

知识工程原理

管纪文 刘大有 等著

(106)

吉林大学出版社

知识工程原理

管纪文 刘大有 等著

吉林大学出版社

内 容 简 介

本书从理论、方法和技术方面详细论述了知识工程这一新领域。全书共分十章,内容包括:知识工程概述,知识表示,基于知识的推理,知识系统结构,知识获取,解释,语言处理接口,知识系统的建造,知识工程工具,以及知识系统的评价等。

本书可作为计算机科学、人工智能、知识工程及其他有关专业的大专院校本科生和研究生的教材或教学参考书,亦可供科技人员参考。

知识工程原理

管纪文 刘大有 等著

责任编辑:王瑞金

封面设计:甘莉

吉林大学出版社出版
(长春市解放大路85号)

吉林省新华书店发行
吉林大学印刷厂印刷

开本: 787×1092毫米 1/16

1988年11月第1版

印张: 19.5

1988年11月第1次印刷

字数: 448千字

印数: 1-1 500册

ISBN 7-5601-0153-4/TP·1

定价: 3.90元

前 言

当今社会是一个知识信息社会。“知识就是力量”已成为人所共知的事实，一个人具备了渊博的知识，就能解决许多复杂的专门问题。本书的论题——知识工程则是研究如何将人的知识赋予计算机的一门综合性边缘学科，它使机器最终能利用知识解决需要人的智能才能解决的专门问题。知识工程是计算机科学和人工智能发展到一定阶段的产物，它使人工智能研究走出实验室，面向实际应用。知识工程以知识的智能处理作为其研究内容，以构造高性能的知识系统作为其研究目标。知识工程的根本目的是构造出具有人类的理解、学习、联想和推理能力的智能机器。

知识工程产生于本世纪60年代中期，80年代获得飞速发展。现在，知识工程产品——专家系统已广泛应用于社会生产和生活的许多方面，并已产生了巨大的社会效益和经济效益。在国外，专家系统及用来建造专家系统的工具正逐步实用化和商品化，一个新的产业——知识工程产业正在形成。在国内，专家系统已成为近几年的一个热门课题，而且，将会得到越来越多的人的广泛关注。本书试图从理论、方法和技术的角度，对现有专家系统研究成果进行比较全面系统地整理和归纳，并把专家系统上升到知识工程的层次来考察和研究。

本书仅作为引玉之砖求教于计算机界和人工智能界的前辈和同志们，以期推动我国的知识工程研究和实用专家系统的发展。

本书是国家自然科学基金资助的项目，由管纪文、刘大有、黄祥喜、张成奇、邱涤虹、洪军、郑方清和马志方等合著而成。第一章和第七章由黄祥喜执笔，第二章和第六章由邱涤虹执笔，第三章和第九章节9.1, 9.2, 9.3和9.7由张成奇执笔，第四章由郑方清执笔，第五章由洪军执笔，第八章和第九章节9.4, 9.5, 9.6和9.8由刘大有执笔，第十章由马志方执笔。管纪文和刘大有对全书各章进行了仔细的修改。

管纪文 刘大有 黄祥喜 张成奇 邱涤虹 洪军 郑方清 马志方

一九八八年七月三十日 于长春

目 录

前 言

第一章	知识工程概述	(1)
1.1	何谓知识工程.....	(1)
1.2	知识工程的产生与发展.....	(2)
1.3	知识工程的基础.....	(6)
1.3.1	知识和知识的工程化.....	(7)
1.3.2	智能问题求解.....	(9)
1.4	知识系统.....	(10)
1.4.1	知识系统的特征.....	(10)
1.4.2	知识系统的类型.....	(12)
1.4.3	知识系统的基本结构.....	(13)
第二章	知识表示	(15)
2.1	概述.....	(15)
2.1.1	什么是知识表示.....	(15)
2.1.2	知识的类型.....	(15)
2.1.3	知识表示简史.....	(16)
2.2	知识表示的主要方法.....	(17)
2.2.1	逻辑表示法.....	(19)
2.2.2	语义网络.....	(20)
2.2.3	产生式系统.....	(21)
2.2.4	框架和脚本.....	(22)
2.2.5	过程表示法.....	(23)
2.2.6	直接表示法.....	(24)
2.2.7	语义原语.....	(24)
2.3	知识表示的发展现状.....	(26)
2.3.1	非规范逻辑.....	(27)
2.3.2	知识表示语言.....	(29)
2.3.3	知识表示系统.....	(31)
2.3.4	KRYPTON (一种新的表示方法) 简介.....	(32)
2.4	问题和结论.....	(37)
第三章	基于知识的推理	(39)
3.1	概述.....	(39)
3.2	基于知识的推理的基础.....	(40)
3.2.1	演绎推理与归纳推理.....	(40)
3.2.2	逻辑推理与似然(Plausible) 推理.....	(41)

