

知识表示方法研究与应用*

年志刚¹, 梁式¹, 麻芳兰¹, 李尚平²

(1. 广西大学机械工程学院, 广西南宁 530004; 2. 广西工学院, 广西柳州 545006)

摘要: 详细阐述了各种知识表示方法的运用形式, 并分析了各自的优点和局限。针对甘蔗收获机械智能设计系统的知识特点, 提出了混合知识表示和神经网络知识表示两种运用于该系统的切实可行的知识表示方法。

关键词: 知识表示; 专家系统; 甘蔗收获机; 智能设计

中图分类号: TP391 文献标志码: A 文章编号: 1001-3695(2007)05-0234-03

Study and Application of Knowledge Expression

NIAN Zhi-gang¹, LIANG Shi¹, MA Fang-lan¹, LI Shang-ping²

(1. College of Mechanical & Engineering, Guangxi University, Nanning Guangxi 530004, China; 2. Guangxi University of Technology, Liuzhou Guangxi 545006, China)

Abstract: The article expatiated application of knowledge expression method of sorts and analyzed advantage and shortage. Finally, it raised mix and nerve network knowledge expression methods which are effectively used in intelligence design system of sugarcane harvester.

Key words: knowledge expression; expert system; sugarcane harvester; intelligence design

智能活动主要是获得知识并运用知识的过程。因此, 知识性是人工智能专家系统的主要特征之一。而造就它的关键技术在于知识的表示、获取和应用。知识的表示方法是至关重要的, 它不仅决定了知识应用的形式, 而且也决定了知识处理的效率和实现的域空间规模的大小, 其成功与否直接关系到智能设计专家系统的水平^[1]。然而, 在不同领域, 专家经验知识又具有各自不同的特点。因此, 知识表示方法的研究历来是建立专家系统首要解决的问题。本文对目前常用的各种知识表示方法的特点进行了研究和比较。结合国家自然科学基金项目——“基于知识的甘蔗收获机械智能设计系统关键技术的研究”, 在充分分析了我国甘蔗收获机械的研发、设计的知识特点基础上, 提出了几种在甘蔗收获机械智能设计领域中切实可行的知识表示形式。同时, 也为其他领域智能设计专家系统的研发提供了有效的参考依据。

1 传统的知识表示方法

智能设计专家系统的工作过程是一个获得并应用知识的过程, 恰当的知识表示对于专家系统的构建具有重要意义。对于专家系统而言, 一个好的知识表示方法应具备以下性质^[2]:

(1) 表达充分性。具备确切表达有关领域中各种知识的能力。

(2) 推理有效性。能够与高效率的推理机制密切结合, 支持系统的控制策略。

(3) 操作维护性。便于实现模块化, 并检测出矛盾的及冗余的知识; 便于知识更新和知识库的维护。

(4) 理解透明性。知识表示便于人类理解, 易读、易懂, 便于知识的获取。

基于以上知识表示方法具备的性质, 目前普遍应用的传统知识表示方法主要有一阶谓词逻辑表示法、产生式表示法、框架表示法、脚本表示法、语义网络表示法等。

1.1 一阶谓词逻辑表示法

谓词逻辑是一种形式语言, 也是目前能够表达人类思维活动的一种最精确的语言, 它与人类的自然语言比较接近, 因此一阶谓词逻辑表示方法也是最早使用的一种知识表示方法。它具有简单、自然、精确、灵活、容易实现等特点。谓词的一般形式为 $P(x_1, x_2, \dots, x_n)$ 。其中, P 是谓词; x_1, x_2, \dots, x_n 是个体(常量、变元或函数)。谓词逻辑适用于表示事物的状态、属性、概念等事实性的知识, 也可以用来表示事物间确定的因果关系, 即规则。

例如: 对于“小李的父亲是教师”, 可以表示为 $\text{Teacher}(\text{father}(\text{Li}))$; 对于“公式 $x < 7$ ”, 可以表示为 $\text{Less}(x, 7)$ 。

