

语义分析

詹卫东

北京大学中文系

zwd@pku.edu.cn

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

Outline

1. 语义分析概述

- 语义分析任务描述
- 语义分析的困难

2. 词汇语义知识的类型

- 关于词义聚类关系的知识
- 关于词义组合关系的知识

3. 在句法分析中附加语义约束

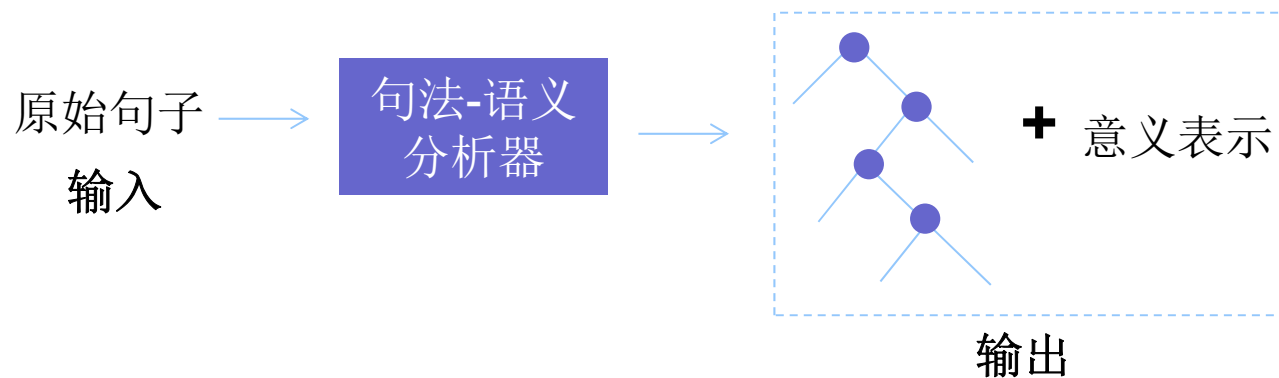
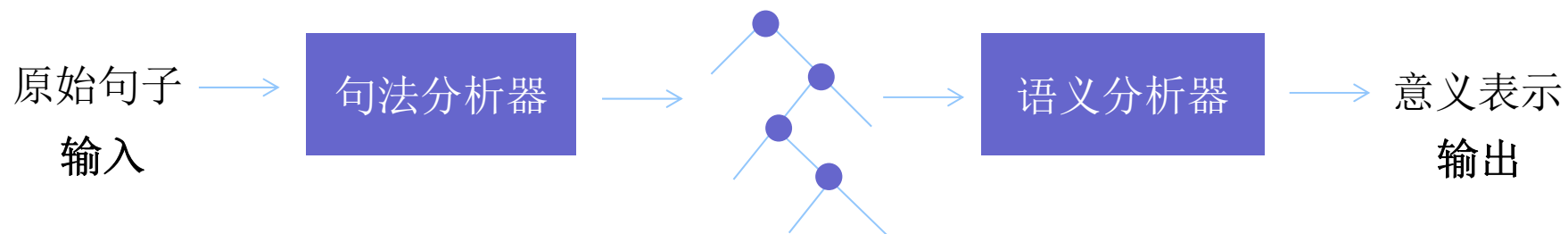
- 特征结构与合一运算
- 在CFG规则上增加合一约束

4. 语义知识描述的复杂性

5. 小结

1 语义分析概述

(1) 句法语义接续的管道模型



(2) 句法语义融合的一体化模型

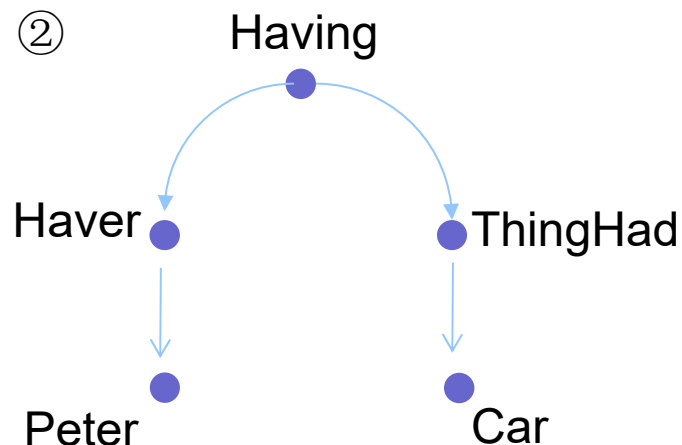
什么是意义 —— 如何表示意义？

□ Peter has a car.

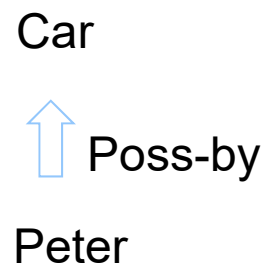
①

$\exists x, y \{ \textit{Having}(x) \wedge \textit{Haver}(\textit{Peter}, x) \wedge \textit{ThingHad}(y, x) \wedge \textit{Car}(y) \}$

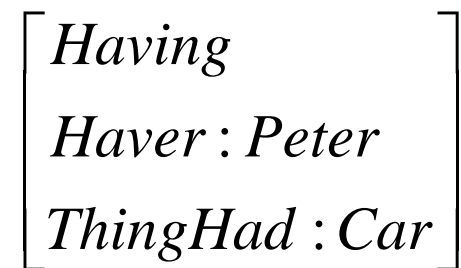
②



③



④



① 一阶谓词逻辑表达式

② 语义网络

③ 概念依存图

④ 框架图 / 特征结构

什么是意义 —— 如何表示意义？

意义 := 符号/形式变换


I. 在一种语言内进行的符号变换

A. 张三打了李四 → B. 李四被张三打了

II. 在不同语言之间进行的符号变换

A. 张三用手打了李四 → B. Zhang San hit Li Si with his hand

III. 在不同性质的符号系统之间进行的符号变换

A.  → B. 这是残疾人通道

A 的 意思 是 B

Paraphrase – 解释意思

- If you paraphrase someone or paraphrase something that they have said or written, you express what they have said or written **in a different way.**

—— Collins COBUILD Dictionary

- para- “beside” + phrazein “to tell”, to tell **in other words**

—— <http://www.etymonline.com/>

“释义”的作用示例：从自然语言问句到SQL查询

■ Q: Which country is Athens in ?

city_table



形式变换：意义解释

SQL:

City	Country	Population
Athens	Greece	1368
Bangkok	Thailand	1178
...

Select Country From city_table Where City = “Athens”

Answer: Greece

语义分析的困难

□ 从自然语言的表达形式到语义解释，存在很大的“不确定性”

1. 一个形式，多个意义（一个形式可以变换成多种形式）
2. 多个形式，一个意义（多个形式可以对应同一个形式）
3. 从形式到形式的变换似乎没有止尽（意味深长）

一个形式可以变换为多个形式：多义句

例1a

人是会死的

苏格拉底是人

苏格拉底是会死的

例1b 鲁迅的作品不适合小学生读

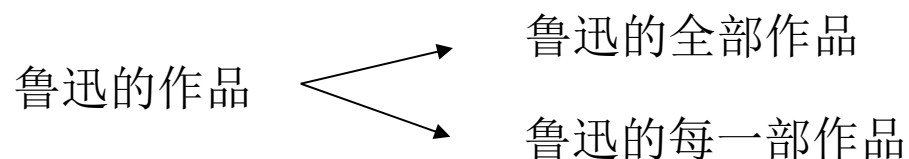
《祝福》是鲁迅的作品

《祝福》不适合小学生读

例1c 鲁迅的作品不可能在一天之内全部读完

《祝福》是鲁迅的作品

《祝福》不可能在一天之内全部读完



多个形式可以变换为同一个形式：同义句

- 例1
- a. 学一食堂 供应 西餐
 - b. 学一食堂 卖 西餐
 - c. 学一食堂 提供 西餐
 - d. 学一食堂 有 西餐
 - e. 学一食堂 做 西餐
 - f. 学一食堂 经营 西餐

供应(学一食堂, 西餐)

- 例2
- a. 董永总是支持七仙女。
 - b. 董永都听七仙女的。
 - c. 董永永远对七仙女点赞。
 - d. 董永从不对七仙女说不。

支持(董永, 七仙女)

意义解读 “无止尽”

A	B
“我买了辆车” →	意思1：我付钱从某处购买了一辆车
	意思2：我拥有了一辆车
	意思3：我可以使用这辆车
	意思4：我的钱都花完了

A 意味着 B {B1,B2,B3,B4,...}

意义解读：预设义、蕴涵义、隐含义

A. 小张跟小王后悔结婚了

序号	例子	意义解读	说明
1	B. 小张跟小王结婚了	预设义	谈话双方都接受的前提 (1) A真B真, A假B亦真 (2) B假则A无所谓真假, B真A未必真
2	B. 小张跟小王都后悔一件事情	蕴涵义 (衍推)	句子的真值条件义 (1) A真则B真, A假B未必假 (2) B假则A假, B真A未必真
3	B. 小张跟小王对婚姻不满意	隐含义	句子的会话含义 (1) A真B一般为真, 但特殊情况也可为假 (2) B假A不失为真, 但说话人明知B假还说A, 违反了语言交际的“合作原则”(量准则、质准则、相关准则、方式准则)

2 词汇语义知识的类型

- ① 关于词义聚类关系的知识
 - 语义特征/义素描述 (semantic components)
 - 语义分类树 (taxonomy)
 - 语义关系网 (semantic network)

- ② 关于词义组合关系的知识
 - 配价理论 (valency theory)
 - 格语法/论元结构理论 (case grammar)
 - 框架语义学理论 (frame semantics)
 - 生成词库论 (generative lexicon theory)

2.1 关于词义聚类关系的知识

语义知识的类型	代表性的语义知识库/工程项目
语义特征（义素）描述	HowNet ^[1]
语义分类树	905语义工程 ^[2]
语义关系网	HowNet WordNet ^[3] / 北大CCD MindNet

[1] <http://www.keenage.com>

[2] http://ccl.pku.edu.cn/doubtfire/Semantics/973_Beida/sem_project_list.htm

[3] <http://wordnet.princeton.edu>

2.1.1 语义特征（义素）描述

→ n个义素

词语 \ 义素	婚姻	成年	男性	人
光棍	—	+	+	+
寡妇	—	+	—	+
女童	—	—	—	+
男婴	—	—	+	+
单身汉	—	+	+	+

2ⁿ个词语

语义特征（义素）描述

→ n个义素

	对象	接触	环境	工具	方式	...
炒	荤、素、面	间接	固体	
烤	荤、面、素	直接	固体	
炸	荤、面、素	间接	油	
煮	荤、面、素	间接	水	
蒸	面、荤、素	间接	汽	
炖			
...						

