

知识工程发展概述

权太范

(哈尔滨工业大学)

〔摘要〕 1977年在美国麻省理工学院召开的第五届国际人工智能会议上,斯坦福大学Feigenbaum教授提出知识工程(Knowledge Engineering)以来,其发展极为迅速。已经广泛地应用在数据处理、诊断、监视、预测、规划,设计等六个方面,并取得了良好的效果。目前正在向钢铁、石油、矿山探测等产业应用发展,已成为计算机、信息以及控制工程界的一个热门。

本文将综述近几年来国外知识工程科学的发展概况及其前景。

一、发展概况

MYCIN 系统是最早诞生的知识工程系统。它是由 Feigenbaum 教授研究小组经三年时间的研究而完成的专门用于诊治血液传染病和脑膜炎的知识工程系统,其程序差不多各占一半。它是用 LISP 智能语言仿真了具有专家医生水平的500条诊治规则,推理机构使用了模糊逻辑和启发累加函数。因此该系统具有高级医生的水平,而且能“学习”专家医生的知识,能模拟医生的思维和诊断推理,给出可靠的诊治建议。用户(医生)只根据临床实验的数据和患者的病历,通过和 MYCIN 系统的对话,可查询病因和治疗方案。该系统的主要特征之一是具有独立的问答程序,因而能跟踪诊断和推理过程。当用户对 MYCIN 系统作出的诊断或治疗方案提出疑问时,该系统能进一步回答其理由。这样,普通的医生可借助 MYCIN 系统做专家医生的诊断工作。因此 MYCIN 系统对用户来说,相当于一个高级专家顾问团了。MYCIN 系统的成功为后期的知识工程研究开辟了方向。

Feigenbaum 教授认为知识工程定义应包括:

(1) 知识工程是以人工智能研究结果的应用为目的的。即它是人工智能的应用分支。

(2) 知识工程中心课题是构造专家系统

(Expert System),以下简称为ES。

(3) 为使 ES 能解决具有专家水平的难题,应开发高性能的计算机软件。

可见,知识工程的研究重点一是应用,二是软件设计。有的学者把知识工程的研究对象简单地说是构造具有专家水平的程序。

知识工程的研究已经经历了三个时期。

第一时期(1965—1975)

标准的 MYCIN 系统出现以前。

第二时期(1975—1980)

在大学研究室开发实验性系统的时代。作为ES规范已出现了MYCIN系统。

第三时期(1980—)

开始在产业部门应用。

美国是知识工程最发达的国家,而斯坦福大学是美国知识工程研究中心。人们把Feigenbaum教授称为美国知识工程之父。过去该大学的知识工程研究经费只是由两家财团提供的,但目前已有好几个财团在投资。美国国防部高级研究计划局(ARPA)每年投资一千万, Schlumberger公司投资五百万美元来从事知识工程研究,开始出现了以知识工程为产品的企业。关于知识工程的研究,日本远落后于美国。但值得注意的是去年(1983),也就是雄心勃勃的第五代计算机开发研究十年计划的第二年,在整个日本控制界掀起了知识工程热。日

本一些重要刊物,例如计测与控制(第9期),系统与控制(第7期),电气学会杂志(第3期),数理科学(第6期)以及运筹学(第6期)都抢先出版了知识工程专集,并呼吁要重视知识工程研究。众所周知,第五代计算机将具有解决问题的功能,学习、认识、理解功能,语言分析和理解的功能及智能对话(即Grace-fully Interacting System)功能等四大功能。因此第五代计算机的研制迫切需要首先开发知识工程的研究。在1981年由日本信息处理开发协会发起的第五代计算机国际会议上,东京大学元罔教授做了“向知识处理系统挑战”的讲演。目前美日之间关于第五代计算机的竞争正在极大地刺激着知识工程的发展。

二、知识工程的研究课题

从数学语言来讲,知识就是人们通过实践并经过大脑的思维活动获得的数据,资料、事件及规则的集合。在知识工程中,通常把知识分为两大类:一是关于事实的知识(Factual Knowledge)。如在教课书、文献中系统的严格的、公认的知识等。二是在实践活动中累积的、带有经验的启发式知识(Heuristic Knowledge)。如专业知识规则、胜于实践的规则、判断、推理等。知识工程的主要对象是第二类知识。我们把现实中存在的问题,如表1所示,可分为具有良好结构和病态结构的问题。

表1 问题的分类

问 题	良好结构的问题	病态结构的问题
特 点	数字的 连续的 重复的 信息是闭合的	符号的 离散的 一次性的 信息是开路的
推理方式	演绎的 控制驱动器 建立模型	归纳 事件驱动器 建立知识库
方式	系统工程 系统科学	知识工程 人工智能

