

# 知识表示方法比较<sup>①</sup>

刘建炜, 燕路峰

(中南大学湘雅医学院 医药信息系, 长沙 410013)

**摘要:** 阐述知识表示的概念, 介绍各种知识表示方法。通过讨论知识表示的执行层、逻辑层、表示应用层三个层面, 构建比较和评估各种知识表示方法在智能系统中的指标框架, 并且进行简单的比较。最后, 通过比较认为该框架可以指导我们选择合适的知识表示方法解决实际问题。

**关键词:** 知识表示; 知识组织; 比较研究

## Comparative Study of Knowledge Representation

LIU Jian-Wei, YAN Lu-Feng

(Central South University, Changsha 410013, China)

**Abstract:** Firstly, the paper describes the concept of knowledge representation and a variety of knowledge representation. Secondly, it discusses tree levels of knowledge representation: implementational level, logical level and epistemological level. For each level, it sets a framework for comparing and evaluating intelligent knowledge representation. Then commonly-used knowledge representation schemes are compared with this framework. It is believed that this framework can be efficiently used for comparison and selection of knowledge representation scheme in problem solving.

**Keywords:** knowledge; representation; knowledge organizing; comparative study

### 1 知识表示的概念

知识表示是知识工程的关键技术之一, 主要研究用什么样的方法将解决问题所需的知识存储在计算机中, 并便于计算机处理。

从一般意义上讲, 所谓知识表示是为描述世界所作的一组约定, 是知识的符号化、形式化或模型化。从计算机科学的角度来看, 知识表示是研究计算机表示知识的可行性、有效性的一般方法, 是把人类知识表示成机器能处理的数据结构和系统控制结构的策略。知识表示的研究既要考虑知识的表示与存储, 又要考虑知识的使用<sup>[1]</sup>。

### 2 常用知识表示方法介绍

#### 2.1 逻辑表示法

逻辑表示法以谓词形式来表示动作的主体、客体, 是一种叙述性知识表示方法。利用逻辑公式, 人们能描述对象、性质、状况和关系。它主要用于自动定理

的证明。逻辑表示法主要分为命题逻辑和谓词逻辑。逻辑表示研究的是假设与结论之间的蕴涵关系, 即用逻辑方法推理的规律。它可以看成自然语言的一种简化形式, 由于它精确、无二义性, 容易为计算机理解和操作, 同时又与自然语言相似<sup>[2]</sup>。

命题逻辑是数理逻辑的一种, 数理逻辑是用形式化语言(逻辑符号语言)进行精确(没有歧义)的描述, 用数学的方式进行研究。我们最熟悉的是数学中的设未知数表示。

例: 用命题逻辑表示下列知识: 如果  $a$  是偶数, 那么  $a^2$  是偶数。

解: 定义命题如下:  $P$ :  $a$  是偶数;  $Q$ :  $a^2$  是偶数, 则: 原知识表示为:  $P \rightarrow Q$

谓词逻辑相当于数学中的函数表示。例: 用谓词逻辑表示知识: 自然数都是大于等于零的整数

解: 定义谓词如下:  $N(x)$ :  $x$  是自然数;  $I(x)$ :  $x$  是整数;  $GZ(x)$ :  $x$  是大于等于零的数。所以原知识表示为:

<sup>①</sup> 收稿时间:2010-07-05;收到修改稿时间:2010-08-03

$(\forall x)(N(x)(GZ(x)\wedge I(x)), \forall(x)$ 是全称量词。

### 2.2 产生式表示法

产生式表示，又称规则表示，有的时候被称为 IF-THEN 表示，它表示一种条件-结果形式，是一种比较简单表示知识的方法。IF 后面部分描述了规则的先决条件，而 THEN 后面部分描述了规则的结论。规则表示方法主要用于描述知识和陈述各种过程知识之间的控制，及其相互作用的机制。

例：MYCIN 系统中有下列产生式知识(其中，置信度称为规则强度)：

IF 本生物的染色斑是革兰性阴性，本微生物的形状呈杆状，病人是中间宿主

THEN 该微生物是绿脓杆菌，置信度为 0.6

### 2.3 框架表示

框架(Frame)是把某一特殊事件或对象的所有知识储存在一起的一种复杂的数据结构<sup>[3]</sup>。其主体是固定的，表示某个固定的概念、对象或事件，其下层由一些槽(Slot)组成，表示主体每个方面的属性。框架是一种层次的数据结构，框架下层的槽可以看成一种子框架，子框架本身还可以进一步分层次为侧面。槽和侧面所具有的属性值分别称为槽值和侧面值。槽值可以是逻辑型或数字型的，具体的值可以是程序、条件、默认值或是一个子框架。相互关联的框架连接起来组成框架系统，或称框架网络。