2ⁿ个词语

语义特征（义素）描述

1. 语义特征的性质

概念义（客观义）

附加义（感情色彩义/语体义/主观义）

2. 获取语义特征（义素）的方式：

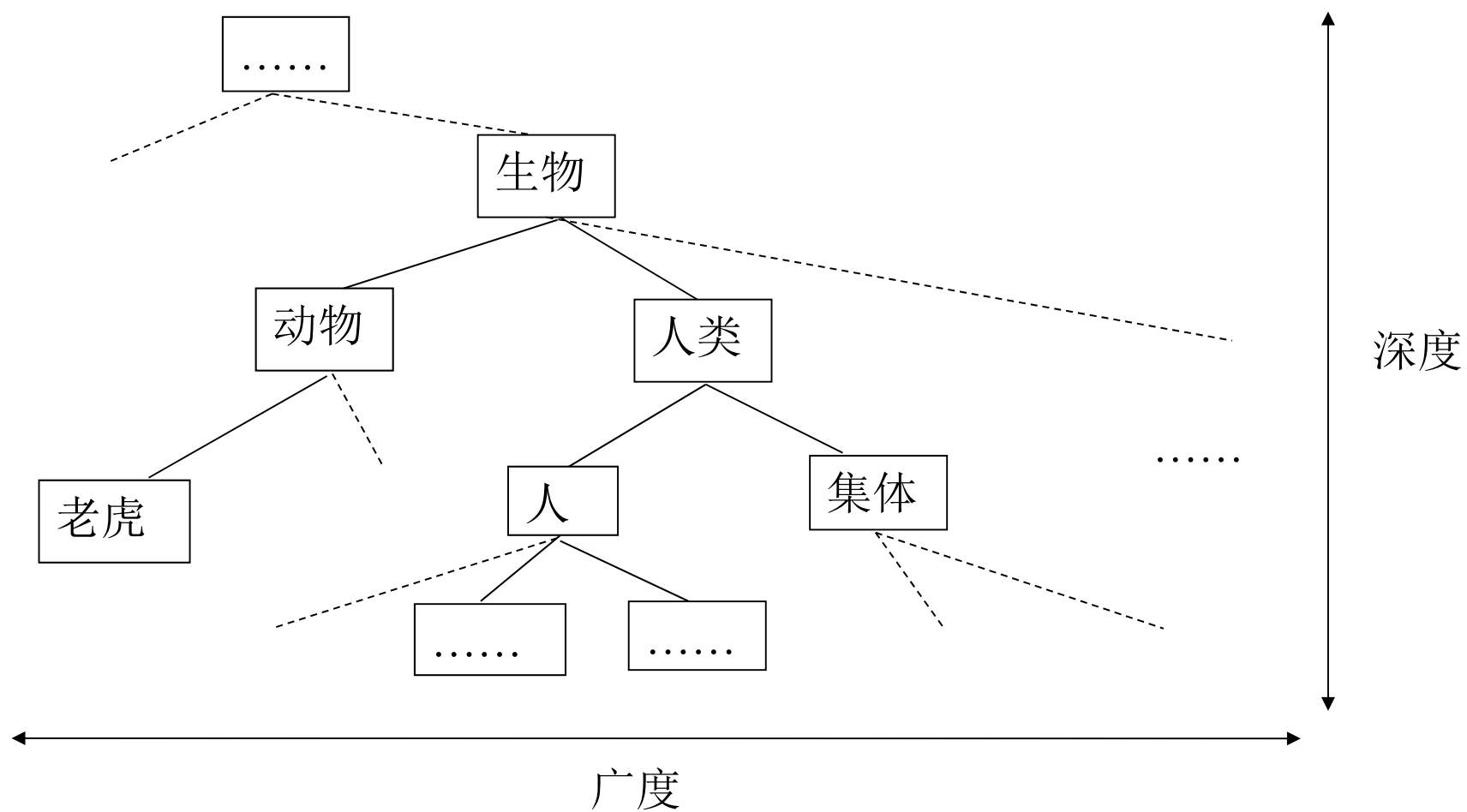
2.1 参考词典释义，提取共同项/对立项特征；

2.2 考察词语在句法格式中使用时的变换差异和对比差异，对具有相同/相反的句法行为的词语，提取共同项/对立项特征；

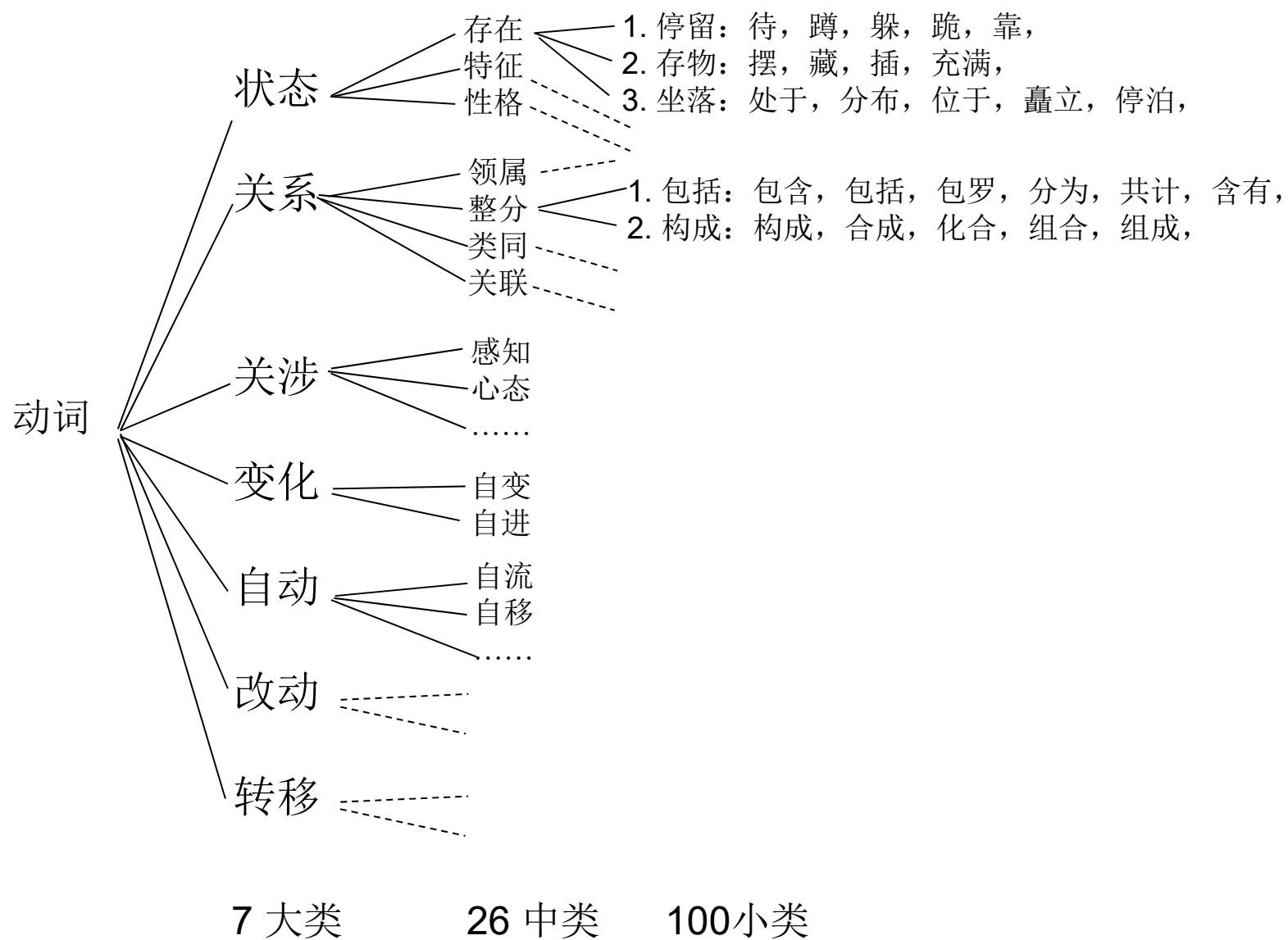
方式 — 方法 — 手段 — 措施 — 伎俩
睿智 — 聪明 — 灵巧 — 狡猾 — 阴险 — 奸诈
见面 — 碰头 — 会晤

死亡，死了，没了，去了，归天，翘辫子，咽气，牺牲，逝世，驾崩，薨，……
die, breathe one's last, go forth, kick the bucket, depart, pass away, expire, decease……

2.1.2 语义分类树



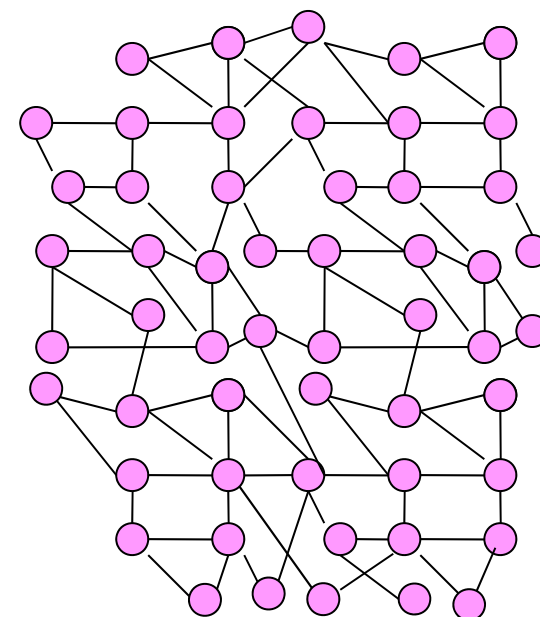
语义分类树



2.1.3 语义关系网

Attribute属性	Goal目标	Possessor领有者
Cause原因	Hypernym上位	Purpose意图
Co-Agent联合施事	Location场所	Size大小
Color颜色	Manner方式	Source源点
Deep_Object深层宾语	Material材料	Subclass子类
Deep_Subject深层主语	Means方法	Synonym同义
Domain领域	Modifier修饰语	Time时间
Equivalent同位	Part部分	User使用者

MindNet



<http://stratus.research.microsoft.com/mnex/InputPath.aspx?l=e&d=d>

语义关系网的自动获取示例

car: a vehicle with 3 or usually 4 wheels and driven by a motor, esp. one for carrying people

文本来源: Microsoft Encarta

```
car
|
|___Hyp>---- vehicle
|
|___Part>---- wheel
|
|___<Tobj ---- drive
|               |
|               |___Means>--- motor
|
|___Purp>----- carry
|               |
|               |___Tobj>--- people
```

Stephen D. Richardson, William B. Dolan, Lucy Vanderwende, 1998, MindNet: acquiring and structuring semantic information from text, In Proceeding of COLING, 1998.

