

一个现代汉语语义知识库的初步实现

陈群秀

清华大学计算机科学与技术系
智能技术与系统国家重点实验室，北京 100084

E-mail: cqx@s1000e.cs.tsinghua.edu.cn

- 摘 要：目前，世界各国学者都十分重视语言信息处理的知识资源的建设，知识包括词汇学知识、句法学知识、语义学知识、语用学知识乃至常识方面的知识，核心问题是语义学知识。在语义学知识中，词汇的语义知识是最基本最重要的语义知识。本文首先提出汉语词汇语义知识表示的一种解决方法：对动词、形容词从格框架角度用论旨网格描述其语义、语用知识，对名词从槽关系角度用槽关系联想表达式（槽类型、槽序）描述复杂定语与被修饰的中心词的语义关系，对名词还从语义场角度描述上下位语义关系。然后，本文着重介绍清华大学计算机系与中国人民大学、北京语言大学、北京青年政治学院共同承担的四个语言工程“现代汉语述语动词机器词典”、“现代汉语述语形容词机器词典”、“现代汉语名词槽关系系统”、“信息处理用现代汉语语义分类词典”的研究和实现情况。最后，论文还建立了现代汉语语义知识库平台。

1 前言

2 现代汉语述语动词机器词典的研究和实现

3 现代汉语述语形容词机器词典的研究和实现

4 现代汉语名词槽关系系统的研究和实现

5 信息处理用现代汉语语义分类词典的研究与实现

6 现代汉语语义知识库平台的实现

1 前言

当前，语言信息处理的竞争很大程度上取决于支撑的知识资源的竞争。

目前，世界上各国学者十分重视语言信息处理的知识资源的建设，知识包括词汇学知识、句法学知识、语义学知识、语用学知识乃至常识方面的知识，核心问题是语义学知识。

在对述语动词、述语形容词、名词词组作了大量的分析和研究的基础上，我们决定：对述语动词、述语形容词用论旨网格方法、从词法、句法、语义、语用多角度全面描述动词和形容词的信息，重点是语义信息；对名词的定语从语类、语序、定语与中心语的语义关系（槽类型、槽序）等多角度进行描述；同时对名词还从分类的角度描述名词的上下位关系。即：对汉语述语动词、述语形容词从格框架角度用论旨网格方法描述其语义、语用知识；对汉语名词的定语从槽关系角度以槽关系为主、以语类、语序为辅、用槽关系联想表达式（槽类型和槽序的表达式）描写复杂定语与被修饰的中心名词的语义关系；对名词还从语义场角度（即义类角度）描述上下位语义关系和语义特征。基于此，清华大学与中国人民大学、北京语言大学、北京青年政治学院共同承担了“现代汉语述语动词机器词典”、“现代汉语述语形容词机器词典”、“现代汉语名词槽关系系统”、“信息处理用现代汉语语义分类词典”四个语言工程的研究和实现的项目。本文下面将着重介绍这四个语言工程的研究和实现情况。



2 现代汉语述语动词机器词典的研究和实现

现代汉语述语动词机器词典（以下简称动词机器词典）的设计思想有以下几个：

1. 以原则参数语法作为理论指导，以论旨网络（ theta-gird ）方式对每个动词的组合关系从“论元属性（ argument ）”、“论旨属性（ thematic property ）”、“句法范畴（ category ）”（即论旨角色的语类）、“论旨角色的句法功能（ syntactic function ）”作详尽描写；
2. 从论旨角色语义约束的角度，建立汉语句词性概念的分类体系（ thesaurus ）以确定名词性概念的聚合关系和上下位关系；
3. 把主要以语言工作者的语感为依据的传统词典学编辑方法同主要以从机贮语料库中获取的大量例证为依据的计算词典学编辑方法结合起来，以使动词机器词典的研究和建立真正立足在丰富和客观的语言事实基础上。



现代汉语述语动词机器词典工作单描述的信息包括“词形”、“拼音”、“动词类型”、“论元数目”、“义项数目”、“义项序号”、“释义”等词汇信息，包括论旨模式的“基本式 1”及“句例”、“变换式 1”及“句例”、“基本式 2”及“句例”、“变换式 2”及“句例”、“基本式 3”及“句例”、“变换式 3”及“句例”、“论旨名称”、“语类”、“句法功能”、“语义分类”、“语义特征”、“论旨标记”、“论旨实例”等论旨属性（语义、句法、语用信息），也包括动词的“否定式”、“时态”、“语义指向动词的后状（可充当动补成分的动词和形容词）”、“抽象意义的趋向动词”、“论旨模式的扩展式”等其他句法和语义信息。即动词机器词典包含有关动词的词法、句法、语义、语用等丰富的信息。

动词的分类（“动词类型”）将汉语动词分为他动词、自动词、外动词、内动词、系属动词、领属动词六类；“论元数目”系指该动词的基本论旨模式的必要论元个数；“论旨名称”是指该动词该义项下基本式 1、基本式 2 和基本式 3 中必要论旨角色的称呼。论旨角色除必要论旨角色之外，还包括论旨模式扩展式中的可选论旨角色。论旨角色共定义了“施事”、“当事”、“领事”、“系事”、“受事”、“客事”、“分事”、“与事”、“同事”、“结果”、“基准”、“数量”、“范围”、“工具”、“材料”、“方式”、“依据”、“原因”、“目的”、“时间”、“处所”、“方向”等 22 个。



实例 1

[词形]：走 [拼音]：zou3 [制作者]：陈群秀 [工作单号]：6

[动词类型]：他动词 [论元数目]：2 [义项数目]：10** [义项序号]：10*

[释义]：* 比喻意义的“走”。

[基本式 1]：施事 + 走 + 客事 [基本式句例 1]：中国坚定不移地走改革开放的道路。教师队伍的建设走上规范化、法制化的轨道。

[基本式 2]：施事 + 走 + 方向 [基本式句例 2]：中国足球队走向世界。国产电梯正走向世界。三大民族史诗走上书架。帆船时代走向终结。

[论旨名称]：施事 [语类]：{ N | NP | VP } [句法功能]：主语

[语义分类]：{ 超类 | 事 | 物 | 时空 } [论旨实例]：宇宙、国民经济、工业、中国足球队、国产电梯、帆船时代、作品

[论旨名称]：客事 [语类]：{ N | NP } [句法功能]：宾语

[语义分类]：{ 空间 | 抽象物 } [论旨实例]：道路、轨道、正轨

[论旨名称]：方向 [语类]：{ N | NP | V | VP | S | A }

[句法功能]：状语

[语义分类]：{ 空间 | 抽象物 | 具体物 } [论旨标记]：向；上；朝

[论旨实例]：世界、书架、终结、光明、实施、法治、失败、稳定、卖国主义

[否定式]：不 V 没 V 没有 V 别 V V 不了

[时态]：要 1V 将要 V 即将 V 就要 V 快要 V V 着 V 着呢 正在 V 正在 V 呢 V 下去 V 下来 V 起来 2 V 下来 2 V 完 曾 V 过 曾经 V 过