如图1所示,系统工程主要是通过建模解决具

有良好结构的问题,而知识工程是通过建立知识库,以推理为主要工具,解决病态问题的。

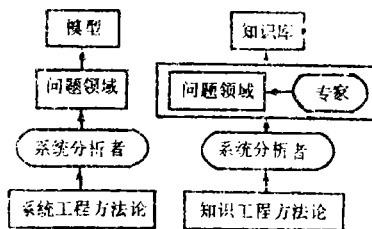


图1 系统工程和知识工程比较

知识工程系统通常如图二所示,由知识库、推理机构两部分组成的。

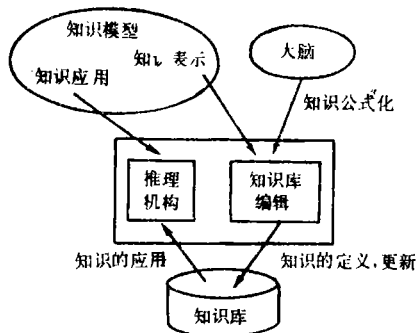


图2 知识系统结构

知识工程的研究课题可分为以下三个方面。

(i) 知识的表示(Knowledge Representation)问题:为了存取知识,在计算机内存中如何用数据构造知识。知识的表示主要考虑:①表示知识的能力 ②表示形式的模块性和柔软性 ③易懂。

表示知识的语言主要有:①函数型语言(Functional Language) LISP、推理逻辑型语言PROLOG (Predicate Logic Language) ②产生式规则(Production Rules) ③框架(Frame) ④网络。

(ii) 利用知识的推理问题:如何应用所整理的知识很好地解决问题。实际上是推理机构(Inference Engine)设计问题。按方向、推理可分为正向推理和逆向推理。前者是由已知初始条件出发,向目标推理的方法,后者则是由目标状态出发,根据某种规则对目标说明

原因的方法。推理主要有演绎推理和归纳推理。演绎推理的基本形式是三段论。计算机的动作,机理与人的思维尽管不同,但就其推理规则而言,都是一样的。

(iii) 最重要的是知识的获取(Knowledge Acquisition)问题:如何自动或半自动地把所需要的知识输入到计算机程序中去。关于知识获取问题的研究主要有:一是获取知识的辅助系统研究,二是学习系统(获取知识的自动化)研究。目前只限于研究前者,其模型如图三所示。

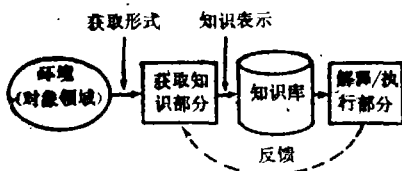


图 3

知识获取问题研究课题主要有①如何系统化、结构化知识,并送到知识库 ②如何保持知识库内的知识完全性和一致性关系 ③如何提高知识库本身的学习能力。

目前研究这些课题的主要工具有:①智能语言 ②使用方便的屏幕编辑等程序环境 ③表示知识的语言以及使用方便的TSS 计算机系统。

三、知识工程的中心课题——专家系统

随着科学技术的不断发展和知识的膨胀,即使是优秀的专家也不可能把本专业的全部知识都吸收和消化。于是人们自然地提出能否把时间和空间上分散的专家知识存入到计算机,建立一个具有专家水平的系统?实践证明,就特定问题而言,完全可以做到使计算机所具有的知识远远超过专家,而且比专家更好地解决问题。在这里,获取某种规则,并公式化来利用的行为叫做专业知识的模型化(Expertise Modeling)并把模型化过程中构造的程序叫做专家系统。ES是一种计算机软件系统,它能在一定的程度上模拟专家的思维,判断,起到专家的作用。ES主要由装有专家知识的知识库

和推理机构组成。用计算机语言而言,知识库就是数据库,而推理机构相当于计算机本身。系统的知识库有三大模型:关系数据库、网络数据库、阶统数据库。其结构如图四所示。

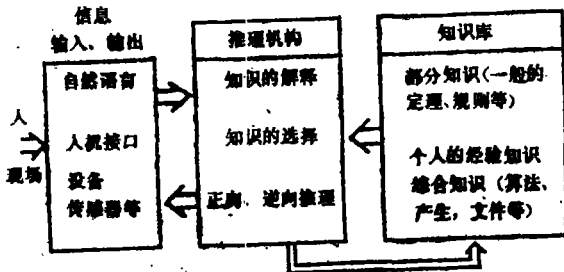


图 4 专家系统结构图

ES的主要研究对象是非确定性问题。主要包括:①用非数字性参数描述的问题。如离散事件系统,生物刺激反应过程等 ②虽然可用数字参数描述,但因问题过于复杂,无法提出数字模型或足够的统计知识。如体温、血压、血液成分等。专家系统的核心问题是如何活用专家知识。即知识库及其管理。为此,ES不但具有说明推理,解决问题的机能,而且具有根据专家的思维过程的模型进行灵活会话的机能。图五表示了ES、用户、设计者三者之间的关系。由此可见,专家系统具有如下特点。