2.2 关于词义组合关系的知识

理想目标：描述任意词语之间的组配关系

语义知识理论	代表性的语义知识库/工程项目
配价语法	剑桥综合语言数据库（ILD）， 北大配价语义词典 ^[1]
格语法/论元结构理论	现代汉语述语动词机器词典， ^[2] 美国宾州大学Propbank ^[3]
框架语义学	美国伯克利加州大学FrameNet ^[4]
生成词库论	Brandeis Semantic Ontology (BSO) ^[5]

[1] http://ccl.pku.edu.cn/doubtfire/Semantics/CUP_ILD/Introduction-to-ILD.htm
http://ccl.pku.edu.cn/ccl_sem_dict/

[2] 林杏光 主编（1994）《现代汉语述语动词机器词典》，北京语言学院出版社1994年版。

[3] <http://verbs.colorado.edu/~mpalmer/projects/ace.html>

[4] <https://framenet.icsi.berkeley.edu/fndrupal/>

[5] <http://www.cs.brandeis.edu/~arum/publications/cpabso.html>

- **配价 (valence) 这一概念借自化学。目的是说明一个动词能跟多少个名词性成分发生组合关联。**

游泳	[某人] 游泳	1价动词
吃	[某人] 吃 [某食物]	2价动词
送	[某甲] 送 [某乙] [某物]	3价动词

配价数

语义：动作行为跟 x 类事物有意义联系。

句法：动词周围有 x' 个空位安放跟它有意义联系的名词。

V

零元（价）动词

___ V

一元（价）动词

___ V ___

二元（价）动词

___ V ___ ___

三元（价）动词

.....

X 元（价）动词

X 价动词

□ _____ 为(向, 对) _____ V

- 董永 总为 别人 着想
- 大家 向 董永 看齐
- 董永 对 七仙女 发火

□ _____ V _____ VP

- 那个犯人 企图 逃跑
- 董永 觉得 不舒服

□ _____ V _____ NP _____ VP

- 董永 告诉 七仙女 晚上六点开会

从动词的配价到形容词、名词的配价

形 — 名 配价

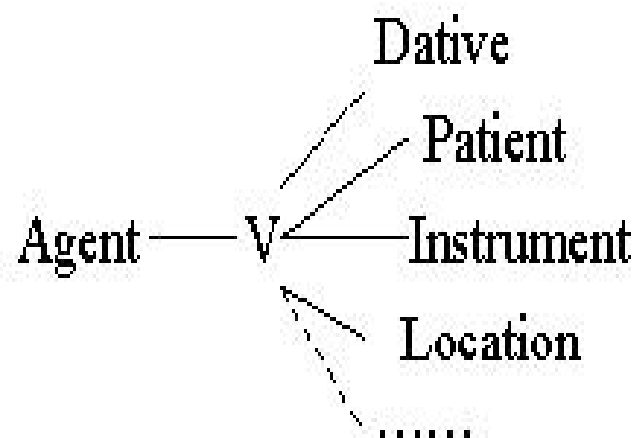
优秀	[某人] 优秀	1价形容词
友好	[某甲] 对 [某乙] 友好	2价形容词

名 — 名 配价

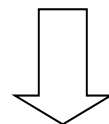
质量	[某物] 的 质量	1价名词
态度	[某甲] 对 [某乙] 的 态度	2价名词

论元类型 - 论旨角色/题元

□ 反映动词所能结合的名词的不同语义角色



主语 + 动词 + 宾语



施事 + 动作 + 受事
施事 + 动作 + 工具
处所 + 状态 + 当事
.....

— 句法模式

细化

— 语义模式

论旨角色的定义方式

□ 原型施事 (proto-typical agent)

题元 \ 语义特点	自 立	使 动	感 知	变 化
施事	+	+	+	—
感事	+	—	+	—
主事	+	—	—	+

老王 包饺子
老王 认识
老王 长疖子了

□ 施事、感事、主事句法上的分布

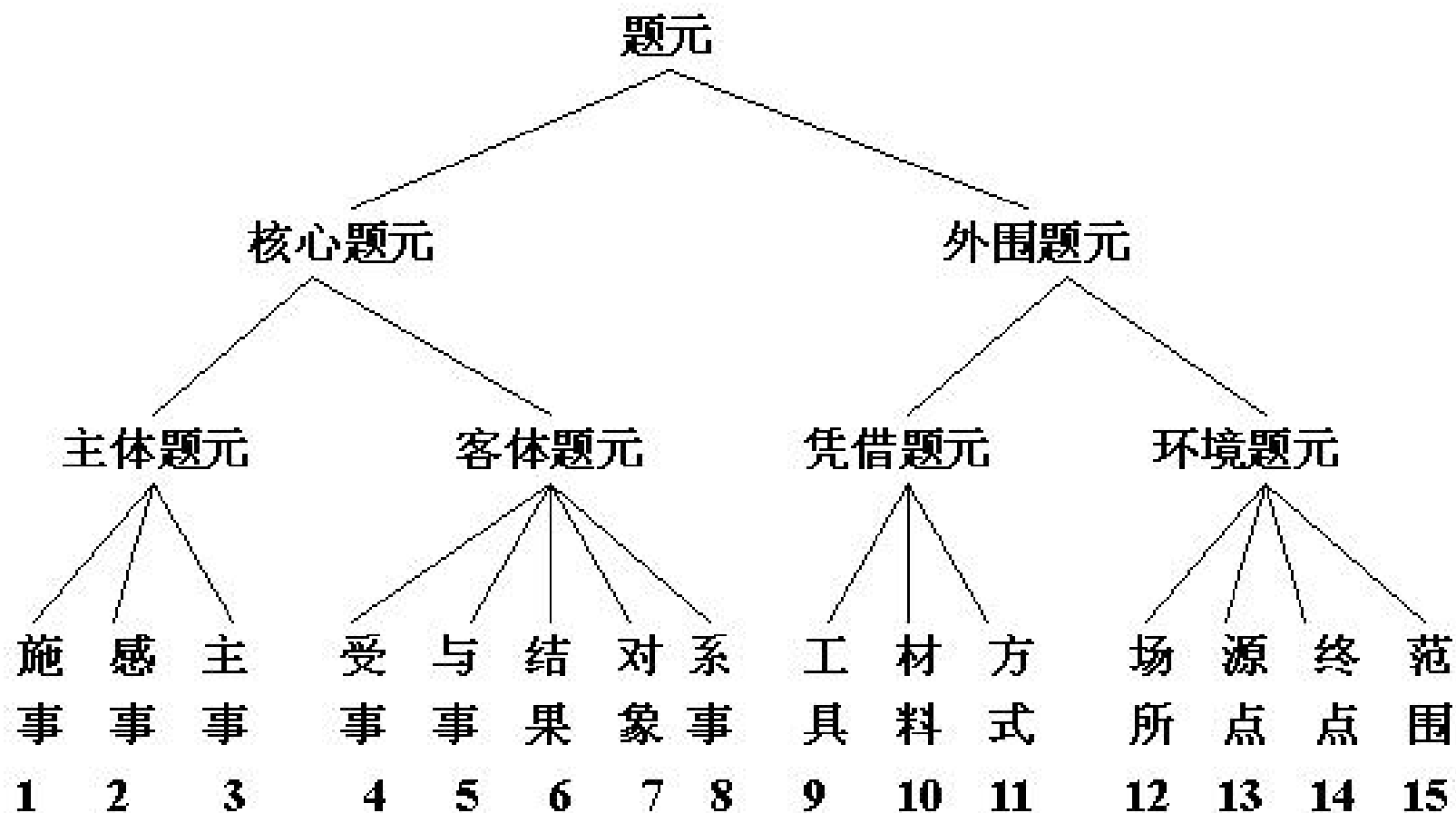
题元 \ 句法特点	作基础句的主语	出现在“不 VP”之前	出现在“没有 VP”之前
施事	+	+	+
感事	+	+	—
主事	+	—	+

施事的句法定义示例

- 1) 作基础句的主语。例如：“弟弟哭了” “小张吃了一碗汤面”；
- 2) 作“把”字句的主语。例如：“哥哥把校长给得罪了”；
- 3) 作“被、由”等介词的宾语。例如：“后勤工作由老刘负责”；
- 4) 作“V—V、V了V”等重叠形式的主语。例如：“你试（一）试”；
- 5) 作“不VP”和“没有VP”等否定形式的主语。例如：“小沈不去”；
- 6) 不在宾语位置上出现，除非句首是处所性成分。
例如：“*笑了小明”；