[动词补充成分]：V 得动 [形容词补充成分]：急

[备注式 1]：扩展式：处所



实例 2

[词形]：有 [拼音]：you3 [制作者]：陈群秀 [工作单号]：2

[动词类型]：领属动词 [论元数目]：2 [义项数目]：9 [义项序号]：1

[释义]：表示领有、具有（跟“无”或“没”相对）。

[基本式 1]：领事 + 有 + 分事 [基本式句例 1]：这个模特儿有双大眼睛。手有五个手指头。

[基本式 2]：领事 + 有 + 客事 [基本式句例 2]：弟弟有辆小汽车。天津市领导对这个问题有个认识过程。

[变换式 2]：客事 + 领事 + 有 [变换式句例 2]：计算机学校已有了。

[论旨名称]：领事 [语类]：{ N | R | NP } [句法功能]：主语

[语义分类]：Y [论旨实例]：宇宙、万事万物、人、天津市领导、手

[论旨名称]：分事 [语类]：{ N | NP } [句法功能]：宾语

[语义分类]：{ 部件 } [论旨实例]：眼睛、心、手、花、手指头、果

[论旨名称]：客事 [语类]：{ N | NP } [句法功能]：宾语

[语义分类]：{ 事 | 物 | 时空 } [论旨实例]：汽车、书、信心、认识过程、时间、事

[否定式]：没 V

[时态]：将 V 要 1V 将要 V 即将 V 就要 V 快要 V V 了 1 已经 V 已经 V 了 V 过 曾 V 过 曾经 V 过

[备注式 1]：扩展式：同事、数量、范围、依据、原因、时间、处所



实例 3 :

[词形]: 打 [拼音]: da3 [制作者]: 王玲玲 [工作单号]: 20

[动词类型]: 他动词 [论元数目]: 2 [义项数目]: 24 [义项序号]: 1

[释义]: 用手或器具撞击物体。

[基本式 1]: 施事 + 打 + 受事

[基本式句例 1]: 父亲打了一些野果子。我们把一排排铆杆打入隧道。

[变换式 1]: 施事 + 受事 + 打 受事 + 施事 + 打

[变换式句例 1]: 父亲把树上的野果子打下来了。树上的野果子他都打下来了。

[论旨名称]: 施事 [语类]: { N | R | NP } [句法功能]: 主语

[语义分类]: { 人类 | 动物 } [论旨标记]: 被; 叫; 让

[论旨实例]: 猎户、庄户人、狗、狐狸

[论旨名称]: 受事 [语类]: { N | NP | R } [句法功能]: 宾语

[语义分类]: { 植物 | 人工物 } [论旨标记]: 把; 将

[论旨实例]: 松果、核桃、枣

[否定式]: 不 V 没 V 没有 V 别 V V 不了 V 不到

[时态]: 要 1V 就要 V V 上了 V 开了 V 着 V 着呢 正在 V 正在 V 呢 V 了 1 已经 V 了 V 完 V 过 曾 V 过 曾经 V 过 V 来着

[动词补充成分]: V 得了 [形容词补充成分]: 急 快 V 得好

[抽象意义的趋向动词]: 下来 下去 进去

[备注式 1]: 扩展式: 结果、基准



实例 4 :

[词形]: 打 [拼音]: da3 [制作者]: 王玲玲 [工作单号]: 27

[动词类型]: 他动词 [论元数目]: 2 [义项数目]: 24 [义项序号]: 9

[释义]: 编织。

[基本式 1]: 施事 + 打 + 结果 [基本式句例 1]: 姑娘在打毛衣。

[变换式 1]: 施事 + 结果 + 打 结果 + 施事 + 打

[变换式句例 1]: 姑娘把毛衣打好了。这件毛衣被姑娘打大了。

[论旨名称]: 施事 [语类]: { N | R | NP } [句法功能]: 主语

[语义分类]: { 人类 } [论旨标记]: 被; 叫; 让

[论旨实例]: 女工、工人、妹妹、她

[论旨名称]: 受事 [语类]: { N | NP } [句法功能]: 宾语

[语义分类]: { 人工物 } [语义特征]: [+ 织物]

[论旨标记]: 把; 将

[论旨实例]: 毛衣、袜子

[否定式]: 不 V 没 V 没有 V 别 V V 不了

[时态]: 要 1V 就要 V V 上了 V 开了 V 着 V 着呢 在 V 正在 V 正在 V 呢 V 了 1 已经 V 了 V 完 V 过 曾 V 过 曾经 V 过 V 来着

[动词补充成分]: 习惯 V 得了 [形容词补充成分]: 快 [抽象意义的趋向动词]: 过去 开 起 起来

[备注式 1]: 扩展式: 与事、基准



实例 5 :

[词形] : 打 [拼音] : da3 [制作者] : 王玲玲 [工作单号] : 34

[动词类型] : 他动词 [论元数目] : 2 [义项数目] : 24 [义项序号] : 17

[释义] : 买。

[基本式 1] : 施事 + 打 + 受事 [基本式句例 1] : 儿子打了一瓶酱油。
[变换式 1] : 施事 + 受事 + 打 受事 + 施事 + 打 [变换式句例 1] : 儿子把酱油打回来了。酱油他打回来了。

[论旨名称] : 施事 [语类] : { N | R | NP } [句法功能] : 主语

[语义分类] : { 人类 } [论旨实例] : 小孩、掏粪工人、他
[论旨名称] : 受事 [语类] : { N | NP } [句法功能] : 宾语

[语义分类] : { 人工物 } [语义特征] : 液体 [论旨标记] : 把 ; 将

[论旨实例] : 酒、油、漆
[否定式] : 不 V 没 V 没有 V 别 V V 不了 [时态] : 要 1V 就要 V V 上了 V 了 1 已经 V 了

[备注式 1] : 扩展式 : 与事、处所



实例 6 :

[词形]: 打 [拼音]: da3 [制作者]: 王玲玲 [工作单号]: 39

[动词类型]: 他动词 [论元数目]: 2 [义项数目]: 24 [义项序号]: 22*

[释义]: 做(体育活动、游戏等)。

[基本式 1]: 施事 + 打 + 受事 [基本式句例 1]: 两个留学生在打网球。打扑克。打枪。打弹子。

[变换式 1]: 施事 + 受事 + 打 受事 + 施事 + 打

[变换式句例 1]: 两个留学生网球打得不错。那个网球他们打飞了。

[论旨名称]: 施事 [语类]: { N | R | NP } [句法功能]: 主语

[语义分类]: { 人类 } [论旨标记]: 被; 叫; 让 [论旨实例]: 球队、队员、学生、工人、农民、小孩

[论旨名称]: 受事 [语类]: N | R | NP } [句法功能]: 宾语

[语义分类]: { 体育用具 | 军事用具 } [论旨标记]: 把; 将

[论旨实例]: 游戏牌、球、麻将、游戏机、枪

[否定式]: 不 V 没 V 没有 V 别 V V 不了

[时态]: 要 1V 就要 V V 起来 1V 上了 V 开了 V 着 V 着呢 在 V 正在 V 正在 V 呢 V 了 1 已经 V 已经 V 了 V 起来 2 V 完 V 过 曾 V 过 曾经 V 过 V 来着

[动词补充成分]: 完 习惯 [抽象意义的趋向动词]: 起 起来

[备注式 1]: 扩展式: 同事、数量



大例7:

[词形]: 考

[拼音]: kao3

[制作者]: 祁素芳 [工作单号]: 17

[动词类型]: 他动词
[号]: 1

[论元数目]: 2

[义项数目]: 3 [义项序

[释义]: 考试。

[基本式1]: 施事 + 考 + 受事 [基本式句例1]: 老师正在考学生。

[变换式1]: 受事 + 被 + 施事 + 考 施事 + 把 + 受事 + 考 [变换式句例1]: 父亲把儿子考住了。我们都被王老师考傻了。

[基本式2]: 受事 + 考 + 范围

[基本式句例2]: 初三正在考语文。

[变换式2]: 受事 + 把 + 范围 + 考

[变换式句例2]: 我把英语考得一团糟。

[基本式3]: 施事 + 考 + 目的

[基本式句例3]: 他在考律师。

[论旨名称]: 施事

[语类]: { N (普名) | NP | R } [句法功能]: 主语

[语义分类]: { 人 }

[论旨标记]: 被 [论旨实例]: 老师、演员、驾驶员、学校

[论旨名称]: 受事

[语类]: { N (普名) | NP | R } [句法功能]: 宾语

[语义分类]: { 人 }

[论旨标记]: 把 [论旨实例]: 学生、儿子、同事

[论旨名称]: 范围

[语类]: { N (普名) | NP } [句法功

能]: 宾语

[语义分类]: { 学科 }

[论旨标记]: 把 [论旨实例]: 数学、作文、地理

[论旨名称]: 目的

[语类]: { N (普名) | NP } [句法功能]: 宾语

[语义分类]: { 社会属性 }

[语义特征]: 学历职称

[论旨实例]: 律师、博

士、会计、演员

[否定式]: 不 V 没 V 没有 V 别 V V 不了 V 不到

[时态]: 将 V 要 1V 将要 V 即将 V 就要 V 快要 V V 上了 V 开了 V 着 V 着呢 在 V 正在 V 正在 V 呢 V 了 1 已经 V 已经 V 了 V 完 V 到 V 过 曾 V 过 曾经 V 过 V 来着

[动词补充成分]: 完 着 到 中 [形容词补充成分]: 错 对 多 够 好 坏 累 乱 少 糟 V 得好 V 得

[抽象意义的趋向动词]: 上 起

[备注式1]: 扩展式: 与事、同事、结果、基准、数量、材料、方式、依据、原因、时间、处所

实例 8 :

[词形]: 喝 [拼音]: he1 [动词类型]: 他动词 [论元数目]: 2
[义项数目]: 2 [义项序号]: 1 [释义]: 把液体咽下去。
[基本式 1]: 施事 + 喝 + 受事 [基本式句例 1]: 孩子喝完奶了。
[变换式 1]: 施事 + 受事 + 喝 受事 + 施事 + 喝
[变换式句例 1]: 孩子把奶喝光了。那瓶奶让他喝光了。那瓶奶他喝光了。
[论旨名称 1]: 施事 [语类 1]: {N (普名) | N (专名) | R}
[语义分类 1]: {人 | 动物} [论旨标记 1]: 让; 给; 被 [论旨实例 1]: 孩子、病人、运动员、狗
[论旨名称 2]: 受事 [论类 2]: {N (普名) | N (专名)} [句法功能 2]: 宾语
[语义分类 2]: {生物耗散物} [语义特征]: 液体 [论旨标记 2]: 把
[论旨实例 2]: 奶、酒、汤、饮料、水、药、敌敌畏 [否定式]: 不 V 没 V 没有 V 别 V
[时态]: 要 1V 就要 V 快要 V V 起来 1V 开了 V 着 V 着呢 在 V 正在 V 正在 V 呢 V 了 1 已经 V 了 V 完 V 过 曾 V 过 曾经 V 过 V 来着
[语义指向动词的后状—有可能充当动词补充成分的动词]: 掉 惯 习惯
[语义指向动词的后状—有可能充当动词补充成分的形容词]: 多 光
[抽象意义的趋向动词]: 下去
[备注]: 扩展式: 与事、同事、结果、基准、数量、工具、材料、依据、时间、处所



3 现代汉语述语形容词机器词典的研究和实现

述语形容词词典工作单的设计遵循下列几条原则：

- 工作单设计以述语形容词为轴心，以每个形容词为中心组织一张工作单的信息；
- 研究方法是“语料库方法 + 联想”。
- “形容词类型”的设计原则是将“从语义角度分类”、“可操作性”和“计算机处理用”几个原则综合考虑，将形容词分为感情形容词、感觉形容词、属性形容词和其他形容词。
- 欲研究和描写的术语形容词的义项选择原则为综合考虑形容词的分布性、形容词的常用频度和典型性。

在形容词的研究中包括词形、拼音、释义、论元数目、义项数目、义项序号等词性和论元属性，也包括形容词类型的语义分类属性。我们的研究把形容词分为感情形容词、感觉形容词、属性形容词、其他形容词四种语义类型。感情形容词是表示感情、心理活动和感情色彩比较强的形容词，例如高兴、愉快、可爱、痛苦、绝望、有趣、可怜、悲哀、伟大、可敬、可恨等词。感觉形容词是指身体、皮肉、骨头、五脏六腑、眼、鼻、嘴、舌、耳等的感官感觉和对周围环境的整体感觉（除去明显的属性之外），例如渴、甜、苦、香、痛、干燥、阴冷、麻木、明亮等词。属性形容词是指根据陈群秀拟定的《信息处理用现代汉语语义分类系统》中在“3.2.5”中表示物体和事物的各种属性的词。例如大、小、宽、窄、深、浅、干燥、薄、厚等物理属性和幼稚、成熟等生理属性。其他形容词是除感情、感觉、属性之外的形容词，例如凑巧、过分、平等、得当、少见、独到、得体等词。

此外，还具体描述了形容词其他一些相关的语义句法信息：同义 / 近义词、反义词、可否作定语（及句例）；可否作状语（及句例）；可否作补语（及句例）；重叠方式；否定方式；可以用以限制的程度副词；可以使用的时态等。



实例 1：

[词形]：高兴 [拼音]：gao1xing4 [制作者]：李千驹

[工作单号]：332 [形容词类型]：感情

[义项数目]：2 [义项序号]：1 [释义]：愉快而兴奋。

[同义/近义词]：愉快|喜悦|快活|快乐|兴奋

[反义词]：伤心|难过|悲哀|忧愁 [基本式 1]：当事 + 高兴

[基本式句例 1] 小明今天特高兴。

[论旨名称]：当事 [语类]：{N (普) |N (专) |NP|R}

[句法功能]：主语 [语义分类]：{人类}

[论旨实例]：全家；他；小明 [可否作定语]：可

[句例]：他家是遇到了什么高兴事儿。

[可否作状语]：可 [句例]：他高兴地期待着她的到来。

[可否作补语]：可 [句例]：我们玩得很高兴。

[重叠方式]：AABB [其他重叠]：[否定式]：不 A

[加程度副词]：很 A A 极了 最 A [时态]：A 起来了

[备注]：基准；时间；处所；范围；方式；原因



实例 2：

[词形]：马虎 [拼音]：ma3hu5 [制作者]：李千驹

[工作单号]：500 [形容词类型]：属性|感情

[义项数目]：1 [义项序号]：1 [释义]：草率；敷衍；疏忽大意；不细心。 [同义/近义词]：草率|潦草

[反义词]：认真|负责|踏实 [基本式 1]：当事 + 马虎

[基本式句例 1] 这人太马虎。 小王做事特马虎。

[论旨名称]：当事 [语类]：{N (普) |N (专) |NP|R|S}

[句法功能]：主语 [语义分类]：{动作行为|人类}

[论旨实例]：这人；小王做事 [可否作定语]：可

[句例]：做事不能抱马虎的态度。

[可否作状语]：可 [句例]：小明马马虎虎地做完作业，就出去玩耍去了。

[可否作补语]：可 [句例]：这个单位的工作搞得太马虎了。

[重叠方式]：AABB [其他重叠]：马里马虎 [否定式]：

不 A [加程度副词]：很 A A 极了 最 A A 透了 [备注]：基准；处所



实例 3 :

