

专家系统间的知识获取与专家系统兼并

朱 浩 (解放军第九十二医院, 福建南平)

摘 要

The narrowness of knowledge domain is considered as the main defect of expert system. In order to overcome the weakness, large scale expert system(LSES)and distributed expert system (DES) have been proposed and developed. But, the building and the knowledge acquisition of LSES and DES are very onerous and difficult. A large scale knowledge base is the key part of LSES. In this paper, a new approach of knowledge acquisition for expert system is proposed and discussed. This method can be used to update the knowledge base of expert system and build large scale expert system.

自DENDRAL产生并被成功地使用近二十年来,专家系统(ES)作为人工智能(AI)的一个分支,不仅在技术上得到了丰富的发展,而且数量和应用领域迅速增加^[1],这主要是由于ES为领域专门化的,不象许多传统软件是领域通用的。但是,目前各领域中现存的ES也出现了一些应引起注意的问题,如ES的安全^[2],已有ES的更新与发展等。另一方面,随着理论与技术的成熟及应用的需求,产生了第三代ES的概念^[3],即大型多学科综合型专家系统。本文提出并浅议已有ES更新与发展及构造大型ES的一种可能的方式——ES间的知识获取与ES兼并。

一、必要性

有几个方面的因素促使我们考虑ES间的知识获取与ES兼并问题。首先分析一下ES应用中的一个基本情况,即ES数/ES覆盖的领域数之比显然远大于1,这就是说,每个应用领域都使用或开发了多个ES,如医学领域,而且开发的各ES尽管用了相同的或不同的程序设计方法,但其知识源是一样的。几乎可以肯定地讲,没有一个ES的知识域(知识库)覆盖了其整个或绝大部分的领域知识,即不存在领域知识完备的ES。至于包含多学科综合性知识的所谓第三代专家系统,其体系结构、知识表示、知识获取及

知识库组织等问题更是尚在探讨中。

1. 从ES继承性的角度看

AI发展中的一个易受指责的问题是,每个新系统难以从先前的系统与经验中获得益处^{[4][5]},即工作缺乏继承性。对ES而言,继承性有二个方面:程序设计上的继承性和知识结构上的继承性。有不少各类ES Shell和ES Tool可以被认为提供了某种程度的程序设计上的继承性,而知识结构上的继承性尚未引起较多注意。我们把知识结构上的继承性定义为:在ES建立和发展过程中,从同一领域的其它ES获取有价值的知识,且这一功能应由ES自行完成。显然,ES间的知识获取其速度比ES通过屏幕交谈向人类专家获取知识快得多,成本也低得多。把这一功能植入ES原型中,则可提高ES知识获取的速度,改变其耗时性和高成本,加强了同一领域ES在知识结构上的继承性,即先前的ES的知识库是现在ES知识库的真子集。

2. 从ES更新的角度看

软件必须更新,这是常识。现已提出了软件再工程的概念,其工作包括,不管有关软件的状况如何,对现有软件的移植、修改、建立文档等,典型说来,离不开要进行再工

程的软件实现的功能^[6]。ES作为一种特殊的计算机软件,更新问题更显突出,因为更新需求不仅来自程序设计方面,而且更来自领域知识的发展。知识决定ES的能力,对ES来说及时的知识更新与增补是十分重要的,否则很可能难以继续胜任工作。ES知识更新的主要途径当然是从原领域专家那里获取新的领域知识,但对那些较早期建立后已与领域专家脱离联系的ES,从新的或其他的同领域ES获取知识未尝不是一种途径。

3. 实现巨系统的一种途径

建造巨型ES是一个诱人的目标,这是向人类智能水平迈进过程中的重要一步,一些研究者正在朝这一目标努力^[7]。巨型ES的基本要素是要有一个巨知识库,知识获取是建造ES的瓶颈,构筑巨知识库更是建造巨型ES的主要难题。建造巨型ES可以有二个方法:人工方法与机器方法。人工方法就是先建立带有多种知识表示方式与多种知识获取机制的巨型ES框架,再由巨型ES框架通过与多学科的领域专家们交谈获取巨量的知识。机器方法则是研究与使用ES间的知识获取与ES兼并,通过多个ES的兼并与融合形成巨型ES。目前的工作多用前一种方法,后一方法似尚未被充分注意。

4. 机器学习

ES间的知识获取与兼并过程包括知识选择、获取、转换及重组,人类在获取知识时也有这些步骤,因此ES间的知识获取行为可看作是机器学习的一种方式而值得加以探讨。

二、几种情形与可行性

1. 几种情形

从简化讨论出发,且不失一般性,我们假定原有的ES均是单一知识表示方式及单一推理方式的。根据复杂程度,我们将ES兼并分为几种情况。

兼并前的两ES:

A. 采用相同推理方式: Aa. 采用相同知识表示方式

Ab. 采用不同知识表示方式

B. 采用不同推理方式: Ba. 采用相同知识表示方式

Bb. 采用不同知识表示方式

兼并后的一个ES:

C. 采用单一推理方式: Ca. 采用单一知识表示方式

Cb. 采用联合知识表示方式

D. 采用联合推理方式: Da. 采用联合知识表示方式。

当两ES间仅作部分知识获取时,获得知识的ES的结果总是Ca,另一个ES结构不变。

对两ES兼并,当前提是Aa或Ba时,不存在知识转换问题,而当两ES的知识域不相交时,可能也不存在知识重组问题(这与知识表示方式有关,在某些情况下仍需重组),这属于最简单的情形。当前提是Ab或Bb时,若结果为Ca,则需作知识转换,若结果为Cb,则不需作知识转换。显然,Aa→Ca或Ba→Ca较为容易。总之,对于结果Ca,主要的困难是不同知识表示方式间的知识转换与重组,对于结果Da,主要的困难是需建立一个元层次机构(元知识与元控制),由元层次对任务进行接收、检查及分配,元层次负责控制与调用联合起来的各子部分。另外,结果Ca的知识库一般是无冗余的,而结果Cb与Da的知识库是有冗余的。

2. 可行性

由上面分析可知,ES间的知识获取与ES兼并的主要困难之一是不同知识表示方式间的知识转换。因此,需考虑的一个重要问题是,不同知识表示方式间的等效性与可转换性。可转换性好,则转换较易行,等效性高,知识获取或兼并的效果便好。AI的发展虽显得百花齐放,各种知识表示与推理方法层出不穷,被广泛应用于各领域中的专家系统,语言理解系统等,但大多数的ES

使用了常用的几种知识表示方法^[1],即产生式规则、框架、语义网络、逻辑表示。因此主要只需考虑这几种表示方式间的知识转换方法,这是可行的。就某一具体领域而言,该领域中ES的知识表示可有二种情况:1.均使用同一知识表示方式,2.使用多种不同知识表示方式。随着ES应用经验与技术的成熟,人们可以比较总结出一些具体领域(如医学、法律、军事等)中的1种或1—2种最佳表示方式。若该领域中的ES大部分使用这种知识表示方式,则ES间的知识获取与ES兼并是易行的。若该领域中的ES使用了多种知识表示方式,则以使用最佳表示方式的ES为主,进行ES间知识获取或ES兼并。

三、应用

从目前来看,ES间的知识获取与ES兼并至少可用于两个方面,即ES更新和构筑大巨型ES。

1. ES更新

ES需要更新(update),这是显然的。更新包括二方面:程序设计方面的更新与知识库的更新。从ES的功能角度讲,知识库的更新是十分重要的。目前由于进入实用和商品阶段的ES还不十分多,这个问题尚不突出。知识库的更新应当有两个知识来源:人类领域专家及同领域的其它ES。从人类领域专家获得新知识固然重要,但某些情况使得我们亦必须重视第二个领域新知识源。以国内开发应用较多的名老中医诊断专家系统为例。名老中医诊断专家系统在功用上有一个与一般ES不同的特点,即保存与传播名老中医积累几十年的中医学知识与经验。很可能,在一个名老中医专家系统建成使用后不久,原领域专家便去世了,对这样的ES能否进行知识更新?答案应该是肯定的。可以从两个方面获取新知识,即同一专业的其它名老中医及同一专业的其它名老中医诊断专家系统。因此,从其它ES获取知识,是ES知识更新的一个必要也是可行的途径。

2. 构筑大巨型专家系统

各领域中的ES为数极多,但有较好使用效益的商品性成果并不多,一个重要原因是许多ES的规模太小,知识域狭窄,能力有限,以至无法投入实际应用。ES是一种特殊的计算机软件,其功能强弱主要取决于所包含知识的多少,而构筑大型知识库是高成本和耗时的。ES间的知识获取与知识兼并为解决这一问题提供了一个新途径。可以把构筑大巨型ES分成几种情况:

1, 在同一领域内, 1a. 相同知识表示方式ES的兼并

1b. 不同知识表示方式ES的兼并。

这样构成的大巨型ES具有较丰富的专业知识与较强的工作能力,这可看作是ES的纵向生长。

2, 不同领域内, 2a. 相同知识表示方式ES的兼并

2b. 不同知识表示方式ES的兼并。

这样构成的大巨型ES具有多学科的综合知识,是一种第三代ES,这可看作是ES的横向生长。

四、我们的尝试

在文[9]中我们提出了一种新的能较贴切表示临床医学知识的知识表示方法——关联索引式知识表示法及相应的双向混合推理,发展了相应的知识获取机制及知识库维护方法,因此它又是一个ES Shell,通过和领域专家交谈,可方便地同时建立几个专家系统。基于这一方法,建立了消化系统疾病和神经系统疾病的医学诊断治疗专家系统。该知识表示及推理方法适用于整个临床医学领域,能够表达与操作多层次的知识。建造一个覆盖整个或大部分临床医学领域的诊断治疗(以至病因、护理等)专家系统是个诱人的目标,有着很大的实际意义,但从众多领域专家获取知识直接建立一个巨知识库是十分复杂和困难的。因此,我们采用通过ES与领域专家交谈,分别建立一个个独立的各

(转第74页)

设计专家思维的另一特征。关于逻辑思维现已有较为成熟的理论可循,至于人类如何进行形象思维,目前人们还知之甚少,但有一点可以肯定:对事物进行模糊类比,是人们把握客观形象的一种手段。因此,我们认为:采用模式匹配的方法去寻求体现形象思维过程是一条有望实现的途径。对此,有两方面的工作要做:

(1) 确定能独立全面表征某事物类的指标;

(2) 拿类比事物集合与实际情况相比较,并依据模糊数学理论中的择近原则等进行模式匹配。

四、专家系统与数据库管理系统的有机结合

在工程设计中,需要调查分析大量的原始数据资料,这就产生了数据的管理及处理问题。而解决这一问题实质上涉及到数据库管理系统和专家系统两种技术的结合。由此又产生一个问题,如果专家系统与数据库有机地结合在一起,那么整个体系中的各功能应如何分散到ES逻辑设计部分和DBMS逻辑设计部分上。例如数据模型、查询语言等其它原由DBMS提供的功能是否要从DBMS中取消一部分,而代之以ES提供?反之原有ES提供的功能是否也需要这样做。对此,我们认为:ES和DBMS的搜索机制是截然不同的。为此,应根据它们各自的搜索机制的特

点来划分两者的性能,即ES的逻辑设计仅仅用在那些确实需要ES搜索机制的功能的地方,否则,则应用较简单的搜索机制。对此,DBMS的逻辑设计应该最大限度地担负起责任,力图减少ES搜索周期中所涉及的知识。为此,我们提出以下假想模式:

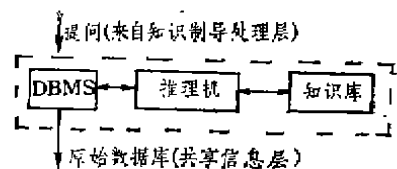


图2 信息处理模块的假想模式

以上我们从四个方面阐述了工程设计类专家系统的基本特征。由此我们相信:随着专家系统理论的日趋完善和一些工程设计类专家系统的相继开发成功,一种具有鲜明特色的工程设计型专家系统定会得到更好的发展。

参考书目

- [1] 王俊善, 吴秀清, 一个用于构造型专家系统的问题求解策略, 计算机研究与发展 1990.4
- [2] 程建群, 秦彤, 胡铭曾, 支持多专家系统知识库的知识组织, 计算机研究与发展 1990.8
- [3] 欧阳为民, 专家数据库引论, 计算机科学 1990.3.
- [4] 施欣: 专家系统的不确定性推理研究, 中国模糊数学与模糊系统五届年会宣读论文。

(接第77页)

临床领域分支ES, 然后让各分支ES相互兼并, 最终形成一个巨ES。我们已设计了基于此种知识表示的知识获取与知识库兼并算法, 作为各子ES的一个功能部分, 这属于ES兼并最简单的情形, 即 $Aa \rightarrow Ca$ 。由于产生式规则是一种最常用的知识表示方法, 建立的医学ES也较多, 接下来我们将考虑采用关联索引式知识表示与采用产生式规则知识表示的ES间的知识获取与知识库兼并, 以及基于规则知识表示的ES间的相互知识获

取与知识库兼并。

ES作为一种逐步走向成熟的AI技术和计算机软件, 人们对它提出了愈来愈高的要求, 如要求处理与解答复杂程度更高的问题。单一的ES已无法胜任这样的工作, 因此ES已向分布式发展, 采用将复杂问题分解为各个不同的相对简单的问题的策略, 或根据不同的问题启用分布式ES中不同的结点ES。但另一方面, ES也可以向集成式发展, 即采取ES间的知识获取与ES兼并, 由此构筑大巨型ES, 以克服知识领域的狭窄, 胜任复杂性更高的问题求解, 或提高实际工作能力。(参考文献略)