

隐含型语句函子的求解方法*

陆汝占 靳光瑾

上海交通大学计算机科学与工程系 200030

南京大学计算机软件新技术国家重点实验室

摘要 空语类理论可用来表示含隐含型缺省成分汉语语句。本文介绍基于空语类理论的求解隐含型汉语语句函子的方法, 求解出 p 空语类变元的语义指称, 从而得到完整的逻辑函子。

关键词 p -空语类, 逻辑函子

A METHOD FOR EXTRACTING THE LOGICAL FUNCTORS IN CHINESE SENTENCES WITH UNDERSTOOD NOUN PHRASES

Lu Ruzhan Jin Guanjin

(Department of computer Science and Engineering Shanghai Jiao Tong University)

Abstract The paper proposes a method based on empty category theory for solving the problems of extracting logical functors in Chinese sentences, which with understood noun phrases in verb structures. In this approach, denotations of p -empty category variables can be obtained, and so can the complete logical functors.

Key word p -empty category, logical functor.

汉语表达有简短、精炼的特点。利用情景和上下文可以在语句中出现不少缺省成分。省略也是汉语语法变化中最容易引起争议的方面之一。缺省成分的分析 and 理解是计算机处理汉语中最为困难的问题之一。出现缺省成分的语句, 对于本族人来说, 使用和理解都已习惯。对于外族人和计算机来说, 要理解就要回答两个问题: (1) 句中什么地方出现缺省。(2) 缺省的成分所指的是什么, 如何找回。所谓计算机理解汉语, 最基本的就是如何自动求解这两个问题。

* 本文受国家自然科学基金资助。项目编号: 69573020

从语用的角度看，有些词语是不必说出的或者难于说出确切的。说/听话时最关心的是语用焦点的信息以及它们的变化。变化的差异必须在言语中表达出来的，前后句子间其余相对不变化的成分，就可以共享，可以省略不出现。利用说话人所处情景、上下文（甚至非语言手段）可以意识到这些缺省成分，从而完整地理解。

从句法角度看，缺省成分的出现是有规律性的，本族人依据语感，外族人只有靠学习句法理论，计算机要依赖句法知识来完成理解。能提供关于缺省成分分布规律的句法知识主要包括：（1）句型结构（2）动词分类和配价理论（以及某些名词、形容词的配价理论）（3）词义。尤其值得注意的重点在于动词的规律，因为缺省成分绝大部分在于：（1）缺省动词（系词）（2）缺省动词的配价成分（即必有成分）（3）进一步与动词词义有关。这是说，从词汇的角度看，动词与其配价成分之间的语义配价关系在动词分析义素之后，这些配价成分与义素动词之间是否继承这种配价关系。此外，语义上外延和内涵的区分，也会影响语义指称。

缺省问题实际上是从句法平面映射到语义平面的问题。汉语语法学家关心的是“有没有一种科学的、精密的、形式化的方法能把句子的表层结构与它的语义解释联系起来”，“为沟通结构和语义这两个平面、为语法学的精密化、定量化与形式化提供一个途径”（陆俭明）。计算语言学家更关心的是基于上述句法研究成果建立汉语的计算分析模型，以便让计算机理解和处理汉语语句中的缺省。沈阳《现代汉语空语类研究》基于动词配价理论，将空语类分为三类：移位、隐含、省略。由此解释几乎覆盖上述所列各种常见句型中所出现的缺省现象，从而在句法和语言两个平面上建立一种基本对应关系。

本文以该理论为参照起点，假定已经知道语句中的缺省位置和类型，求解相应的空语类变元指称什么，如何找回。思路和求解策略如下：

1. 求解过程应该是可构造性的计算过程，因此要符合Frege叠置原理，即允许复合。

2. 直接计算最基本的原子组成。最基本的结构是什么？是词、短语还是句呢？应该是最小的意义自足的主谓结构，这就是范晓所说的动核结构，沈阳所说的基本句位（最小句位）先解决单个基本结构的缺省成分求解。

