

利用专家系统开发工具设计电子信息系统

白素怀 (西安电子科技大学)

张焱 (中国银行深圳分行)

摘要:本文介绍一种通用型专家系统开发工具 HEARSAY-III 的结构及特点,并阐述了如何使用 HEARSAY-III 工具开发一个“电子信息分析”专家系统。文章给出了“电子信息分析”专家系统的知识表示办法及不精确推理算法及模块组成。

一、引言

现代科学技术发展方向一是智能化,90年代是人工智能(Artificial Intelligence-AI)的时代,21世纪将是智能时代。

AI是对人类智能的模拟、延伸和扩展。这种智能包括:思维、感知、行为等方面的智能。AI研究的目的是如何通过模拟人的智能来提高现有机器智能水平以及如何设计和制造出具有更高智能水平的机器。

知识工程(Knowledge Engineering-KE)是 AI 技术的应用工程,是以知识表示、知识推理、知识获取、知识存储、知识管理及知识利用等技术为基础的各种 AI 应用工程。

专家系统(Expert System-ES)是典型的知识工程系统,是 AI 技术领域发展中发展迅速、应用广泛,目前最活跃的一个重要分支。ES 应用 AI 和 KE 的理论、技术、方法和工具来模拟、延伸和扩展在某一专业领域内专家的智能,用以指导该领域工作实践。

ES 已成为在 AI 领域中应用最广的一个分支,并在实际应用中取得了很大的成功。ES 的开发和应用是随着 ES 开发工具的发展而得以深入进行,ES 开发工具在 ES 开发过程中起着加助剂的作用。

所谓 ES 开发工具是指一级程序系统,它能简化建造 ES 的工作步骤。建造和开发一个 ES 是件复杂而艰巨的过程,一般需几年或更长时间才能完成。随着 ES 应用领域的不断扩大,知识工程师不断地寻求建造 ES 的新途径。70 年代末,知识工程师开始着手 ES 开发工具的研究工作,各种建造 ES 的开发工具随之相继问世。这样可大大缩短了建造 ES 的周期,减少了建造 ES 的资金,促进了 ES 向实用化发展的进程。

到目前为止世界上已推出的 ES 开发工具已有近百种,其中已商品化的有三十多种。典型的骨架系统有:EMYCIN 系统、RAS 系统、EXPERT 系统等。通用型系统有:HEARSAY-III、M.I、ROSIE、KES 等系统。组合型系统有:AGE、ADVISE 等系统。我国一些院校及科研单位在引进、消化的基础上,汉化了一些专家系统工具的同时,也研制出自己的 ES 开发工具,如 ESDEI、MIKE、TOES、GMES、EIPIE 等,现已得到广泛的应用。

二、HEARSAY-III 的结构与特点

HEARSYS-III 属于通用型 ES 开发工具。它是由美国加利福尼亚 VSC 信息科学院的 L.D.Erman,R.Balier 等人实现的。该工具旨在开发、调试、学习和理解的基础上,选出领域的专门技能和问题求解原理,开发出适用于该领域问题求解的 ES。

HEARSYS-III 的结构如图 1 所示。

1.黑板数据结构

问题求解状态数据存放在一个全局数据库—黑板中,知识源可修改黑板,从而逐步导出问题的解,知识源之间的通信和交互都是通过黑板来完成。黑板上的数据是树型结构分层组织的,只有知识源才允许修改黑板上的数据。

2.知识源

知识源由求解问题所需要的领域知识组成,可以划分为多个不同层次的知识源。知识源的任务是提供能导出问题解的信息,一个知识源取得黑板上一组当前信息,并用它存储的特定知识修改信息。知识源用黑板上信息来进行相互间的交换,以产生问题的解。知识源仅仅修改黑板或控制数据结构,并且对求解状态的修改是显示且可见的。

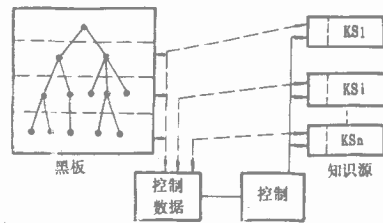


图 1 HEARSYS-III 的结构

注:——控制流——数据

3.控制机制

控制模块用来监督黑板上的修改,并决定下一步要进行的操作。各种信息对控制模块而言都是全部可用的,信息在黑板上也可以分开保存。解是一步一步的构造出来的,在解形成的每一个阶段都可以使用任何类型的推理步骤(数据驱动、目标驱动等),因此知识源的调用序列是动态的且是实时的,而不是事先规定的。

