


第三讲 汉语的句法规则系统 (4) 附录

詹卫东

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

当代形式语法理论概观

- 1) TG, GB, MP,
- 2) LFG, GPSG, HPSG,
- 3) PATR-II, DCG, FUG,
- 4) 树邻接语法 (TAG)
- 5) 链语法 (Link Grammar)
- 6) 范畴语法 (Categorial Grammar)
- 7) 依存语法 (Dependency Grammar)
- 8) 词语法 (Word Grammar)
-

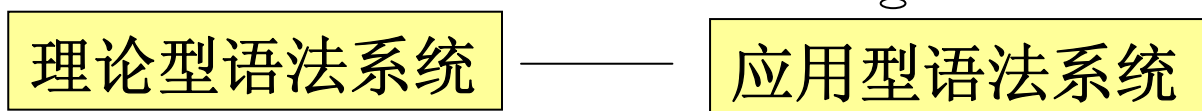


当代形式语法理论的不同阵营

当代形式语法理论体系的分类

分类角度：

(1) 从目标来看：



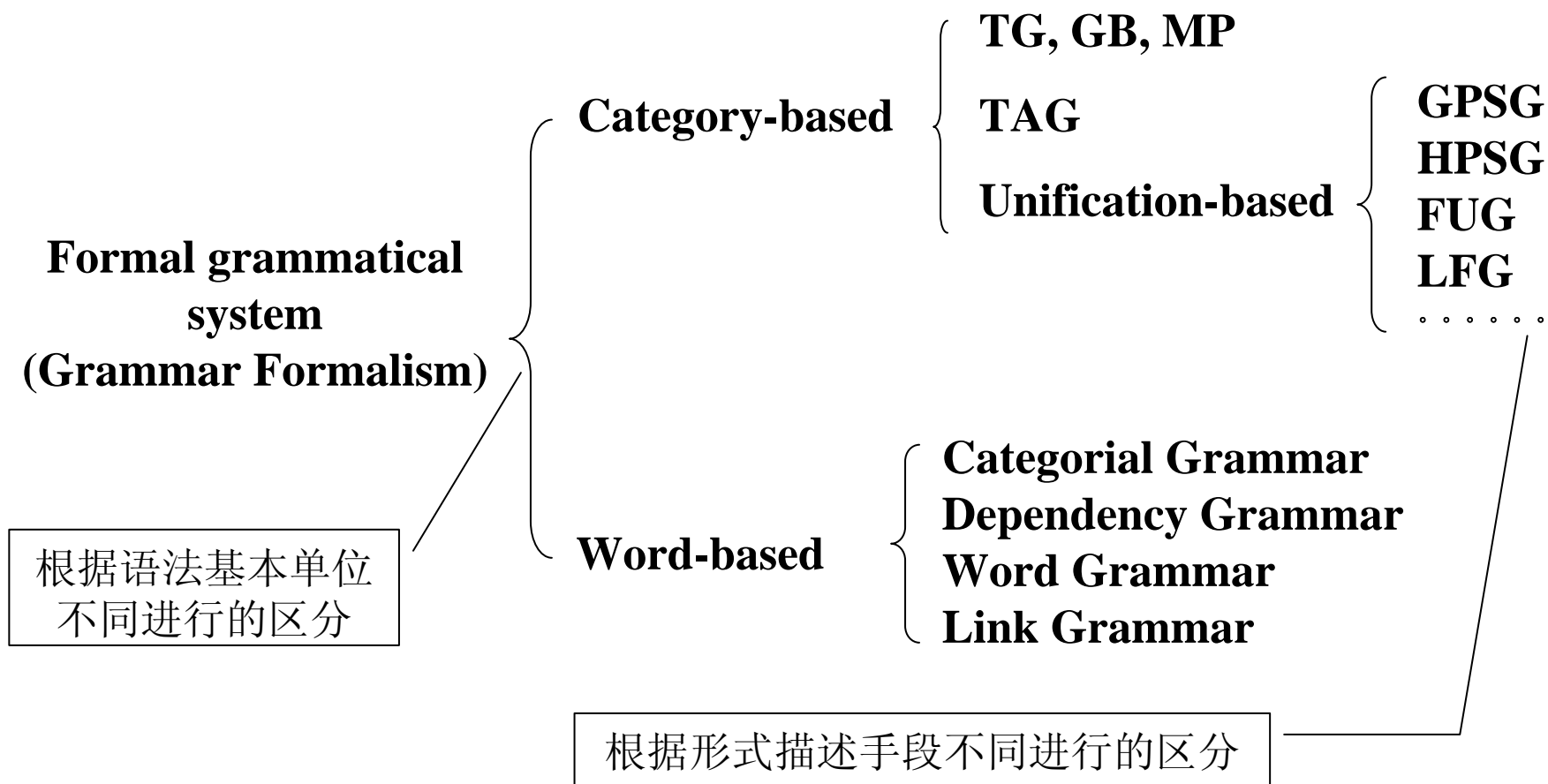
做了语法模型
干什么用呢？

(2) 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

剪刀! 锤子! 布! 这
么多工具, 用哪个
建模型呢?

一个根据语法基本单位不同和形式描述手段不同得出的分类层级



理论型语法系统

采用转换规则

TG(1957), GB(1981), MP(1992)

— Chomsky一脉

1假设深层结构;
2仅用原子范畴

LFG(1982)

GPSG(1985)

HPSG(1987)

— 非Chomsky派

1不假设深层结构;
2广泛采用复杂特征

不引入转换规则

应用型语法系统

- **PATR-II** (Stuart M. Shieber, 1984, 1986)
CFG + [特征结构 / 合一式 / 函数] (相互独立)
- **DCG** (F.C.N. Pereira & D.H.D. Warren, 1980)
语法规则可以直接转换为Prolog程序语言
- **FUG** (Martin Kay, 1979, 1985)
基于FD的规则 + 条件测试 (融合)
- **Lingware** (德国Sail Labs公司语言知识规则描述语言)
Lisp程序语言风格的规则

LFG（词汇功能语法）

- ◆ Lexical Function Grammar
- ◆ Kaplan, R. and J.Bresnan, 1982, *Lexical-Functional Grammar : A Formal System for Grammatical Representation*, In J.Bresnan ed. *The Mental Representation of Grammatical Relations*, MIT Press 1982.
- ◆ <http://www-lfg.stanford.edu/lfg/bresnan/>
- ◆ <http://www.parc.xerox.com/istl/members/kaplan/>

基本特点

- ◆ 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。
- ◆ 分析结果：
 - C-结构（constituent structure）：对句子的句法结构树描述
 - F-结构（functional structure）：对句中各成分的功能描述