例：用框架表示下述地震事件：[虚拟新华社 3 月 15 日电]昨日，在云南玉溪地区发生地震，造成财产损失约 10 万元，统计部门如果需要详细的损失数字可电询 62332931。另据专家认为震级不会超过 4 级，并认为地处无人区，不会造成人员伤亡。

解：如下表：

表 1 框架表示法举例

Frame (框架名) : 地震		
报导时间 (槽名): Value:3 月 15 日 (槽值)	地震发生时间: Value:3 月 14 日	地点: Value:云南玉溪
财产损失: Value:nul Default:约十万元 If needed: (侧面名) 电询 6233293 (侧面值) If added:核实	震级: Value:≤4 级	人员伤亡: Value:nul If needed:电询自然 灾害系统中心 If added:check

### 2.4 面向对象的表示方法

面向对象的知识表示方法是按照面向对象的程序设计原则组成一种混合知识表示形式，就是以对象为中心，把对象的属性、动态行为、领域知识和处理方法等有关知识封装在表达对象的结构中。在这种方法中，知识的基本单位就是对象，每一个对象是由一组属性、关系和方法的集合组成。一个对象的属性集和关系集的值描述了该对象所具有的知识；与该对象相关的方法集，操作在属性集和关系集上的值，表示该对象作用于知识上的知识处理方法，其中包括知识的获取方法、推理方法、消息传递方法以及知识的更新方法。

### 2.5 语义网表示法

语义网络是知识表示中最重要的方法之一，是一种表达能力强而且灵活的知识表示方法。它通过概念及其语义关系来表达知识的一种网络图。从图论的观点看，它是一个“带标识的有向图”。语义网络利用节点和带标记的边构成的有向图描述事件、概念、状况、动作及客体之间的关系。带标记的有向图能十分自然的描述客体之间的关系。

例：用语义网络表示下列知识：中南大学湘雅医学院是一所大学，位于长沙市，建立时间是 1914 年。

解：用语义网络表示表示如下图 1：

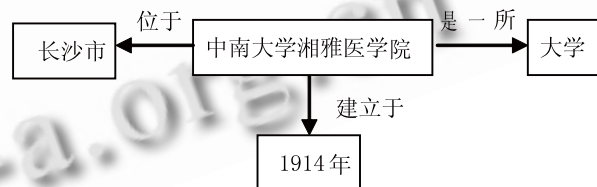


图 1 语义网表示法示例

### 2.6 基于 XML 的表示法

在 XML(eXtensible Markup language, 可扩展标记语言)中，数据对象使用元素描述，而数据对象的属性可以描述为元素的子元素或元素的属性。XML 文档由若干个元素构成，数据间的关系通过父元素与子元素的嵌套形式体现。在基于 XML 的知识表示过程中，采用 XML 的 DTD(Document Type definitions, 文档类型定义)来定义一个知识表示方法的语法系统。通过定制 XML 应用来解释实例化的知识表示文档。在知识利用过程中，通过维护数据字典和 XML 解析程序把

特定标签所标注的内容解析出来,以“标签”+“内容”的格式表示出具体的知识内容。知识表示是构建知识库的关键,知识表示方法选取得合适与否不仅关系到知识库中知识的有效存贮,而且也直接影响着系统的知识推理效率和对新知识的获取能力。

### 2.7 本体表示法

本体是一个形式化的、共享的、明确化的、概念化规范。本体论能够以一种显式、形式化的方式来表示语义,提高异构系统之间的互操作性,促进知识共享。因此,最近几年,本体论被广泛用于知识表示领域。用本体来表示知识的目的是统一应用领域的概念,并构建本体层级体系表示概念之间的语义关系,实现人类、计算机对知识的共享和重用。五个基本的建模元语是本体层级体系的基本组成部分,这些元语分别为:类、关系、函数、公理和实例。通常也把 Classes(类)写成 Concepts。

将本体引入知识库的知识建模,建立领域本体知识库,可以用概念对知识进行表示,同时揭示这些知识之间内在的关系。领域本体知识库中的知识,不仅通过纵向类属分类,而且通过本体的语义关联进行组织和关联,推理机再利用这些知识进行推理,从而提高检索的查全率和查准率<sup>[3]</sup>。

上面简要介绍分析了常见的知识表示方法,此外,还有适合特殊领域的一些知识表示方法,如:概念图、Petri、基于网格的知识表示方法、粗糙集、基于云理论的知识表示方法等,在此不做详细介绍。在实际应用过程中,一个智能系统往往包含了多种表示方法。

## 3 构建比较和评估的框架指标

知识表示方法在人工智能领域的应用最终体现为软件产品,如决策支持系统(Decision Support System, DSS)。DSS 三部件结构如图 2。根据此结构我们自上到下将该系统分为三个层面:表示应用层、逻辑层、执行层。根据软件的特性和知识表示方法的实际应用,再在每个层面针对性的构建一些可取的指标用来比较和评价各种知识表示的方法,最终得到知识表示方法评价的框架体系。表示应用层的指标主要参考的是各