袁毓林（2003）《一套汉语动词论元角色的语法指标》，《世界汉语教学》2003年第3期。

汉语动词的论旨角色层级系统

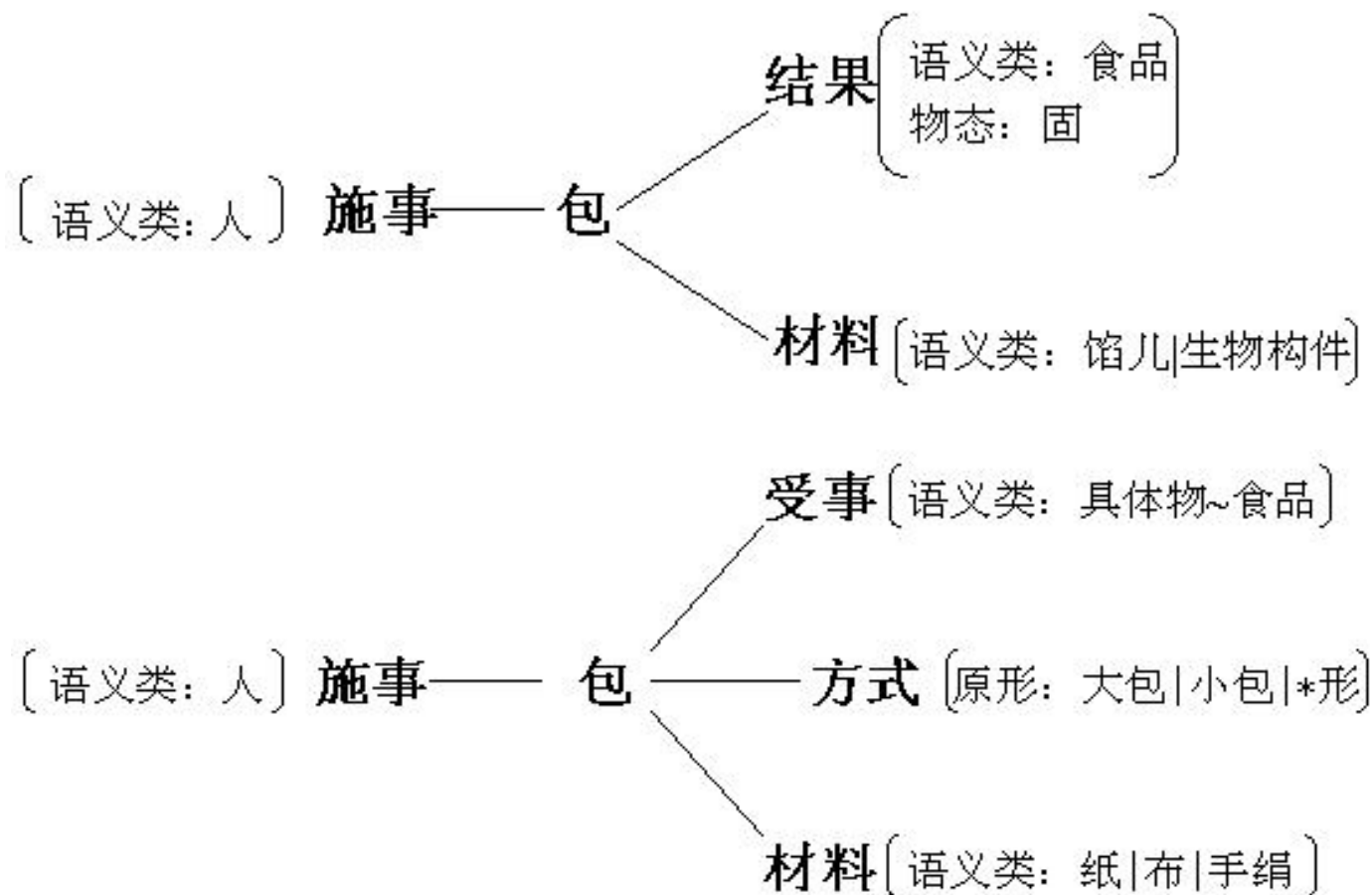


对论旨角色的选择限制(selectional restriction)

□ 动词的各个论旨角色在语义上满足什么样的约束条件

吃	施事论元	[+动物]	张三吃西瓜	
	受事论元	[+固体][+食物][+药物]	* 张三吃思想	
			* 空气吃鱼	
想	对象论元	[+人]	想妈妈	想1 = 思念
	对象论元	[+抽象物]	想答案	想2 = 思索

选择限制（语义组配）



描述论旨角色选择限制的方式

□ 句法属性层次

- 商量: {当事: [...], 复数: Yes, ...}

□ 语义属性层次

- 吃: {施事: [语义类: 人|动物], 受事: [语义类: 食物|药物]}

□ 词例层次

- 嗑: {施事: [语义类: 人], 受事: [词语: *瓜子]}

□ 正面描述与反面描述

- 变卖: {施事: [语义类: 人|集体], 受事: [语义类: ~人]}

小结

配价描述的内容	取值（具体的描述方式）
配价数	<ol style="list-style-type: none">1. 动词跟 x 类名词有语义联系；2. 动词周围有 x 个空位放置名词；（x 一般取值 0-3 ）
论旨角色/题元	<ol style="list-style-type: none">1. 跟动词有语义关联的 x 类名词分属哪些类型（T_1, \dots, T_i）2. T_i 能够出现在动词周围的哪些空位上
动词对其论旨角色的选择限制	<p>充当T_i的名词要满足哪些条件？</p> <ol style="list-style-type: none">1. 语法（形式）特征 语义属性（类别）特征2. 包容性条件 排除性条件

语义知识的应用举例：帮助确定多义词的义项

例1 爷爷这两天给孙女**想**名字，没有什么头绪，**想**放弃。这时候，爷爷就很**想**奶奶。如果奶奶在就好了。

想1：思考

想2：打算

想3：思念

例2 吃**饭**比吃**面**耐饱

饭1：米饭

饭2：食物（饭+菜）

◆ 语义知识可以附加在句法规则上发挥作用

3 在句法分析中附加语义约束

CFG规则：基于原子（简单）范畴的知识描述体系

$S \rightarrow NP VP$

- | | | |
|---------------|--------------------|---|
| 1. those dogs | bark in the yard. | ✓ |
| 2. a dog | bark in the yard. | ✗ |
| 3. those dogs | barks in the yard. | ✗ |
| 4. a dog | barks in the yard. | ✓ |
| | | |
| 5. 我们 | 找到了 | |
| 6. 答案 | 找到了 | |
| 7. 小偷 | 找到了 | |

语言成分（NP，VP）组合时有各种约束限制。CFG规则不方便描述这些约束限制，因此，需要引入更为灵活的知识表达机制。

特征结构（feature structure）是对CFG规则中简单范畴的扩充。

特征结构是有限个“特征-值”（attribute : value）对儿的集合。

$$\left(\begin{array}{l} \text{attribute}_1 = \text{value}_1 \\ \text{attribute}_2 = \text{value}_2 \\ \dots \\ \dots \\ \text{attribute}_n = \text{value}_n \end{array} \right)$$

空特征结构：不含任何特征的特征结构，记作 []

特征结构允许“嵌套”：一个特征结构是另一个特征的值

□ “特征值”有两种情况：

- (1) 特征值是一个字符串值或数值等简单类型。一般把这种特征值称为原子 (atom)。
- (2) 特征值是另一个特征结构，这就是所谓的特征结构的“嵌套”。

[词语:听听
词性:动词
重叠:是
音节:2]

a. 简单特征结构

[主语: [词语:董永
词性:名词
数:单数]
谓语: [述语: [词语:喜欢
词性:动词]
宾语: [词语:七仙女
词性:名词
数:单数]]]

b. 复杂特征结构（嵌套）

特征结构允许“共享”：两个特征共享一个值

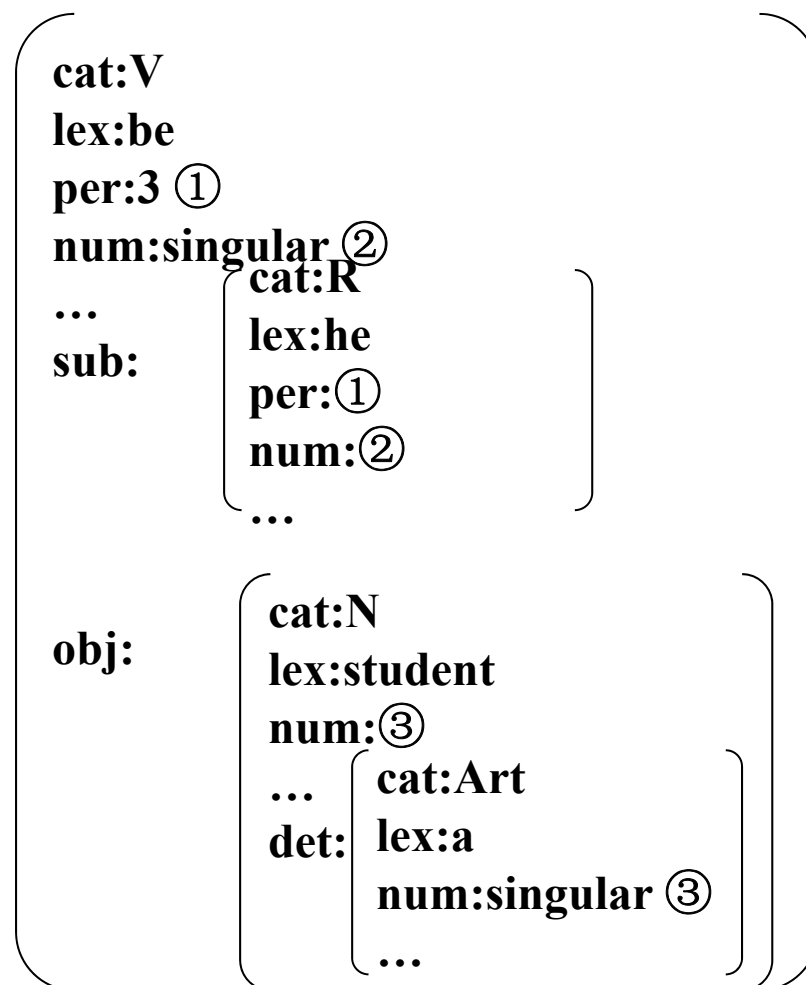
- 一个特征结构中的不同特征可以共享同样的特征值。这种特征结构也称为“重入式特征结构”（reentrant structure）。
- 在重入式特征结构中，共享特征值用数字索引标示。

例： He is a student.

