

“第二届汉语词汇语义学研讨会”，北京大学计算语言学研究所，2001.5.14-18

面向组配约束的现代汉语 语义词典的构建

北京大学 詹卫东

2001年5月14日

本报告介绍国家重点基础发展规划项目（“973”项目）——“图像、语音、自然语言理解与知识发掘”之子课题——“面向中文信息处理的现代汉语动词论旨结构系统和汉语词语语义分类层级系统研究”的总体规划情况。分为三部分：

- 1 背景；
- 2 框架；
- 3 实施；

背景

- 应用驱动：

消解句法分析歧义；多义词义项判定；
相似度计算；

- 定位：

在句一级水平，以描述词语间组合约束为主，
强调语义与句法的对应性

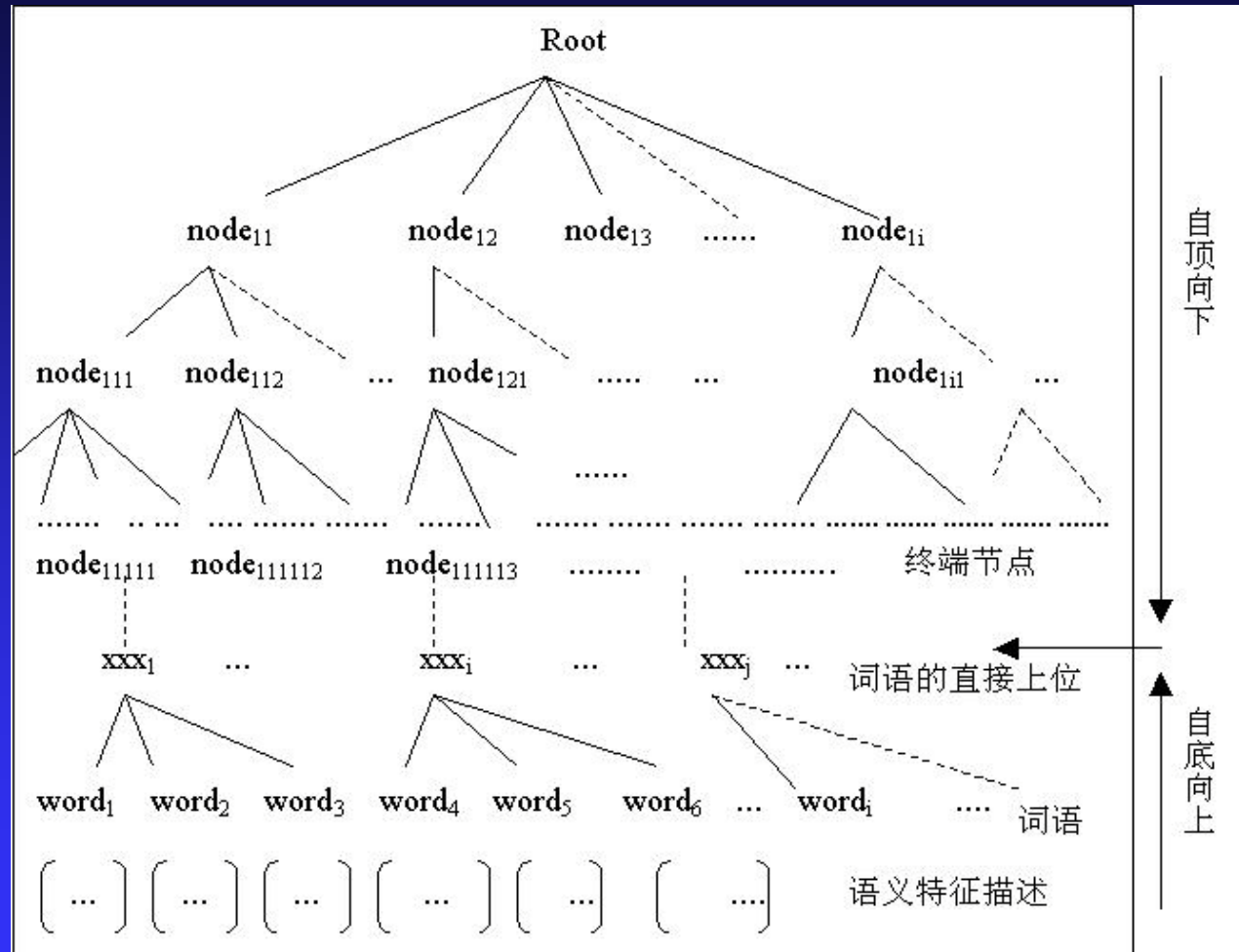
- 原则：

够用 — 能用 — 好用

框架

- 自顶向下与自低向上相结合来构造汉语实词的语义分类树
 - 采用语义类与语义特征相结合的方式来描述名词的语义信息
 - 以谓词为中心的实词组配约束信息描述，包括：
 - (I) 动词的论元数信息、论旨数信息
 - (II) 动词的论旨角色信息
 - (III) 动词对其论旨角色的选择限制信息
 - (IV) 动词的论旨角色在句法结构上的投射信息
(论旨角色的句法配位)
- (以上有关动词的组配约束描述，部分也适用于形容词)

语义分类树



N:58 - 16; A:33 - 3; V:14 - 1 (终端节点 - 非终端节点)

语义分类

■ 从句法类到语义类：一次细化

修理汽车的工人 - v + n + 的 + n

寻找烈士的遗物 - v + n + 的 + n

[[修理汽车]的工人] - v + 交通工具 + 的 + 人

[寻找[烈士的遗物]] - v + 人 + 的 + 具体物

语义特征

■ 从语义类 到 语义特征：二次细化

关心烈士的家属 $\left\{ \begin{array}{l} \text{[关心 [烈士的家属]] ?} \\ \text{[[关心烈士]的家属] ?} \end{array} \right.$

[V + [人 + 的 + 人 {关系: Yes}]]