方法的表示充分性方面的指标,逻辑层主要参照的是各方法的推理充分性和软件本身的操作方便性指标,执行层主要参照的是计算机可实现性和实现的高效性指标。

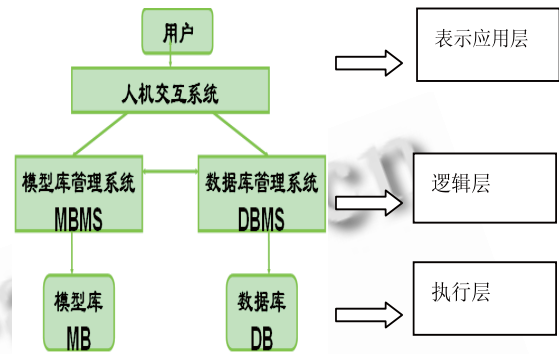


图 2 DSS 的三部件结构

### 3.1 表示应用层

表示应用层,是知识库的基础,是系统开发之前由领域业务专家制定出来的知识标准,包括知识、元知识等;是抽象的知识的集合;是解决实际问题的基础模型。如:要表示的知识的内容和范围,知识表示的单位(框架、产生式、类等),用户界面等,该层应该考虑以下几个方面:

- ① 自然性。即使用方便和理解方便,这些知识要求和自然语言表示相近,使用户能够迅速理解其中的概念。
- ② 可描述知识的类型。知识类型含:陈述型、规则型、控制型、元知识型等。
- ③ 可表示知识的范围。不同领域的知识表示方法要求不同,复杂的知识表示既要表示知识本身,还要表示知识和知识的关系:如时间、空间、部分的关系等。
- ④ 针对的用户。不同的用户,知识表示方法的选择是不一样的,面向专家或是知识工程师的系统,可能精确性要求高,面向普通用户的系统可能要求实用简单方便。
- ⑤ 知识表示单位。不同的表示方法采用不同表示单位。如,对象、本体、产生式、逻辑表达式等都是知识表示单位。

### 3.2 逻辑层

介于执行层和表示应用层之间的抽象层，把执行层中无意义的信息变为逻辑上有意义的信息，供使用者、维护者去操作和维护。其核心思想是：对表示应用层提供一个统一的数据模型，即提供各类知识和元知识结构及其对此的各类维护操作和柔性扩展；对于底层则屏蔽具体的数据库的关系模型，并利用映射原理建立透明的知识转化机制，并对底层数据库进行合理有效的管理。它主要涉及知识表示方法中的语义变化和操作；是知识表示方法实现的关键步骤。如元知识管理系统、规则库管理系统、模型库管理系统、数据库管理系统等属于该层。该层应注意这几方面：

① 模块性。当知识发生变化时，该知识及涉及的知识关系会随着发生改变，同时不影响其他知识及关系的变化。要求模块独立性强、共享性高、重用性好

② 操作和维护方便。当增加、删除、修改知识、元知识时操作方便、快捷。能够被用户和维护者快速的使用是知识表示方法的一个目的。

③ 推理能力。主要是知识推理方面的比较，要求推理的结果一致，支持多种方法的推理。

④ 整个框架易于用户理解。

### 3.3 执行层

执行层，不是用户直接使用的层面，而是由开发者设计的，涉及信息处理的执行效率。知识表示形式的理论变为应用的关键就是通过计算机程序实现，该层面应该主要考虑表示方法在计算机内部的算法及其优化。因此，对物理数据库、算法、计算机硬件等都有要求。好的知识表示方法系统该层应该注意这几方面：

① 空间高效性。物理空间高效性：数据库的使用是否高效、稳定、冗余少；这个由于不同类型的数据库会造成一些影响。逻辑空间高效性：一个被执行的概念所包含的属性、方法是否足够强大且计算机方便实现；这同时涉及表示应用层对知识体系的定义和控制。

