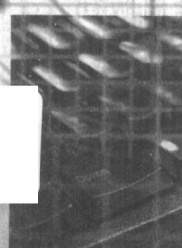
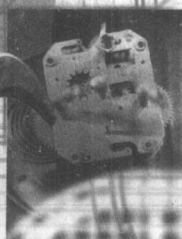
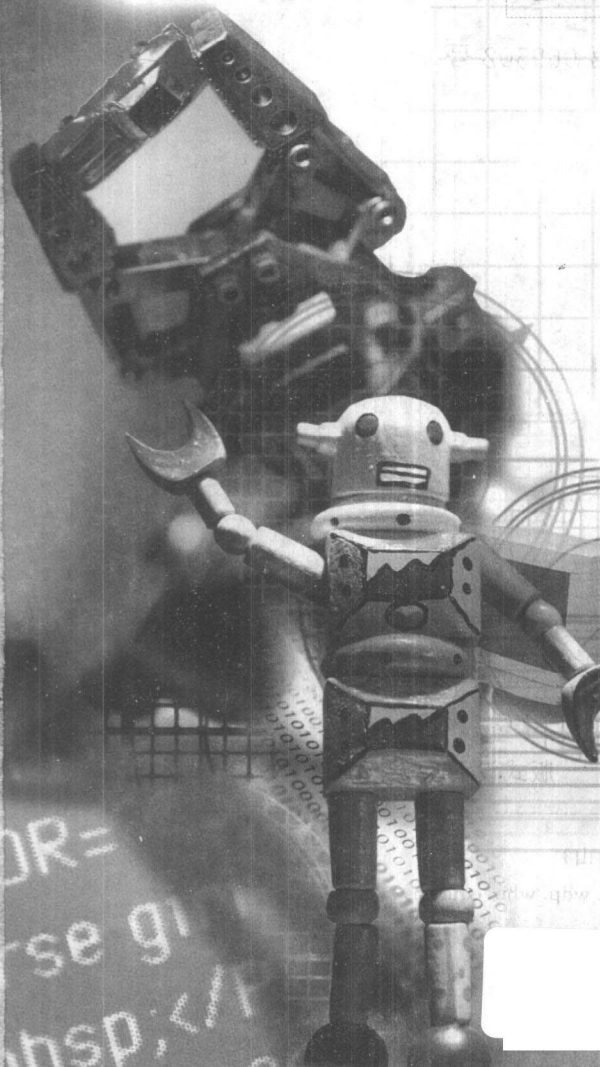


RENGONG ZHINENG YUANLI

人工智能原理

朱福喜 汤怡群 傅建明 编著



武汉大学出版社

图书在版编目(CIP)数据

人工智能原理/朱福喜,汤怡群,傅建明编著. —武汉: 武汉大学出版社, 2002. 2

ISBN 7-307-03355-0

I. 人… II. ①朱… ②汤… ③傅… III. 人工智能 IV. TP18

中国版本图书馆 CIP 数据核字(2001)第 069502 号

责任编辑: 瞿 嵘 责任校对: 卢 建 版式设计: 支 笛

出版: 武汉大学出版社 (430072 武昌 珞珈山)

(电子邮件: wdp4@whu.edu.cn 网址: www.wdp.whu.edu.cn)

发行: 新华书店湖北发行所

印刷: 湖北省通山县印刷厂

开本: 787×1092 1/16 印张: 26 字数: 348 千字

版次: 2002 年 2 月第 1 版 2002 年 2 月第 1 次印刷

ISBN 7-307-03355-0/TP · 114 定价: 32.50 元

版权所有, 不得翻印; 凡购我社的图书, 如有缺页、倒页、脱页等质量问题, 请与当地图书销售部门联系调换。

前 言

人工智能(Artificial Intelligence)作为研究机器智能和智能机器的一门综合性高技术学科,产生于20世纪50年代,曾经在20世纪80年代和90年代经历了一个轰轰烈烈的研究和发展时期,并且取得过不少令人鼓舞的成就,至今它仍然是计算机科学中备受人们重视和非常具有吸引力的前沿学科。

使计算机程序具有智能,能够模拟人的思维和行为,一直是计算机科学工作者的理想和追求。尽管人工智能的发展道路崎岖不平,自始至终充满了艰辛,但不畏艰难的从事人工智能研究的科学工作者们并没有放弃过对这个理想的追求。尽管计算机科学其他分支的发展也非常迅猛,并不断出现些新的学科领域,但是当这些学科的发展进一步深化时,人们不会忘记这样一个共同的目标:要使计算机更加智能化。所以不同知识背景和专业的人们都密切关注人工智能这门具有崭新思想和实用价值的综合性学科,并正从这个领域发现某些新思想和新方法。