带“修饰”的Rewriting Rule

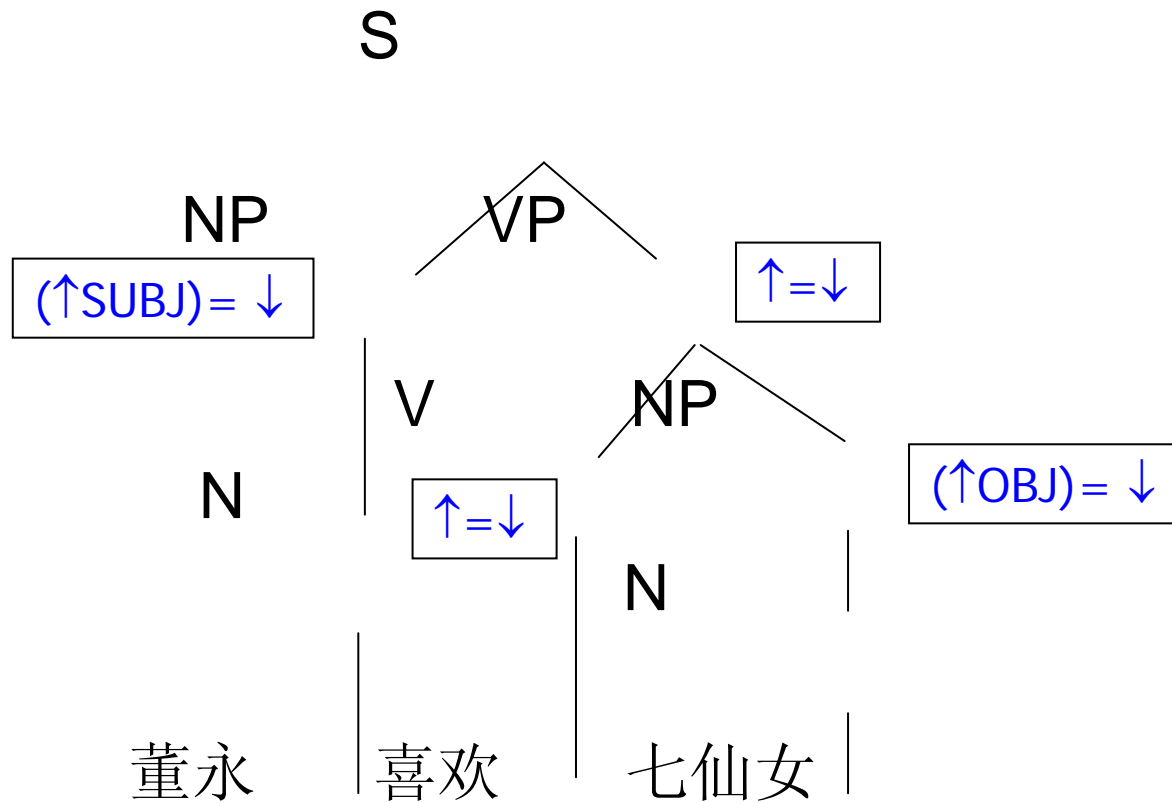
S →	NP	VP
	(↑SUBJ)= ↓	↑=↓
VP →	V	NP
	↑=↓	(↑OBJ)= ↓
VP →	V	
	↑=↓	
NP →	N	
	↑=↓	

↑ 表示父
亲节点；
↓ 表示当
前节点；
↑SUBJ 表
示父亲节
点的SUBJ