[词形] : 凉爽 [拼音] : liang2shuang3 [制作者] : 李千驹

[工作单号] : 120 [形容词类型] : 感觉

[义项数目] : 1 [义项序号] : 1 [释义] : 清凉爽快。

[同义/近义词] : 凉快 | 清凉 [反义词] : 闷热 | 炎热

[基本式 1] : 当事 + 凉爽 [基本式句例 1] 这几天真凉爽。 天气十分凉爽。 夏日凉爽。 厅内凉爽。

[论旨名称] : 当事 [语类] : {N (普) | NP | N (处所) | N (时间) } [句法功能] : 主语

[语义分类] : { 气象现象 | 有限空间 | 季节 }

[论旨实例] : 天气 ; 夏日 ; 厅内 [可否作定语] : 可

[句例] : “雪花”牌落地扇吹来阵阵凉爽的清风。

[可否作状语] : 否 [可否作补语] : 否 [否定式] : 不 A

[加程度副词] : 很 A A 极了 最 A [备注] : 方式 ; 基准



实例 4 :

[词形]: 绿 [拼音]: lv4 [制作者]: 李千驹 [工作单号]: 923 [形容词类型]: 属性

[义项数目]: 2 [义项序号]: 1 [释义]: 草和树叶茂盛时的颜色。 [同义/近义词]:

[反义词]: [基本式 1]: 当事 + 绿 [基本式句例 1] 对面山坡上的那片草全绿了。

[论旨名称]: 当事 [语类]: {N (普) |NP} [句法功能]: 主语 [语义分类]: {植物}

[论旨实例]: 草; 树叶 [可否作定语]: 否 [可否作状语]: 否 [可否作补语]: 否

[否定式]: 不 A [加程度副词]: 很 A A 极了 最 A A 透了

[备注]: 基准



实例 5 :

[词形] : 热 [拼音] : re4 [制作者] : 李千驹

[工作单号] : 1121

[语义分类] : 属性 | 感觉 [论元数目] : 1

[义项数目] : 9 [义项序号] : 2

[释义] : 热能多的 ; 温度高的或感觉温度高。 [反义词] : 冷

[基本式 1] : 当事 + 热 [基本式句例 1] 三伏天很热。

屋子里太热。 我很热。

[论旨名称] : 当事 [语类] : {N (普) | N (专) | NP | R}

[句法功能] : 主语

[语义分类] : {物 | 有限空间} [论旨实例] : 三伏天 ; 屋子里 ; 我

[做定语句例] : 热水 | 热气

[否定式] : 不 A

[加程度副词] : 很 A A 极了 最 A

[备注] : 基准



4 现代汉语名词槽关系系统的研究和实现

本文作者研究名词槽关系在继承语言学的研究成果基础上又有新的思路，既从定语出现的位置、定语的语类出发研究定语，又采用定义 70 个槽类型新颖方法，从研究定语与中心词的语义关系的新的角度来研究汉语定语与名词中心词的关系，是寻求研究汉语名词定语的新突破口。论文作者认为，研究名词槽关系，关系在于研究和揭示定语与中心词的语义关系。在对定语与中心词的语义关系研究基础上，还需对大量名词进行工程性描写。目的在于研制一个能提供名词前定语的语类、定语与中心词语义关系类型（槽类型）、定语与中心词语义关系类型的顺序（槽类型顺序、即槽序）的现代汉语名词槽关系系统。

21 现代汉语名词槽关系系统研究中名词槽关系系统的研究路线、槽类型的研究和设计遵循以下几个原则：

- （1）工作单设计以名词为轴心，即以每个名词为中心组织一张工作单的信息；
- （2）研究方法是“语料库方法 + 联想”，即把主要从机贮语料库中获取的大量例证为依据的计算词典学方法同主要以语言工作者的经验、语感为依据的传统词典学方法结合起来，以使名词的槽关系研究真正建立在丰富的客观的语言事实基础上，同时发挥语言工作者的经验、语感进行联想和归纳。
- （3）“语类”和“槽类型”的设计原则是将“必要性”、“够用原则”、“可操作性”和“计算机处理用”几个原则综合考虑；
- （4）欲研究和描写的名词的义项选择原则为综合考虑名词在《信息处理用现代汉语语义分类系统》中的分布性、名词的常用频度和典型性。



在以上原则基础上，名词槽关系系统工作单描述的项目有词形、拼音、语义分类（填写该名词在“信息处理用现代汉语语义分类体系”中的上位或上上位义类）、论元数目（系指该名词的配价、暂空）、义项数目（该词所具有的义项数目总数）、义项序号（该义项在该词义项数中的序号）、释义、词组实例 1、定语 1、语类（系指该词组实例中该定语的句法范畴）、槽类型（该定语与该词组实例的中心名词的语义关系类型，有 70 种类型）、定语 2、语类、槽类型、定语 3、…、定语 11、语类、槽类型、槽类型顺序、可能顺序、词组实例 2、定语 1、语类、槽类型、…、定语 11、语类、槽类型、槽类型顺序、可能顺序、词组实例 3、定语 1、语类、槽类型、…、定语 11、语类、槽类型、槽类型顺序、可能顺序、槽关系表达式联想（综合三个词组实例的可能顺序、发挥语言工作者的语感和经验，归纳出的该名词定语可能有的槽关系类型顺序表达式）。



槽类型（槽类）系指槽关系的语义类型，即词组实例 i 的第 j 个定语与中心名词之间的语义关系。在对若干名词的若干定语进行分析之后，本文作者设计槽类型系统体系结构的基本思想或依据是：定语对中心词的修饰作用无非是说明中心词的属性、状态或关系。因此槽类型的顶层（第 0 层）应该有属性、状态、关系，但从系统简洁性考虑，第 0 层就不具体列入设计。属性包括有事域、事情、领属、来源、去向、数量、顺序、频度、指量、空间、时间、内容结构情节等共有的属性，也包括物理属性、化学属性、具体物社会属性（人除外，归入生物属性中的人的社会属性）、生物属性等具体物部分共有的属性和抽象物属性；状态包括事情状态、具体物状态、时空状态；关系指社会关系、位置关系等关系。因此第一层槽类型有 21 种，第二层槽类型为状态、关系和部分属性的细化，有 49 种。例如，物理属性可从材料、颜色色调、亮度透明度、形状、大小、物态、重量、温度、湿度、浓度纯度、气味味道、音质、硬度强度密度、光洁平整度、新旧完整度、导电导热耐磨性、弹性延展性松紧性、型式构件等方面来描述；化学属性可从酸碱性、可燃性、耐磨蚀性、化合性质等方面描述；具体物社会属性（人除外）可从品牌、质量、价值、价格、功能、类型等方面来描述；抽象物属性又从抽象物状态、抽象物评价、抽象物类型、抽象物一般属性等方面描述；关系属性包括社会关系、位置关系、时间关系、逻辑关系、因果条件目的关系、比较关系、部分全体关系、包含关系等。根据这一设计思想，将槽类型设计为二层 70 个槽类型，第一层有 21 个槽类型，第二种有 49 个槽类型。



