

# 属性分析说略\*

陈小荷

南京师范大学文学院 江苏 210097

**摘要:** 中文信息处理的关键是语义的形式化, 属性分析在语义形式化的研究中可扮演重要角色。本文探讨属性词语和表值词语的语言学性质, 属性分析的基本内容、目标和研究方法, 并从中文信息处理的角度论述这一理论性课题的研究意义。

**关键词:** 属性词语; 表值词语; 属性分析

## On Semantic Attributes

Chen Xiaohe

School of Chinese Language and Culture, Nanjing Normal University, Jiangsu 210097

**Abstract:** The crucial issue of Chinese information processing is the formalization of meanings in which attribute analysis can play an important role. In this paper we discussed the linguistic features of attribute-words and value-words, the primary content, goal and method of attribute analysis, clarified the significance of this theoretical research in Chinese information processing.

**Keywords:** Attribute-Word; Value-Word; Attribute Analysis

### 一. 属性分析

中文信息处理的实践表明, 仅靠语法约束是不能解决汉语语句的自动分析问题的, 语义分析应该在自动分析中发挥更大的作用。语义分析需要形式化, 用一套代码来表示每个实词的词义分类或词义特征, 并且建立起使用这套代码的规则。在中文信息处理领域, 语义形式化的主要努力是编写机器可读的语义词典(董振东, 1998; 陈小荷, 1998), 或用语料库统计方法对词语进行自动聚类(李涓子, 1997)。

我们认为, 语义分析的重点应从动词转移到名词和形容词。大多数语言学家把动词看做句子的中心, 许多人对动词的配价进行了多方面的研究。名词和形容词的配价是新的研究课题。重点转移的理由是: 名词数量大, 总体出现频率高, 名词词义分析的基础相对薄弱。动词词义分析中包含有对每个配价的语义要求, 配价一般由名词性短语充当, 如果不做名词词义的分析, 从语言信息处理的角度看, 动词的配价分析就不能实际发挥作用。至于形容词, 其重要性主要不在于其配价的多少, 而在于它可以为名词和动词的词义分析提供客观依据。

本文特别指出, 属性分析可在语义形式化的研究中扮演重要角色。涉及属性分析的词

---

\* 本文的一些观点的形成过程中, 得到董振东先生的许多启发和指导, 特此致谢。

语主要有以下三类：

(一) 实体词语，表示具体或抽象的客观事物实体，包括事件、时间和地点。例如“人”、“桌子”、“思想”、“决赛”、“2003年3月9日”、“韩国棋院”。

(二) 属性词语，表示事物的某种属性。例如“高度”、“颜色”、“含量”、“比率”、“科技含量”、“市场占有率”。从最宽泛的意义上说，名词和名词性短语表示事物，属性也是一种事物。但大多数名词和名词性短语表示实体，属性词语则表示用以观察或描述这些实体的角度或论域。借用数据库的术语来说，实体是记录(Records)，属性是字段(Fields)。

(三) 表值词语，表示事物在某种属性上的取值，例如“高”、“低”、“1米”表示具体事物在高度这一属性上的取值，“深刻”和“肤浅”表示思想在抽象深度这一属性上的取值。从语法上看，表值词语涉及形容词、数量短语、区别词和名词等多种词类。例如，表示规模属性的“大型”和表示式样属性的“新式”是区别词，表示颜色属性的“红色”是名词，“红”是性质形容词，“红彤彤”是状态形容词。有些属性经常用实体词语来表示属性值，例如“2003年3月9日”作为时间属性的一个值，“韩国棋院”作为地点属性的一个值。有些属性经常用动词性短语来表示属性值，例如“就餐”是“饭桌”的用途属性取值，“主张和平解决”和“主张军事打击”是对某个国际问题的立场属性取值。

属性可用来区分不同的实体类别。例如“桌子”、“操场”和“江”都有面积属性，但“桌子”有高度，“江”有深度，“操场”没有这些属性。又如，“思想”有抽象的深度、广度和先进性等属性，“感情”有抽象的温度、浓度和距离等属性。说某类实体有无某种属性，意思是该类实体可以或不可以从某个角度去描述。

属性值也可用来区分不同的实体类别。例如，“红旗”在颜色属性上取值为红色，区别于一般的“旗”。“热情”作为实体，在抽象温度上取值为“热”，区别于一般的“感情”，作为属性值，则是抽象温度的一个取值。“浓情”、“亲近的感情”则表示在抽象浓度和抽象距离上已经取值为“浓”和“近”的两类抽象实体。