特征； =
表示合一

词汇信息描述（特征结构）

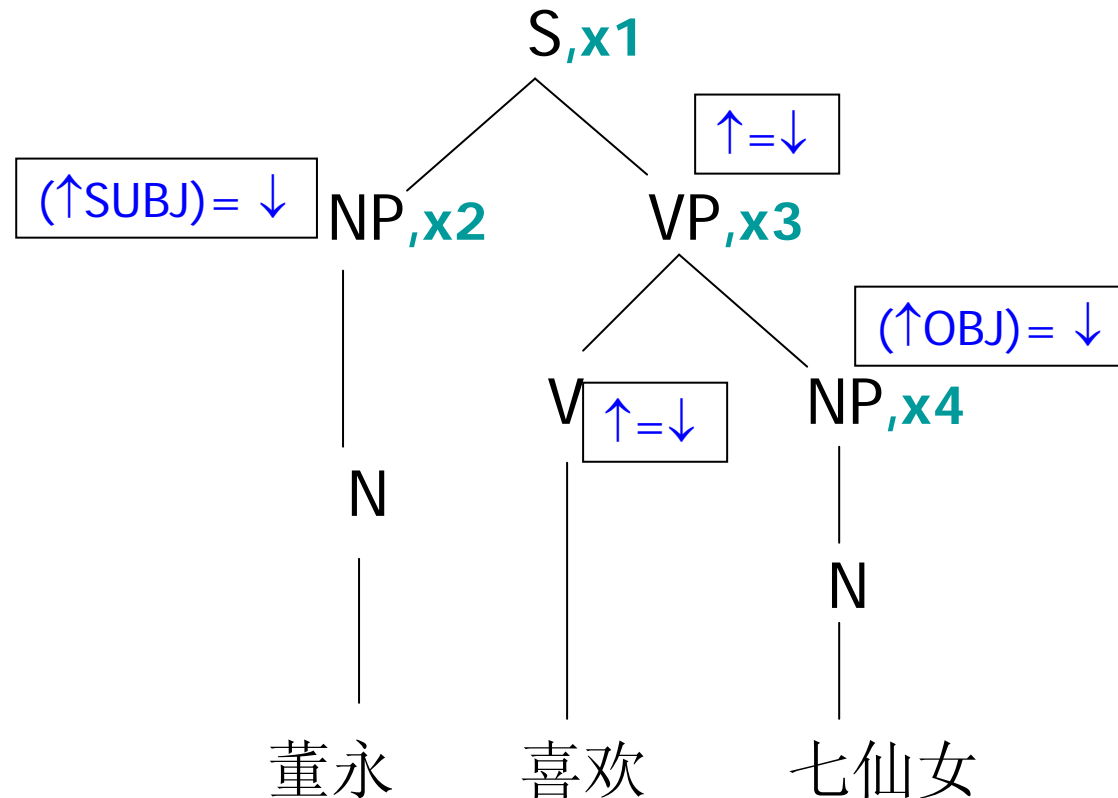
董永	N	(↑LEX)='董永' (↑SEM)=HUMAN (↑PERS)=3
七仙女	N	同上
喜欢	V	(↑PRED)='喜欢<(↑SUBJ), (↑OBJ)>' (↑SUBJ SEM)=HUMAN (↑OBJ SEM)=HUMAN