辖域是为了明确空语类变元的语义指称关系中的约束概念，是相对于什么作用域（Scope）而言的，称为辖域。如果将动词看成约束词，那么辖域是动词对名词性成分的约束位置。沈阳（1994）按三个不同阶的动词与名词性词语的句法形式分为三个抽象的基本句法，称为最小句位。加上动词短语作主语宾语的形式，共五个。最小句位对应的抽象逻辑形式，称为函子。其中V为约束词，管辖变元x,y,z为直观起见，写成前缀形式。函子与谓词不同在于出现的项，前者为名词项，后者为个体项。

最小句位：Sp1: [NP1 V1] 函子：V¹ (x)

Sp2: [NP1 V2 NP2] V² (x,y)

Sp3: [NP1 V3 NP2 NP3] V³ (x,y,z)

Spd1: SP1-3 V NP 函子， x ↔ 命题

Spd2: NP V SP1-3 函子， y/z ↔ 命题

最小句位所在的句称为句域，简称句。基本句位是抽象形式，具体有形词语出现的句子是句位的一个实例（Instance），称为句位实例。出现在函子中的项是动词支配的必有成

分。至于说明的可有成分（状语、补语）出现在扩展函子中。“昨天张三在会上批评李四”
函子 V^2 批评（张三,李四），扩展函子（昨天） \nearrow （在会上） $\nearrow V^2$ 批评（张三,李四）。

基本句位中用变量 e 、 t 、 p 代替NP后得到句位式子，称为空语类句。一个具体语句句法分析后按动词配价理论抽象出相应的空语类句式，每个这样的句式中出现基本句位的变异，所以语句可抽象为空语类句式。从空语类句式中提取出函子（可能含空语类变元），求解出代换和解。

变量 e 表示省略型空语类。

①出现情况：在动词V的基本句位中可充当某个NP出现的有形词语，在句位实例中任何位置上不出现，包括V的最小辖域及其句域内都不出现。句中是省略成分，在句位中用 e 表示。

②句法上，这个省略的有形词语可以“补出来”，补回的项从句域外寻找。

③语义上，最小辖域内的 e 是自由的，可代入。代以句域外某个合适的词语项。

变量 t 表示移位型空语类。

①出现情况：在动词 V 基本句位中，本可充当某个NP出现的有形词语，未在句位实例中的相应位置上出现，却出现在句域内的其他位置上。此时，句中移位前的用 t 表示。 t 表示最小辖域内的空缺，因此

②句法上，如果试将移位的词语设法移回 t 位置上，就有两种可能：

a.可以移回，此时的 t 在最小辖域内是自由的，移回去就意味着 t 允许代入最小辖域外的项，b.不可以移回。

③语义上也有相应的区分。

a.可移回的 t ，最小辖域内自由，可代入最小辖域外但是句内之项。

b.不可以移回，不能代以辖域外的项，因此此类 t 在最小辖域内不是自由的，是“束缚”的 (obligative)，但是 t 的语义指称与移位项的指称是相同的。

变量 p 表示隐含型空语类。

①出现：在动词V的基本句位中，本可以充当某个NP出现的有形词语，在句实例中任何位置上不出现，包括最小辖域及句域内都不出现，所以仅是形式上用变元 p 表示。

②句法上， p 位置上绝不能补回任何项。这个 p 位纯粹是由于动词的基本句位形式上语义分析上的需要设定的。实际语句绝不会在此位置上出现任何词语。这个 p 位在句法上也是紧凑格式，如连动句型、动结补句型，“打算”后出现 p 。

③语义上， p 是束缚的， p 语义指向与某个词语所指相同，但两者无任何句法上的必然联系，这词语可在句内、句外，上下文、说话情景之中隐含的，需要理解分析的。

空语类变元取值

作为变元就有一个取什么值的问题。空语类变元 (e、t、p) 也有取值和语义指称的问题, 包括①值是什么? ②如何取? 怎样的取值过程?

