

知识工程学入门

〔日〕上野晴树 著

[illegible]

科学出版社

73.27237

111

知识工程学入门

〔日〕上野晴树 著

李 东 译

科学出版社

内 容 简 介

本书是以介绍知识工程学的基本概念及方法为目的的入门性读物。书中主要介绍了知识工程学和专家系统的理论、概念,并介绍了产生式系统和框架系统两种主要的知识表现模型,讨论了关于专家系统的研制和发展等问题。对于学习、了解人工智能和知识工程学的理论,具有一定的参考价值。

本书曾获1985年日刊工业新闻社优秀科技图书奖。

本书可供大专院校计算机专业的师生、有关的专业工作者、工程技术人员和知识工程学初学者阅读、参考。

上野晴樹
知識工学入門
オーム社, 1985

知识工程学入门

〔日〕上野晴樹 著

李 东 译

责任编辑 曾美玉

科学出版社出版

北京朝阳门内大街137号

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1989年1月第一版 开本:787×1092 1/32

1989年1月第一次印刷 印张:7 3/8

印数:0001—2,660 字数:162,000

ISBN 7-03-000686-0/TP·45

定 价: 4.80 元

译 者 的 话

知识工程学是人工智能研究的重要领域之一，而专家系统是知识工程学应用的一个主要方面。

自从 60 年代第一个专家系统问世以来，经过 20 多年的研究探索，如今，有关专家系统的理论和方法都已日趋成熟。目前，在诸如医疗诊断、故障诊断、法律咨询、气象预报、计算机辅助设计等许多应用领域中，各种专家系统越来越广泛地被用来代替专家进行复杂的智力工作，在发展社会生产力的过程中显示了它所具有的巨大潜力。

此外，专家系统又是人工智能学科的一个重要的研究领域。由于专家系统虽然是面向应用的一种方法，但是却融合了一系列人工智能的研究成果，所以又把它称为应用性人工智能。专家系统在许多应用领域中所取得的成功也体现了对人工智能进行研究的深远意义。目前，人工智能研究的一个主要目标就是要使得计算机能够具有一定程度的智能，摆脱以往单纯作为计算工具进行数据、信息处理的被动局面，而使计算机所处理的对象提高到知识的水平上来。这就需要对知识表现、推论方法、知识获取等一系列有关的重要课题进行深入的研究。基于这种原因，著名的人工智能专家费根包姆 (Feigenbaum) 在 60 年代就提出可以把对知识的表现、推理、处理、获取等一系列问题综合成为知识工程学 (Knowledge Engineering) 的研究，这种研究在 70 年代得到了很大的发展，进入 80 年代后，对知识工程学的研究更加快了步伐，而且有很大一部分成果已经投入应用，成为当前发达国家的计

计算机行业中所注目的发展方向之一。

日本对人工智能的研究,近年来发展得极为迅速。特别是从1982年日本成立第五代计算机研究中心,确定以人工智能的综合研究成果来设计制造第五代电子计算机之后,人工智能的研究热潮在有关的企业、研究机关和大学中更加高涨了。目前,PROLOG等人工智能程序设计语言已成为日本最流行的程序设计语言。在大学中,不仅计算机类专业的学生要学习人工智能课程,许多非计算机类的学科也纷纷开设了人工智能课程。日本的各大计算机公司也争相研制和生产专家系统、自然语言翻译系统等等。可以说,日本的人工智能研究已达到了世界先进水平。

在我国,目前也正在兴起对人工智能的研究。许多大学都开设了有关课程,在一些应用领域中也出现了一批投入应用的专家系统。尽管如此,我国的人工智能研究从研究的历史到研究的深度、广度都较外国有一定的差距。为了尽快地提高我国人工智能研究水平,培养和造就具有现代科技知识的四化建设人才,为广大读者提供内容丰富的人工智能读物,引导广大科技工作者和高等院校师生尽快掌握这门学科的基础知识已成为当务之急。

本书是以知识工程学的初学者为对象,以介绍知识工程学的基础概念及方法为目的的入门性人工智能读物。作者不是以教科书的形式来编写本书,因而没有使用抽象的数学语言来叙述,而是力图通过深入浅出的一般性叙述,使读者在读后能够把握住知识工程学的主要脉络。全书文笔简明流畅、叙述清晰易懂,极适宜广大计算机专业工作者、大专院校师生以及有关方面的读者阅读。