3.2.3	知识表示与推理	(42)
3.3	推理机	(44)
3.3.1	控制与推理的效果和效率	(44)
3.3.2	数据驱动的控制	(46)
3.3.3	目标驱动的控制	(48)
3.3.4	混合控制	(50)
3.3.5	元控制	(55)
3.4	不确定性推理	(56)
3.4.1	模型方法	(56)
3.4.2	工程方法和控制方法	(75)
3.4.3	结束语	(76)
3.5	问题和讨论	(77)
第四章	知识系统的结构	(78)
4.1	概述	(78)
4.2	不确定性知识和数据的处理	(80)
4.2.1	不确定性的来源	(80)
4.2.2	对不确定性的处理	(81)
4.3	时变数据的处理	(81)
4.4	大型解空间的处理	(83)
4.4.1	搜索空间的分解	(83)
4.4.2	搜索空间的抽象	(83)
4.4.3	处理大空间的其它方法	(87)
4.5	多知识源的处理	(88)
4.6	提高系统效率的方法	(90)
4.7	几种典型系统结构	(92)
4.7.1	元知识系统结构	(92)
4.7.2	分布式结构	(93)
4.7.3	多知识表示系统	(95)
4.7.4	多层系统结构	(96)
4.8	问题与结论	(98)
第五章	知识获取和机器学习	(99)
5.1	概述	(99)
5.2	知识获取	(99)
5.2.1	知识获取的几个主要阶段	(99)
5.2.2	知识获取的方式和途径	(102)
5.2.3	知识获取的主要技术和手段	(103)
5.3	机器学习	(105)
5.3.1	示教学习	(106)

5.3.2	实例学习	(108)
5.4	几个典型的机器学习系统	(111)
5.4.1	KAS	(111)
5.4.2	TEIRESIAS	(118)
5.4.3	META — DENDRAL	(130)
5.5	问题与结论	(135)
第六章	解释	(136)
6.1	概述	(136)
6.2	解释的结构	(137)
6.2.1	认识论问题	(137)
6.2.2	用户模型	(138)
6.2.3	论证方法	(138)
6.3	解释的主要方法	(138)
6.3.1	预制文本和执行追踪法	(139)
6.3.2	策略解释法	(141)
6.3.3	自动程序员法	(147)
6.4	问题和结论	(153)
第七章	知识系统的人机接口	(155)
7.1	知识系统的语言接口的抽象结构	(155)
7.1.1	语言接口的功能和特征	(155)
7.1.2	语言接口的结构和分类	(155)
7.1.3	会话处理、翻译和综合	(156)
7.1.4	用户模型	(158)
7.2	自然语言接口	(159)
7.2.1	概述	(159)
7.2.2	关键字—模式匹配技术	(162)
7.2.3	词法分析	(163)
7.2.4	句法分析	(168)
7.2.5	语义分析	(174)
7.2.6	语用分析	(178)
7.2.7	语言生成	(181)
7.3	问题与结论	(181)
第八章	知识系统的建造	(183)
8.1	概述	(183)
8.2	领域问题的分析	(184)
8.3	知识获取阶段	(185)
8.4	专家系统设计	(185)
8.5	ES的实现阶段	(187)

