NP 的全部功能信息是 VP 的第二宾语的功能信息。

词汇功能语法

LFG的句法规则举例



词汇功能语法

- LFG的词法规则
- 词法规则由词典信息提供。词法规则在LFG中有重要地位，它带语法功能结构的预示信息。例如：

a:	DET,	(↑ SPEC) = A
		(↑ NUM) = SG
girl:	N,	(↑ NUM) = SG
		(↑ LEX) = 'GIRL'

"(↑NUM) = SG" 表示 "其父结点具有的功能NUM(数)的值为SG (单数)"。

- LFG 把词汇按词的不同意义立项，词汇项所含的信息具有语法范畴和功能注释。功能注释的形式与语法规则的功能注释完全一致

格语法

- 格语法(case grammar)是美国语言学家Charless J. Fillmore于1966年提出的
- 基本观点：诸如主语、宾语等语法关系实际上都是表层结构上的概念，在语言的底层，所需要的不是这些表层的语法关系，而是用施事、受事、工具、受益等概念所表示的句法语义关系。这些句法语义关系，经各种变换之后才在表层结构中成为主语或宾语
- 格的定义：格语法中的格是“深层格”，它是指句子中体词(名词、代词等)和谓词(动词、形容词等) 之间的及物性关系(transitivity)，如：动作和施事者的关系、动作和受事者的关系等，这些关系是语义关系，它是一切语言中普遍存在的现象

格语法

- 这种格是在底层结构中依据名词与动词之间的句法语义关系确定的，这种关系一经确定就固定不变，不管经什么操作、在表层结构中处于什么位置、与动词形成什么语法关系，底层上的格与任何具体语言中的表层结构上的语法概念，如主语，宾语等，没有对应关系
- 例如：

The door opened

The key opened the door

The boy opened the door

The door was opened by the boy

The boy opened the door with a key

the boy:施事格

the door:客体格

the key：工具格

格语法

- 格表：

1.施事格(Agentive):动作的发生者

2.工具格(Instrumental):对动作或状态而言作为某种因素而牵涉到的无生命的力量或客体

3.承受格(Dative):由动词确定的动作或状态所影响的有生物。如，He is tall.

4.使成格(Factitive):由动词确定的动作或状态所形成的客体或有生物。或理解为：动词意义的一部分的客体或有生物。如：John dreamed a dream about Mary.

5.方位格(Locative):由动词确定的动作或状态的处所或空间方位。如：He is in the house

格语法

6.客体格(Objective):由动词确定的动作或状态所影响的事物。如：He bought a book

7.受益格(Benefactive):由动词确定的动作为之服务的有生命的对象.如：He sang a song for Mary.

8.源点格(Source)：由动词确定的动作所作用到的事物的来源或发生位置变化过程中的起始位置.如：He bought a book from Mary.

9.终点格(Goal)：由动词确定的动作所作用到的事物的终点或发生位置变化过程中的终端位置。如：I sold a car to Mary.

10.伴随格(Comitative)：由动词确定的与施事共同完成动作的伴随者。如：He sang a song with Mary.

格语法

- 用格语法分析语义：格框架约束分析
- 格框架中可以有语法信息，也可以有语义信息，语义信息是整个格框架最基本的部分
- 一个格框架可由一个主要概念和一组辅助概念组成，这些辅助概念以一种适当定义的方式与主要概念相联系。一般地，在实际应用中，主要概念可理解为动词，辅助概念理解为施事格、受事格、处所格、工具格等语义深层格

格语法

- 例：In the room, he broke a window with a hammer

```
[BREAK
    [ Case-frame:
        [Agentive: HE
         Objective: WINDOW
         Instrumental: HAMMER
         Locative: ROOM ]
        [MODALs:
            Time: past
            Voice: active ]]]
```