人工智能的研究范畴不只局限于计算机科学和技术,而是一门涉及心理学、认知科学、思维科学、信息科学、系统科学和生物科学等多学科的综合性的技术学科,目前已在知识处理、模式识别、自然语言处理、博弈、自动定理证明、自动程序设计、专家系统、知识库、智能机器人等多个领域取得举世瞩目的成果,并形成了多元化的发展方向。近几年来,计算机网络,尤其是Internet的迅猛发展,多媒体、分布式人工智能和开放分布式环境下的多智能体(Multi-Agent)及其知识挖掘等计算机主流技术随之兴起,使得人工智能研究更加活跃,其研究和应用领域更加广阔,并正朝着健康和成熟的方向发展。

必须看到,尽管人工智能取得了以上所述的许多成果,但是比起人工智能刚刚兴起时许多专家的预想还相差甚远,很多在当时过于乐观的设想并没有实现,探究其原因也许要追根到目前人类对自身的思维规律和智能行为研究仍然处于探索阶段,因此,人工智能研究要比这些专家的预想艰难复杂得多,甚至到今天,对机器能否实现智能仍有争论。这种状况正如Lovelace女士100多年前曾经说过的:

在考虑任何新颖课题时,常常存在一种倾向,先是过高估计已发现是有趣或值得注意的东西。接着,当发现所研究的概念已超过曾一度保持不变的那些概念时,作为一种自然的反应,就会过低估计该事件的真实状况。

此时,我们必须清楚地认识到:人工智能研究道路的曲折和艰难以及许多尖锐的争论并不表明人工智能学科没有前景,它只是向我们表明理解人类认知和智能的机制、探索“智力的形成”是人类面临最困难、最复杂的课题之一。摆在人工智能学科面前的任务是极其艰巨和复杂的,这需要广大的计算机科学工作者不畏艰难,勇于探索,辛勤耕耘,共同开创人工智

能的美好未来。

本书共分 12 章,主要阐述人工智能的基本原理和技术,包括问题空间和搜索技术、机器人规划技术、知识表示和处理方法、多种典型推理方法、专家系统开发的实用技术、机器学习、机器人视觉和自然语言处理原理和方法等,最后两章是在国内同类教材较少出现的遗传算法和次协调逻辑,它反映了人工智能研究的最新方向和成果。本书参考了许多较新的国外同类教材和其他文献,力图保持前沿性和实用性,强调基本概念和基本观点,注重理论和实际相结合。

本书作为大学本科学习人工智能的教科书,在一学期内可能无法讲完全书,但可以介绍完问题求解、知识表达、推理等基本方法与技术。若附加某些专题文章的选读也可以作为研究生教材和计算机专业工作者了解人工智能的自学用书。

本书第一章至第五章、第七章、第十章至第十二章由朱福喜副教授执笔,第六章由傅建明博士执笔,第八章和第九章由汤怡群教授执笔,全书最后由朱福喜统稿。

在本书的编写过程中,主要参考了由 E. 丽奇编著并由李卫华、汤怡群等编译的《人工智能引论》,参考了刘娟博士、金涛博士的博士论文,在此向他们表示衷心的感谢。在编排录入工作中得到了王瑛、朱利娟、朱丽达的大力支持。同时本书的出版得到了武汉大学出版社的支持和帮助,在此一并表示深深的感谢。

由于时间和水平所限,难免有错,恳请读者批评指正。

作 者

2000 年 8 月于珞珈山

目 录

第一章 序 论	(1)
1.1 什么是人工智能?	(1)
1.2 AI 的产生及主要学派	(2)
1.3 人工智能、专家系统和知识工程	(4)
1.4 人工智能的技术特征	(6)
1.5 AI 模拟智能成功的标准	(7)
1.6 人工智能应用系统	(8)
第二章 问题求解与搜索方法	(13)
2.1 问题与问题空间	(13)
2.1.1 如何把问题求解定义为问题状态空间的搜索	(13)
2.1.2 问题特征分析	(16)
2.2 盲目搜索方法	(21)
2.2.1 宽度优先搜索(Breadth-first Search)	(21)
2.2.2 深度优先搜索(Depth-first Search)	(21)
2.2.3 分支有界搜索	(22)
2.3 启发式搜索方法	(22)
2.3.1 启发式信息的表示	(22)
2.3.2 几种最基本的搜索策略	(27)
2.4 图搜索策略	(31)
2.4.1 一个通用的图搜索算法	(31)
2.4.2 Λ 算法与 A^* 算法	(34)
2.5 问题归约与 AO^* 算法	(44)
2.5.1 问题归约求解方法	(44)
2.5.2 与/或图搜索	(45)
2.5.3 与/或图搜索的特点	(46)
2.5.4 与/或图搜索算法 AO^*	(47)
2.5.5 对 AO^* 算法的进一步观察	(49)
2.5.6 用 AO^* 算法求解一个智力难题	(50)
2.6 约束满足法	(55)
2.6.1 约束满足问题	(55)
2.6.2 约束满足问题求解	(56)