属性值的另一个重要作用是区分事物的不同个体。专有名词和词组式的专名是名称属性的取值，但个体还经常凭指代词和表值词语来相互区别。事件也是一种实体，一类事件有哪些基本属性，每个属性应如何取值，当由对该类事件感兴趣的用户来确定。在信息提取中，我们主要关心如何识别文本中的个体事物(命名实体识别)，如何从文本中抽取表达属性和属性值的词语。如果说，单纯的语言学研究主要关注事物的类别，那么语言信息处理还应关注事物的个体，特别是个体在每一属性上的取值。

袁毓林(1994)指出，事物可有多方面的属性，属性有强弱之分，表示强属性的名词可以缺省，表示弱属性的名词不能缺省；如果形容词专表某种属性，属性名词也可以不出现；一般规律是，整体属性强于局部属性，心理属性强于物理属性；属性的强弱可能依具体事物而有所不同。这些论述对于名词和形容词的语义分析极有价值。但是，语言信息处理中的属性分析还有许多基础工作需要做，包括：表值词语的分类和识别，属性名词的界定，属性词语的分类和识别，属性三角(实体、属性和属性值)之间的关系，等等。

## 二. 表值词语的分类

表值词语可从语义上分为两大类，一类是定性的表值词语，另一类是定量的表值词语。

定量的表值词语由数量短语充当，其中量词可提示表值词语所对应的属性。量词为度量词时，表值词语所对应的物理属性一般是比较明确的，称作度量值，例如下面这些量词：

“公里”、“米”、“毫米”等等，对应于距离、长度、宽度、厚度、高度、深度等属性，这些属性从物理意义上看是同类的，但在自然语言里有不同的认知意义。

“年”、“月”、“日”、“天”、“点（钟）”、“小时”、“分”、“秒”等不是度量词，但所对应的物理属性（时长或时点）也是明确的。

有些度量词所对应的属性有不同的物理意义，例如“度”可以对应于温度、角度、电量等不同属性，需参考相关的实体词语才能确定。

定性的表值词语可分为四种，第一种是对立值，通常由两个意义相反的形容词来充当，例如“高”和“低”（高度），“重”和“轻”（重量），“大”和“小”（体积、规模、年龄等）。两个对立的形容词都可以加副词来表示更多的取值，形成一个表值序列，如“最高”、“非常高”、“很高”、“高”、“不很高”、“不高”、“不很低”、“不低”、“低”、“很低”、“非常低”、“最低”。表值序列中的各个值虽然很难量化，但可以比较其大小。对立值中往往有一个形容词是“无标记”的，从属性分析的角度看，无标记的形容词不表示属性值，而是表示属性，例如“深”表示深度属性（“水深两米”，“水有两米深”）。

第二种是单个值，也是由形容词充当，但语言中不存在另一个意义相反的形容词，例如“渴”表示人或动物缺水的属性，“累”表示人疲乏的属性。加副词后所形成的表值序列较短，例如“非常累”、“很累”、“不累”等。

第三种是分级值，例如职称属性的取值分“高级”、“中级”、“初级”，年龄属性的取值分“老年”，“中年”，“青年”，“幼年”等，形成一个表值序列。跟前两种不同的是，这种表值序列中的值都由单词充当。

第四种是枚举值，例如颜色属性的取值分“红”，“绿”，“黄”，“蓝”，“白”，“黑”等，虽然物理学家可以用光波波长来说明这些值也是一个序列，但这不是常识，常识是把这值看成一个集合。颜色值的集合是有限的。

如果不跟具体事件联系起来，某些属性的取值的集合是无限的或近乎无限的，例如名称、地点、用途、立场、方法等属性。这些属性的取值通常是实体词语，因为属性词语本身表示某种关系，而关系就主要是存在于实体跟实体之间。例如，名称属性表示实体与符号之间的关系，用途属性表示实体与事件之间的关系，立场属性表示人的个体或群体与事件之间的关系。如何从文本中分析出这些属性的取值，是一个既重要又困难的问题。

从构词上看，有的形容词可同时提示多种属性。例如，“严明”表示在严格性属性上取值为“严”并且在明确性属性上取值为“明”，“温和”表示在具体或抽象温度上取值为“温”并且在适合度属性上取值为“和”。但是，并非每个词素都提示不同的属性。例如，“宽广”并非提示宽度和广度两个属性，而是提示具体或抽象的面积属性（“水域宽广”，“胸怀宽广”）。