空语类变元的值就是一个词语项, 空语类变元按取值过程不同可分为两类。一类称为实变元, 它的值是通过具体的代换赋给的, 按形式理论的说法, 它可用形式系统内的表示形式赋值的。从约束关系看, 这类变元是自由 (free) 的, 例如 e 变元, 部分 t 变元在句法上人们都是可以移回原位的。实际上, 赋值过程是 $e :=$ “那两件衣服”, 计算表达式为 $\theta \cdot \psi$, 其中 $\theta: \{e \rightarrow \text{那两件衣服}\}$ 。ψ 为表达式, 计算 (赋值) 结果就是将表达式 ψ 中的 e 的所有出现, 统统代入 θ 中规定的项。另一类称为形式变元, 它们在句法上是不允许直接移回来的, 值不是通过代换赋给的, 而是“间接”指向关系, 由解释得到的。按形式理论的说法, 他无法用形式系统内的表达式赋值, 要用系统外的元语言 (模型) 来解释其语义值, 从约束关系上看它们是束缚 (obligative) 的。实际上仅仅是一个语义解释过程, 例如 p 变元, 我们可以有 $|p| \rightarrow | \text{那两件衣服} |$, 就是说, 如果在具体环境下解释了“那两件衣服”是什么, 也就解释了 p 是什么, 即 p 与该词项同指, 但是我们没有“p := 词语项”这样的赋值过程。形象的说, 实变元的值是“拿得到”的, 形式变元的值不是“拿得到”的, 仅仅是“影射到”的。这就是说, 赋值的变元当然与所赋值的表达式同指, 但同指的不一定能被赋给该表达式。

空语类型	p (隐含空语类)	t (移位空语类)	e (省略空语类)
约束类型	形式变元	实型变元	
约束关系	束缚	自由	
指称形式	被解释同指	被代换同指	
句法手段	不可移回, 仅影射	可移回	
指称范围	句内, 句外	句内	情景, 上下文

句法上 P 空语类变元出现及其同指项出现的结构有两类: 从属结构和并合结构。从属结构式 S1[S2] 和 S1[VP] 形式。S1 为外层大句, S2 为嵌入的内层小句, P 变元出现在内层, 同指性出现在外层。外层大句中动词子类特征决定 P 的指向关系。S1[VP] 形式 VP 是动词短语或者小句, 作为语句 S1 的一个成分: 主语、状语、定语。逻辑上 VP 表示一个命题 Q, 它又作为 S1 中的一个论元或这一个修饰成分。并合结构式指动结补结构、“得”字句, 内容丰富, 受篇幅限制, 简略介绍。

A. “进行”, “打算”类动词: 特征: Va, 二价, 与 Vb 有共享的施事论元。空语类句式: NP1 Va [P VPb], 语义同指规则: $|P| \rightarrow |NP1|$

例1 小李打算学德语

空语类句式: 小李打算 [P 学德语]

函子: V^2 打算 (小李, Q), $Q \leftrightarrow V^2$ 学 (P, 德语), 由同指规则: $|P| \rightarrow | \text{小李} |$

例2 这个问题我们一定进行处理

空语类句式：这个问题我们一定进行[P 处理 t]

函子： V^2 进行（我们，Q）， $Q \leftrightarrow V^2$ 处理（P，t），同理，先确定|P|→|我们|。t空语类变元，由相应得规则在句内求解移位项，得到 θ ：{t→这个问题}（t空语类见另文）

B. 递系动词子类，使令动词以及“鼓励”等这类动词与内层动词构成论元之间有传递关系的结构。空语类句式：NP1 Va NP2 [p Vb] 语义同指关系规则|P|→|NP2|，特征：动词Va的价 ≥ 2 ，第二论元与Vb的第一论元（p空语类）同指，语义类型：人/机构。

递系结构 NP1 Va NP2 [VPb]，有三种处理方法：

(1) Va，二价，把 NP2 VPb看成一个小句Q，所以语句逻辑式：

$Va(NP1, Q)$ and $Q \leftrightarrow VPb(NP2)$

这样处理的优点是价与句法一致，逻辑结构简洁。缺点是 Va 的论元含命题变元，成为高阶式，其实这儿可以避免用高阶。动词Va的对象原本是NP2，可以用 $Va(NP1, NP2)$ 显露出它与NP1的关系，这对于进一步分解语义是有价值的，但现在这个NP2包含Q中，未分离出来。此外，也不能充分利用空语类理论。

(2) Va看成三价，语句逻辑式： $Va^3(NP1, NP2, Q) \wedge Q \leftrightarrow VPb(P)$ 。

优点：用上空语类理论，显露出NP1语NP2的关系。

缺点：Va的价与动词配价理论不一致，这样会使升价的情况过多。

(3) V1仍为二价，仍用上空语类理论。但需定义一种V1与V1p的关系表示。

逻辑式： $Va(NP1, NP2)$ VPb(p)

