



知识工程 原理

曹立明 陈石麟 周强 编著

*Zhishi
Gongcheng
Yuanli*

中国矿业大学出版社

知识工程原理

曹立明 陈石麟 周 强 编著

中国矿业大学出版社

内 容 提 要

本书以知识工程的三个基本研究课题(知识表示、知识利用和知识获取)为核心,阐明了有关的基本概念、原理和技术,以及如何应用这些知识设计、实现和维护一个系统。特别对专家系统的构造、种类、工具及建造的过程做了全面论述,帮助读者理解和掌握知识工程原理。

本书可作为理工科大学计算机及其它相关专业的高年级本科生和研究生的教材,也可供教师和从事专家系统开发的软件人员参考。

责任编辑 孙树朴

技术设计 关湘雯

知识工程原理

曾立明 陈石麟 周强 编著

中国矿业大学出版社出版发行

新华书店经销 江苏省赣中印刷厂印刷

开本 850×1168毫米 1/32 印张 6.125 字数 153 千字

1995年12月第一版 1995年12月第一次印刷

印数:1-2000册

ISBN 7 - 81040 - 457 - 1

TP·22

定价:9.00元

前 言

自 1965 年开发的第一个专家系统 DENDRAL 问世以来,专家系统的研究和开发取得了迅猛发展,各种不同领域的专家系统不断涌现,它们正被广泛地应用于军事、科技、企业和卫生等部门,引起人们极大重视。

专家系统的深入研究必然要涉及知识的表示、理解、学习、获取和利用等问题,这些就是知识工程要研究的内容。知识工程主要研究在计算机中组织知识,建立高质量知识库,使计算机获取与搜索到有用的知识以及使用知识求解问题的原理、技术和方法。因此,知识工程的研究必将为专家系统的建造打下良好基础。

本书共分六章,分别阐述了知识、问题和搜索、知识表示、知识利用、知识获取和专家系统的基本概念、原理和实现技术。

本书力求叙述简练,尽量对每个概念给出明确的定义,并通过大量实例加以说明。对知识表示、推理、专家系统的设计等重要的或在学习和运用中容易遇到困难的方法都用 PROLOG 语言实现,帮助读者掌握要点,使读者能在自己的应用领域较自如地运用这些技术开发专家系统。

本书曾作为研究生教材,几经修改,最后由曹立明、陈石麟和周强三位副教授合作完成。书中有不足和错误之处,敬请读者批评指正。

作者

1995 年 10 月

目 录

前言.....	(1)
第一章 知识工程引论	(1)
第一节 什么是知识工程.....	(1)
第二节 知识工程的研究内容.....	(2)
第三节 知识及其分类.....	(3)
第四节 知识系统.....	(5)
第五节 智能工程.....	(8)
第六节 知识工程发展方向	(10)
第二章 问题和搜索	(12)
第一节 问题	(12)
第二节 广度优先搜索	(16)
第三节 回溯和深度优先搜索	(17)
第四节 启发式搜索	(20)
第五节 显式图和隐式图	(24)
第六节 A* 算法	(26)
第七节 面向对象方法描述问题	(31)
第三章 知识表示	(35)
第一节 概论	(35)
第二节 产生式系统	(36)
第三节 谓词逻辑表达	(39)
第四节 PROLOG	(41)
第五节 语义网	(58)
第六节 框架	(63)

第七节	框架系统及其实现	(70)
第八节	脚本	(77)
第四章	知识利用	(80)
第一节	归纳推理、演绎推理和类比推理	(80)
第二节	正向推理和反向推理	(82)
第三节	元推理	(87)
第四节	不精确推理	(90)
第五节	非单调推理	(91)
第六节	模糊推理	(95)
第五章	知识获取和学习	(103)
第一节	概论	(103)
第二节	机器学习方法	(111)
第三节	知识获取要考虑的问题	(121)
第六章	专家系统	(122)
第一节	引论	(122)
第二节	专家系统的构造	(124)
第三节	专家系统的种类	(126)
第四节	诊断的原理和方法	(129)
第五节	智能决策支持系统	(133)
第六节	几个著名的专家系统	(139)
第七节	专家系统的设计与实现	(146)
第八节	专家系统开发工具	(168)
第九节	黑板模型	(173)
第十节	分布式专家系统	(181)
第十一节	专家系统的建造	(182)
参考文献		(187)