本书曾获1985年日刊工业新闻社优秀科技图书奖。作者上野晴树教授,是日本人工智能学会理事,是日本人工智能

研究界有代表性的专家之一。

本书是译者在日本东京工业大学学习期间翻译的。在翻译过程中，曾与中国人工智能学会常务理事郭荣江先生多次讨论，并承他在从事研究的百忙之余仔细审校了全部译稿。本书在出版过程中，北京计算机学院吕景瑜副教授对译稿作了认真的审阅，并提出了宝贵意见；徐虹、李双同志也给予了热心的帮助，在此特向他们一并致谢。

本书未附原书中的索引。

由于译者水平有限，时间又很仓促，译文中的不妥之处在所难免，恳切希望广大读者批评指正。

前 言

本书并不是一本介绍有关知识工程学的技术说明书，作者的意图在于向读者介绍一些贯穿在一般的知识工程学中的最本质的内容。技术，总是在变化着的，尤其在知识工程学领域，技术的发展就尤其迅速，以至于人们常用日新月异这样的词来形容它。从表面上看的确如此，但实际上，多年来知识工程学的本质或其基本内容并没有发生令人惊异的变化。尽管在这一领域不断地出现具有划时代意义的新技术，但我们只要仔细想一下，就会发现这只不过是因为已经具备了使这种新技术得以发芽生长的土壤而已。无论是在 1965 年使知识工程学得以创生的 DENDRAL 系统的方案，还是在 1975 年提出的对以后的研究开发产生巨大影响的框架理论，它们都是以突破以往传统方法的某种界限为基本出发点的。

进入 80 年代，知识工程学开始向产业界普及，也有它的原因。主要原因是过去在大学和研究机关中基于经验开发的研制专家系统使用的语言或软件工具已经可以由市场提供，从而能够很方便地研制专家系统。但是到目前为止所开发的大部分专家系统仍未跳出供学习之用的试制品范围，可作为实用系统使用的例子还比较少。尽管如此，人们对知识工程学的期望却似乎超过了它的现实发展水平，从而产生了不少混乱的见解。这也许是由于社会上对于人工智能或知识工程学进行了过分的商业化的解释和评论而引起的。因此，许多学者都希望从学术角度或是从工程学的角度对这门学科以冷静的态度予以中肯的解释。本书正是从这个角度出发，以回

答知识库系统或专家系统的基本问题作为其核心内容。所以，在相当程度上是基于笔者个人见解的。另外，本书的内容并未包罗知识工程学的全部内容，只是想通过本书使读者能对知识工程学有自己的认识，然后通过阅读其它文献和亲身的实践经验来弥补其欠缺部分。

知识工程学又称为应用人工智能，它确实有将人工智能研究的成果应用于实际的工程学性质。专家系统就是其典型的例子。另一方面，我们也不可忽视，知识工程学本身作为一门独立的学科也正在迅速发展之中。知识表现、知识利用以及知识获得是基本的重要课题，而这些课题又是以各自的基础学科，包括认知科学、逻辑学以及用于实现知识库系统或专家系统的方法论等所构成的。本书中有意识地介绍了这方面的内容。另外，本书读者对象是相当广泛的，虽然是以大学中本专业的学生和该领域的研究人员为主要对象，但是产业、贸易部门的有关专业人员和与本领域有关的非专业人员也能充分理解。因其没有拘于细节的论述，一般地浏览后亦能把握住本书的主要内容。

本书共分六章。第一章和第二章是对知识工程学和专家系统的基本内容的简明易懂的介绍，即使对初步涉及该领域的读者也不难理解。在第三章和第四章中比较具体地解释了代表性的知识表现模型——产生式系统及框架系统。这两章中重点放在有关的基本概念上，因而没有程序设计经验的读者也能理解。第五章是根据作者经验，介绍有关专家系统的研制开发方面的内容，对于计划试制应用专家系统的人来说可以提供某种程度的具体的信息，而且它又可作为理解知识工程学的应用内容的较为具体的介绍。第六章则从今后的课题的角度出发，就对象模型和同数据库的结合这两个问题展开论述，这样做对该领域的有关研究者也许可提供些有意义