依存语法

- Robinson提出的依存关系四大公理

1.一个句子中只有一个成分是独立的

2.除独立成分外，句子中其他成分都必须依存于某成分

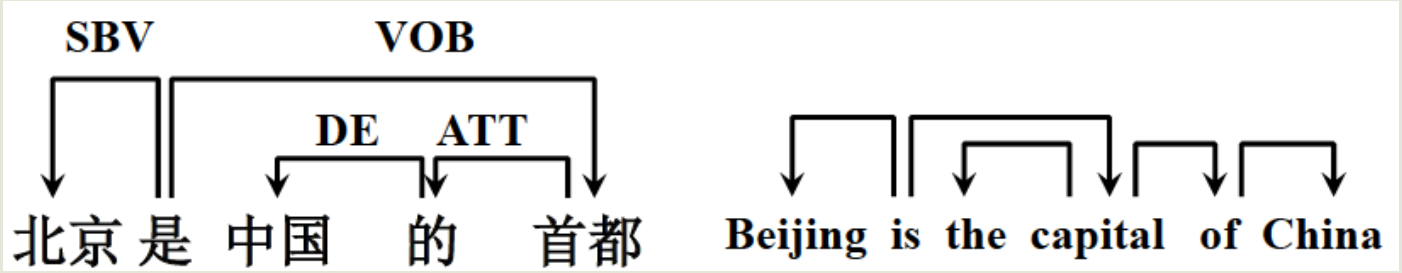
3.句中任何一个成分都不能依存两个以上的其他成分

4.如果A成分从属于B成分，而C成分在句中处于A和B之间，则C成分或者从属于A，或者从属于B，或者从属于A，B之间的某个成分

- 这四条公理相当于对依存图和依存树的形式约束为：单一父节点、连通、无环、可投射，由此保证句子的依存分析结果是一棵有根的树

依存语法

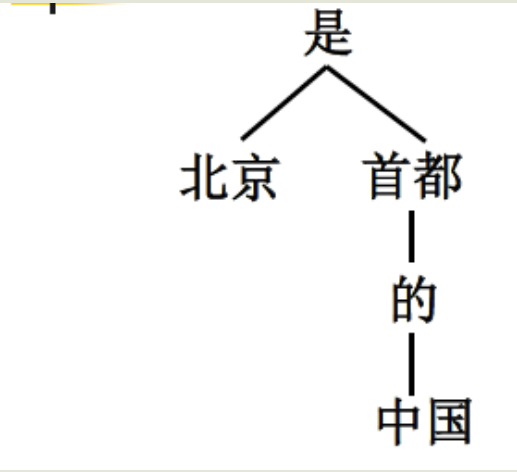
- 在依存语法理论中，“依存”就是指词与词之间支配与被支配的关系，这种关系不是对等的，而是有方向的。处于支配地位的成分称为支配者(governor、regent、head)，而处于被支配地位的成分称为从属者(modifier、subordinate、dependency)



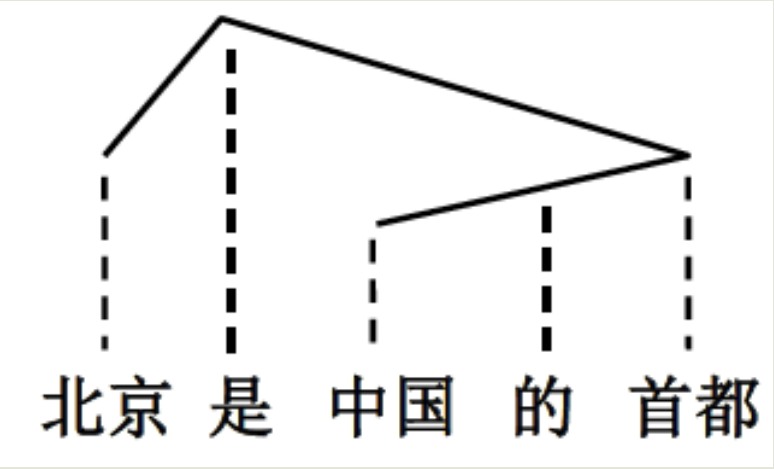
- 两个有向图用带有方向的弧(或称边，edge)来表示两个成分之间的依存关系，支配者在有向弧的发出端，被支配者在箭头端，我们通常说被支配者依存于支配者