①：动词的“per”特征取值为3.
主语（sub）的“per”特征跟动词的“per”特征共享值。

②：动词的“num”特征取值为单数. 主语（sub）的“num”特征跟动词的“num”特征共享值。

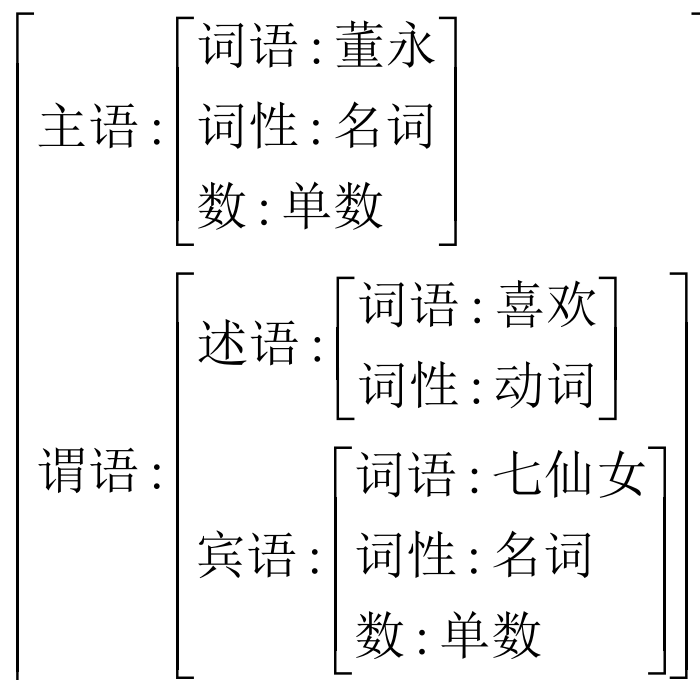
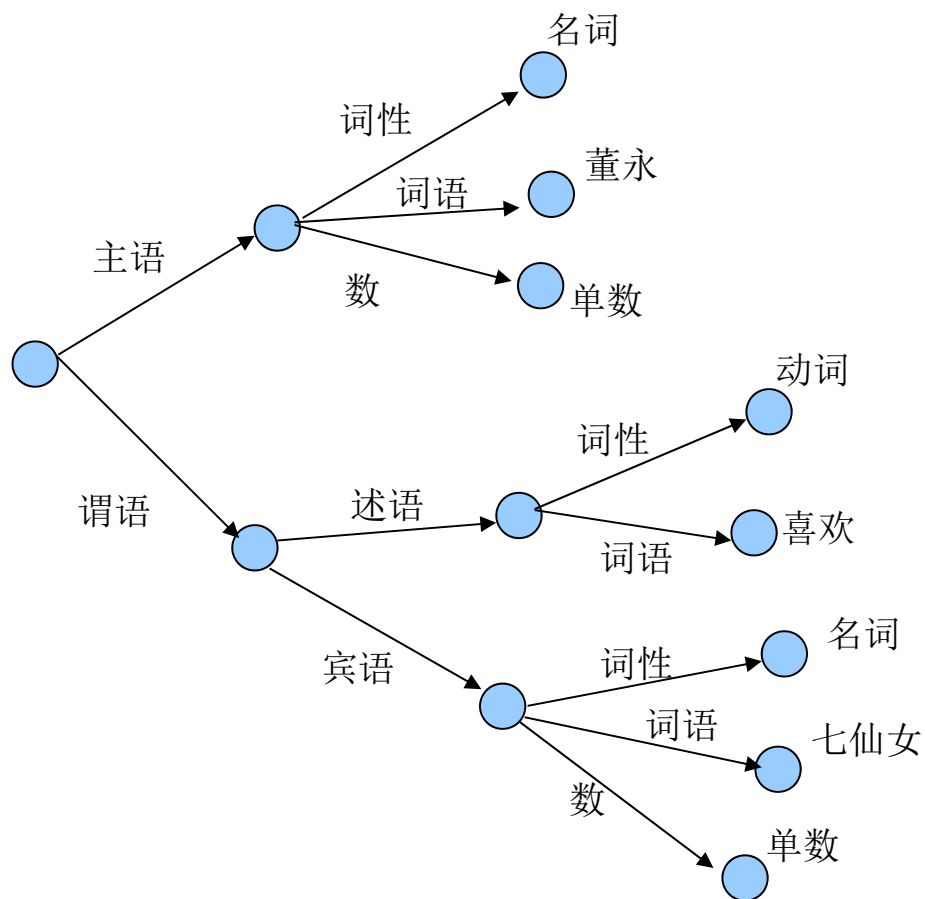
③：冠词的“num”特征取值为单数. 其修饰的名词的“num”特征跟冠词的“num”特征共享值。



特征结构的图表示法

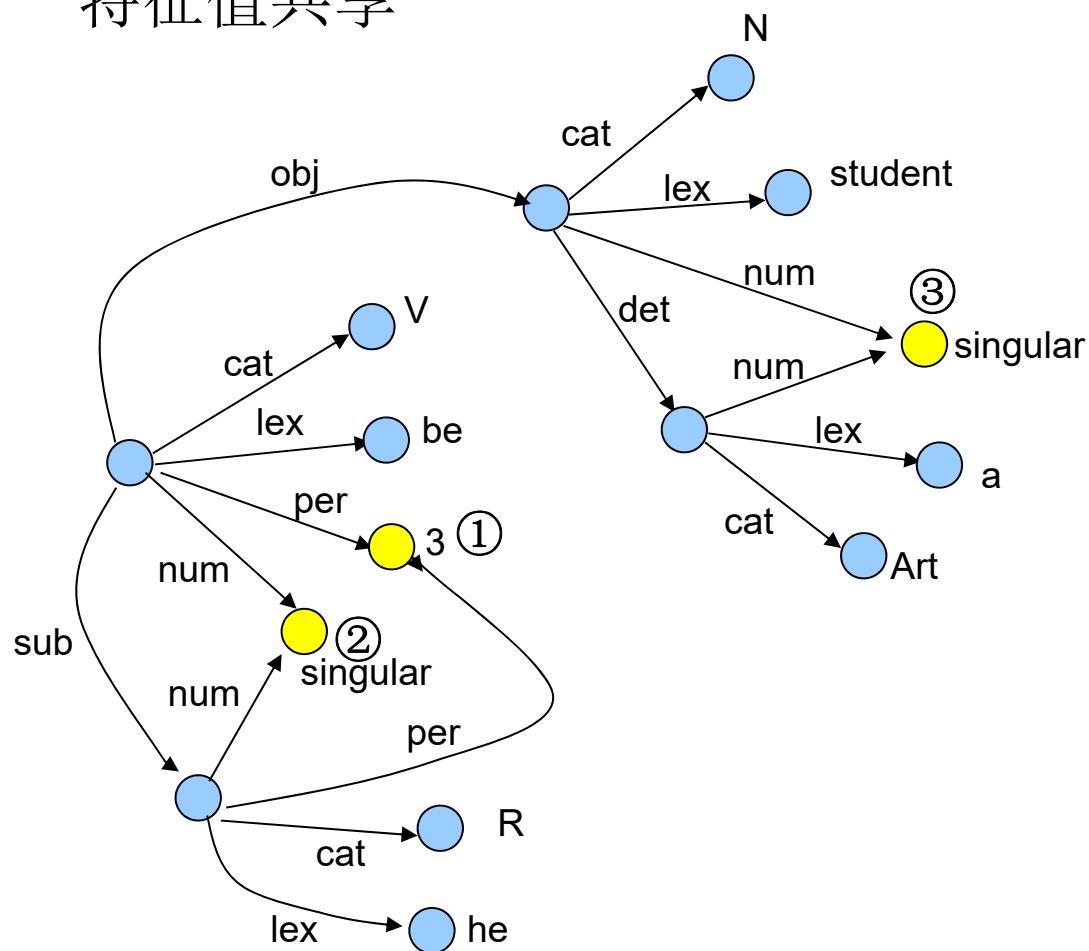
特征结构可以表示为有向无环图（**Directed Acyclic Graph**）
边（**edge**）表示特征；节点（**node**）表示特征值

例：董永 喜欢 七仙女

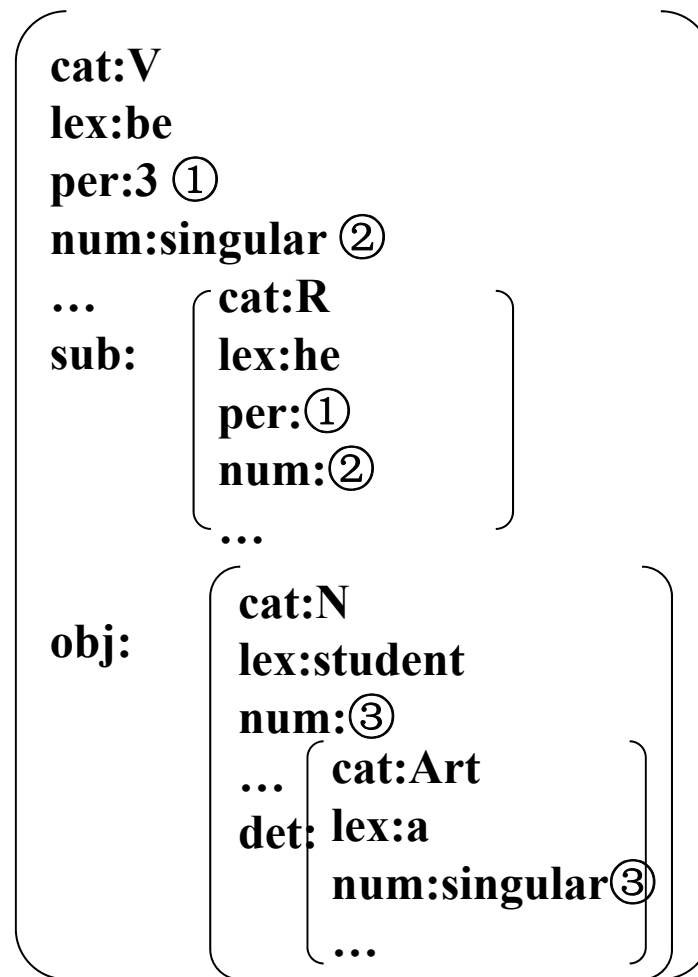


特征结构的图表示法（续）

特征值共享



例： He is a student.