三、利用 HEARSAY-III 开发“电子信息分析”ES

我国电子信息战领域是一个多传感器数据汇集系统,同时也汇集来自雷达、雷达侦察、通信侦察、遥感侦察等系统的数据呈报,充分利用这几方面的有用信息,互为补充,消除模糊,及时且准确地识别各种目标体系,这样可以提高决策的可靠性、准确性和实时性。

1.知识表示

电子信息分析的要求知识表达方法能使现出系统的

动态变化特征反其实时性。

(1)事实表达。根据实时采样系统的特点,可以以时间点为标尺,将所有的数据和信息以“属性-值”对作为基础表示成表结构的形式。例如从传感器获取的系统某一动态特性的数据集合表示为

$$S_i = ((V_{ij}, t_{ij}) \dots (V_{in}, t_{in}))$$

其中: S_i 表示第 i 个动态特性的数据集合; V_{ij} 表示特性 S_i 的第 j 个值; t_{ij} 表示获得的第 j 个特性的值 V_{ij} 的时刻,又称为时标,可通过系统时钟来确定。这样就为表达出系统状态变化及操作序列的时序关系打下了基础。

上述表达方式便于各个推理行为对系统状态进行动态访问。系统的动态特性(传感器测量值等)可通过查询方式采集,而操作指令则通过中断方式获得,每实行一次操作,即对主机发出一次中断,输入此时的操作指令。

(2)时间的表达。以 J.E.ALLEN 的间隔代数学为基础,用时间区间来表示时序逻辑,并把区间定义为实数对,与具有一定持续时间的事件一一对应。例如,某事件 A 发生在时间 $(T1, T2)$ 内,则可定义一个相应时间区间 $I = \langle T1, T2 \rangle$ 。对于表示各时间点之间以及时间点与时间区间之间的关系,也可采用相似的描述形式。

定义了这些关系后,为表达任意两个时间点或时间区间之间的可能关系,可以利用关系传递公式。例如:

- 1) Before(i_1, i_2) \wedge Before(i_2, i_3) \rightarrow Before(i_1, i_3)
- 2) Meets(i_1, i_2) \wedge During(i_2, i_3) \rightarrow (Overlaps(i_1, i_3) \vee During \vee Starts(i_1, i_3))

对时间点或时间区间不断地运用关系传递公式的过程,就是对时序关系的推理,来建立系统动态变化或操作序列之间的时间关系。

同时还可以定义两个谓词,以表达事件与时间的关系:

HOLDS(P, i)表示特性 P 在时间点或时间区间 i 上成立。

OCCURS(Pr, i)表示过程 PR 在时间区间 i 中成立。

这样在表达事件的时序上就有了一个较为完整的途径。

2.推理控制机制

本系统是一个多种知识综合处理的复杂系统,而黑板结构正是解决综合资源利用的一种有效方法。黑板结

构由知识源、黑板和控制机组成。知识源由求解问题所需要的领域知识组成,可以划分为多个不同层次的知识源。黑板实际上是一个倒全局数据库,用以存放问题求解的状态数据和中间解。知识源用黑板上的信息来进行相互间的交互,以产生问题的解。控制机制则是监督控制知识源适当地响应对黑板的修改。以黑板这个全局数据库作为各个知识源进行间接通信的媒介,可以利用多知识源进行多路不精确推理,共享数据与结果,减少重复查询和重复推理。同时,黑板结构还易于将复杂问题进行层次化分解,按照层次组织规则块;且知识源具有相对独立性,故可先设计知识源再考虑与黑板机构的联系,使系统设计变得容易。

电子信息战领域是一个多传感器信息的汇集,这些信息本身就带有模糊性和不精确性,处理这类信息在推理机制上采用不精确性推理。较常用的有:MYCIN 不精确推理、主观 BEYES 方法、模糊集理及证据理等算法。本系统采用 MYCIN 不精确推理算法。

信任增长度

$$MB(h, e) = \begin{cases} 1 & ; \text{若 } P(h) = 1 \\ \frac{\max[(h/e), P(h)] - P(h)}{1 - P(h)} & ; \text{否则} \end{cases}$$

可以看出若 $P(h/e) > P(h)$ 时,在 e 存在的条件下,对 h 信任增加。

不信任增长度

$$MD(h, e) = \begin{cases} 1 & ; \text{若 } P(h) = 0 \\ \frac{\min[(h/e), P(h)] - P(h)}{-P(h)} & ; \text{否则} \end{cases}$$