这 70 个槽类型具体设计为：(1) 事域；(2) 事情；(3) 领属；(4) 来源；(5) 去向；(6) 数量属性；(7) 顺序属性；(8) 频度属性；(9) 指量属性；(10) 空间属性（包括(11) 位置属性、(12) 距离属性、(13) 空间范围属性）；(14) 时间属性（包括(15) 时点、(16) 时段、(17) 历时）；(18) 物理属性（包括(19) 材料属性、(20) 颜色色调属性、(21) 亮度透明度属性、(22) 形状属性、(23) 大小属性、(24) 物态属性、(25) 重量属性、(26) 温度属性、(27) 湿度属性、(28) 浓度纯度属性、(29) 气味味道属性、(30) 音质属性、(31) 硬度强度密度属性、(32) 光洁平整度属性、(33) 新旧完整度属性、(34) 导电导热耐磨性属性、(35) 弹性延展性松紧性属性、(36) 型式构件属性）；(37) 化学属性（包括(38) 酸碱性属性、(39) 可燃性属性、(40) 耐腐蚀性属性、(41) 化合性质属性）；(42) 具体物社会属性（人除外）（包括(43) 品牌属性、(44) 质量属性、(45) 价值属性、(46) 价格属性、(47) 功能属性、(48) 具体物类型）；(49) 生物属性（包括(50) 生理属性、(51) 心理属性、(52) 人的社会属性）；(53) 事情状态；(54) 具体物状态；(55) 时空状态；(56) 抽象物属性（包括(57) 抽象物状态、(58) 抽象物评价、(59) 抽象物类型、(60) 抽象物一般属性）；(61) 内容结构情节属性；(62) 关系属性（包括(63) 社会关系、(64) 位置关系、(65) 时间关系、(66) 逻辑关系、(67) 因果条件目的关系、(68) 比较关系、(69) 部分全体关系、(70) 包含关系）。



实例 1 : (3-1)

[词形] : 瓶子

[拼音] : ping2zi5 [制作者] : 刘智颖 [工作单号] : 3

[语义分类] : { 容器包装用具 }

[论元数目] : [义项数目] : 1 [义项序号] : 1

[释义] : 容器，一般口较小，颈细肚大，多用瓷或玻璃制成。

[词组实例 1] : 那些从拣破烂儿那里买来的布满灰尘的现将用于存放废弃液体的旧的空可乐瓶子

[定语 1] : 那些

[语类 1] : [DL]

[槽类型 1] : [指量属性]

[定语 2] : 从拣破烂儿那里买来 (的)

[语类 2] : [VP]

[槽类型 2] : [来源]

[定语 3] : 布满灰尘 (的)

[语类 3] : [VP]

[槽类型 3] : [具体物状态]

[定语 4] : 现将用于存放废弃液体 (的)

[语类 4] : [VP]

[槽类型 4] : [功能属性]

[定语 5] : 旧 (的)

[语类 5] : [A]

[槽类型 5] : [新旧完整度属性]

[定语 6] : 空

[语类 6] : [A]

[槽类型 6] : [具体物状态]

[定语 7] : 可乐

[语类 7] : [N]

[槽类型 7] : [功能属性]

[槽类型顺序] : [指量属性][来源][具体物状态][功能属性][新旧完整度属性][具体物状态][功能属性]

[可能顺序数] : [来源][指量属性][具体物状态][功能属性][具体物状态][新旧完整度属性][功能属性]



(实例 1 3-2)

[词组实例 2]：我的一个喝水用的长形的、大些的、非常干净并带有美丽花纹的褐色
玻璃罐头瓶子

[定语 1]：我 (的)	[语类 1]：[R]	[槽类型 1]：[领属]
[定语 2]：一个	[语类 2]：[ML]	[槽类型 2]：[数量属性]
[定语 3]：喝水用 (的)	[语类 3]：[VP]	[槽类型 3]：[功能属性]
[定语 4]：长形 (的)	[语类 4]：[NP]	[槽类型 4]：[形状属性]
[定语 5]：大些 (的)	[语类 5]：[AP]	[槽类型 5]：[大小属性]
[定语 6]：非常干净	[语类 6]：[AP]	[槽类型 6]：[具体物状态]
[定语 7]：(并) 带有美丽花纹 (的)	[语类 7]：[VP]	[槽类型 7]：[具体物

状态]	[定语 8]：褐色	[语类 8]：[N]	[槽类型 8]：[颜色色调属性]
	[定语 9]：玻璃	[语类 9]：[N]	[槽类型 9]：[材料属性]
	[定语 10]：罐头	[语类 10]：[N]	[槽类型 10]：[功能属性]
	[槽类型顺序]：[领属][数量属性][功能属性][形状属性][大小属性][具体物状		

态]	[具体物状态][颜色色调属性][材料属性][功能属性]		
	[可能顺序数]：[领属][功能属性][数量属性][大小属性][形状属性][颜色色调属		

性]	[具体物状态][具体物状态][材料属性][功能属性]		
----	----------------------------	--	--



(实例1 3-3)

[词组实例3]: 目前已拒绝回收以前的那批带“优”字的空啤酒瓶子

[定语1]: 目前已拒绝回收 (的) [语类1]: [VP] [槽类型1]: [具体物状态]

[定语2]: 以前 (的) [语类2]: [N (时间)] [槽类型2]: [历时]

[定语3]: 那批 [语类3]: [DL] [槽类型3]: [指量属性]

[定语4]: 带“优”字 (的) [语类4]: [VP] 槽类型4]: [内容结构情节属性]

[定语5]: 空 [语类5]: [A] [槽类型5]: [具体物状态]

[定语6]: 啤酒 [语类6]: [N] [槽类型6]: [功能属性]

[槽类型顺序]: [具体物状态][历时][指量属性][内容结构情节属性][具体物状态]

[功能属性][可能顺序数]: [具体物状态][指量属性][历时]

[内容结构情节属性][具体物状态][功能属性]

[槽关系表达式联想]: [领属][来源][指量属性] | [数量属性][具体物状态]*

[形状属性][大小属性][颜色色调属性][新旧完整度属性]

[材料属性][功能属性]



实例 2 : (3-1)

[词形] : 愿望 [拼音] : yuán4wàng4 [制作者] : 舒琴 [工作单号] : 114

[语义分类] : { 感觉心境 } [论元数目] : [义项数目] : 1 [义项序号] : 1

[释义] : 希望将来能达到某种目的的想法。

[词组实例 1] : 中国人民渴望已久，至今仍未得以实现的承办两千年奥运会的美好而强烈的愿望

[定语 1] : 中国人民渴望已久， [语类 1] : {S} [槽类型 1] : [抽象物状态]

[定语 2] : 至今仍未得以实现(的) [语类 2] : {VP} [槽类型 2] : [抽象物状态]

[定语 3] : 承办两千年奥运会(的) [语类 3] : {VP} [槽类型 3] : [内容结构情节属性]

[定语 4] : 美好而强烈(的) [语类 4] : {AP} [槽类型 4] : [抽象物状态]

[槽类型顺序] : [抽象物状态][抽象物状态][内容结构情节属性][抽象物状态]



(实例 2 3-2)

[词组实例 2]：中国与南亚各国人民共同的进一步发展贸易合作关系的这一迫切愿望

[定语 1]：中国与南亚各国人民 [语类 1]：{NP} [槽类型 1]：[领属]

[定语 2]：共同(的) [语类 2]：{A} [槽类型 2]：[抽象物状态]