特征结构间的包孕关系 (subsumption)

- 特征结构F1包孕F2，记作 $F1 \supseteq F2$ ，当且仅当
 - (1) 若特征 $f \in F1$ ，则 $f \in F2$ ，并且
 - (2) 若f的值是特征结构，则 $value_{F1}(f) \supseteq value_{F2}(f)$
 - (3) 若f的值是简单原子，则 $value_{F1}(f) = value_{F2}(f)$
- 空特征结构包孕任何特征结构

$$[Number \quad SG] \supseteq \begin{bmatrix} Number & SG \\ PERSON & 3 \end{bmatrix}$$

$$[Number \quad SG] \not\supseteq [Number \quad PL]$$

特征结构的合一运算 (Unification)

- 合一运算：将两个独立的特征结构F1，F2组合为一个新的特征结构F3，满足条件：

$$F1 \subseteq F3 \qquad F2 \subseteq F3$$

- 如果F1和F2是相容的（compatibility），则F1跟F2合一成功，合一的结果F3中包含了F1，F2所包含的信息。
- 如果F1和F2是不相容的，则F1跟F2合一失败。
- 合一的含义是：对两个特征结构进行类似于集合求并的一种运算，从而可以在“小”的特征结构基础上形成“大”的特征结构，这种运算非常适于刻画“小”的语言单位发展为“大”的语言单位的过程中的信息增加。
- 合一运算可以检查参与合一的两个特征结构是否相容。

合一运算示例

$$E = \left[\begin{array}{l} \text{Agree: } \left[\text{Number: Singular} \right] \textcircled{1} \\ \text{Subject: } \left[\text{Agree: } \textcircled{1} \right] \end{array} \right]$$

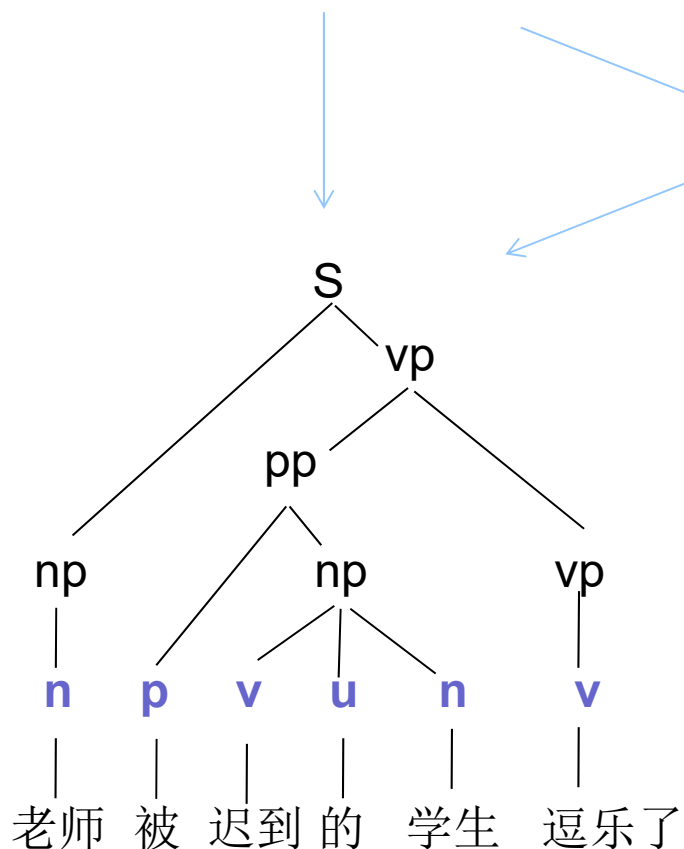
$$F = \left[\begin{array}{l} \text{Subject: } \left[\text{Agree: } \left[\text{Person: 3} \right] \right] \end{array} \right]$$

$$E \cup F = \left[\begin{array}{l} \text{Agree: } \left[\begin{array}{l} \text{Number: Singular} \\ \text{Person: 3} \end{array} \right] \textcircled{1} \\ \text{Subject: } \left[\text{Agree: } \textcircled{1} \right] \end{array} \right]$$

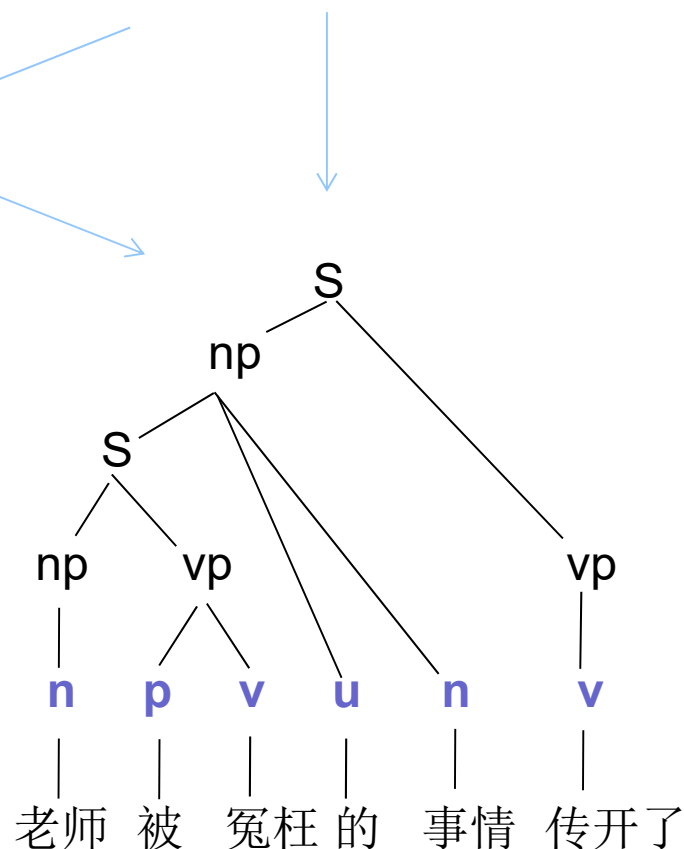
合一成功

在句法分析中附加语义约束

1) a 老师被迟到的学生逗乐了

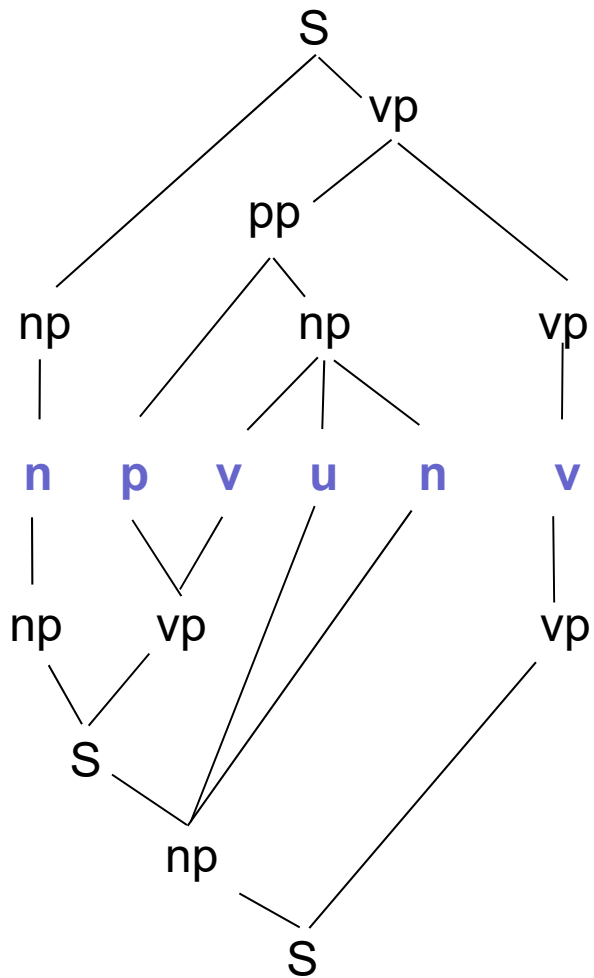


b 老师被冤枉的事情传开了



基于CFG规则的句法分析

a 老师被迟到的学生逗乐了



b 老师被冤枉的事情传开了

1. $S \rightarrow np\ vp$
2. $vp \rightarrow pp\ vp$
3. $pp \rightarrow p\ np$
4. $np \rightarrow v\ u\ n$
5. $np \rightarrow n$
6. $vp \rightarrow v$

- n → 老师 | 学生
v → 迟到 | 逗乐
p → 被
u → 的

7. $vp \rightarrow p\ v$
8. $np \rightarrow S\ u\ n$

- $v \rightarrow$ 冤枉 | 传开

引入“特征结构”对语言对象进行细致区分

谓词: $\left[\begin{array}{l} \text{词语: 迟到} \\ \text{词性: 动词} \\ \text{配价数: 1} \end{array} \right]$
题元: $\left[\begin{array}{l} \text{施事: } \left[\begin{array}{l} \text{语义类: 人} \\ \text{词性: 名词} \end{array} \right] \end{array} \right]$

谓词: $\left[\begin{array}{l} \text{词语: 冤枉} \\ \text{词性: 动词} \\ \text{配价数: 2} \end{array} \right]$
题元: $\left[\begin{array}{l} \text{施事: } \left[\begin{array}{l} \text{语义类: 人} \\ \text{词性: 名词} \end{array} \right] \\ \text{受事: } \left[\begin{array}{l} \text{语义类: 人} \\ \text{词性: 名词} \end{array} \right] \end{array} \right]$

“冤枉”的特征结构与“逗乐”相同

谓词: $\left[\begin{array}{l} \text{词语: 逗乐} \\ \text{词性: 动词} \\ \text{配价数: 2} \end{array} \right]$
题元: $\left[\begin{array}{l} \text{施事: } \left[\begin{array}{l} \text{语义类: 人} \\ \text{词性: 名词} \end{array} \right] \\ \text{受事: } \left[\begin{array}{l} \text{语义类: 人} \\ \text{词性: 名词} \end{array} \right] \end{array} \right]$

谓词: $\left[\begin{array}{l} \text{词语: 传开} \\ \text{词性: 动词} \\ \text{配价数: 2} \end{array} \right]$
题元: $\left[\begin{array}{l} \text{施事: } \left[\begin{array}{l} \text{语义类: 人} \\ \text{词性: 名词} \end{array} \right] \\ \text{受事: } \left[\begin{array}{l} \text{语义类: 事} \\ \text{词性: 名词} \end{array} \right] \end{array} \right]$