② 算法是否可以优化。也可以称为时间高效性，即查找、推理、控制的过程的算法是否高效。

③ 是否容易通过编码实现。这是实现的本质要求。不能通过代码实现的知识表示方法在系统中应用的可能性大大减小。

## 4 结果和意义

### 4.1 各种知识表示方法的比较

下面将应用前文讨论的框架分析几种常用知识表示方法，评判的等级是参考相关的文献，结合第二部分对知识表示方法的讨论，再加上自己的分析判断的出的。结果见表2。

表2 知识表示方法比较较框架表

	逻辑产生式	语义网	框架	面向对象	XML	本体
<b>表示应用层</b>						
自然性	很好	很好	很好	很好	很好	很好
知识表示单位	逻辑公式	规则	网络图	框架	对象	对象
可描述的类型	陈述型	规则型、控制型	陈述型	陈述型、规则型、控制型	兼有	兼有
可描述的范围	好	好	好	粗糙	好	好
针对的用户	初学者	初学者	初学者	专家	初学者	专家
<b>逻辑层</b>						
模块性	好	好	差	一般	很好	很好
操作维护方便	较差	差	一般	差	很好	很好
推理能力	一般	好	一般	差	好	一般
整个逻辑框架易于理解	需要简要说明	易于理解	易于理解	易于理解	易于理解	相对难理解
<b>执行层</b>						
编码容易实现	容易	容易	容易	容易	容易	较复杂
空间高效性	涉及数据库的选择和设计，也涉及知识、元知识在实际过程中的应用情况，需根据具体案例分析。					
算法可优化	根据具体应用的要求选择最合适的算法和处理方式。					

通过比较，传统的知识表示方法已经很难适应大量信息的处理和表示，新的知识表示方法即将得到更多的应用，如：基于XML的知识表示方法将应用于万维网信息处理、知识管理系统、数据交换、数据和

信息管理等系统中;基于本体的知识表示方法将广泛应用于知识库系统、决策支持系统、人工智能系统等系统中。

#### 4.2 构建指标框架的意义

通过构建指标框架,明确各种知识表示方法的优缺点;选择一种或多种知识表示方法更加有效的解决实际问题;帮助构建和评估知识表示方法比较的规范;在旧的知识表示方法的基础上创造新的更符合框架指标的知识表示方法。

#### 4.3 知识表示方法的发展趋势

随着计算机技术的发展,智能系统理论不断进步,将会涌现出更多的适合特定领域的知识表示方法,如基于本体的表示方法,基于云理论的知识表示等。随着知识表示方法不断被普通用户接受和使用,满足用户的个性化需求将是知识表示方法发展的客观要求,即更加灵活的把抽象的概念更形象、更快捷、更方便的通过计算机展现出来。此外,知识表示、知识获取、知识应用三个方面将互相进步、互相促进发展。

#### 4.4 建议和意见

在实际解决问题的过程中,智能系统中往往不是单一的采用一种知识表示方法,而是选择多种适合特定领域知识的表示方法,互相弥补不足,发挥各自优势。即使针对同一类的问题,由于不同的目的、不同的应用,采用的表示方法也不一样。所以我们在选择表示方法之前,必须通过严密的需求分析,选择合适的一种或多种知识表示方法,来满足求解和描述问题。

智能系统中,关于各种知识方法比较和评价标准的研究还比较少,多做这方面的理论或实际研究,将有助于知识表示方法的发展和创造出更合适的知识表示法。

#### 参考文献

- 1 蒋云良.知识表示综述.湖州师专学报,1995,(5):18—22.
- 2 徐宝祥,叶培华.知识表示的方法研究.情报科学,2007,25(3):690—694.
- 3 张攀,王波.专家系统中多种知识表示方法的集成应用.微型电脑应用,2004,20(6):4—6.
- 4 谢和平,等.分形几何.重庆:重庆大学出版社,1992.
- 5 赵莹,高隽,陈果,冯文刚.一种基于分形理论的多尺度纹理特征提取方法.仪器仪表学报,2008,4(29):177—181.
- 6 Sarkar N, Chaudhuri BB. An Efficient Approach to Estimate Fractal Dimension of Textural Images. Pattern Recognition, 1992,25(9):1035—1041.
- 7 于子凡,杜贵君,林宗坚.图像盒子维数特征计算方法改进.测绘科学,2006,1(31):87—89.
- 8 张涛,孙林,黄爱民.图像分形维数的差分盒方法的改进研究,2007,5(13):55—57.
- 9 梁东方,李玉梁,江春波.测量分维的“数盒子”算法研究.中国图象图形学报,2002,5(3):246—250.
- 10 韩杰,陆桂华.测量分维的矢量计盒算法研究.中国图象图形学报,2008,3(13):525—530.

(上接第241页)

## 5 小结

总的来说,分形布朗运动模型法精确性好,可测分形维数适用范围宽,但地毯覆盖法计算量大,尤其用8邻域来覆盖,运算时间过长。盒子维和多尺度分数维法计算时间适中,但当图像的分数维很高时会低估图像的FD,因而有一定的使用限制。结合适用范围和计算量及准确度三方面来考虑,布朗运动模型和盒子维法较好。

#### 参考文献

- 1 Pentland P. Fractal-based description of natural scenes. IEEE Transactions on Pattern analysis and Machine Intelligence, 1984,6(6):661—674.
- 2 杨光俊.分形的数学(上,下).昆明:云南大学出版社,2002.76—189.272—301.
- 3 Mandelbrot BB, Van Ness J. Fractional Brownian motion, fractional noise and applications. SIAM Review, vol. 10, 1968.