一阶谓词逻辑表示法的局限性在于它难以表达不确定性知识和启发性知识, 推理方法在事实较多时易于出现组合爆炸, 且推理过程冗长、效率低。

1.2 产生式表示法

产生式表示法又称为产生式规则表示法。“产生式”这一术语是由美国数学家波斯特(E. Post)在1943年首先提出来的。它具有自然、灵活、清楚、模块性好、通用性强等优点, 目前已成为人工智能中应用最多的一种知识表示模式。产生式的基本形式为^[3]: $\text{if}(\text{前提}1) \ \&(\text{前提}2) \ \&\dots \ \text{then}(\text{结论}1) \ \&$

收稿日期: 2006-03-02; 修返日期: 2006-04-30 基金项目: 国家自然科学基金资助项目(50365001)

作者简介: 年志刚(1977-), 男, 黑龙江人, 硕士研究生, 主要研究方向为机械 CAD、CAE、CAM 智能设计(Nzg_vs@163.com); 梁式(1948-), 男, 广东人, 教授, 硕士, 主要研究方向为机械 CAD、CAE、CAM; 麻芳兰(1976-), 女, 广西人, 讲师, 博士研究生, 主要研究方向为智能设计; 李尚平(1956-), 男, 广西人, 教授, 博士, 主要研究方向为机械 CAD、CAE、CAM。

(结论 2) &...。其中,前提亦称前件、条件;结论也可以为要执行的操作。整个产生式的含义是:如果前提被满足,则可推出结论或执行所规定的操作。例如:if(动物为胎生 & 哺乳喂养) then (该动物为哺乳动物)。

产生式表示法的局限在于它不能表达具有结构性的知识;且由于求解过程是一个匹配—冲突消解—执行反复进行的过程,工作效率不高;另外,在求解复杂问题时容易引起组合爆炸。

1.3 框架表示法

框架(Frame)理论是在 1975 年由美国著名人工智能学者明斯基(Minsky)首先提出的,它是描述对象属性的一种数据结构。在框架表示法中,框架被看成是知识表示的基本单元。不同的框架之间可以通过属性之间的关系建立联系,从而构成一个框架网络,充分表达相关对象间的各种关系。它的突出特点是善于表示结构性知识;具有良好的继承性,不仅减少框架网络表示知识的冗余,而且较好地保证了知识的一致性。

框架的一般形式^[4]如下:

```

框架名: 框架名的值
槽名 1: { 侧面名 11: 侧面名 11 值
          约束: 约束条件
侧面名 12: 侧面名 12 值
          约束: 约束条件
...
侧面名 1n: 侧面名 1n 值
          约束: 约束条件 }
{ 槽名 2: { 侧面名 21: 侧面名 21 值
          约束: 约束条件
侧面名 22: 侧面名 22 值
          约束: 约束条件
...
侧面名 2n: 侧面名 2n 值
          约束: 约束条件 }
...
槽名 n: { ... }

```

一个框架由若干个被称为槽的结构组成;每一个槽又可根据实际需要分为若干个侧面。一个槽用于描述对象某一方面的属性;一个侧面用于描述相应属性的一个方面,每一个方面又可以给出具有的约束条件。

例如:一个简单描述学生基本情况的框架表示法:

```

框架名: 学生
  姓名: 单位(姓,名)
  年龄: 单位(岁)
  性别: 范围(男,女)
         缺省条件: 男
  健康状况: 范围(健康,一般,差)
         缺省条件: 一般

```

框架表示法的主要不足之处是不善于表示过程性的知识,通常要与其他表示方法相结合使用。

1.4 脚本表示法

脚本(Script)表示法是由 Roger C. Schank 在文献[5]中提出的。其结构类似于框架,用于描述固定的事件序列,如去医院看病、到餐厅就餐、到影院看电影等。它与框架相比,区别在于:框架是一种通用的结构;而脚本形式比框架形式应用范围窄,对某些专门知识(如理解故事情节等)更为有效。

1.5 语义网络表示法

语义网络是奎廉(J. R. Quillian)于 1986 年最先提出的。它是一种通过概念及其语义关系来表示知识的一种网络图。

一个语义网络就是一个带有标志的有向图。其中,有向图的节点表示各种事物、概念、属性、动作、状态等;有向弧表示它所连接的节点间的某种语义联系,每个节点可以带有若干属性。它具有灵活、自然、易于实现、善于表示结构性知识等优点。一个最简单的语义网络形式是一个三元组:(节点 1, 弧, 节点 2)。图 1 即是它的一个有向图表示,通常称为一个基本网元。