从词义上看，一些形容词可提示不同的属性。例如，“大”和“小”可提示体积、规模、年龄等属性，“深”和“浅”可提示具体深度和抽象深度（“水很深”，“这本书很深”）。一般来说，复合形容词的词义比较确定，例如“深刻”、“深邃”、“肤浅”等只提示抽象深度。

属性分析的一个基础工作是在语义分析词典中列出每个表值词（量词、形容词等）所提示的属性以及该词所表示的属性值的性质，为计算机理解奠定基础。具体来说，需要说明

每个表值词提示的是哪些属性，如果并非单纯的属性，那么多义还是复合，多义的需要依据语境来消歧，复合的则可以同时并存。需要说明每个表值词所表示的是哪一种属性值，如果是度量值，应该指明同类度量值之间的换算关系。如果是对立值，应该确定哪一个是无标记的，然后就可以用统一的代码来表示对立值。例如，在说明了所对应的属性之后，“高”、“大”、“深”等所有的无标记值都可以编码为“A”，而“低”、“小”、“浅”等所有的有标记值则都可以编码为“非A”，词义分析的重点显然在属性词语，而对立值的表值词语无需多做解释。对于分级值，应该指明表值词在序列中的位置。至于枚举值，由单词充当的应该列入词典，由短语充当而无法列举的，研究重点在于动态识别。

### 三. 属性名词的界定

属性词语可以只由一个名词充当，称为属性名词。属性名词的界定，关键是如何把属性名词跟表示抽象实体的名词区分开来。

现代汉语有一种嵌套的主谓结构，其中的谓语部分本身也是主谓结构，形成“大主语+小主语+小谓语”格式。例如，“他作风严谨”、“老师态度温和”。从语义上看，大主语是实体词语，小谓语是表值词语。但小主语是不是属性词语（属性名词）呢？“作风”、“态度”等名词到底是表示主体为人的一种属性，还是表示跟人相关的一种抽象实体呢？这个问题还值得深入研究。

我们注意到，并非出现在这个位置上的名词都是属性名词。许多事物名词也可以出现在这个位置上，例如“他口渴”、“这间教室讲台很高”，其中“口”和“讲台”显然不是属性名词。

但我们必须承认，当大主语是实体词语，小谓语是表值词语时，这个位置的确是属性名词经常出现的位置之一。正因为如此，许多本不表示属性的词语在这个位置上也可以“提示”某种属性。“他口渴”，其中“口”是人体的一个部件，提示口的干渴程度这一属性。类似地，“这间教室讲台很高”，其中“讲台”是教室的一个部件，提示讲台高度这一属性。

一个实体常常具有多方面的属性，表现在语言中就是一个实体词语常常可以被不同类的形容词来陈述或修饰。例如，作为教室部件的“讲台”有高度、表面积、时新程度等属性，可以受“高”、“低”，“大”、“小”，“新”、“旧”等形容词陈述或修饰，属性就是这些形容词的分类依据。如果直接把“讲台”看成表示“教室”的一个属性的名词，会带来两方面的问题：第一，上述形容词没有了分类依据，但它们显然不是同一个属性的表值序列。第二，这些形容词还可以陈述或修饰别的实体名词（例如“房子”），无法把这些名词都看成属性名词，否则就必须说这些形容词每修饰一个实体名词就有一个不同的义项。虽然有些形容词的确是多义的，例如“大”和“小”，但之所以多义，并非由于它们能陈述或修饰不同的实体名词，而是由于它们能提示不同的属性（体积、规模、年龄等）。因此，我们应该确定一个原则：从表值词语（特别是形容词）出发来提取属性，一个表值序列或集合只提取一种属性。由此可以推论：如果一个名词可以跟多个表值序列或集合发生联系，就不是一个属性名词。