依存语法

- 依存树：用树表示的依存结构，树中子节点依存于该节点的父节点
- 依存投射树：实线表示依存联结关系，位置低的成分依存于位置高的成分，虚线为投射线



依存树



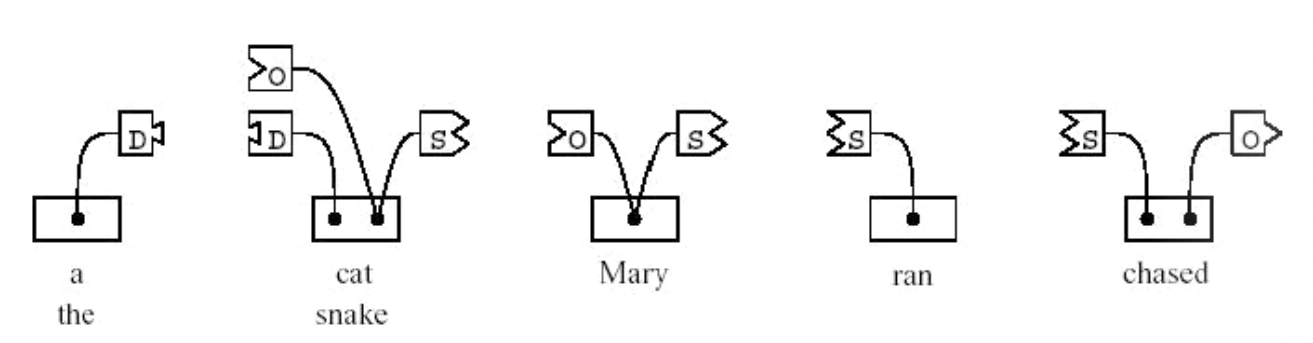
依存投射树

链语法

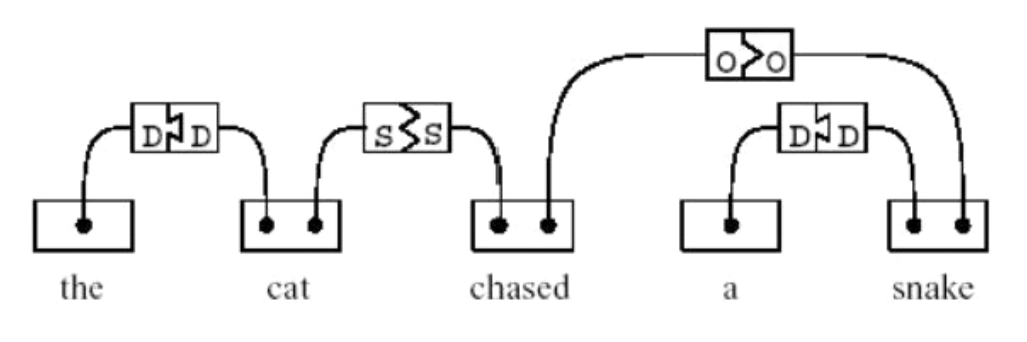
- 链语法(Link Grammar)由美国CMU计算机学院的Daniel Sleator和美国Columbia University音乐系的Davy Temperley共同提出
- 链语法不是建立在树结构的基础上，而是将语言知识完全落实到词汇基础上，通过词语的链接属性，来对句子进行分析
- 链语法对句子的分析结果表现为句中词汇间的链接关系，即不是树结构，而是图结构
- 跟其他形式语法系统相比，链语法是持强词汇主义观点的形式语法系统。它并不强调语言成分的层次组合关系，而是从词汇的局部着眼，力图揭示一个句子中任意两个词之间是否有联系，以及是什么联系

链语法

- 词典中带链接的词



- 由词链接而成的句子



链语法

- 一个语言的链语法就是该语言中的单词的集合，并且对每个单词都定义了它的链接要求(linking requirement)。单词的链接要求可以通过一个或若干个链接表达式(formula of connector)指定
- 例子：