的信息。

最后还要指出一点是：知识工程学还是一门非常年轻的学科，作者也缺乏经验和理解。另外，限于时间关系，本书未能得以充分地完善。这些都望读者见谅。此外，在本书出版过程中，欧姆出版社给予了非常大的帮助，作者对他们在3年中非常耐心的工作表示感谢。

如果本书能成为对知识工程学有兴趣的广大读者的合适的人门读物，并对这门学科的发展能够略有贡献的话，那将是作者最大的喜悦。

上野晴树

1985年4月

目 录

译者的话	i
前言	iv
第一章 知识工程学概论	1
1.1 发展智能式计算机的必要性	2
1.2 何谓人工智能	8
1.3 专业知识的利用	10
1.4 知识	15
1.5 知识库系统	21
1.6 与以往程序的区别	30
1.7 知识工程学的应用领域	34
第二章 专家系统	38
2.1 MYCIN: 专家系统的一例	38
2.2 专家系统的结构	46
2.3 MYCIN 类专家系统的推理机构	52
2.4 分析型决策与设计型决策	55
2.5 新型程序与传统程序的比较	58
2.6 专家系统的课题	60
2.6.1 专家能力的极限问题	61
2.6.2 两个课题	65
第三章 产生式系统	67
3.1 产生式系统的历史	67
3.2 产生式系统的基本结构	70
3.3 AND/OR (与/或)树	74
3.4 数据库	78
3.5 产生式规则	80

3.6	竞选的消解	82
3.7	产生式系统的利弊	84
3.8	小结	87
第四章	框架系统	89
4.1	明斯基框架理论概要	90
4.1.1	基本的思想——框架	90
4.1.2	景物理解与框架	94
4.1.3	用框架理论表现教室	99
4.2	NUDGE: 一个基于框架的日程安排系统	104
4.2.1	NUDGE 的人-机会话的例子	106
4.2.2	日程计划的处理	107
4.3	框架型通用知识表现语言的结构	117
4.3.1	框架的结构及知识的表现单位	118
4.3.2	推理的控制和消息	123
4.4	小结	128
第五章	专家系统的研制	130
5.1	问题的选择	130
5.1.1	选择的条件	130
5.1.2	问题的选择方法	133
5.2	问题的领域	135
5.2.1	从问题性质角度观察的专家系统	135
5.2.2	从应用领域观察的专家系统	138
5.3	知识表现模型的选择	141
5.3.1	知识表现模型的类别	142
5.3.2	规则模型	144
5.3.3	黑板模型	147
5.3.4	因果网络模型	149
5.3.5	框架模型	149
5.3.6	语义网络模型	151
5.3.7	知识表现模型的选择	153
5.4	专家系统的开发语言	156
5.5	开发专家系统的进展方法	159

5.6 从开发过程中得到的经验	161
5.6.1 作为知识获取手段的知识表现模型	162
5.6.2 COMEX 的经验	169
5.6.3 专家系统开发语言的优、缺点	170
第六章 对象模型	173
6.1 对象模型	173
6.1.1 何谓对象模型	175
6.1.2 问题的背景	177
6.1.3 对象模型的作用	180
6.1.4 用于对象模型的知识表现方式	184
6.1.5 小结	189
6.2 知识库和数据库的接近问题	189
6.2.1 研究的动向	189
6.2.2 数据库和知识库	191
6.2.3 各种接近形态	197
6.2.4 在知识获取中 DB 的作用	200
6.2.5 使用关系数据库的知识库的实现方法	203
6.2.6 小结	205
术语集	207
参考文献	218

第一章 知识工程学概论

作为知识工程学的概论部分，本章主要说明以下一些问题：什么是“给计算机以智能”？人工智能与知识工程学的关系如何？何谓“智能”？何谓“知识”？何谓“智能式的问题解决方式”？究竟什么样的活动才算是“智能式”？怎样用计算机进行知识处理？知识型程序和传统程序在本质上有什么区别？知识型程序擅长处理什么？不擅长处理什么？等等。本章不仅要回答这些问题，同时也想使读者在读完本章后，对第二章以后的详细论述有个概念性的基础。因为本书是以今后打算学习知识工程学和想对此领域的基本概况有所了解的读者作为对象的，因而一部分说明可能有些冗长。