8.6	几个典型的知识系统	(188)
8.6.1	一个演示型的简单知识系统	(188)
8.6.2	生油条件评价专家系统	(207)
8.6.2.1	工具系统 MES 1 的知识表示	(207)
8.6.2.2	MES 1 中不确定性问题的描述	(211)
8.6.2.3	MES 1 中的不确定性推理	(215)
8.6.2.4	MES 1 生成的专家系统的工作过程	(217)
8.6.2.5	MES 1 的元推理机——任务规则解释模块	(218)
8.6.2.6	MES 1 的目标推理机——方法解释模块	(223)
8.6.2.7	MES 1 的抽象知识库与知识获取	(228)
8.6.2.8	系统行为解释子系统	(231)
8.6.2.9	生油条件评价专家系统框图	(232)
8.6.2.10	关于 MES 1 的讨论	(234)
8.7	问题与结论	(235)
第九章	知识工程工具	(238)
9.1	概述	(238)
9.2	骨架工具系统	(239)
9.2.1	EMYCIN	(240)
9.2.2	EXPERT	(242)
9.3	工具语言	(244)
9.3.1	OPS 5	(244)
9.3.2	HEARSAY - III	(247)
9.4	综合工具	(249)
9.5	辅助工具	(256)
9.6	集成工具系统 3MT	(262)
9.7	设计工具应考虑的一些问题	(271)
9.8	问题与结论	(272)
第十章	知识系统的评价	(274)
10.1	概述	(274)
10.2	评价的内容	(275)
10.3	质量评价	(278)
10.4	评价	(282)
10.5	评价工具	(285)
10.6	求精评价实例	(287)
10.7	理论研究的评价实例	(291)
10.8	问题与结论	(293)
参考文献 (一)		(295)
参考文献 (二)		(297)

第一章 知识工程概述

本章作为全书的开头,对知识工程作了概貌性描述。首先综述了知识工程的含义、目标和特点,并将知识工程的研究内容分为基础研究、知识系统的开发研究、知识工程环境研究及与智能机和自动化相关课题的研究四个方面。第二节全面叙述了知识工程学科的产生背景和发展现状。智能问题求解和知识的工程化构成了知识工程的两大基石,因此,第三节对这两个问题作了详细讨论。知识工程的目标是构造出高性能的知识系统,这构成了本章第四节的主题。这一节对知识系统的主要类型——专家系统的特征、种类和结构作了扼要介绍。

通过本章的学习,读者将对知识工程这门新学科有一个比较全面的了解。本书以后各章对第一章的内容作了深化。

1.1 何谓知识工程

知识工程(以下简称KE)这个概念,产生于本世纪70年代中期。当时专家系统的研究和开发已取得一定成果,但在方法论上还存在许多问题,迫切需把专家系统的建造提高到工程的高度来认识。为此,美国人工智能(以下简称AI)研究的开创者之一J. McCarthy提出了“认识论工程”的概念,试图概括建造专家系统的有关技术和方法。之后不久,英国AI研究的先驱D. Michie将此概念的含义加以扩充,并改用“知识工程”这一名词(Michie, 1979)。1977年,美国著名AI专家E. Feigenbaum在第五届国际人工智能会议(以下简称IJCAI)上,以《人工智能的艺术:知识工程的课题及实例研究》为题,对知识工程作了全面论述(Feigenbaum, 1977)。从此,知识工程这个术语便为全世界广泛使用。对专家系统和知识工程的发展作出开创性贡献的Feigenbaum教授也被誉为“知识工程”的鼻祖。

按照大多数人的理解,知识工程是研究如何使用AI的原理和方法构造专家系统的一门工程性学科。与构造专家系统有关的技术、方法和工具都是知识工程的研究内容。知识工程是AI的应用分支,或叫做“应用AI”。什么是专家系统呢?专家系统是一个智能程序,它能对那些需要专家知识才能解决的应用难题提供专家水平的解答(Michie, 1979; Feigenbaum, 1977; Hayes-Roth et al, 1983)。这里,专家知识一般是指专家解决问题所使用的经验性、判断性知识(或称启发式知识)。

我们觉得,从学科发展的角度看,应对上述理解加以扩充。实际上,目前国内外的研究正在对上述的“知识工程”含义进行扩充。比如对专家系统中知识的考虑已不局限于专家的经验性、判断性知识,还包括了常识性知识及专门领域的“基本原理”。关于这一点,Lenat和Feigenbaum作了清楚而全面的论述(Lenat, Feigenbaum, 1987)。事实

上,早在1984年, Hayes-Roth在《基于知识的专家系统》一文中对这个问题就作了讨论(Hayes-Roth, 1984)。

我们认为,最为重要的一点是,知识工程的研究目标应是构造出高性能的知识系统。专家系统只是知识系统的一种类型。知识系统所包括的知识不论是从广度还是从深度讲都应超过现有专家系统所含的知识。自然,他们的问题求解能力也会更强。其次,知识工程一方面用成熟的AI技术和其他相关技术构造知识系统,另一方面也研究AI的核心问题,如知识表示、推理、知识获取和学习等。第三,知识工程技术构成了智能机(日本称作“知识信息处理系统”)的核心技术。

更具体地讲,知识工程的研究应包括以下四个方面:

- ①基础研究;
- ②实用知识系统的开发研究;
- ③知识工程环境研究;
- ④与智能机和自动化相关课题的研究。

基础研究主要包括知识工程中基本的理论和方法的研究。比如关于知识的本质、分类、结构和效用的研究,关于知识表示方法和语言的研究,关于知识获取和学习方法的研究,关于推理和控制机制的研究,关于解释和接口模型的研究,以及关于认知模型的研究等等。

系统开发研究强调解决在建造实用知识系统过程中碰到的实际技术问题,强调实用技术。它以知识系统的广泛实用化和商品化为最终目标。研究的主要内容有:实用知识获取技术,知识系统体系结构,实用知识表示方法和知识库结构,实用推理技术,实用解释技术,实用接口设计技术,知识库管理技术,知识系统调试、分析与评价技术,知识系统的硬件实现技术等等。

知识工程环境研究主要是为实际系统的开发提供一些良好的工具和手段。好的环境可以缩短知识系统的研制周期,提高知识系统的研制质量,使知识系统的研制从手工作坊方式转变为工业化生产方式,加速知识系统的商品化进程。环境研究主要包括:知识工程的基本支撑硬件和软件,知识工程语言(包括知识描述语言和系统结构设计语言),知识获取工具(特别是半自动或自动知识获取系统),骨架工具系统,知识库管理工具(如知识库一致性、完备性的检查工具,知识库性能测试工具,知识库操作语言),接口设计工具和解释工具以及上述工具的集成化工具(或称综合工具)等等。

1.2 知识工程的产生与发展

如果我们以专家系统的开发作为知识工程历史的主线,则可以将知识工程的历史分为以下三个阶段。

一、萌芽期 (1965年以前)

AI专家们通过许多实验研究发现,要想使AI系统能解决现实世界中的复杂问题,它不能光有强有力的一般问题求解方法,还需要领域中的专门知识。尽管计算机在许多方面

超过人类,如速度和一致性,但这并不能弥补它的无知。要想使高速计算机工作得象人类专家那样好,必须为它提供人类专家所具有的那些专门知识。正是在这样的背景下,1965年,美国斯坦福大学 E. Feigenbaum, R. Lindsay 等人研制出世界上第一个专家系统 DENDRAL (Barr et al, 1981), 这是将一般问题求解策略与专家的专门知识和经验有效结合起来解决现实问题的初步尝试,它标志着“知识工程”这个新领域的正式诞生。

二、形成期(1965~1977年)

这个时期研究的主流是开发了许多专门领域的专家系统,并由此提出了一些经验方法和工具。最典型的例子当推1976年由 E. Shortliffe 等人研制的 MYCIN 系统,这是一个用于细菌感染疾病诊治的系统。该系统由于它的结构清晰性和高性能而获得许多知识工程专家和医学专家的承认。在工具方面,最著名的工作应推 R. Davis 于1976年研制的知识获取工具 Teiresias。这是一个与 MYCIN 系统配套的系统。

1977年, Feigenbaum 在 IJCAI-5 上所作的特邀报告中,对1965~1977年专家系统的发展作了全面总结和评价,并由此提出了知识工程的概念,用大量的证据说明了知识工程已经成为 AI 中一个最有发展前途的子领域。Feigenbaum 的这个报告对后来专家系统的大发展起了很大的促进作用。

三、发展期(1978年至现在)

这个时期的特点,我们想从基础研究、系统开发和环境研究三个方面来分别论述。

1. 知识工程的理论和方法不断丰富

知识表示、推理和系统结构是专家系统研究的三个核心问题。近几年,对这些问题的研究不仅强调了系统实现,还有人在理论和方法上作了大量探索。

在知识表示方面,除了对原有知识表示方法(全面的论述见 Barr et al, 1981)作更深入的研究之外,加强了知识表示本质的研究,加强了新表示方法的开发及原有表示法的综合应用研究,加强了表示语言和表示系统的研究。出现了功能表示法(Levesque, 1984),面向对象的表示法(Tokoro et al, 1984)及各种非规范逻辑表示法,如模糊逻辑、模式逻辑、过程逻辑、内省逻辑、缺值逻辑等。

在推理方面,研究了分类学推理,基于模型的推理,并行推理,归纳推理,定性推理,非单调推理,不确定性推理及各种推理控制策略。近年来对不确定性推理更是给予了极大的关注。关于这一课题的研究, Cohen 和 Gruber 在1984年写的一篇文章中作了很好的总结。他们把不确定性处理方法归纳为三种:

①工程方法:在构造系统时避开不确定性的影响。设计者预测可能对系统性能产生影响的不确定性因素,在问题形式化过程中设法消除其影响。一般是通过相关性假设和完备性假设来达到这一目的的。

②控制方法:通过识别领域中引起不确定性的某些特征,利用控制策略减少不确定性的影响或说明不确定性的来源。主要方法有相关性制导的回溯,最小冒险规划,机缘控制,启发式搜索等。

③并行确定性推理方法:这种方法将推理过程分为两个半独立的过程,一个象没有

不确定性那样处理,另一个确定由前一个过程所得结论的确定程度。这种方法又可细分为三类。(a)可信度方法,包括 MYCIN 系统的“确定性因子”模型, PROSPECTOR 系统的主观 Bayes 模型,可能性理论模型,证据论模型及发生率计算模型。(b)推理维持方法,这种方法是专门为处理由不完备知识所引起的不确定性而提出的。在并行确定性推理的意义上,该方法的第一个推理过程是以相信假设命题为基础的,第二个过程提供仔细撤回假设(如果该假设是错误的)的机制。(c)注记理论(endorsement),由 P.R. Cohen 于1983年提出。它是用非数字信息表示和处理不确定性的一种方法。注记就是对不确定性因素进行明显的标记。该理论需要考虑对可信假设和不可信假设推理的注记,注记的排序,注记的传播,以及利用启发式知识消除不确定性等问题。

系统结构方面, M. Stefik 等人1982年发表的一篇文章(即“专家系统组织导论”),对现有专家系统组织原理和方法分十二种情况作了全面而深入的讨论。我们这里重点介绍一下近一个时期大家讨论的几种结构。

①元知识系统结构:元知识的概念由 Davis 于1976年提出,但目前还没有严格的定义。概括地讲,元知识就是关于知识的知识,即(a)关于知道什么的知识,(b)关于如何使用知道知识的知识。元知识系统可看成是由元级和目标级组合而成的两级系统。元级系统组织和利用元知识进行高层推理,指导目标级系统的运行。目标级系统就是目前通常意义上的专家系统,它利用专门领域的知识解决问题。

②分布式系统结构:由多个物理上分离的处理节点组成,每个处理节点都是一个较小规模的知识系统,它们或只具有部分描述整个问题状态的数据,或只有部分领域知识,或两者都是不完全的。大多利用“黑板”结构。分布式结构的优点是能提高实时反应速度,增强可靠性、灵活性及问题求解能力,降低解决问题的成本。

③多推理机结构:一个系统包括多个推理机。各个推理机以松耦合方式协同工作。

④多知识表示结构:在一个系统中,根据各种知识的特点和使用情况,用与其相应的表示方法来表示这些知识。同时存有多种知识表示形式。

⑤多知识表示、多推理机结构。

⑥多层系统结构:包括表层、深层两层系统。表层系统包含经验性、判断性知识,深层系统包含原理性知识。这两个层次的关系是:表层知识直接用于解决问题或制导深层知识的使用。深层知识可用来制导表层知识的使用,也可用于弥补表层知识的不足,即在使用表层知识不能解决一个问题时,再使用相关的深层知识;还可用来直接进行问题求解。关于表层系统、深层系统的含义及两者之间的关系,现在还无统一的理解,本书后面还将讲到这一点。

知识获取、知识库的管理和维护、用户接口及解释等都是在实现专家系统的实用化、商品化过程中要解决的重要问题。近几年,对这些问题,人们也作了许多探讨。

知识获取方面,目前已探讨了四种方法:

①非正式会谈:口头备忘录是获取关于专家求解问题的知识的一种重要途径。但要慎重使用。

②机器归纳:应用归纳算法和知识从问题域抽取的大量实例中归纳出规则。这种方法目前只能归纳出简单的规则,还存在许多问题有待解决。比如,往往得不到完整的训练

实例数据库；需要人提供构成规则的描述空间；要经过长期反复的试验，得到的规则缺乏灵活性；得出的规则往往是极其复杂和难理解的，从而影响系统的透明度；等等。

③辅助工具：可分为三类。(a) 知识库编辑程序和接口，它可以简化向系统输入知识的任务，进行语法和部分语义检查以减少出错机会。典型的例子有KAS；(b) 解释设施，专家利用它可以对求解问题中所使用的规则进行分析和修改。典型的例子有EMYCIN。(c) 知识库调试和求精系统：分辅助方式和自动方式两种。典型的例子有SEEK，SEEK2，TEIRESIAS等。