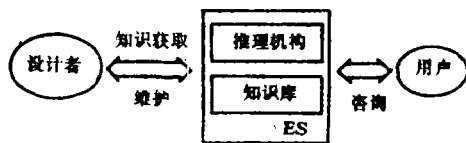


图 5

- ①利用了专家的专门知识。
- ②具有实用价值,且性能好,效率高。
- ③对象是目前难于解决的问题。
- ④容易改造系统和扩充新的知识。
- ⑤能跟踪解决问题的推理过程。
- ⑥和用户可进行自然会话。

ES已广泛应用在医学,机器故障诊断、飞行器设计,信号处理等十一个方面。其中医学是

ES 应用的传统领域,已经历了两代的发展。1977年前为第一代,其特点是单科病的诊治。典型的系统有 CASNET (用于青光眼)、MYCIN (用于传染病)、PIP (用于肾脏病)等。这些系统应用受到一定的限制,而且速度较慢。(如 MZCIN 系统进行一次一系列咨询需要二十分钟)。1977年至今为第二代。其特点是用于多科病的诊断。典型系统有 EXPERT, EMYCIN, AGE 等。也开始出版了计算机诊断的专著。因此 Feigenbaum 教授提出对于 ES 来说,八十年代将是黄金时代。

四、知识工程的前景

今后自动控制研究方向可考虑下面两个方面。

①提高机器的智能,实现高度的机械化。这是以往研究方向的继续。

②借用计算机使机器具有人的记忆、思维、判断能力。这是新的研究领域。这要求用知识工程的概念、理论、方法设计控制系统。它是继解放人类体力劳动的蒸汽革命,电气革命之后的又一个解放人类脑力劳动的划时代科学技术革命。它的发展,将引起控制系统的一系列变革。例如,控制系统能处理目前无法处理的难于用数学语言描述的模糊系统;输入、输出量不仅是数字量,也可以是语音、图象、文字符号等非数字量;系统与人之间可建立灵活的会话接口装置以及系统结构由识别——推理——决策——执行的软结构代替现有的检测——比较——计算——操作的硬结构等。知识工程已在各个领域显示出年轻的生命力。它的前景可分下面三个方面叙述。

①降低成本,提高经济效益。

如被用来确定有机化合物分子结构式的 DENDRAN 专家系统只需要价格为五千美元的数据处理装置,而且判断能力往往超过年轻博士的水平。但早先需要用价值十万美元的数据

处理装置。

②提高计算机竞争能力。

随着科学技术的迅速发展,计算机的改进和更新周期越来越短。据预计,八十年代到九十年代计算机价格将大幅度地下降。因此为提高竞争能力、计算机应具有知识情报处理能力。

③最重要的效益——学习新的知识。

科学技术的发展,迫切需要工程技术人员不断地学习新的知识,特别是在书本上难以学到的许多优秀专家的知识。这些工作可由 ES 承担。也可以用 ES,把知识传授给学生,并且辅导和答疑。

但是知识工程必竟是一门年轻的科学,它还需要解决一系列问题。如:

①硬件设备不足。它需要开发新的计算机系统及大量的硬件设备。但因造价太贵受到一定的限制。

②知识工程的方法和技术还不完善。特别是难于建立企业所需要的大规模系统。

③研究人员不足。尤其是至今还没有知识工程学的具体教程。

④知识获取问题仍然是个难题。

随着新的技术革命的出现,知识工程研究将有新的突破。

附言:科学院系统所韩京清同志对本文初稿提出了宝贵意见,为此谨表谢意。

参考文献

- (1) Barr, A and Feigenbaum, E. A: The Handbook of Artificial Intelligence, 2, Los Altos, 1982
- (2) Stefik, M et al, The Organization of Expert System, Art. Intell, Vol. 18, No. 2, P. 135—173, 1982
- (3) 知识工学特集. 计測と制御. Vol. 22, No. 9, P. 745—802, 1983
- (4) 特集: 第五世代计算机——知识情报处理システム, 数理科学, Vol. 20, No. 4, P. 5—55, 1982.
- (5) 蒋新松等: 参加第七届国际人工智能会议的总结, 国外自动化, 2期, P. 1~6, 1982