知识工程学这一领域开辟以来不过只有 20 多年的历史，1970 年以后，人们逐渐重视起这门学科的实用性，而开始在产业界得到应用则是 80 年代以后的事情。因此，它的概念、技术方法、术语等至今都尚未完全定型，很难估计这门学科的一些分支的实用可能性，将来究竟会如何发展也难以预料。因此，在这一章中，将尽可能地抓住本领域中的基本理论进行介绍，同时力求不仅说明这门学科的发展现状，而且贯穿介绍笔者的观点和看法。望读者在阅读本书时，不要单纯地注目各个分散的项目及其内容，而要在内心把握住实质性的问题来读它。

1.1 发展智能式计算机的必要性

(1) 软件的开发能够更方便吗？

电子计算机与其它机械的本质区别，可以说是在于其“按程序动作”这一点上。对电子计算机来说，如果只有硬件，它不能做任何工作，但一旦给它配备了指定工作顺序的程序，它就可以准确高速地进行极其复杂的逻辑运算。近年来，由于VLSI（超大规模集成电路）等技术的进步，硬件的能力飞速提高，价格反而急剧下降，而且，人们预计这种情形今后将持续下去。从这种发展状况来看，可以预想今后将会有各种各样的电子计算机进入我们日常生活的各个角落，正如人们所说的那样，今后离开了电子计算机就无法维持现代的信息化社会了。

但随之带来了几个问题，其中的一个问题是为了给计算机准备这种“工作顺序”，就必须给各种使用计算机的工作都配备上程序。软件产业正是由此应运而生，近年来更是膨胀式地发展起来。有人预测过，如果目前对软件的需求增长情况一直继续下去的话，到21世纪初时，对软件程序员的需要人数将超过地球上的全体人口的数目！当然，这种现象是不可能出现的，因为人们一定会设法提高软件生产的性能和效率，避免这种危机的来临。另外，从软件开发的现状来看，人们计算过，大约90%左右的程序编写工作量是在对已开发软件的扩展和改进上。这些工作大多是由于计算机系统本身的变动、使用计算机的具体业务的变动或因用户的要求变化而进行的对软件的扩展和改进。总之，程序不得不时常随环境的变动而变动，软件开发行业的现状正是这样维持着一种人力难以应付的状况。

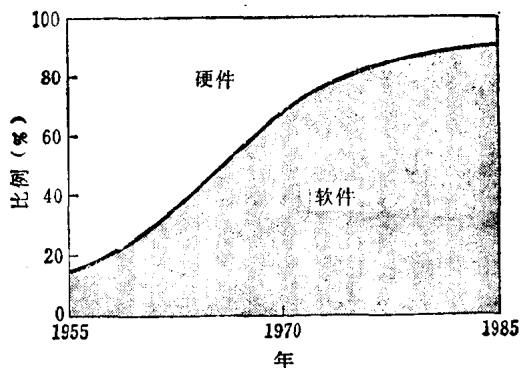


图 1.1 硬件与软件的价格构成比的变化 (B. W. 伯艾姆)

图 1.1 是 B.W. 伯艾姆在 1973 年统计的计算机系统中硬件和软件价格比的年增长趋势图。当时他就预测：在 70 年代后期，软件价格的比例将超过系统整体的 80%，80 年代前期将超过 90%，从那以后的发展结果来看，大体是符合他当时的预测的。提高软件的生产效率和软件的可靠性问题，已成为软件工程学的中心课题。但是以往的软件工程学并没能根本解决这个问题。

在这里，我们将用委托人来解决问题的方式与用计算机来解决问题的方式加以比较。如果我们委托的问题解决者对问题的领域有比较充分的专业知识的话，我们只须简单地对他说明问题的性质和目的即可，相反，如用计算机来解决问题时，我们却必须将解决问题的具体步骤完全以程序的形式表现出来，再输入到计算机中。这时，根据所用程序语言的不同，所花费的时间会有相当差异。但问题并不在这里，在实际编写程序时，由于误解、表现方法的错误、输入的误差等，常常需要花费许多预想不到的时间。而人解决问题时，即使有些要求提得不够完全，工作者也会主动地在工作中加以完善，必