④学习系统。包括机械学习，示教学习，从实例中学习，类比学习等。

解释方面，早期的方法有预制文本法和执行追踪法。MYCIN系统使用了这种解释方法。近年人们又提出了几种新的方法。如Hasling 1984年提出的策略解释法将控制知识明确表示出来，与领域规则完全分开。在一般策略的层次上产生解释。Swartout 1983年提出的自动程序员解释法将自动程序设计思想用于解释。

用户接口方面，探讨了三种接口类型。一种是“用户适应类”接口：用户在与系统通信中不断调整自己的行为，以适应系统中固定的用户模型。另一种是“系统适应类”接口：系统动态地适应用户，以用户认为最方便的形式和用户通信。第三种是上述两类接口的综合。目前大部分系统使用的接口都属于第一种。近年，特别重视自然语言会话模型、用户模型及接口设计工具的研究。在1985年的国际人工智能会议上还对用户模型展开了专题讨论，提出了参数模型法、状态机描述法、递归模型法和计划模型法等。专家系统的自然语言接口设计工具国外已有商品出售。在我国，由于汉语自然语言理解还有许多问题有待解决，建造商品化的汉语接口尚需大量艰苦细致的工作。

近年，知识工程的研究内容也开始拓广。已有人探讨了百科知识系统(Lenat et al, 1983)、一般知识系统、知识网等问题(Hayes-Roth, 1984)。

根据上面的介绍，我们可以说，知识工程的理论和方法确实很丰富了。但也要看到，知识工程作为一个领域，离成熟还很远很远。

2. 专家系统作为知识系统的主要类型，其开发范围愈来愈广，而且正在走向实用化和商品化。

70年代，专家系统的开发仅限于医学、化学、数学等少数领域，而且主要是为了研究的目的。现在，专家系统的开发几乎遍及社会生产和生活的每一个领域。Waterman 1986年写的《专家系统指南》一书中给出了一份详细的专家系统名录。下面我们根据本书的材料介绍几个已商品化的专家系统。

① SEQ：该系统帮助分子生物学家进行多种核苷酸序列分析。它最初是由斯坦福大学作为MOLGEN计划的一部分开发的，后来(1981年)由Intellicorp公司进一步发展成商品。它用LISP语言实现。

② ISA：该系统根据现在和将来计算机设备的配发情况排定用户订购计算机系统的订单。1984年由DEC公司开发。用OPS 5实现。

③ TIMM/TUNER：该系统帮助用户调整VAX/VMS计算机系统，以适应经常变化的计算机环境。1983年由General Research Corporation公司开发。这里的TIMM是一个商品化自动知识获取系统。

④ R1 (XCON): VAX 计算机配置系统。

⑤ ACE: 它是一个发现电话网中的故障并提出修理和维护建议的系统。1983年由贝尔实验室开发。用 OPS 和 Franz Lisp 实现。

⑥ MACSYMA: 数学专家系统。

3. 知识工程工具的发展最引人注目。不仅数量大、种类多、功能强, 而且正在逐步走向实用化和商品化

根据 Waterman 的《专家系统指南》一书提供的材料, 到1986年为止, 开发的知识工程工具总共有95个。其中商品化的工具有30个。对这些商品化工具, 按其所使用的知识表示方式具体列举如下:

①系统建造辅助工具: ACLS, EXPERT-EASE, IN-ATE/KE, PICON, PLUME, RULEMASTER, TIMM, T. 1 等;

②基于框架的语言: KEE, SRLT 等;

③基于逻辑的语言: APES, DUCK, PROLOG 等;

④面向对象的语言: FLAVORS, SMALLTALK-80 等;

⑤面向过程的语言: COMMON LISP, INTERLISP, INTERLISP-D, LISP, ZETALISP 等;

⑥基于规则的语言: ART, KES, M. 1, OPS5, OPS 83, PERSONAL CONSULTANT, RADIAL, SAVOIR, S. 1 等。

上述工具的详细情况可从《专家系统指南》(Waterman, 1986) 中得知。

上面我们从三个方面对知识工程发展期的特点作了讨论。介绍的基本上都是国外的情况。国内知识工程研究起步较晚, 但发展也是很快的。象高等院校的吉林大学、浙江大学、清华大学、北京大学、南京大学、武汉大学, 北京航空学院、西安交通大学、中科院的数学所、计算所、自动化所、沈阳自动化所、合肥智能所等许多单位, 在这方面都做了大量扎扎实实的工作。在大家的努力下, 我国的知识工程研究必将在一个较短的时间内接近并赶上世界先进水平。