单词	链接表达式	说明
a	D+	D是链接冠词（Determiner）和名词的链
the	D+	+ 表示向右链接
cat	D- & (O- or S+)	O是链接动词和宾语（Object）的链
snake	D- & (O- or S+)	S是链接动词和主语（Subject）的链
chased	S- & O+	& 表示逻辑上的“并且”关系
ran	S-	- 表示向左链接
Mary	O- or S+	or 表示逻辑上的“或”关系
dog	{@A} & D-	@A表示该单词可以有多个A链接

链语法

- 一个链接表达式由链接子(connector)、二元操作符(&,or)以及圆括号（规定了组合符号的优先顺序）组成
- 一个链接子由链名(name)和链接方向(direction)两部分组成
- 链名是一个符号串，用于标记两个单词之间的关系
- 链接方向有两个，向左(-)和向右(+)
- 如果一个链接表达式由多个链接子组合而成，其中有并列关系的链接子之间是有顺序要求的，比如cat的链接表达式是：

D-&(O-orS+)就不能写成(O-orS+)&D，如果写成后者，就会出现先有O链，再有D链

(实际上就是匹配的优先级，便于后续处理)

链语法

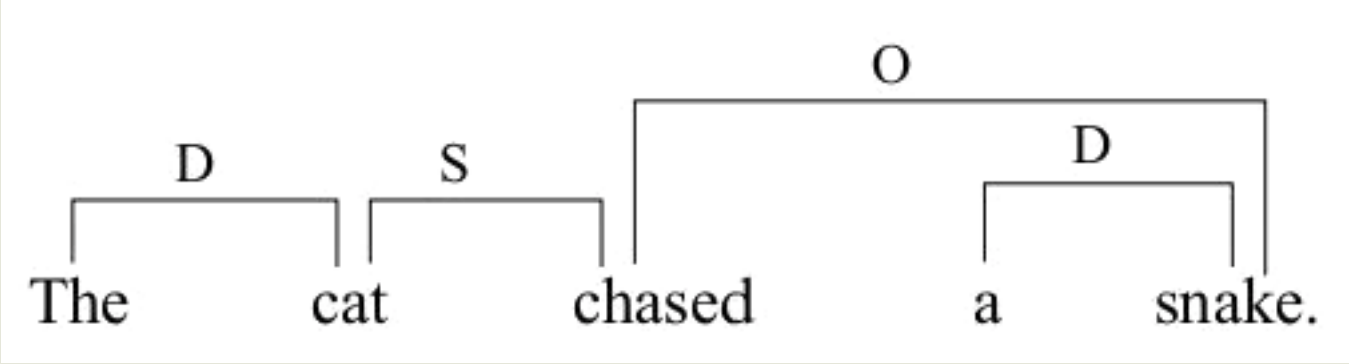
- 单词串中某个单词如果有一个向右的链接子，例如X+，而另一个单词有一个向左的链接子X-，那么这两个链接子相互匹配，在这两个单词之间就可以画一条X链。这时，我们说链接子X+或X-得到了满足或者说存在一个链接，满足了链接子X+或X-
- 链接表达式X&Y要被满足，则链接必须同时满足链接子X和Y
- 链接表达式XorY要被满足，则链接必须至少满足链接子X和Y中的一个

链语法

- 对于一个由单词组成的串S，如果根据一部链语法，S中所有单词的链接要求都得到满足且每个链接要求都只被满足一次，并且所有的链接符合下面4条元规则的要求：
- 平面性(Planarity)：链与链之间互相不交叉
- 连通性(Connectivity)：所有的单词应该链在一起，形成连通图
- 顺序性(Ordering)：链接表达式中靠前的链接子跟距离该单词较近的单词链接，链接表达式中靠后的链接子跟距离该单词较远的单词链接
- 排他性(Exclusion)：一对单词之间不能同时有两个链接
- 就说S是LG所定义的语言中的句子。使得S合法的全部链接称为一个链接集(Linkage)，链接集就是链语法分析句子的结果

链语法

- the cat chased a snake



结语

- 文法在NLP中常被成为formalism
- 根据语言规律或者NLP的应用需求，可以设计众多的formalism
  - GPSG
  - Category Grammar 范畴文法
  - GB theory 管辖约束
  - 配价文法
- 无论那种系统描述，针对的都是知识描述形式，系统的知识构建未见突破性进展。