[定语 3]：进一步发展贸易合作关系(的) [语类 3]：{VP} [槽类型 3]：[内容结构情节属性]

[定语 4]：这一 [语类 4]：{DL} [槽类型 4]：[指量属性]

[定语 5]：迫切 [语类 5]：{A} [槽类型 5]：[抽象物状态]

[槽类型顺序]：[领属][抽象物状态][内容结构情节属性][指量属性][抽象物状态]



(实例 2 3-3)

[词组实例 3] : 傅老昨天在会上转达的一些国外的跨国大公司表示想与中国业界密切合作、深受大家欢迎的积极愿望

[定语 1] : 傅老昨天在会上转达的 [语类 1] : {S} [槽类型 1] : [来源]

[定语 2] : 一些国外的跨国大公司表示想与中国业界密切合作
[语类 2] : {S} [槽类型 2] : [内容结构情节属性]

[定语 3] : 深受大家欢迎 (的) [语类 3] : {VP} [槽类型 3] : [抽象物状态]

[定语 4] : 积极 [语类 4] : {A} [槽类型 4] : [抽象物评价]

[槽类型顺序] : [来源][内容结构情节属性][抽象物状态][抽象物评价]

[槽关系表达式联想] : [来源][领属][抽象物状态]* [空间范围属性][抽象物状态]*

[内容结构情节属性][指量属性][数量属性][抽象物评价]



实例 3: (3-1)

[词形]：小偷 [拼音]： xiao3tou1 [制作者]：肖榕 [工作单号]： 5

[语义分类]：{人物} [论元数目]： [义项数目]： 1 [义项序号]： 1

[释义]：偷东西的人。

[词组实例 1]：那几个从新疆来的经常在公共汽车站出没的以外地人为偷窃对象的屡次漏网的小偷

[定语 1]：那 [语类 1]：{D} [槽类型 1]：[指量属性]

[定语 2]：几个 [语类 2]：{ML} [槽类型 2]：[数量属性]

[定语 3]：从新疆来 (的) [语类 3]：{VP} [槽类型 3]：[来源]

[定语 4]：经常在公共汽车站出没 (的) [语类 4]：{VP} [槽类型 4]：[具体物状态]

[定语 5]：以外地人为偷窃对象 (的) [语类 5]：{VP} [槽类型 5]：[具体物状态]

[定语 6]：屡次漏网 (的) [语类 6]：{VP} [槽类型 6]：[具体物状态]

[槽类型顺序]：[指量属性][数量属性][来源][具体物状态][具体物状态][具体物状态]3

[可能顺序数]：[具体物状态][具体物状态][指量属性][数量属性][来源][具体物状态]3



(实例 3 3-2)

[词组实例 2] : 公安局刚刚抓获的一个与当地犯罪集团有密切联系的四处流窜的小偷
[定语 1] : 公安局刚刚抓获 (的) [语类 1] : {S} [槽类型 1] : [来源]

[定语 2] : 一个 [语类 2] : {ML} [槽类型 2] : [数量属性]

[定语 3] : 与当地犯罪集团有密切联系 (的) [语类 3] : {VP} [槽类型 3] :
[具体物状态]

[定语 4] : 四处流窜 (的) [语类 4] : {VP} [槽类型 4] : [具体物状态]

[槽类型顺序] : [来源] [数量属性] [具体物状态] [具体物状态]

[词组实例 3] : 电视上通缉的那个身高一米七左右、头发略为卷曲、带农村口音的小偷

[定语 1] : 电视上通缉 (的) [语类 1] : {S} [槽类型 1] : [具体物状态]

[定语 2] : 那个 [语类 2] : {DL} [槽类型 2] : [指量属性]

[定语 3] : 身高一米七左右 [语类 3] : {S} [槽类型 3] : [生理属性]

[定语 4] : 头发略为卷曲 [语类 4] : {S} [槽类型 4] : [生理属性]

[定语 5] : 带农村口音 (的) [语类 5] : {VP} [槽类型 5] : [人的社会属性]

[槽类型顺序] : [具体物状态] [指量属性] [生理属性] [生理属性] [人的社会属性]

[槽关系表达式联想] : [[指量属性] | [数量属性]] [来源] [内容结构情节属性]


[生理属性] * [人的社会属性] * [具体物状态] *



5 信息处理用现代汉语语义分类词典的研究与实现

研究名词词汇意义的一个方法是语义场方法，即研究词的聚合关系、上下位关系。语义场体现义位的关系和区别，也体现语义的系统性。这个系统性是客观世界的系统性在语义中的反映。研究汉语分类语义场必须先研究汉语语义分析体系。分类的意义和作用在于：分类是人类对事物认识的一种结果，也是人类对事物认识的一种手段。分类法刻划事物简洁、清晰、信息密度大，反映了词汇意义中最基本也是信息量最大的属性。当然并不是说分类可以解决全部的问题。“任何简单的分类、单一的关系描述、完全静态的分析都是无济于事的，必须把多层次、多类型、多关系、多变化这些性质综合加以考虑。”不过分类体系描述既可节省存储空间（下位结点可以继承所有上位结点的语义特征），又可避免单个义项描写时遗漏语义特征或不同人描写出现的分歧。

作者十几年来在研究参考了国内外十来个分类体系和词典优缺点的基础上加上自己的见解初步确定了一个信息处理用现代汉语语义分类体系，并在此基础上、设计和实现一个信息处理用现代汉语语义分类词典管理系统软件，用来对几万词条进行实际分类和代码化组成一个语义分类词典库。词典管理系统主界面包括“创建索引文件”、“批量添加词”、“归类回收站”、“导出节点下的词库”、“整个语义词典词数的统计”、“语义分类各层节点数统计”、“叶结点数统计”等功能，而“语义树的管理功能”设计了强大齐全的语义树管理和修改功能，主要包括“添加一个语义分类节点”、“删除当前语义分类节点”、“修改一个语义节点的名称”、“拆分当前语义分类节点”、“移动当前语义分类节点”。“词典库管理功能”设计为主要包括词的查询功能（含“查询语义分类节点下的词”、“查询词的语义分类”、“查询同一语义分类节点的其他词”、“显示暂时删除词选项”）和词的增删改功能（含“添加一个词”、“删除若干词”、“修改词及属性”）。



我们还完成了对 7 万多汉语词语 11 万多义项进行机助标注的工作形成语义分类词典。目前正在对该语义分类词典进行校对修正工作。

信息处理用现代汉语语义分类词典管理软件功能强大齐全，人机界面友好。例如，用户想查询语义分类节点“超类”下有多少词，可在系统界面左框中用鼠标左键点击语义树的“超类”分枝，而又想同时查询词汇“中国”的语义分类，则可在“查询语义分类节点下的词”输入框中输入汉字“中国”然后点击“查询”按钮，则系统返回（如图 1）



语义分类词典管理系统

系统(S) 特殊功能(Q) 查询(F) 词的操作(W) 树的操作(T) 统计(C) 查看(V) 帮助(H)

S1 S2 S3 AW DW MW 增 删 改 分 移 ?