C-结构



从C—结构到F—结构

- 从成分结构导出功能方程;
- 从功能方程求解功能结构;



从成分结构导出功能方程

1. $x1.SUBJ = x2$
2. $x1=x3$
3. $x3.OBJ = x4$
4. $x2.LEX = \text{“董永”}$
5. $x2.SEM = HUMAN$
6. $x2.PERS = 3$
7. $x3.PRED = \text{“喜欢}<(x3 SUBJ),(x3 OBJ)>”}$
8. $x3.SUBJ.SEM = HUMAN$
9. $x3.OBJ.SEM = HUMAN$
10. $x4.LEX = \text{“七仙女”}$
11. $x4.SEM = HUMAN$
12. $x4.PERS = 3$

. 相当于“的”

从功能方程求解功能结构

$$x_1 = x_3 = \left[\begin{array}{l} PRED = \text{喜欢} < SUBJ, OBJ > \\ SUBJ = x_2 \\ OBJ = x_4 \end{array} \right]$$

$$x_2 = \left[\begin{array}{l} LEX = \text{董永} \\ SEM = HUMAN \\ PERS = 3 \end{array} \right]$$

$$x_4 = \left[\begin{array}{l} LEX = \text{七仙女} \\ SEM = HUMAN \\ PERS = 3 \end{array} \right]$$

“董永喜欢七仙女”的功能结构

$$FS(x_1) = \left[\begin{array}{l} PRED = \text{喜欢} < SUBJ, OBJ > \\ \\ SUBJ = \left[\begin{array}{l} LEX = \text{董永} \\ SEM = HUMAN \\ PERS = 3 \end{array} \right] \\ \\ OBJ = \left[\begin{array}{l} LEX = \text{七仙女} \\ SEM = HUMAN \\ PERS = 3 \end{array} \right] \end{array} \right]$$

LFG检查句子合法性的标准

- 一个属性只允许有一个值（唯一性）；
- 每个属性都应该有值（完备性）；
- 不该有的属性不应该有值（一致性）；

GPSG（广义短语结构语法）

- Generalized Phrase Structure Grammar
- Gerald Gazdar, Ewan Klein, Geoffrey K. Pullum, Ivan A. Sag, 1985, *Generalized Phrase Structure Grammar*, Oxford, England, Blackwell Publishing and Cambridge, 1985.
- <http://www.cogs.susx.ac.uk/lab/nlp/gazdar/gazdar.html>

基本特点

- 语法模型：
基本规则 -> 合格性检查 -> 树结构
- 基本规则：
 - 直接支配规则（Immediate Dominance Rule）
 - 元规则（Meta Rule）
- 合格性检查：
 - 特征共现限制（Feature Co-occurrence Restriction）
 - 线性顺序描述（Linear Precedence Statements）
 -