好处：既显露了NP1、NP2的关系，又用上p空语类理论，与句法理论中的传递关系一致。VPb修饰Va的关系不仅刻划了这种句型的特征，而且也表示了“叫、让、鼓励、命令”这一类动词与后面事件、行为、动作的关系。本文采用第三种办法。

从词汇意义上说，这种关系是表示意愿、命令等意识上的关系，并不是实体关系。当前逻辑理论分析最基本的在于解释实体关系，用函子表这些实体关系。意识上的关系是高一层的。

例3 父母嘱咐他换件衣服出门。

空语类句式：父母嘱咐他[[p换件衣服][p出门]]

函子： V^2 嘱咐(父母,他),{ V^2 换(p, 一件衣服), V 出门(p)}

由规则得 |p|→|他|，函子关系 V^2 嘱咐 \bowtie (V^2 换, V 出门)。

例4 领导安排他做行政工作。

空语类句式：领导安排他[p做行政工作] 函子： V^2 安排(领导,他), V^2 做(p, 行政工作)

由规则得 |p|→|他|，函子关系 V^2 安排 \bowtie V^2 做。

C. “陪同”类动词，由词义决定了同指的项应指向复数（复合多个所指）。“它陪妻子上街”，不能单从句法[VP]VP得到表示式： $V^2_{\text{陪同}}(\text{他}, \text{妻子}) \leftrightarrow Q, V_{\text{上}}(Q, \text{街})$ ，这与原句意思不符，应表示成： $V^2_{\text{陪同}}(\text{他}, \text{妻子})$ 并且 $V_{\text{上}}(\{\text{他}, \text{妻子}\}, \text{街})$ 。这是因为词义项：陪同：陪伴着一同，取外延义： $[V^2_{\text{陪同}}(\text{他}, \text{妻子})] = \{\text{他}, \text{妻子}\}$ 。这类语句的空语类句式： $NP1 Va NP2 [P VPb]$ ，同指规则： $|P| \rightarrow |[Va(NP1, NP2)]|$ ，函子： $V^2_{\text{陪同}}(\text{他}, \text{妻子})$ ， $V^2_{\text{上}}(P, \text{街})$ ，其中 P 按上所指。

D. 含P变元的VP作主语、宾语、状语、定语的情况，如果一律都作为小句处理，比较烦琐，不如直接简化处理。本文略。

E. 并合结构中主要是补语结构，外套大句Va，内层小句VPb的关系。可以归纳为五种情况，用图示意 P 变元出现。

模式(1) — (5)：

- (1) $\square Va \square$ (2) $\square Va$ (3) $\square Va ? \downarrow$ (4) $\square Va \square$ (5) $\square Va \square e1 Va e2$
 $\uparrow? Vb$ $\uparrow? Vb$ $\uparrow? Vb \square$ $\uparrow? Vb ? \uparrow$ $\uparrow? Vb$

例示：

例10 他开着窗户睡觉
 他 开 窗户
 $\uparrow?$ 睡觉
 (模式1)

例11 小姑娘急哭了
 小姑娘 急
 $\uparrow?$ 哭
 (模式2)

例12 我听懂了你的话
 我 听 ? \downarrow
 $\uparrow?$ 懂 你的话
 (模式3)

例13 我倒了杯水喝
 我 倒 水
 $\uparrow?$ 喝 $\uparrow?$
 (模式4)

例14 他喝酒喝醉了
 他 喝 酒 (e1喝e2)
 $\uparrow?$ 醉
 (模式5)

其中模式(5)是模式(1)的变异，加上重置动词的e变元代换。

P 变元的同指规则：

(1) VPb的第一论元必为P，指向Va的第一论元。

(2) VPb第二论元Va第二论元出现P变元（两者互斥，不兼有），则两者进行合一运算，即得P所指。

参考文献

[1]. 胡裕树, 范晓, 动词研究, 河南教育出版社, 1995
 [2]. 沈阳, 现代汉语空语类研究, 山东教育出版社, 1994