第一章 知识工程引论

第一节 什么是知识工程

在政治、经济、科技、文化等一切领域,人们遇到大量形形色色的问题。解决问题需要知识和经验。知识工程就是研究如何获取、表达、利用知识和经验,进行分析、决策、规划、设计和实施,以期在现有的条件和规定的时间内合理地解决问题。

“知识工程”一词最初由斯坦福大学的费根鲍姆教授在1977年“第五届人工智能国际会议”的演讲中提出的。他说:“知识工程是人们积极利用经验知识以解决问题作为前提并与人工智能应用有关的一个研究领域。”人工智能是计算机科学的一个分支,主要研究使用计算机模拟人类的智能活动。因此,知识工程的研究和计算机密切相关,它要求问题最终用计算机解决。所以,知识工程主要研究如何在计算机中组织知识,建立高质量知识库,如何使计算机获取与搜索到有用的知识,以及使用知识来解决问题。

在不同的领域和行业,有不同的描述问题的语言或“行话”、不同的解决问题的方法。知识工程的产生同人工智能面向解决实际问题,特别同专家系统的发展紧密相连。知识工程研究通用的问题表达和解决办法。

知识工程所要进行的信息处理是以知识为对象,它比通常以数据为主的信息处理涉及的问题要广泛得多、复杂得多。在产业界,解决问题和系统控制的对象发生了巨大的变化。过去是封闭的、小规模的系统控制,而现在的控制对象已逐步扩大为大规模的

系统。知识工程的研究对象主要是复杂的特大系统。特大系统一般都是非结构化的,因此知识工程处理问题的方式和结构化系统完全不同。表 1-1 揭示了两种不同结构问题的特征和处理方式。

所谓非结构化问题是使用的知识范围、知识间的相互关系,很难预先从总体上加以掌握并且很难在形式上描述的问题。由此可见,知识工程在理论和应用领域都有广阔的发展前景。

表 1-1 问题特征和解决方法

问题	结构化	非结构化
特征	1. 数值的 2. 相位的 3. 反复的 4. 闭信息的	1. 记号的 2. 组合的 3. 一次的 4. 开放信息的
逼近方法	系统科学	知识工程
推理方法	1. 理论的 2. 控制驱动的 3. 模型库	1. 经验的 2. 事件驱动的 3. 知识库

第二节 知识工程的研究内容

知识工程的研究内容主要包括三个方面:

1. 知识表示

知识表示是指在某一专门领域内,把教科书式的事实知识和只有特定专家所具有的经验知识形式化,使计算机能够接受和对它们进行操作。它涉及知识库的设计和管理。

2. 知识利用

知识利用是在某一数据结构的基础上,利用形式化的知识解

决问题。这涉及推理机的设计问题。

知识表示和知识利用如同数据与算法,存在着表里一体的关系。

3. 知识获取

知识获取是对于假设的对象问题领域,把事实知识和经验知识移植到知识库。它主要包括支持从专家那里抽取知识、维持知识库的完整性、新旧知识的一致性、知识编辑和机器学习等课题。

围绕着这三方面有一系列软件、硬件系统和方法论、基础理论的研究。例如,知识工程软件系统的研制和开发,智能用户接口,开发工具、环境的研究,智能计算机系统的研制与开发,分布式问题求解,分布式推理机制、控制与协调问题的研究等。

第三节 知识及其分类

知识是事实和关系的集合,它以高度概括的形式揭示研究对象的属性和相互关系,用来解决实际问题 and 从事创造活动。个人或计算机具备的知识质量可用解决问题和创造能力来评判。

人类处理的知识是极其多样的,其中最主要的类别可列举如下:

1. 对象知识

对象知识是关于对象的有关事实的知识。例如:“鲸鱼是哺乳动物”,“企鹅是一种鸟,但企鹅不会飞”,等等。对此,需要有表示对象、对象集合、类别和描述对象的方法。

2. 事件知识

事件知识是关于过去发生、现在发生或将要发生的事件的知识。对此,不仅要把事件本身符号化,并要表示事件发生的时间、地点、状态、性质和因果关系。例如:“昨天下午他因病没有参加比赛。”

3. 行为知识

行为知识是行为者之间、行为者和行为对象之间在事件中相互作用时所表现的行为的知识。例如,狗在条件反射实验中的各种反应的知识等。

4. 工序知识

工序知识是记述进行某些行为的操作步骤的知识,要求表达起动条件、顺序关系、因果关系等。如描述汽车驾驶和烹调过程时运用的知识。

5. 元知识

元知识是关于知识的知识,用来指导如何选择事实和规则(关系)的知识。这是比解决具体问题所需要的知识更高层次的带有指导意义的战略层次的知识。例如,“物极必反”是元知识。

元知识包括选择、评价、说明、诊断或检查知识去解决实际问题所需要的知识和开拓新领域需要的知识。

在知识工程研究中,通常把问题求解所需要的知识分为三类。

1. 陈述型知识

这类知识叙述系统状态、环境和条件,问题的概念、定义和事实的知识。

前面论述的对象知识、事件知识和行为知识属于陈述型知识。

2. 过程型知识

过程型知识是有关系统状态变化、问题求解过程的操作、演算和行动的知识。

工序知识属于这类知识。

3. 控制型知识

控制型知识是关于如何选择相应的操作、演算和行为的比较、判断、管理和决策的知识。

这三类知识分成不同的层次。控制型知识是指导性的知识,处于最高层,用来指导进行搜索、判断、管理和决策。因此控制型知识

是一种元知识。

知识系统具有的元知识越丰富,它的性能越好,工作效率越高,智能水平也越高。

数据、信息和知识是三个不同层次的概念。

数据是信息的素材,它是从自然现象中搜集的原始素材。人们根据一定的目的按一定的形式加工处理数据就形成信息。

数据通过收集获取,以是或不是的方式表现;信息通过整理、分析和统计数据获取,以 what, which, who, when, how many, how much 方式表现;知识通过对信息进行概括和推理的方式获取,是以 why 和 how 的形式记忆的。

信息有如下特点:

(1) 可塑性:可产生与消失,可修改和复制,可组合迭加或剪裁;

(2) 可相关性:按接受者需要或观点与它物相关;

(3) 可转换性:根据相关性作联想转换。

例如,“苹果落地”这个现象对牛顿是一个极其重要的信息。牛顿把它和地心引力相关,进一步联想转换到天体力学,导出万有引力定律。

智能是比知识更高层次的概念。

智能是能够领会事物的道理,正确运用信息、知识、经验和悟性,进行判断、计划、处理事务以及健全行动的能力。

智能是人内在的能力,是永存的财富。而知识是暂时的、变动的、不完全的、不精确的,而且不断老化。

第四节 知识系统

知识工程实现的系统称为基于知识的系统,简称知识系统。它是为了完成一项任务或若干任务而去处理知识的系统。

要说明一个知识系统必须提供下面两点：

- (1) 系统的基本知识库的规格说明；
- (2) 系统要执行任务的规格说明。

知识库中的知识是高度结构化的符号数据，用于表示数据元素和数据元素构成的方法及其关系的模型。知识库的性能取决于知识的质量(结构化、完整性、有效性和一致性)以及应用这些知识的方法。

知识系统实现的一般方法包括：

- (1) 以某种合适的数据结构表示知识库；
- (2) 有一个解释程序用在这些数据结构上执行各种操作，提取信息；
- (3) 一些过程的合成，它们调用解释程序提供适于完成全局系统任务的信息。