可以看出若 $P(h/e) < P(h)$ 时,在 e 存在的条件下,对 h 不信任增加。

可信度

$$CF(h, e) = MB(h, e) - MD(h, e)$$

根据 $MD(h, e)$ 与 $MB(h, e)$ 具有互斥性,若 $P(h, e) = P(h)$,则表示 e 与 h 具有独立性。则

$$CF(he) = \begin{cases} MB(h, e) - 0 = \frac{P(h/e) - P(h)}{1 - P(h)} & ; \text{若 } P(h/e) > P(h) \\ 0 & ; \text{若 } P(h/e) = P(h) \\ 0 - MD(h, e) = \frac{P(h) - P(h/e)}{-P(h)} & ; \text{若 } P(h/e) < P(h) \end{cases}$$

讨论:

(1) $CF(h/e) > 0$

说明 e 的出现增加 h 为真的概率,即 $P(h/e) > P(h)$, e 出现增加 h 为真的可信度。

(2) $EF(h/e) < 0$

说明 E 的出现减少 h 为真的概率,即

$P(h/e) < P(h)$, e 出现增加了 h 为假的可信度。

(3) $CF(h/e) = 0$

说明 e 出现时对 h 没影响,即

$P(h/e) = P(h)$

可以看出,每个断言 e 都有一个与之对应的确定性度量 $CF(e)$ 。如果知道 e 为真,则 $CF(e) = 1$;如果知道 e 为假,则 $CF(e) = -1$;如果 e 一无所知,则 $CF(e) = 0$ 。

对一般规则:

IF E_1 AND E_2 AND AND E_n THEN H(CF)

其中: $E_i (i = 1, 2, \dots, n)$ 是证据

h 是一个或多个结论。

CF 是可信度。

含义是当证据 E_1, E_2, \dots, E_n 都存在时,结论 H 以 CF 的可信度成立,且取值范围为 $-1 < X < 1$ 。

3. “电子信息分析”ES 组成模块

本系统利用 HEARSAY-III 工具开发的, HEARSAY-III 工具只为用户提供了一个系统框架,也就是黑板模型结构如图 1 所示,而其中的黑板划分,知识源划分及控制机制都是用户自行设计,这也正是以通用开发工具的优点。

使用 HEARSAY-III 工具的基本框架,把电子信息分析的划分为雷达识别知识源、通信识别知识源、遥感侦察知识源等多个子系统。与这种分布式模块化结构相适应,作为全局数据库的黑板也被划分若干个区,分别存放各知识源有关状态要素。此外黑板上还存放开关变量,供黑板控制模块调用知识源。整个系统相应结构框架如图 2 所示。

基于上述框图,系统的求解过程如下:

(1) 将信息源的脉冲信号(模拟量)进行滤波、合成处理,将数据送入信息预处理子系统,该子系统进行时间、空间相关分析,将数据存入数据文件中。

(2) 控制机制从数据文件中录入数据,并把这些数据

存放到黑板的参数层数据区。然后利用元规则产生调度队列,调度相应的知识源。

(3) 知识源被激活后,进行条件部分的匹配,然后作不精确推理,获得结论的可信度,修改黑板。

(4) 随着黑板的不停变化,控制机制实时地激活相应知识源,又对黑板作新的修改,增、删直至最后得出综合信息。

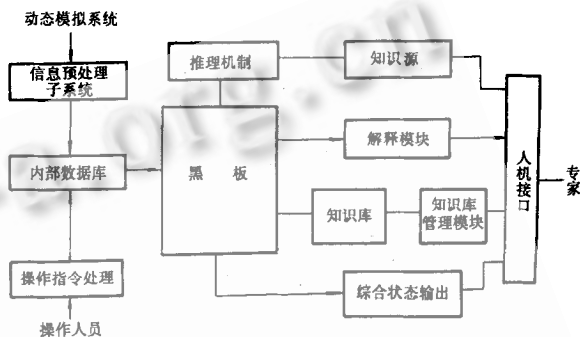


图 2 “电子信息分析”ES 框图

四、新一代 ES 开发工具

新一代 ES 应提高 ES 工具的灵活性,以面向对象的方法设计 ES 工具,使用户能灵活选用功能自由嵌入集成 ES。

新一代的开发工具应具有相当的灵活性,它能提供几种知识表示方式、推理机制、知识获取方式(包括一些机器学习模块、知识获取策略)和一些系统支撑模块(如语义检查、解释等),由知识工程师自由选用一些模块,集成 ES,并允许知识工程师能将自己编写的功能模块嵌入系统中。

参考文献:

[1]李卫华:《专家系统工具》,第一版,北京气象出版社 1987 年

[2]王申康:《新一代专家系统及其在农业中的应用前景》,《第二届全国农业知识工程学术会议论文集》,1991 年,杭州

[3]夏国平:《专家系统开发工具》,第一版,北京,清华大学出版社,1990 年