例如:猎狗与狗的语义网络表示如图 2 所示。



图 1 语义网络基本网元

图 2 语义网络基本网元示例

语义网络表示法的局限性主要表现在它对知识表示的非严格性,不能保证不存在二义性;以及由于其自身的灵活性和非严格性所带来的知识处理的复杂性。

2 知识表示方法的改进及发展

传统的知识表示方法在运用上尽管比较成功,但每种表示法均存在着一定的局限性,不能完全满足知识所具备的四项性质。随着现代科学的不断发展,专家系统的应用已经遍布各个科学领域。知识的类型多种多样,包括分类知识、事实知识、关系知识、统计知识、判断知识、经验知识、模糊知识和控制知识等。因此,知识表示方法必须进行改进和发展,才能适应新科技的发展。

2.1 传统知识表示方法的改进

(1) 产生式表示法的改进

从产生式表示法的基本形式上分析,可以将产生式变换成推理规则+计算规则^[6]。同时,为了解决模糊知识的处理,还可以在计算规则中引入可信度因子 r 。

推理规则: if (前提 1, 前提 2, ..., 前提 n) then 结论

计算规则: 结论- \rightarrow 值 = FUNC[(前提 1- \rightarrow 值, 前提 2- \rightarrow 值, ..., 前提 n - \rightarrow 值), r] ($0 < r < 1$)

这种增强型的产生式系统,不但具有普通系统的优点,而且每条推理规则都对应一条计算规则,函数 FUNC 与推理规则一起变化。当可信度 $r = 1$ 时,即可实现灵活的模糊推理。同时,还可以引入规则树理论。规则树是产生式规则表示法中引入决策树技术的一种知识表示结构。这样就可以将前提和结论用一个二维表进行表示,便于计算机的存取,从而实现强大的计算和推理能力。

(2) 混合知识表示方法

在知识表示的应用过程中,人们逐渐发现很难找到一种单一的表示方法能有效地解决领域内的所有知识表示。因此,混合知识表示方法将以往成熟的传统知识表示方法有效结合起来进行运用,成为了许多专家探讨的课题。目前,比较成熟的混合知识表示方法有^[1]:谓词逻辑、产生式规则和过程式的结合;产生式、框架和过程式的结合;语义网络与框架的结合;语义网络与产生式规则的结合等。这种方法有效地克服了单一表示方法的局限性,并可充分发挥各种方法的长处。

2.2 面向对象的知识表示方法

近年来随着面向对象技术的深入发展,面向对象的知识表示方法逐渐广泛应用于专家系统中,具有封装性、模块性、继承性、易于维护和可扩展性好等优点。广义上,对象可以理解为客户世界的任何事物。按照面向对象方法学的观点,一个对象

的形式可定义为四元组 对象 $:: = (ID, DS, MS, MI)$ 。其中 ID 代表对象标志符、DS 代表数据结构、MS 代表方法集合、MI 代表消息接口。这种方法可以将知识抽象为对象的内部状态和静态特征属性进行封装和隐藏,而知识的处理方法表示为对内部状态和特征属性的操作,并由消息接口与外界发生联系。它既可以实现灵活的推理机制,又使知识的修改操作局限在对象的内部,容易实现知识库的一致性和完备性。此外,由于对象属性的独有,就可以把整个系统抽象为类,每一个对象即为子类。其不但实现了多学科专业知识的层次划分,还有效地降低了系统的开发难度。

2.3 模糊技术的知识表示方法

在各种实际领域中,严格精确与确定的知识并不多见,大量知识带有一定的模糊性和不确定性。可以说,不确定性是智能问题的本质特征。因此,智能系统能力的提升更能体现在模糊理论的研究上。目前,常用的应用于知识表示的模糊理论主要有:基于概率论经典理论的知识表示,但它要求直接给出知识的概率,因此其应用受到限制;基于可信度的不确定理论是肖特里菲(E. H. Shortliffe)等人在确定性理论上提出的,因其直观、可信度的计算比较简便,而得到广泛应用;文献[7]中所阐述的云模型理论已得到普遍应用,云模型是用语言值表示的定性概念与其定量表示之间的不确定性转换模型;基于扎德(Zadeh)提出的模糊集合论、模糊逻辑和可能性推理的模糊元图、模糊关系矩阵等都取得了很好的应用。