知识系统将沿着多功能、通用、综合的方向发展。

专家系统是一种典型的知识系统。它是一个已被赋予知识和才能的计算机程序，用以解决某种特定的、难度较高的问题。

(1) 这类问题是现实世界提出的需要由专家来分析和判断的问题；

(2) 问题通过模拟专家推理过程来解决，并且要求得到的结论和专家相同。

专家系统是一个高级问题求解系统，可具有解释、判断、预测、决策、指导、监控、分析、咨询、规划或设计等功能。

专家系统最基本的构成如图 1-1 所示，包括知识库、推理机、数据库和人机接口。

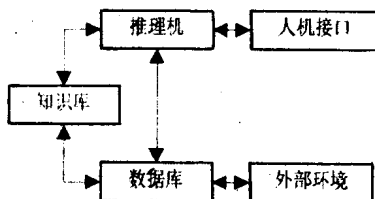


图 1-1

- (1) 知识库

知识库是用产生式规则或其它结构记录的某一领域的理论知识和专家经验知识。

(2) 推理机

推理机是为了完成解决问题的控制部分。它的功能是找出与解题有关的知识,用这些知识一步步地构成解题推理路线,作出可能的解释或最佳推论。

(3) 数据库

数据库指容纳来自外界的输入以及推理结果的部分。

(4) 人机接口

人机接口是通过精心设计好的问题进行人机对话,获取必要信息的系统。其目的是使计算机的功能尽可能地“自然”。它采用的语言尽可能接近人类的自然语言或者接近某一领域的特殊语言。

在知识库中必须存放高质量的启发式规则。这是专家系统是否有效的重要因素。启发式规则是发掘了专家在求解问题时使用的启发式方法而编写的规则。专家新颖的启发式方法并不是以 if-then 的形式表示的,而建造一个知识系统的重要步骤之一是把专家的启发式知识转化为规则。知识系统的功能就是在知识库中体现启发式规则。

启发式方法也称为探试方法。“探试”(heuristic)一词起源于希腊文 eureka(“探索”),用来表示一种经验法则或善于推测的法则。这是一种简化的方法,用来化简或限制大型问题空间的搜索。启发式方法不象普通的算法那样能够绝对保证获得预期效果,但是这种方法给出的结果多半是特别有用的。

费根鲍姆强调“知识工程是把人们积极利用经验知识解决问题作为前提…”也就是说,知识工程是以研究启发式方法作为前提的。

知识库中存贮的知识可划分为深层知识和表层知识。基本理

论、重要原理、公理和领域的事实称为深层知识。用来解决实际问题的经验知识称为表层知识或启发式知识。它通常指一个特殊领域或任务所包含的限定事实、理论及大量启发式规则。

专家系统必须具有解释功能。“解释”是一个特殊推理或动作过程的说明。在知识系统中,这是帮助用户理解系统如何工作的一套技术。很多知识系统允许用户提问“Why”,“How”,“What”,或要求系统“Explain”等。对每种提问,系统以显示它的假设和内部推理过程作为回答。

专家系统的研究分成两大类:

(1) 基础研究:计算机技术和自动化专业人员进行的对专家系统技术本身的研究,包括专家系统中的有关理论和由于应用引起的一些新问题的研究,解决应用领域中有关共性的专家系统开发工具的研究。

(2) 应用研究:解决实际领域的具体问题而进行的专家系统的开发与设计。

第五节 智能工程

1987年,加拿大 ALBERTA 大学人工智能研究室主任 M·RAO 教授首次提出分布式集成智能系统理论(IDIS),有效地克服了传统单一领域专家系统的局限性。IDIS 是一种大规模智能集成环境,即把各种不同的专家系统、神经元网、数据库、计算机软件包和图形处理系统等进行有机集成,解决复杂问题。