现在用这个原则来检验“作风”、“态度”这两个名词。大主语总是表示人的个体或群体的名词（从略），这两个名词作小主语，身份有待确定，小谓语通常是形容词：

作风：优良，严谨，廉洁，漂浮，浮躁，……

态度：积极，消极，主动，被动，热情，冷淡，认真，马虎，严肃，随便，……

上列每一行形容词都明显不属于同一个表值序列。例如，出现在“态度”之后的形容词可以分为积极性（积极、消极），主动性（主动、被动），感情流露程度（热情、冷淡），认真程度（认真、马虎），严肃程度（严肃、随便），果断程度（果断、犹豫），温和程度（温和、凶狠）等好几个表值序列，也就是说，“态度”其实是一个抽象实体，可以从不同的角度（属性）来加以描述。这种抽象实体犹如人（个体或群体）的一个部件，“他态度热情”跟“他口渴”、“他衣服漂亮”本质上并没有什么不同：大主语和小主语都是实体词语，两者之间是整体和部件关系，小谓语直接陈述小主语，其间的属性词语是隐含的。虽然“他态度热情”可以简化为“他热情”，“他衣服漂亮”不能简化为“他漂亮”，但前一种简化实际上也有信息损失，因为“他热情”可能是“他态度热情”（临场表现），也可能是“他为人热情”（秉性如此），区别仅在于损失大小而已，信息损失过大时，人们便认为不能简化。

把“作风”、“态度”等看做是人的一个抽象部件的名称，似乎有些费解。不过，根据认知语言学的理论，语言既是认知活动的工具，也是认知活动的产物，人类从认知自己的身体开始，通过隐喻等手段来认知其他抽象实体。汉语中许多表示抽象实体的词语，如“胸怀”、“心地”、“脑筋”本义都跟人体部件相关，甚至所用的表值词语也跟具体事物相同，例如“胸怀宽广”（比较“水面宽广”），“胸怀坦荡”（比较“田野坦荡”），虽然我们可以区分出具体面积和抽象面积，具体的可见性和抽象的可见性，等等。

属性虽然不是实体，但也是一种事物。因此，当属性名词出现在小主语位置上时，有可能是把属性作为一种抽象事物来加以陈述，其中隐含着另一个属性。例如，“两者差距明显”。差距作为一种属性，其表值词语是“大”、“小”等等，而“明显”、“不明显”等等是显示度属性的表值词语。又如，“那件衣服颜色鲜艳”，是把颜色作为一种抽象事物，来讨论其亮度和悦目程度等属性。

综上所述，“大主语+小主语+小谓语”格式可以表示以下几种属性关系：

第一，实体+属性+属性值，其中属性是属于实体的，例如“这间教室面积大”；或者把属性作为一种事物来陈述，属性值跟属性不直接相关，中间隐含另一属性，例如“那件衣服颜色鲜艳”。

第二，实体+部件+属性值，其中属性值是部件的某一属性的取值，例如“这间教室讲台很高”和“他作风严谨”。

#### 四. 属性词语的语义分析

从概念意义上看，基本属性有一维、二维、三维的空间属性，重量、湿度、温度、亮度、硬度、强度、力度、数量之类的物理属性，以及用途、品质、规模、范围之类抽象属性，等等。

以一维空间属性为例，“公里”、“米”、“毫米”是其度量值，汉语中有许多相关的属性名词，这些名词的词义中折射出人类对自己身体和周围空间的感知结果。“长度”和“宽度”分别表示二维平面中较长和较短的一维，这种二维平面通常与地面平行。“高度”通常是三维空间中垂直的一维，所谓垂直，就是垂直于地面，或者说跟人体的正常的直立姿势平行。

“厚度”是三维物体特别短的一维。“距离”是两点之间的长度，跟“长度”等属性名词不同的是，这两个点通常不在人们所认知的同一个实体之上，换言之，中间是“空”的。“高度”也有类似的差别：若谈论一个实体的高度，表值词语一般是“高”和“矮”；若谈论两点之间的垂直距离，表值词语一般是“高”和“低”。一些属性名词跟运动有关：“路程”是运行轨迹的长度；“深度”是从入口到底部的长度，具有这些属性的实体，如巷子、江河、思想等等，其入口（巷口、水面、词句等）是可见的，其底部则是隐蔽的，需要加以探测或探究。

一些属性是复合的。比例就是一种最常见的复合属性。例如，“浓度”是某种实体的数量跟所在液体的数量之比，“纯度”是符合需要的实体的数量跟所在整体的数量之比，“密度”是实体数量跟所在整体的长度（或面积、体积）之比，充盈度（“满”，“空”）是实体体积跟容器的容积之比，忙闲程度（“忙”，“闲”）可以理解为事件数量跟完成这些事件所需时间的长度之比。复合属性并非都是比例，例如，及时性（“早”，“晚”）可以理解为两个时点的差，“运量”是客货数量与路程的乘积，等等。