1.3 知识工程的基础

知识工程和大多数工程领域一样, 既有理论, 也有实践; 既是一门科学, 也是一项技术, 还是一种方法论。目前, 这门学科已表现出下述特点:

①由于知识系统求解通常需要人的智能才能解答的问题, 所以它们显示了与大多数人工的或自然的智能问题求解系统相同的性质。

②知识系统的组织和设计是一个复杂的工程问题, 必须有大量的一般原理作为指导。知识工程尽管出现的时间不长, 但在这方面已给出了许多有用的思想。

③虽然知识能用来产生智能性动作, 但是它自身并没有给出认识和发掘这种潜在能力的方法。因此, 在建造一个实际系统时, 知识工程师必须将知识工程化, 即把它转换成可供知识系统直接使用的内部表示形式。下面的1.3.1和1.3.2两小节详细讨论知识的工程化和智能问题求解, 这两方面构成了知识工程的两大基石。1.4节全面叙述知识系统。

1.3.1 知识和知识的工程化

知识工程中的基本假设是“知识就是力量”。只有大量的专门知识才能使知识系统在解决复杂的难题时达到高性能，而具体的推理方法只是为系统使用这些知识提供了某种形式化的手段。那么，什么是知识呢？

一. 知识

一般地说，知识是特定领域中的描述、关系和过程 (Bernstein, 1977)。

描述——知识库中的描述识别和区分对象及对象的类。它是用某种语言写成的。这种语言的基本元素是一些原始特征或概念。描述系统一般还包括在专门应用中使用和解释上述描述的规则和过程。

关系——它是知识库中的一种特定描述。它表示知识库中对象间的依赖性和关联性。典型关系有分类学关系（层次关系等）、定义性关系和经验性联想关系（如医生看病所使用的因果模型）。

过程——在推理或问题求解时规定要执行的操作。

简言之，知识包括符号描述和过程两大部分。前者刻画一领域中的定义性和经验性关系，包括上面的描述和关系。实际上描述是关于对象及其类的个体描述，关系是关于对象间的关系的描述。后者（即过程）是对符号描述进行各种操作。

知识和技巧不同。技巧意味着有正确的知识并能有效地使用它。任务的技巧性执行要涉及许多在任务的非技巧性执行中不存在的特征，如较高的效率、较少的错误、较高的自适应性和较好的完美性。这些特征反映了知识和工程两个方面。知识工程的目标，实际上就是建造技巧性计算机系统，它首先从知识源提取知识，然后组织和使用它。

知识的工程化，是构造知识系统的关键所在，也是知识工程学科研究的核心问题之一。目前的知识工程技术，只能将人类知识的一小部分工程化，而对人们在解决复杂问题时常用的许多重要知识却无能为力。表1.1给出了工程化知识的可能类型。左边表示目前知识系统可以使用的知识类型，右边表示目前尚在研究中的知识类型。下面对表1.1右边所列知识类型作些解释。

表 1.1 工程化知识的可能类型

现 在 实 用 的	不 可 用 的
定义和分类学知识	一般问题求解知识
离散描述	类比
简单的约束和不变量	朴素 物理学
经验性联想图	元表示知识
感觉结构	元知识
演绎方法	基本原理
简单归纳方法	知识的编译和逆编译
很简单的实体模型	
很简单的启发性搜索	

①对一般问题求解知识,知识工程师通常将其直接纳入知识系统本身的结构和推理机中,而未能明显表示出来。

②知识系统不具备类比能力,主要是因为迄今为止还没有任何有效的形式手段使知识系统能用类比来辅助问题求解。

③朴素物理学知识指人们日常使用的常识性知识,连小孩都具备一些这样的知识。

④元表示知识指示了表示知识的最佳方式。

⑤元知识描述并指示如何使用知识。

⑥基本原理构成任何一个学科领域的最核心部分。知识系统要想最有效地解决某专门领域中的问题,必须知道该领域的基本原理。

⑦知识编译是指把知识的高级外部形式转换成可由系统直接使用的内部形式,知识的逆编译则是指把知识的内部形式(系统形式)转换成可为用户所理解的高级形式。追求速度的系统必须编译自己的知识,追求理解知识的系统必须维持知识的一种高级形式或对编译过的知识实行逆编译。