要时只须重新问一下委托者,加以确认即可。

这里有两个值得注意的问题。一个是问题解决者已掌握了有关问题领域的足够的知识。另一个是人们可以用语言(自然语言)来进行指挥和应答,如果同时可以用图形的话,人们之间的交流就更为方便了。而就目前的技术水平来看,让计算机具备这样的功能还是相当困难的,虽然将来也许在某种程度上会具有这种功能,但目前,只有对问题和功能加以足够的限制,才能存在这类可供应用的技术。简单的音声应答装置和文字、文章的识别装置就是这样的例子。

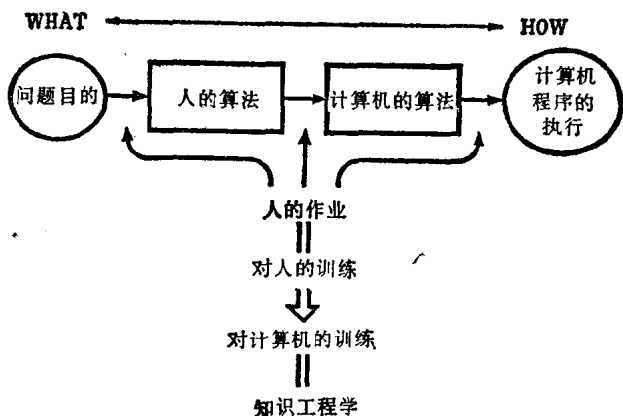


图 1.2 程序开发的过程

图 1.2 给出了从给定某个问题到用计算机解决为止的处理过程。首先将问题用人的算法(解决问题的逻辑或步骤)来表示,表现形式可以是数学公式或图表,也可以是人们头脑中考虑的比较抽象的理解模型。对这些表现形式来说,即使在逻辑上极其明显简单,往往也很不适于让计算机处理。所以下一个步骤,就是必须将上面的算法变换成适合计算机处理的算法。这就需要将数据的表现形式、输入输出的规格、处理步

骤的严格记述等明确地写出来。这就是用计算机语言写的可由计算机执行的计算机程序。

这里应注意的是：在这样的作业过程中，直到用程序设计语言来描述问题为止，全部工作都是由人来完成的。而且一般来说，人的算法和计算机的算法具有相当的差异，这就是所谓“瓶颈”之处。更进一步来说，在初期阶段的人使用的算法只有该问题领域的专家能做成，这种算法常常难以被计算机理解，而且这种算法经常并不全部明确地给出。越是关系到人的智能的决策活动，这种倾向就越强。从变换到计算机的算法起，就是计算机技术工作者的工作了。实际上，如果他们对问题本身和问题领域的理解尚未十分成熟的话，作为其工作成果的程序，其问题解决能力也是不完善的。一个很能说明问题的例子是，心电图自动分析程序的编写者，在他的程序完成时，他的关于心电图的知识已经可以接近专科医生的水平。这里的问题是，第一，这样的程序的编写不但需要很长的时间，而且由此得到的知识和经验只限于程序编写者个人，难以将其传播。第二，以后对程序的改进也是很困难的事情，因为程序的构造和算法都与人的算法和表现有很大的差异。

就目前编写程序状况来看，编写者大多处于一种“为完成程序竭尽全力”的被动状态，正因如此，对于程序编写者的训练就显得越发重要。人们常说，技术水平越低就越需要人力，由此看来，编写程序的工作目前还处在相当低的阶段。试想如果计算机能够接受那种“面向人的算法”的话，不但能大幅度地减轻人的负担，而且也能减少对人的训练，取而代之的是“对计算机的训练”，亦即让计算机更加“智能化”。换言之，就是由“HOW”（即教给计算机手段、方法）向“WHAT”（即只向计算机提出问题或表示人的某种目的）的接近。这也是智能计算机或智能程序研究的目标之一。此研究领域跟人工智