引入“合一约束”对CFG规则进行细致区分

a 老师被迟到的学生逗乐了

b 老师被冤枉的事情传开了

1. $S \rightarrow np\ vp$:: \$.语态=vp.语态, IF vp.语态=被动 THEN vp.受事=np^[1]
2. $vp \rightarrow pp\ vp$:: IF pp.lex=被 THEN vp.配价数>1, vp.施事=pp.宾语, \$.语态=被动^[2]
3. $pp \rightarrow p\ np$
4. $np \rightarrow v\ u\ n$:: v.施事=n^[3]
5. $np \rightarrow n$
6. $vp \rightarrow v$
7. $vp \rightarrow p\ v$:: IF p.lex=被 THEN v.配价数>1, \$.语态=被动^[4]
8. $np \rightarrow S\ u\ n$

上面合一约束式中n、v、np、vp等代表规则中的节点，\$代表根节点。

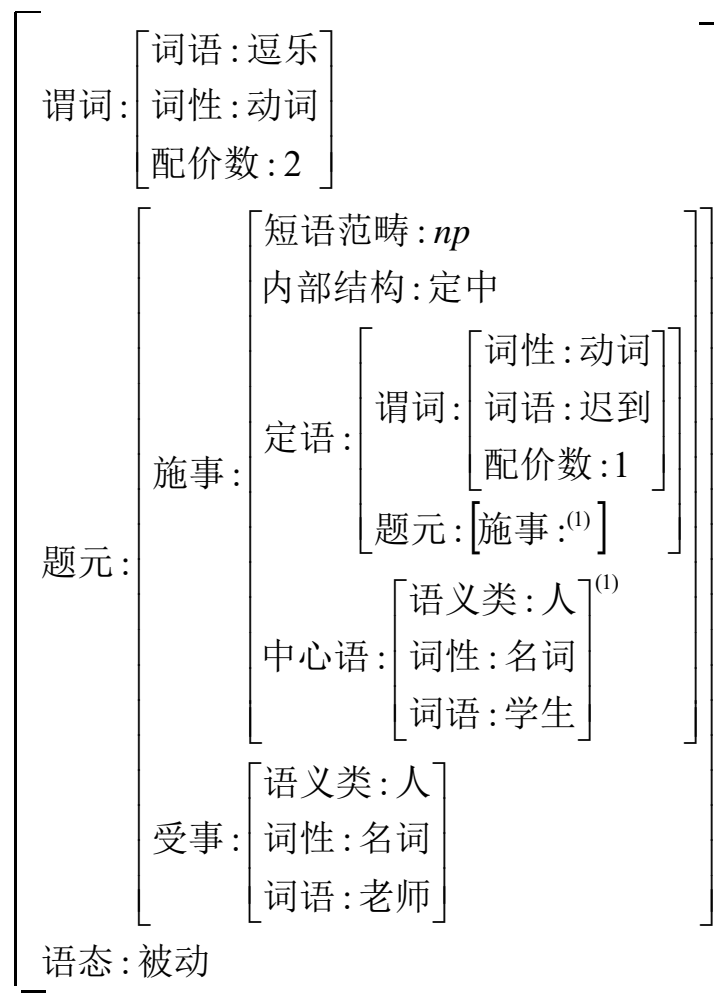
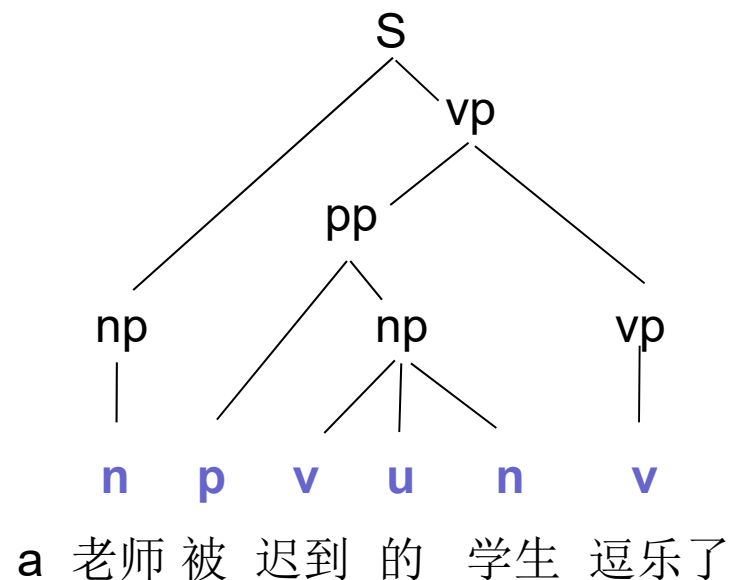
[1] a句应用规则1后，可以得到：“逗乐”的受事为“老师”

[2] a句应用规则2后，可以得到：“逗乐”的施事为“迟到的学生”

[3] a句应用规则3后，可以得到：“迟到”的施事为“学生”

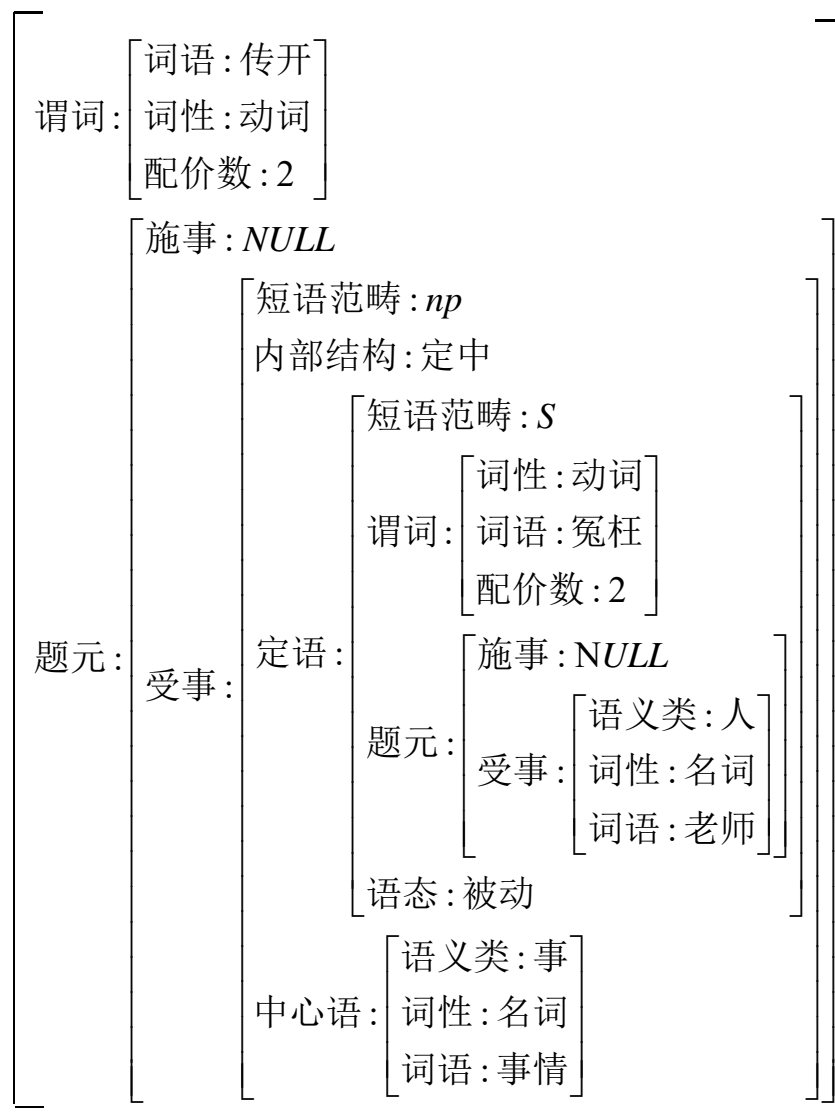
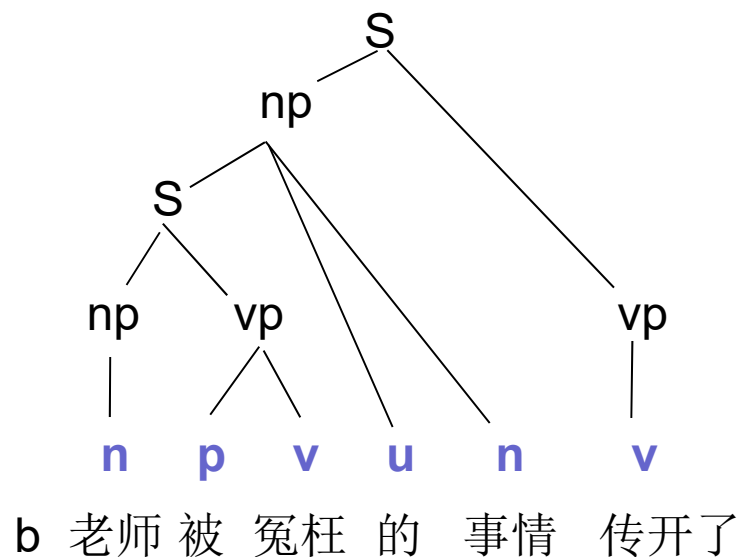
[4] b句应用规则4、1后，可以得到：“冤枉”的受事为“老师”

基于CFG规则及特征合一约束的句法分析



注释(1) “迟到”的施事跟“逗乐”的“施事”（中心语）共享一个值

基于CFG规则及特征合一约束的句法分析



4 语义知识描述的复杂性

- (1) 命题义（超常论元结构）
- (2) 模态（情态）义
- (3) 羡余否定
- (4) 语义指向
- (5) 名词指称义
- (6) 代词指代义
- (7) 焦点义
- (8) 构式义
- (9) 隐喻义
- (10) 句际关系义