2.4 神经网络知识表示方法

神经网络是反映人脑结构及其功能的一种抽象数学模型。它是由大量神经元节点互连而成的复杂网络,用以模拟人类进行知识表示、储存和推理的行为。具有代表性的模型主要有感知器、多层影射 BP 网络、GMDH 网络、RBF 网络、Hopfield 反馈神经网络以及双向联想记忆等。目前,神经网络与模糊理论的结合发展很快。可以预见,随着技术的深入发展,神经网络与模糊理论融为一体,必将推动人工智能研究与应用进入新的发展时期。

3 知识表示方法在甘蔗机械智能设计系统中的应用

3.1 甘蔗收获机械知识特点

甘蔗收获机械属于复杂的机械产品,除具有机械设计知识的一般特点外,还具有自身的特点。主要包括以下几个方面:

(1) 设计知识的模糊性。它包括:系统应用环境状态知识描述的模糊性,如甘蔗地的坡度、土壤的硬度等;系统状态知识描述的模糊性,如砍蔗刀盘的厚薄、砍蔗效果的好坏等;设计数据和参数知识的模糊性,如关键部件材料的应力要求、砍蔗的破头率要求、动力的分配等。

(2) 设计知识的层次性。主要体现在整机与部件、部件与零件以及相似零部件的关系上。

(3) 设计知识的耦合性。机械零部件的装配关系、动力分配、运动协调等之间的联系,表现为设计变量之间的约束关系。

(4) 创新性知识的复杂性。这部分知识主要来源于基于不同应用环境所进行的虚拟样机设计得到的知识。由于仿真约束前提的多样性,导致得到知识的差异性和复杂性。

3.2 基于面向对象的混合知识表示方法运用

要充分表达甘蔗收获机械知识的特点,并排除创新性知识的差异性,满足智能设计系统的能力要求。采用单一的知识表达方法很难实现,可以利用面向对象语言的特点,采用基于面向对象的混合知识表示方法。它将知识抽象为对象类和一个个子类,并在类与子类、子类之间利用产生式、框架、语意网络和过程体等来表达不同特征的知识,混合知识的表达模型如图 3 所示。

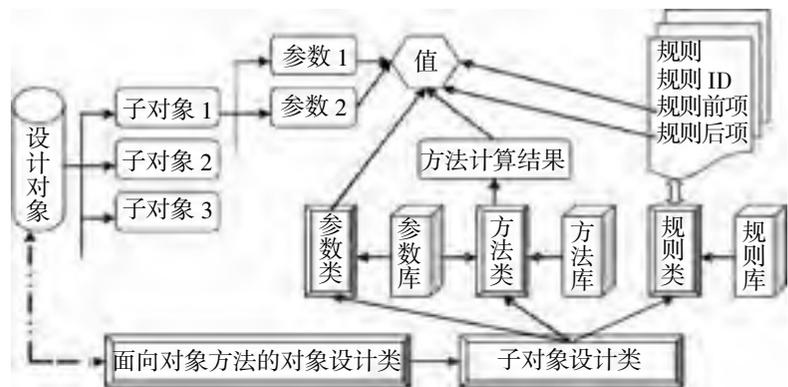


图 3 混合知识表达模型

通过图 3 的知识表达模型,将设计对象层次细分为各个子对象,抽取各个子对象的设计参数进行设计。参数值可以是直接提取参数库的值,或是经方法库内的某种方法计算的结果,或是规则库内某个规则的前项或经推理机运作得出的规则后项。由于面向对象的方法将对象类、参数类、规则类等分别进行了封装和隐藏并且各封装类还可以根据知识的特点进行子类的划分和封装。不同的子类采用不同的知识表示方法,如结构性的知识采用框架或是语意网络、规则性知识采用产生式表示法、模糊性知识采用增强的产生式表示或是基于模糊理论的知识表示等。通过这种形式可以将各种知识表示方法混合运用,充分发挥了各自的优点,既提高了智能系统的工作效率,又方便了知识的管理和系统的维护。

3.3 神经网络知识表示方法运用

神经网络具有优良的自组织、自学习和自适应能力等优点。将神经网络嵌入到智能设计系统内可以很好地解决智能设计评价系统的决策和模糊推理的困难,同时又可以利用智能系统良好的解释机能来弥补神经网络中过程表达的缺陷。