这些功能系统可用不同语言单独开发。整个系统的管理和控制由“元系统”完成,它负责各功能系统的通讯、选调和运行等工作。

“元系统”具有自己的数据库、知识库、推理机和多介质界面,并用“元知识”进行控制。它能有效地处理各种知识和各种信息(符

号、数值、图像),并能综合运用各种计算机语言。

分布式集成智能系统的结构如图 1-2 所示。

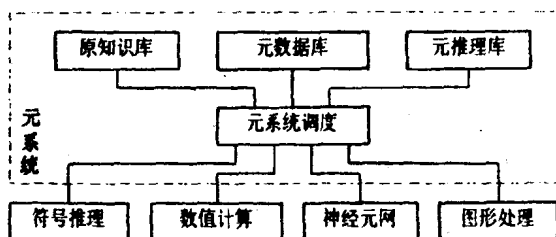


图 1-2

国外在这方面的研究包括：

- (1) 工业过程的在线监督和控制；
- (2) 操作优化；
- (3) 生产调度；
- (4) 减少环境污染和事故处理的智能操作支持系统(IOSS)；
- (5) 开、停车自动控制系统；
- (6) 概念设计自动化系统——使控制系统的设计工作从选取调节参数和系统结构到实施这些方案均能自动进行。

美国有的汽车公司可以根据用户需求输入数据,最后全过程自动化输出用户要求定制的汽车。

从工业生产自动化发展来看,大致分为四个阶段:

- (1) 劳动密集型生产；
- (2) 设备密集型生产；
- (3) 信息密集型生产；
- (4) 知识密集型生产。

目前正进入第三阶段,其重要标志是柔性制造技术的发展,特点是数据信息的处理占统治地位,其中 CAD/CAM 将是关键领

域。

第四阶段包括 CIMS(计算机集成制造系统)、CIPS(计算机集成过程系统)技术,集成了 CAD、CAM、CAPP(计算机辅助生产计划)、CAT(计算机辅助检测)、人工智能等各种技术,使系统能完成从销售、管理、设计到计划、调度、制造和检测等一系列任务。

长期以来,在新技术和它们的应用之间脱节很大,原因之一是对工业应用没有足够重视。智能工程的目的就是要在理论研究和实际应用之间,在人工智能和工程之间架起一座桥梁。

智能工程是一个新的工程和技术领域,它是一个由实际工程问题和人工智能及计算机科学相互交叉和渗透而形成的边缘性领域。它包括两方面内容:

- (1) 将人工智能技术应用于工程问题;
- (2) 从工程背景出发,开展理论基础和应用技术研究。

智能工程和知识工程的区别在于:

智能工程要解决的是工程问题,强调将各个不同领域的知识集成起来,以解决实际问题。

知识工程主要研究知识本身;知识表示、知识获取和知识利用的理论、技术和方法。

第六节 知识工程发展方向

知识工程是在专家系统的基础上发展起来的,随着专家系统的广泛应用和日益深入到各个领域,它向知识工程提出了新的要求。

知识工程的发展方向主要包括下列五个方面:

1. 建立大型知识库

专家系统要解决的问题越来越复杂,要求的知识覆盖面也越来越大,如何建立大型知识库的问题就成为突出的矛盾。美国对此

提出了 CYC 知识共享计划,日本正在实施 EDR,即电子字典研究(Electronic Dictionary Research)计划。

2. 自动获取知识问题

这是如何将研究成果转化为实际可用知识的方法研究。

自动获取知识一直是困扰人们的“瓶颈”问题,目前没有任何突破的进展,很多专家系统的开发者只得暂时避开它,由知识工程师完成后,把获取的知识装入知识库。但从知识工程研究与发展观点,必须逐步解决这个问题。

3. 解决易使用性问题

使用户能更方便、自然地使用专家系统。如系统采用语言输入和多媒体(文字、图形、声音和图像)输出。

4. 专家系统和知识库维护问题

大型知识库的建立必然波及系统维护问题。一般情况,软件维护目前占软件费用的 60%~80%。

一种好的方法是,一开始就由编程人员、设计人员和用户一起协同合作进行系统设计,有关文档全部存放在关于这个软件的知识库中。以后对软件的维护和修改可通过这个知识库进行,它相当于一种在线的文档,可省一大笔维护开支。

5. 建立异构的分布式知识库系统

它通过网络将不同的知识库连起来,采用多专家系统机制进行知识通讯及协同运行。