派生属性是对属性主体或相关事件加以限制而得到的。这种限制不一定明显反映在构造上，例如“年龄”是以年为单位的时间长度属性，属性主体一般为人的个体，相关事件是人的个体的存在。但许多属性词语在构造上有一定表现，例如“车速”是速度，其属性主体确定为车辆，“航速”也是速度，其相关事件确定为航行，也就间接地提示了属性主体为舰船或飞行器。又如，“含量”是某种实体在一整体中的数量属性，“科技含量”则将被包含的实体确定为“科技”。比例是同处于一个整体中的两种实体的数量之比，“市场占有率”则规定相关事件是“占有”，事件的相关客体是“市场”。从语义上看，表示派生属性的词语有四种基本的构造模式：

第一，主体+属性（车速，树龄，气压，体温，室温，气温，粮田面积）；

第二，事件+属性（产量，含量，占有率，周转量，载重量，婚龄，工龄，学龄，对抗性，射程，居住面积，使用面积）；

第三，客体+事件+属性（水稻产量，科技含量，市场占有率，货物周转量）；

第四，事件+属性值+属性（标注正确率，度量误差）。

在不引起歧义时，“标注正确率”可说成“正确率”，“度量误差”可说成“误差”，但其实都隐含着某一相关事件。“安全系数”虽然没有相应的复杂表达，但的确应该理解为“保证安全的系数”。

对于属性词语，语言信息处理的任务一是要能识别，二是要能理解。“这个县水稻产量很高”，其中“县”、“水稻”和“产量”是连续出现的三个名词，不进行属性分析，就没有把握将“水稻产量”识别为一个短语，也无法将这个句子跟另一个意思基本相同的句子（“这个县生产了很多水稻”）联系起来。

真实文本中的属性词语是一个开放集，难以一一枚举。另外，有些词语序列是不是属性词语，要看具体的上下文。因此，需要动态识别属性词语。例如，“高科技含量”在下面的句子中是属性词语：

要大力提高产品的高科技含量。

其中，“产品”是主体，“高科技含量”是其属性。“高科技含量”这个属性词语的构造模式是客体+事件+属性，意思是含高科技的数量，“高”是客体“科技”的某个属性的取

值。但是，“高科技含量”在另一个句子中不是属性词语：

加紧开发高科技含量、高附加值的产品。

其中，“产品”是主体，“科技含量”和“附加值”是其属性，这两个属性的取值都是“高”。这里“高科技含量”可变换为“科技含量高”，前一句的“高科技含量”不能这样变换。用语法术语来说，前一句的形容词“高”修饰名词“科技”，后一句的形容词“高”修饰名词短语“科技含量”。虽然同一个语言现象可以从不同角度（句法或语义）来解释，但是解释的深度和实用价值还是有所区别的。

属性词语的构造模式包含事件时，事件时态和模态是属性分析的重要内容。例如，“工龄”和“教龄”都是以年为单位的时间长度属性，分别与工作、教书事件相关，事件时态都是已然。“婚龄”和“学龄”则是以年为单位的时点属性，分别与结婚、入学事件相关，事件模态是允许或义务。<sup>①</sup>只有分析到这个深度，才能让计算机理解，“他有20年工龄和10年教龄”是什么意思，“调整婚龄”和“学龄儿童”是什么意思，才能让计算机从“他参加工作20年了”，“孩子有8岁了”这样的句子中自动提取工龄属性和属性值，判断“孩子”属于学龄儿童，等等。

## 五. 属性三角

陈小荷（1998）提出了一个属性分析的框架，其要点是建立属性与表值词语的联系，事物与属性的联系，但没有展开论述。

属性和属性值都是概念，其语言实现则分别是属性词语和表值词语。词语的概念意义的分析是语言信息处理的紧迫任务。但是，这种分析不能脱离现实的词语，不能随心所欲地构造一些概念（或语义分析特征），搞出一个百科词典式的概念层次体系。属性分析是词义分析的一种方法，它把属性作为联系实体和属性值的桥梁，通过对属性的系统化的归纳来建立实体词语的描述体系，从而实现词义分析从动词到名词的重点转移。属性三角是指属性（以及属性词语）、属性值（以及表值词语）和实体（以及实体词语）之间的关系，应首先分析表值词语，分辨歧义，得到属性值，接着由属性值归纳出属性，然后用属性和属性值作为基本概念来描述实体词语所反映的概念意义。