由于神经网络处理数字信息,而工程领域的知识形式丰富,有描述性知识、图表知识、结构性知识等,神经网络应用的困难主要表现为如何将工程领域的知识转换成数字模式。针对这一要求,需要引入模式转换模块,完成逻辑信息向数字信息的转换。图 4 是一种简单的神经网络处理模型。



图 4 神经网络处理模型

由于神经元处理信息的节点值取值域一般为 $[0, 1]$ 或 $[-1, 1]$,输入转换模块必须同时将信息进行归一化处理:对于数值变化方式可采用线性插值;其他模式信息可以进行定量转换,具体量化标准根据专家的经验取定。对于输出转换模块则采用相应的逆变换。图 5 为图 4 模型在甘蔗收获机砍蔗机构中知识处理的具体应用实例。(下转第 286 页)

采用基于 Socket 的 KQML (Knowledge Query and Manipulation Language) 作为主体之间的通信语言。自定义通信语言中的行为源语作为消息头,具体的通信内容作为消息体,利用本体来理解消息的语义,以便采取正确的动作处理消息。若需要增加额外的服务设施,通过集中式服务来解决主体通信中的定位问题。集中式服务主要由系统管理主体负责。主体创建时,会得到一个全局唯一的标志符,并向它注册其名字、能力和位置等信息;当主体之间需要合作时,可把消息发送给系统管理主体,由它充当消息缓冲区,找到相应的主体并把消息传递过去或由系统管理主体提供另一方主体的当前地址,双方主体直接通信^[6]。

如果 MAS 系统只部署在一台机器上,上述方法是很有用的;如果需要部署在多台机器上,可以再设计一类本地协调主体运行在每台机器上,各台机器上的主体都向本地协调主体注册。主服务器上运行全局唯一的系统管理主体。当主体之间需要合作时,先查找本地协调主体,看是否有中意的合作伙伴,若有则在本地通信,没有则本地协调主体通过 Socket 通信向系统管理主体查询。由于本地协调主体承担了本地主体的管理,减轻了全局管理的压力,而且本地主体的通信不必查询系统管理主体,加快了本地主体通信的速度。

(2) 任务协调

主体内部的任务触发机制是源于一个目标 (Goal) 的生成。一个 Goal 生成之后,主体自动检查自身是否能够完成该任务以及是否占有完成该任务所需的全部资源。若没有,则把所缺少的资源以子目标 (SubGoal) 的形式向相关主体发出询问,寻求合作。模型中任意两个主体 A 和 B 之间的交互行为执行算法为: A 向 B 发送请求服务的消息; B 收到之后向 A 发送 Accept 消息,并在自己的任务集中寻找与之相吻合的操作;如果没有相吻合的计划,则向 A 发送一个 Reject 消息,否则执行其操作。完成之后,用 Result 操作将结果信息发送给 A。因此在确定 Agent 内部的任务处理之后,要描述整合以本体为信息流的任务之间的相互关系,明确子目标,从而反映出代理之间的任务协调关系。图6是监视主体接到一个MT的连接

请求时所做的任务协调过程。

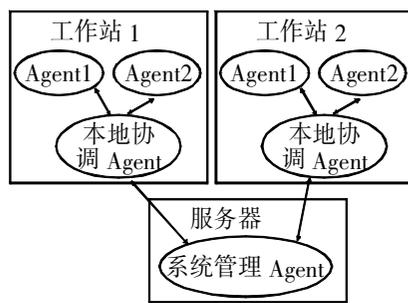


图5 Agent 基于消息传递的通信机制

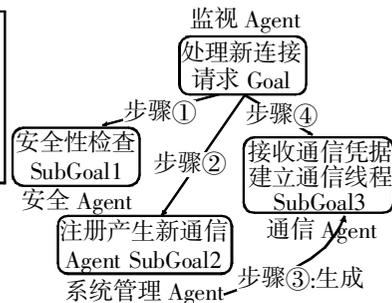


图6 新通信连接的任务协调过程

3 结束语

本文提出并在多智能主体的通信服务器上开发了一整套面向物流服务的车辆监控与指挥调度系统。在网 GPS 车辆有 600 多台, GPS 型号有四种, 经过五矿货运广东公司的反复测试, 与其原有的车辆监控系统对比, 丢包率由原来的 15% ~ 20% 减为 3% ~ 5%, 处理一条消息的平均延时由原来的 386 ms 缩短为平均 117 ms。特别是对报警类紧急信息, 系统响应时间大幅度缩短, 取得了良好的运行效果。