HPSG（中心词驱动的短语结构语法）

- Head-Driven Phrase Structure Grammar
- Pollard, Carl and Ivan A. Sag. 1987. *Information Based Syntax and Semantics*. CSLI Lecture Notes, No.13, The University of Chicago Press, Chicago.
- Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. The University of Chicago Press, Chicago.
- Sag, Ivan A. & Thomas Wasow, 1999, *Syntactic Theory: A Formal Introduction*, CSLI Publications, Stanford, California.
- <http://ling.ohio-state.edu/research/hpsg/>
- <http://hpsg.stanford.edu/>

基本特点

- 强调中心词在短语结构规则中的作用
 - 中心语 — 补足语规则 (Head - Complement Rule)
 - 中心语 — 指示语规则 (Head - Specifier Rule)
 - 中心语 — 修饰语规则 (Head - Modifier Rule)
- 产生式规则 + 特征结构 + 合一运算
- 基于中心词的属性特征传递 (Head Feature Principle, ...)
- 以同样的形式化方式表达句法知识和语义知识

PATR-II

- Shieber, Stuart M., 1984, *The Design of a Computer Language for Linguistic Information*, In Proceedings of Coling84(10th), Stanford University, Stanford, California.
- Shieber, Stuart M., 1985, *Using Restriction to Extend Parsing Algorithms for Complex-feature-based Formalisms*, In Proceedings of the 22nd Annual Meeting of the ACL, University of Chicago, Chicago, Illinois.
- Shieber, Stuart M., 1986, *An Introduction to Unification Based Approaches to Grammar*, CSLI Lecture Notes Series, No.4, Stanford University.
- <http://www.eecs.harvard.edu/~shieber/>

基本特点

- 在独立的**CFG**基础上增加特征结构与合一运算
- 在规则中引入附加范畴（比如通配符 `*. %`）
- 在合一式中引入函数（比如 `morph_gen()`）
- 引入逻辑表达符号增强规则表达能力（比如多个合一式之间可以有析取关系）

DCG

- Definite Clause Grammar
- Pereira, F.C.N., and D.H.D. Warren, 1980, *Definite Clause Grammars for Language Analysis - A Survey of the Formalism and a Comparison with Augmented Transition Networks*, In *Artificial Intelligence*, Vol.13, pp231-278, 1980.

基本特点

标准DCG中为 -->

- DCG语法可以方便地翻译为Prolog程序语言

C
F
G

S -> NP VP
VP-> V NP
NP-> N
N->董永

s(P1,P3) :- np(P1,P2), vp(P2,P3).
vp(P1,P3) :- v(P1,P2), np(P2,P3).
np(P1,P2) :- n(P1,P2).
is_n(董永).

D
C
G

₁董永₂喜欢₃七仙女₄

→

word(董永,1,2).
word(喜欢,2,3).
word(七仙女,3,4).

分析“董永喜欢七仙女”

`n(From,To) :- word(Word,From,To), is_n(Word).`

`v(From,To) :- word(Word,From,To), is_v(Word).`

`?- s(1,4).`

`** Yes.`

增加一条规则，描述短语内部结构

`s(P1,P3,s(NP,VP)) :- np(P1,P2,NP), vp(P2,P3,VP).`

`s(1, 4, s(np(n(董永)),vp(v(喜欢), np(n(七仙女))))).`

得到短语内部结构的分析结果

FUG（功能合一语法）

- Functional Unification Grammar
- Martin Kay, 1979, *Functional Grammar*, In Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society.
- Martin Kay, 1985, *Parsing in Functional Grammar*, In D.Dowty, L. Karttunen, and A. Zwicky eds, *Natural Language Parsing*, Cambridge University Press, Cambridge, 1985.

基本特点

- 弱化线性序结构关系
- 强调功能结构
- 对所有语言单位统一采用**FD**形式描述
- 适合于生成（**Generation**）