- 1 超类
- 2 事
 - 2.1 事域
 - 2.2 事情
- 3 物
 - 3.1 具体物
 - 3.1.1 生物
 - 3.1.1.1 人类
 - 3.1.1.1.1 人
 - 3.1.1.1.1.1 人称
 - 3.1.1.1.1.2 人名
 - 3.1.1.1.1.3 人物
 - 3.1.1.1.1.4 人间
 - 3.1.1.1.1.5 神鬼
 - 3.1.1.1.2 集体
 - 3.1.1.2 动物
 - 3.1.1.2.1 兽
 - 3.1.1.2.2 鸟
 - 3.1.1.2.3 鱼
 - 3.1.1.2.4 虫
 - 3.1.1.2.5 两栖动物
 - 3.1.1.2.6 爬行动物
 - 3.1.1.2.7 软体动物
 - 3.1.1.2.8 家禽家畜
 - 3.1.1.2.9 宠物
 - 3.1.1.2.A 其他动物
 - 3.1.1.3 植物
 - 3.1.1.4 微生物
 - 3.1.2 非生物
 - 3.1.3 现象物
 - 3.1.4 现象物的源和因
 - 3.1.5 具体物集合类名
 - 3.2 抽象物
- 4 时空
- 5 部件
- 回收站

词形	词性
沧海	n
沧桑	n
大自然	n
东西	n
六合	n
乾坤	n
社会	n
世界	n
事物	n
万事万物	n
万籁	n
物	n
物事	n
物体	n
物质	n
宇宙	n
漩涡	n
瀚海	n
寰球	n
寰宇	n

查询词:中国,结果如下:

3.1.1.1.2.1 国家地区

3.2.5.1.3.3.1.4 国籍

4.2.1.1.4.2.1.1 国家地区

就绪 1 超类 当前节点下的叶结点数:0当前节点下共有20个词:

开始 My... Fl... 语... Xr... 李... Ve... 语... 词... Nou... 语...