部门主任 - ? 主任部门

警察叔叔 - ? 叔叔警察

论元与论旨

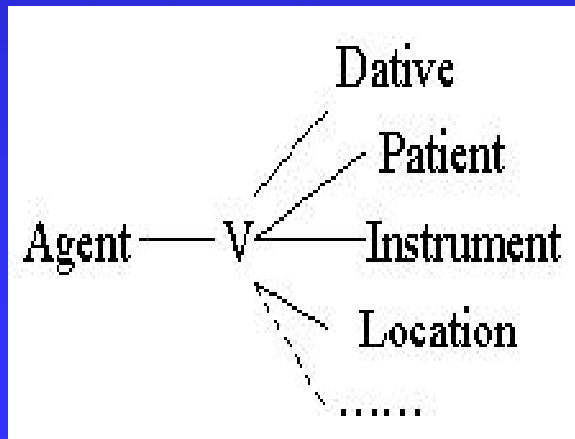
■ 论元

反映动词所能投射的**基本句法结构**的性质

—— V ; —— V —— ; —— V —— 。 ……
一元 ; 二元 ; 三元 。 X元

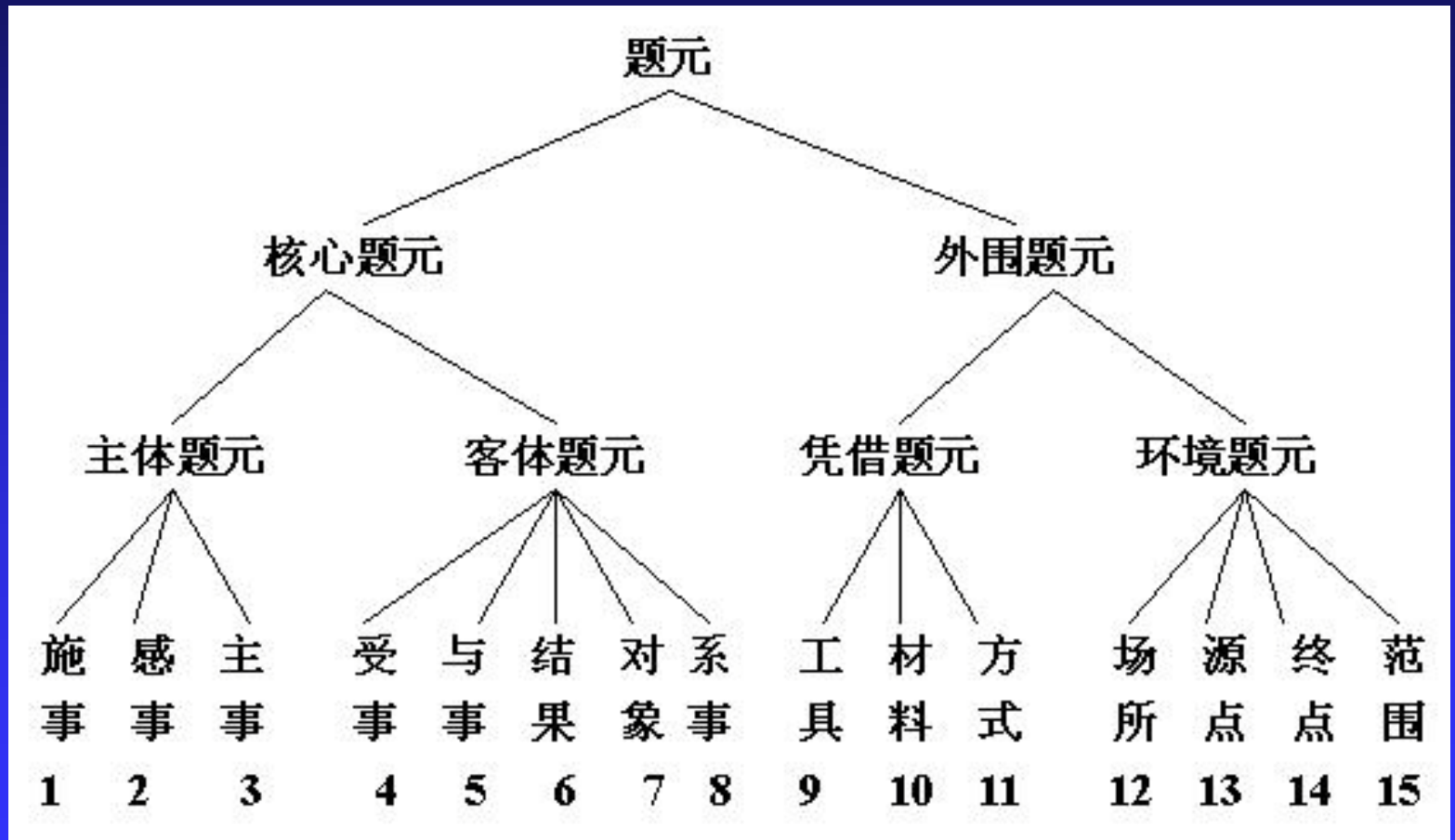
■ 论旨

反映动词所能结合的**名词语义角色**的性质



不 多
多 乎
也 哉

汉语动词的论旨系统



论旨角色示例

- 原型施事（proto-typical agent）典型性渐减的三个小类

语义特点 \ 题元	自立	使动	感知	变化
施事	+	+	+	-
感事	+	-	+	-
主事	+	-	-	+

- 施事、感事、主事句法上的分布

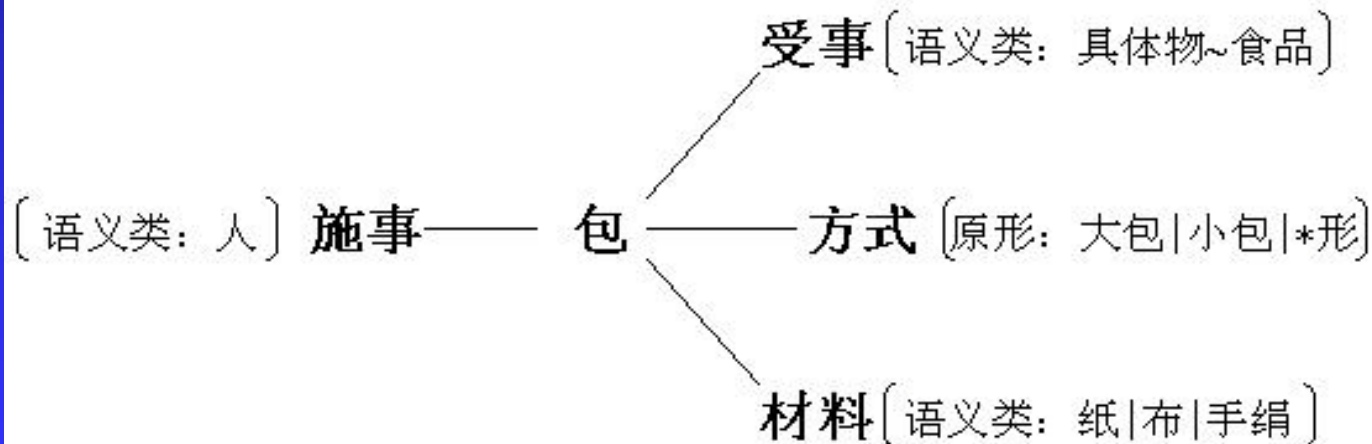