通过语料库统计来直接建立名词形容词之间的词义关系，在理论上是不可靠的，实践中也是行不通的。形容词作为一种表值词语，它与属性词语和实体词语的词义关系有本质差别。一些形容词是多义的，能提示多种不能共存的属性。例如，“大”和“小”所能陈述或修饰的名词很多，有的是属性名词，有的是实体名词，所提示的属性可能是面积、体积、规模、年龄、数量，等等。我们不能因为“年纪”和“纸”都能与“大”或“小”发生语法关系便认为这两个名词属于同一个词义类别，也不能根据同样的理由认为“纸”和“房子”有共同的词义特征。此外，语料库统计很难克服数据稀疏的缺陷，例如，在一个极大规模的语料库中也可能统计不到“罪犯”跟“温和”的同现，但根据实体概念的描述体系，“罪犯”属于人类，人类有抽象部件“神情”，“神情”有抽象温度属性，“温和”便是该属性的一个取值，因此可以推论出“罪犯”跟“温和”同现的可能性。在这种需要深度分析

<sup>①</sup> 《现代汉语词典》中，时长和时点意义的“龄”分别解释为“岁数”和“年限”。

的问题上，人类应比机器更能有所作为。

另一种做法是通过研究者的内省来直接建立名词的词义体系，例如构造概念分类体系，或者是用概念作为词义特征描述的基元，或者两者兼而有之。这种做法的危险在于可能脱离人类认知事物的实际，不能反映普通人关于名词的概念意义的常识。分类也好，特征描述也好，关键是要有依据，依据应来自语言事实，但又不是对语言事实的照相式的反映。属性分析正好提供了这样一种研究手段：分类或特征描述依据的是属性和在某些属性上已经确定的取值，属性有无则依据跟实体名词相关的表值词语的分析。从纯科学角度看，“绳子”、“纸”和“房子”表示的都是三维物体，但从语言事实来看，人类对这些实体的认知有不同的意象作为基础。绳子的主要属性是长度和粗细，粗细属性意味着它是一种长条形的实体，与柱子类似，长度属性则意味着它的典型存在方式是水平的，区别于有高度属性而无长度属性的柱子。纸的主要属性是面积和厚度，面积属性意味着它是一种平面式的实体，与土地类似，但土地不具有厚度属性。房子的主要属性是高度和面积（蕴含着容积），是一种立体式的内空的实体，高度属性暗示了它的正常姿态是直立，这是许多三维实体（包括人体）的共同属性，只有在这个基础上才能理解什么叫“倒下”。有高度属性的实体可以“倒下”，长条形实体可以“切断”，有厚度的平面式实体可以“撕开”，诸如此类的常识都可以通过属性分析而十分自然地推断出来。

属性分析可以在词义消歧、信息提取和机器翻译等应用领域发挥作用。词义消歧的前提是多义词义项设置合理，目前的语文词典在这一点上还做得不很完善。例如，《现代汉语词典》“新”下列七个义项，除最后一个表示姓氏的之外，其他六个义项连人都很难区分，更遑论机器。按照本文提出的观点，“新”是一个表值词语，关键在于确定它所联系的属性：一是事物个体的磨损度（磨损度小），二是事物个体或类别出现的时点（时点晚）。机器翻译方面，词对词的翻译是初级阶段采用的方式，否定这种方式，一是要做句法的转换，二是要对词语做属性分析。例如，直接在汉语和英语的形容词之间建立对应关系是不可取的，词语层次上的对应关系太复杂。属性是概念层次上的，应该确定每个形容词在各自语言里所对应的属性，以及属性所描写的事物，然后通过查表来进行翻译。例如，从英语的“*he has a big question*”和汉语的“他有一个重要问题”，可归纳出重要性属性，“question”和“问题”具有这个属性，该属性在汉语和英语中的表值词语分别是“重要”、“重大”和“big”、“important”等等，从而可以避免“他有一个大问题”之类的硬译。

## 参考文献

- [1] 陈小荷，一个面向工程的语义分析体系，《语言文字应用》，1998年第2期
- [2] 董振东，语义关系的表达和知识系统的建造，语言文字应用，1998年第3期
- [3] 李涓子、姬东鸿、黄昌宁，基于组合实例的双向优化聚类，《语言工程》，清华大学出版社，1997年
- [4] 袁毓林，一价名词的认知研究，中国语文，1994年第4期
- [5] 袁毓林，语言的认知研究和计算分析，语言文字应用，1996年第1期
- [6] 张敏，认知语言学与汉语名词短语，中国社会科学出版社，1998年