参考文献:

[1] 原仓周, 柳重堪, 张其善. 大规模车辆监控通信服务器的设计与实现 [J]. 北京航空航天大学学报, 2004, 30 (3) : 232-235.
 [2] 陈斌, 李德华, 姚迅. 一种基于 GPRS 技术的可扩展车辆监控系统的设计与实现 [J]. 计算机应用研究, 2005, 22 (6) : 175-178.
 [3] 原仓周, 柳重堪, 张其善. 大规模车辆监控 SMS/GPRS 通信服务器参数分析 [J]. 小型微型计算机系统, 2005, 26 (5) : 775-778.
 [4] 刘小明, 王飞跃. 基于 Agent 的单路口交通流控制的研究 [J]. 系统仿真学报, 2004, 16 (4) : 853-855.
 [5] WOOLDRIDGE M. 多 Agent 系统引论 [M]. 石纯一, 等译. 北京: 电子工业出版社, 2003.
 [6] 郭中, 王惠芳, 黄永忠, 等. 软件 Agent 的通信模型 [J]. 计算机工程与设计, 2002, 23 (11) : 9-11.

(上接第 236 页)

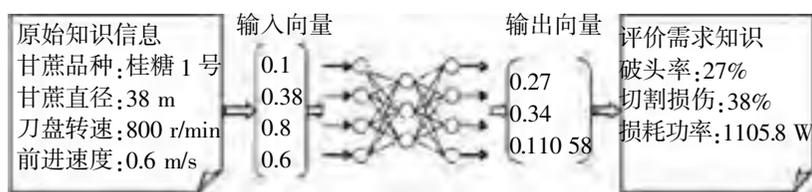


图5 神经网络知识表示方法应用

4 结束语

知识表示是人工智能研究领域的重要问题之一,直接影响着智能系统的深入发展和应用,一直是智能领域的研究热点。在甘蔗收获机械智能设计系统的开发过程中,运用基于对象的混合知识表达方法和神经网络知识表达方法对知识进行处理。经实践证明该方法是可行的、有效的。

参考文献:

[1] 罗燕琪, 陈雷霆. 专家系统中知识表示方法研究 [J]. 电子计算机, 2001 (4) : 28-31.
 [2] 张攀, 王波, 卿晓霞. 专家系统中多种知识表示方法的集中应用

[J]. 微型电脑应用, 2004, 20 (6) : 4-5.
 [3] 李太福, 黄茂林, 赵明富. 旋转机械故障诊断的框架知识表示 [J]. 辽宁工程技术大学学报, 2002, 21 (5) : 633-636.
 [4] 冯豪, 何玉林, 麻芳兰, 等. 面向对象知识表示方法在摩托车设计中的应用 [J]. 重庆大学学报: 自然科学版, 2005, 28 (1) : 1-4.
 [5] RENATO D M. Computer models of speech using fuzzy algorithms [M]. New York: Plenum Press, 1983.
 [6] 谭建龙. 基于数据库的一种推理算法 [J]. 电脑与信息技术, 1999, 7 (6) : 16-18, 51.
 [7] 刘常昱, 李德毅, 潘莉莉. 基于云模型的不确定性知识表示 [J]. 计算机工程与应用, 2004, 40 (2) : 32-35.
 [8] 马小娟. 甘蔗收获机械智能设计知识表达及知识库系统开发 [J]. 中国农机化, 2006 (1) : 50-53.
 [9] 尹朝庆, 尹皓. 人工智能与专家系统 [M]. 北京: 中国水利水电出版社, 2002: 13-55.
 [10] 乔桂玲. 甘蔗收获机械智能设计中数据库及管理系统的研究 [D]. 广西: 广西大学, 2005.
 [11] 李士勇. 模糊控制、神经控制和智能控制论 [M]. 哈尔滨: 哈尔滨工业大学出版社, 1996.