知识工程化具有两个明显的特点:

①知识工程师一点一滴从知识源获取知识并把它综合到知识库中,因而知识工程师是用基本知识单元(即知识片)来构造知识系统的;

②知识系统利用知识求解问题的方式直接影响知识工程师怎样获取、表示和综合知识。

二、知识的阈值

上面我们只是说,知识对知识系统的性能影响极大。那么,知识系统所拥有的知识量与其所具有的问题求解能力之间到底存在一种什么样的数学关系呢?D.Lenat和E. Feigenbaum在“论知识的阈值”(Lenat et al,1987)一文中对此作了重要论述,提出了“知识的原理”和“知识的阈值”等重要观点。

1. 知识的原理

一个系统能显示出高水平的智能,其原因是它具有领域专门知识并能很好地运用这种知识。这里的智能是指在巨大的搜索空间中,系统迅速找到合适解的能力。系统拥有的领域知识愈多,求解问题的能力就愈强。

2 知识的阈值

知识的阈值,更准确地说是知识的数量阈值,它可用来测定系统在一个专门领域中所能达到的问题求解能力。Lenat和Feigenbaum区分了以下三种阈值:

①合式阈值(用W表示)

合成阈值明确表达一个任务所需要的最小知识量。

②能力阈值(用C表示)

能力阈值是求解难题所需要的知识量。具体地讲,如果系统拥有知识量使它足以解决一个领域中的大部分问题(约90%左右),那么,这个知识量就被称作能力阈值。这时的知识系统可以认为达到了该领域的一名人类专家的能力。

③全体专家阈值(用E表示):

这是一个领域的所有人类专家的专门知识的总和。

上述三种阈值与问题求解能力之间的关系如图 1.1 所示。

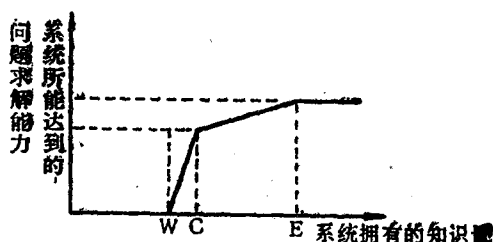


图1.1 知识系统的知识量与其问题求解能力之间的关系

从图1.1中可以看到，只要系统拥有的知识量在点W和C之间，那么，每给系统增加一个知识片，都将使系统的性能明显提高。在系统的知识量达到点C后，系统就能解决一个领域的大部分问题，从而达到了该领域一个人类专家的水平。当系统拥有的知识量超过C点后，再增加知识所带来的性能改善将不如前面那么明显。这时添加的知识尽管有用但不常用，它们只是用于解决比较罕见的问题。如果系统的知识量达到了E点，这时系统就可以解决该领域的几乎所有罕见的和困难的问题。从而系统的能力就超过了该领域任何一个人类专家的能力。

1.3.2 智能问题求解

在问题求解的意义下，所谓知识是指那些能改进问题求解效率和效果的数据，包括表示有效命题的事实，表示似然命题的信念以及在不存在有效算法时作出最佳判断所使用的启发性知识。

专家之所以称其为专家，就是因为他们所具有的上述几类知识不仅数量大而且质量高。在面临一个复杂问题而可用数据又是不确定的时候，专家能够灵活运用自己的知识并在合理的资源消耗前提下找到一种能产生可接受答案（不一定是最佳答案）的方法。知识系统处理的是复杂的智能性问题，因而它们所处的环境也是一种不能肯定找到最优解或正确解的环境，需要在所给答案的质量与所作的努力之间寻求某种平衡。

在这种情况下，知识系统要想达到高性能，改进其问题求解的效率是关键的一环。特别是速度等方面的改进能使知识系统很快地产生可为用户所接受的解，从而使知识系统能在允许的时间内找到更好的解。

表1.2给出了智能问题求解的基本思想。

表1.2 智能问题求解的基本思想

△知识=事实+信念+启发性知识。

△成功=利用一切可用资源找到一个“足够好”的解。

△搜索效率直接影响问题求解的成功。

△改进搜索效率的措施：

1. 利用一切可用的正确知识避免失误及区分问题求解的不同环境；
2. 及时排除无用的候选解类，进而排除无用的求解路径；
3. 使用中间结果消除冗余计算；