语义知识描述的复杂性（续）

（1）命题义（超常论元结构）

- 1 过去他是铁饭碗，干一天是十块钱，不干也是十块钱 ■ 成分省略
- 2 他是猪肉馅儿，我是牛肉馅儿
- 3 张三 不可能 吃 思想
- 4 张三 怎么 可能 吃 思想 ■ 否定语境
- 5 张三 跟李四 比赛 吃 玻璃 ■ 超常现象
- 6 布尔毕达哥拉斯三元数问题：从1到7824的
任何整数能够被染成红色或蓝色，使得满
足 $a^2 + b^2 = c^2$ 的三个整数a、b、c不
全是同一种颜色；而对于从1到7825的整数，
不存在满足条件的染色。 ■ 假想世界

语义知识描述的复杂性（续）

（2）模态（情态）义

- | | |
|--------------------|-------|
| 1. 小张连打了三个喷嚏，感冒了吧。 | 猜测、存疑 |
| 2. 他搞不好已经毕业了。 | 猜测、相信 |
| 3. 看样子他很有经验。 | |
| 4. 他保证已经毕业了。 | 认定 |
| 5. 他能熟练使用两种语言。 | 能力 |
| 6. 这次会议能使用两种语言。 | 许可 |
| 7. 我能骗你吗？ | 可能 |

语义知识描述的复杂性（续）

（3）羡余否定

1. 在问题没有解决之前 = 在问题解决之前
2. 在考试没有结束之前 = 在考试结束之前
3. 年轻人难免不犯错误 = 年轻人难免犯错误
4. 当时差点儿没吓死 = 当时差点儿吓死
5. 天冷，小心别着凉 = 天冷，小心着凉
6. 他们宿舍好不热闹 = 他们宿舍好热闹

语义知识描述的复杂性（续）

（4）语义指向

1. 从小爷爷就教育我要诚实。
2. 当时那孩子就稀里糊涂判给了男方。
3. 今年我和他先后去过广州和福州。
4. 老张看着很无奈

语义知识描述的复杂性（续）

（5）名词指称义

1. 这人升了大官了。
2. 这人都爱往高处走。
3. 在北大图书馆自习室，你找不到一个玩游戏的学生。
4. 在美国的大众新闻媒体机构中，人们找不到一个持社会主义观点的新闻工作者或政治评论家。
5. 区块链技术：“有一个人有仲裁权”是集中式系统，还是分布式系统？
6. 一个计划生育宣传者说到世界上的人口问题的严重性时说 “**A woman will have a baby every three minutes**”。
有听众就问：“Why don't we find her and stop her?”

语义知识描述的复杂性（续）

（6）代词指代义

1. 一个男人临死前对siri说，**tell my wife, I love her.**
于是，他妻子的手机收到了一条来自丈夫手机的短信：**I love her.**
2. 齐齐哈尔民警杨文峰在抓捕一起绑架案犯罪嫌疑人时，
不幸被 **对方** 持枪击中头部，身负重伤。

语义知识描述的复杂性（续）

（7）焦点义

- 重音在“一个”上 剩下的就是几个小伙子跟两个小孩
1. 出事故的车上只有一个老人
- 重音在“老人”上 没有别人了

- 重音在“这道题”上 连这道题你都不会做 → 这道题很容易，你不会做这道题
2. 这道题你都不会做
- 重音在“你”上 这道题连你都不会做 → 这道题很难，你不会做这道题
- 你什么题都不会做
- 谁都不会做这道题

语义知识描述的复杂性（续）

（8）构式义

1. 村委会终于想出了一个不是办法的办法。
2. 武松放着好好的公务员不当，偏要去混黑社会。
3. 特斯拉是天才中的天才。
4. 大学跟高中的差别可大了去了。

语义知识描述的复杂性（续）

（9）隐喻义

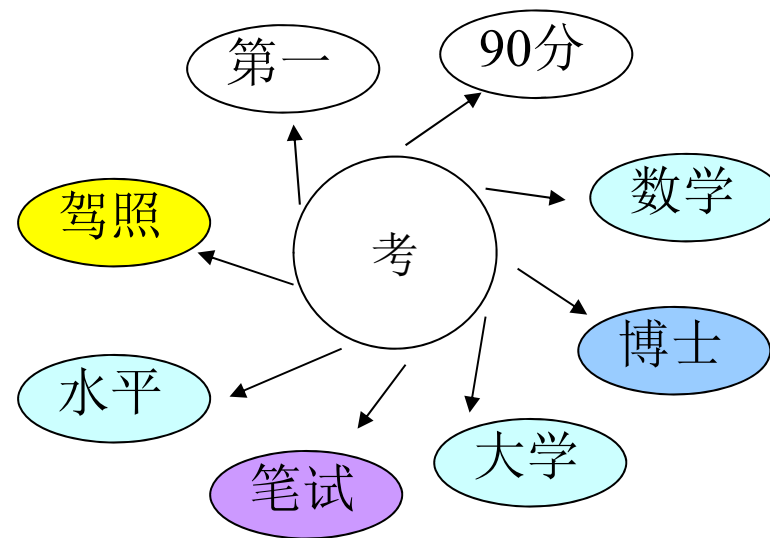
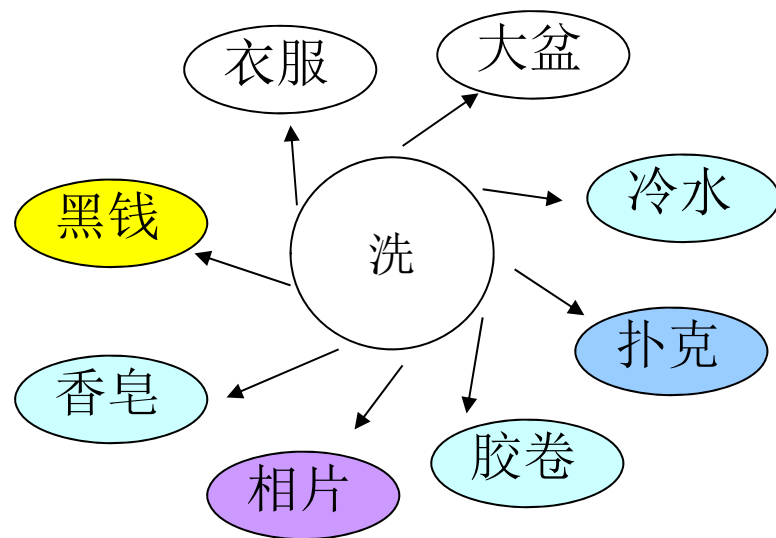
1. 牢骚帝的养生贴士：每天把牢骚拿出来晒晒太阳，你的心情就不会缺钙。
2. 只花了半天就 吃 掉 了 敌人 两个营

语义知识描述的复杂性（续）

（10）句际关系义

1. 谁造成的污染，谁负责治理。
2. 钱什么时候到账，我什么时候开工。
3. 我怎么说，你怎么写。
4. 她走到哪儿，哪儿就是花露水的香味。
5. 老张没有评上教授，新房子不会分给他。

复杂论元结构示例



论旨角色的复杂性示例

复杂论元结构示例

1. 我**飞**了你的**马**。
2. 一下午就**敲**了两百多**字**。
3. 今天老张干嘛去了？ — **蹬煤**去了。
4. 他这回总算**游**了一个**冠军**。
5. 他晚上又要来**催稿子**。
6. 他刚**拖**了**地板**，窗户还没有擦。

结构压缩与论元新增

原始的双事件结构				压缩后的句法结构	
V1	O1	V2	O2	V1	O2
飞	象	吃	马	飞	马
敲	键盘	输入	汉字	敲	字
蹬	三轮车	运送	煤块	蹬	煤
游	(泳)	获得	冠军	游	冠军
催	人	交	稿子	催	稿子
拖	拖把	擦洗	地板	拖	地板
参与	讨论	发表	意见	参与	意见
考	驾驶技术	获得	驾照	考	驾照
考	文化知识	进入	大学	考	大学
洗	胶片	显现出	照片	洗	照片

[+ 获得]

[+ 给予]

[+ 位移]

5 小结

1. 语义的表示形式：框架（特征结构）表示法
2. 词汇语义知识是语义分析的基础知识资源。
3. 特征结构提供了一个统一的知识表示框架，可以把句法约束知识和语义知识用同一套形式方式来加以表达。
4. 合一运算可以跟普通的句法分析算法结合起来，实现句法和语义分析的一体化。

进一步阅读文献

- 冯志伟等译（2005）《自然语言处理综论》第14，15，16章。
- 詹卫东，2001，《确立语义范畴的原则及语义范畴的相对性》，载《世界汉语教育》2001年第2期。
- 徐烈炯，1995，《语义学》，语文出版社
- Geoffrey Leech, 1983, *Semantics: The Study of Meaning*, Penguin Books.
- 蒋严，潘海华，1998，《形式语义学引论》，中国社会科学出版社
- 姚天顺 等，1995，《自然语言理解》，清华大学出版社，第四章，pp50-81，（语义网络）
- 张普，1991，《信息处理用现代汉语语义分析的理论与方法》，载《中文信息学报》1991年第3期。
- Pustejovsky, J. 1995. *The Generative Lexicon*. Cambridge, MA: The MIT Press
- Shalom Lappin, ed., 1996, *The Handbook of Contemporary Semantic Theory*, Oxford: Blackwell.
- VerbNet <http://verbs.colorado.edu/~mpalmer/projects/verbnet.html>
- FrameNet <http://framenet.icsi.berkeley.edu/>
- HowNet <http://www.keenage.com/>
- WordNet <http://wordnet.princeton.edu/>
- MindNet <http://research.microsoft.com/en-us/projects/mindnet/>

复习思考题

1. 请用配价理论分析下列动词的共性和差异。

见面 顶嘴 发火 撤职

2. 下面两个词组都是“n a 的 n”格式，但它们的句法结构不同，请给出基于语义知识的判断条件。

a. 文科优秀的教材

b. 颜色鲜艳的花朵

3. 请设计一个带特征结构合一约束的CFG规则系统，来分析下面两组四个例句的句法结构和语义，帮助计算机读懂这四句话的意思，使得计算机能够针对每个例句，正确回答：“明天谁要去当伴郎？”

1) a 乔峰答应明天去当伴郎

b 乔峰同意明天去当伴郎

2) a 乔峰答应虚竹明天去当伴郎

b 乔峰同意虚竹明天去当伴郎