10:17 星期二 2004-9-28

图1 查询语义节点“1 超类”下的词和查询词“中国”的语义节点代码



若想查询词“辕马”的语义分类，则可在“查询词的语义分类”输入框中输入汉字“辕马”，然后点击“查询”按钮，则系统返回（如下图 2）：

若想查询词汇“蛋”的语义分类，则可在“查询词的语义分类”输入框中输入汉字“蛋”，然后点击“查询”按钮，则系统返回（如下图 3）：



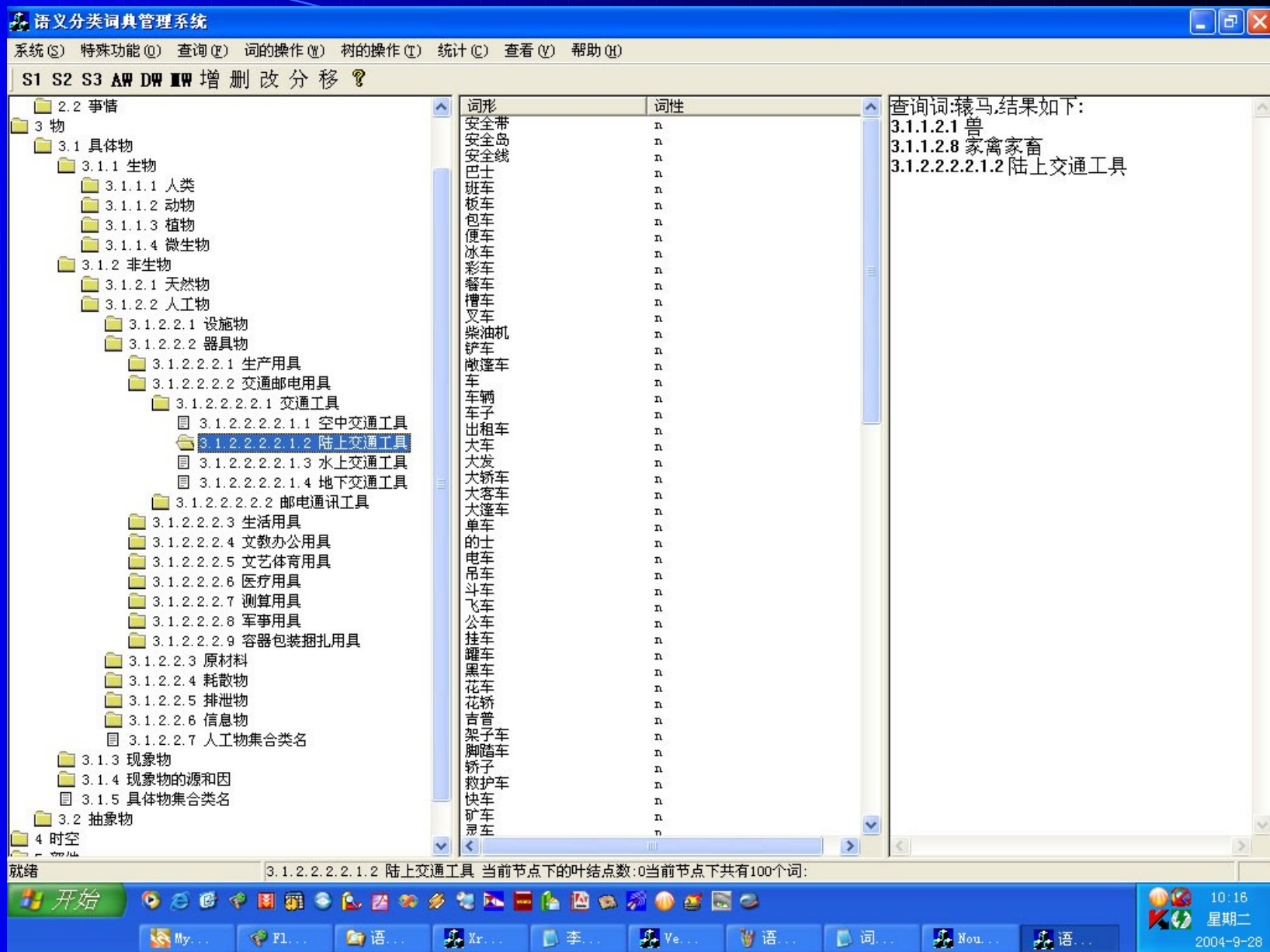


图2 查询词“轘马”的语义节点代码

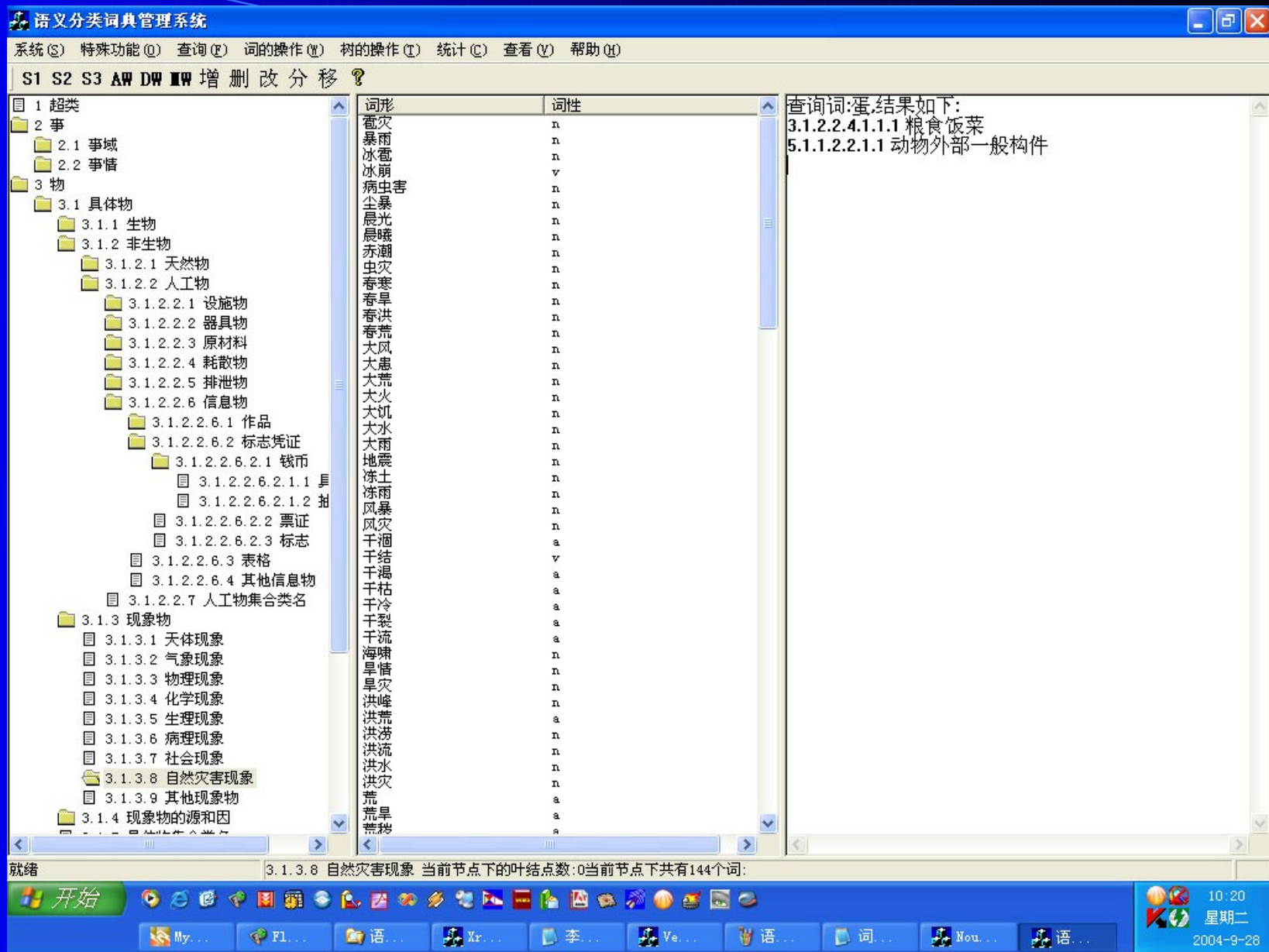


图3 查询词“蛋”的语义节点代码



6 现代汉语语义知识库平台的实现

目前，现代汉语述语动词机器词典，现代汉语述语形容词机器词典、现代汉语名词槽关系系统，现代汉语语义分类词典和现代汉语虚词义类词典已经整合集成一个现代汉语词义知识库平台。整合的原则和目标是：（1）界面的灵活性；（2）库与库之间灵活地切换；（3）跨库的查询；（4）信息的综合输出和定制输出。整个语义知识库可为中文信息处理和语言学本体研究提供主要实词类的语义、句法和部分语用信息。语义知识库平台的总界面如图4。

至今，现代汉语述语动词机器词典已描述了 5648 个汉语动词的 9286 个义项，现代汉语述语形容词机器词典已描述了 1815 个汉语形容词的 2300 个义项，现代汉语名词槽关系系统已描述了 3910 个汉语名词的 6308 个义项，现代汉语语义分类词典已标注了 72966 个汉语词汇的 112057 个义项。现代汉语虚词义类词典的描述正在进行中。



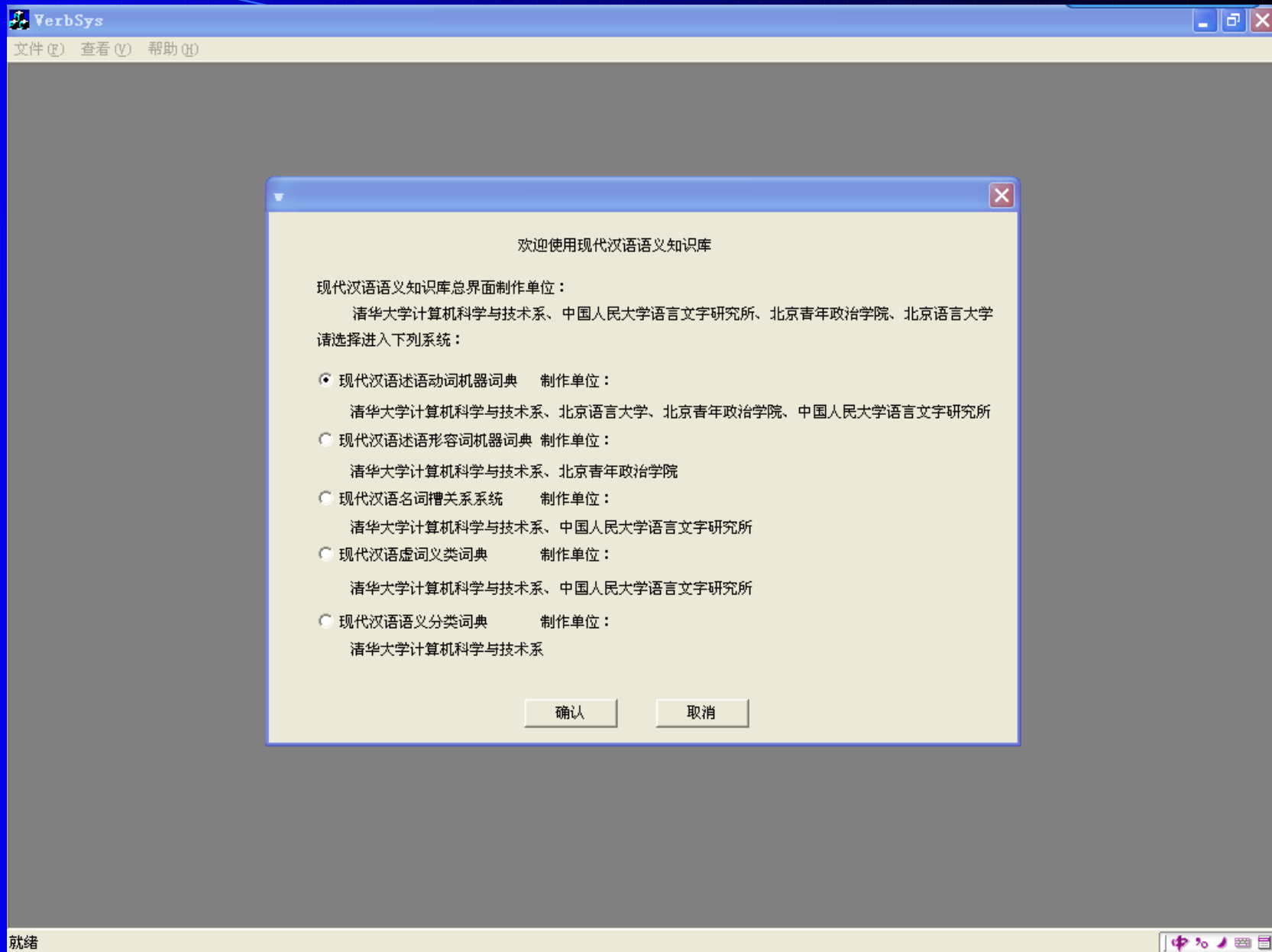


图 4 现代汉语语义知识库总界面



谢谢大家！