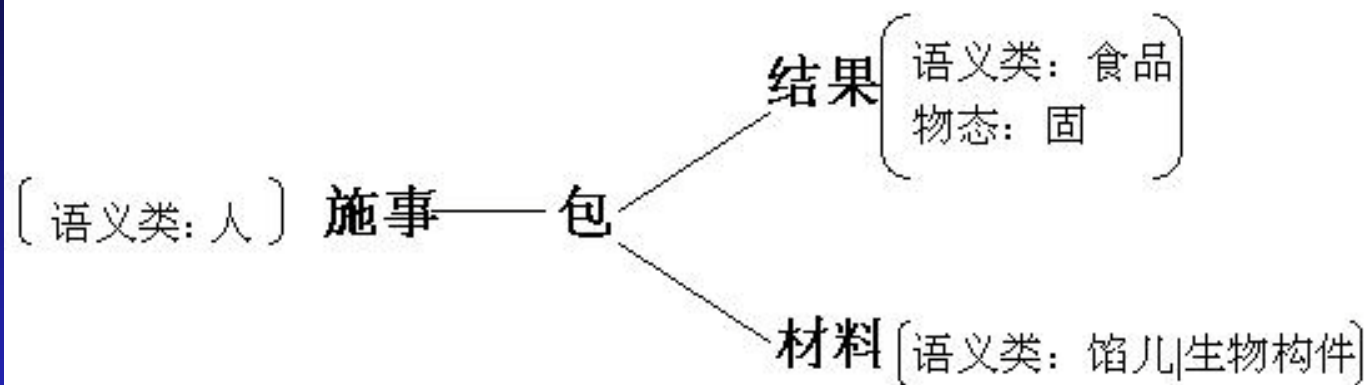
句法特点 \ 题元	作基础句的主语	出现在“不 VP”之前	出现在“没有 VP”之前
施事	+	+	+
感事	+	+	-
主事	+	-	+

施事的句法定义

- (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”；
- (7) 作主语时可以左向出位（left-dislocation）而成为话题，原来的位置上可以是句法空位（syntactic gap），留下的空位也可由续指代词（resumptive pronoun）“他”等来填充。

例如：“我妹妹，〔她〕笑个不停|小刚，〔他〕考上了清华大学”。

动词对其论旨角色的选择限制



把包子 **包** 好点儿 —— 哪个“包”？

动词论旨角色在句法结构上的投射

.....

S3: A + 用 + I + + P

例如：妈妈用小刀切西瓜

S4: I + A + + P

例如：这把刀我切肉

S5: Ma + A + + P

例如：那些白菜我都包饺子了

S6: A + 用 + Ma + + R

例如：姐姐用细毛线织了一件上衣

S7: A + 把 + Ma + + R

例如：妈妈把板栗全包了粽子

S8: A + + I

例如：我切这把大刀

S9: A + + Ma

例如：我包韭菜馅儿，你包肉馅儿

S10: A + + M

例如：我包小包，你包大包

.....

A-施事； I-工具； P-受事； Ma-材料； M-方式； R-结果；

实施

- 专家知识 + 语料库
- 词典填写的辅助软件工具

结 语

更多资料，请访问

http://ccl.pku.edu.cn/doubtfire/Semantics/973_Beida/index.htm

谢谢，请多指教