2.7	手段-目的分析法(Means-End-Analysis)	(58)
2.8	博弈	(62)
2.8.1	概述	(62)
2.8.2	极小极大搜索过程	(63)
2.8.3	α - β 剪枝算法	(67)
2.9	搜索算法的效率分析	(69)

第三章 知识表达与处理方法 (73)

3.1	概述	(73)
3.1.1	知识和知识表示的含义	(73)
3.1.2	AI 中知识表示方法分类	(71)
3.1.3	AI 对知识表示方法的要求	(74)
3.1.4	知识表示要注意的问题	(75)
3.2	逻辑表示法	(76)
3.3	产生式表示法	(77)
3.3.1	产生式系统的组成	(77)
3.3.2	产生式系统的表示	(78)
3.3.3	产生式系统的推理方式	(81)
3.3.4	产生式规则的选择与匹配	(83)
3.3.5	产生式表示的特点	(85)
3.4	语义网络表示法	(85)
3.4.1	语义网络结构	(85)
3.4.2	二元语义网络的表示	(86)
3.4.3	多元语义网络的表示	(86)
3.4.4	连接词和量词的表示	(87)
3.4.5	语义网络的推理过程	(91)
3.4.6	语义网络的一般描述	(93)
3.5	框架表示法	(94)
3.5.1	框架理论	(94)
3.5.2	框架结构	(94)
3.5.3	框架表示下的推理	(97)
3.6	过程式知识表示	(98)

第四章 谓词逻辑的归结原理及其应用 (102)

4.1	命题演算的归结	(102)
4.1.1	基本概念	(102)
4.1.2	命题演算的归结方法	(103)
4.2	谓词演算的归结	(103)
4.2.1	谓词演算的基本问题	(103)

4.2.2	将公式化成标准子句形式的步骤	(103)
4.2.3	合一算法	(104)
4.2.4	变量分离标准化	(106)
4.2.5	谓词演算的归结算法	(106)
4.3	归结原理	(108)
4.3.1	谓词演算基础	(109)
4.3.2	归结方法理论基础	(110)
4.3.3	Herbrand 定理	(113)
4.4	归结过程的控制策略	(120)
4.4.1	简化策略	(120)
4.4.2	支撑集策略	(122)
4.4.3	线性输入策略	(123)
4.4.4	几种推理规则及其应用	(124)
4.5	应用实例	(126)
4.5.1	归结在逻辑电路设计中的应用	(126)
4.5.2	利用推理破案实例	(128)

第五章	进一步的推理方法	(132)
5.1	非单调推理	(132)
5.1.1	单调推理与非单调推理的概念	(132)
5.1.2	默认逻辑	(133)
5.1.3	非单调推理系统 TMS	(134)
5.2	Dempster-Shater(D-S)证据理论	(141)
5.3	不确定性推理	(146)
5.3.1	不确定性	(146)
5.3.2	主观概率贝叶斯方法	(148)
5.4	MYCIN 系统的推理模型	(150)
5.4.1	理论和实际的背景	(150)
5.4.2	MYCIN 模型	(151)
5.4.3	MYCIN 模型分析	(153)
5.4.4	MYCIN 推理网络的基本模式	(154)
5.4.5	MYCIN 确定性因子的评价	(156)
5.5	模糊推理	(156)
5.5.1	模糊集论与模糊逻辑	(156)
5.5.2	Fuzzy 聚类分析	(158)
5.5.3	模糊认知图	(165)
5.6	基于模型的推理	(170)
5.6.1	基于模型推理在诊断中的应用	(170)
5.6.2	基于模型的推理系统的描述	(171)

5.6.3	基于模型的推理实例	(172)
5.6.4	基于模型的推理的优缺点	(173)
5.7	基于事例的推理	(173)
5.7.1	基于事例的推理的基本思想	(173)
5.7.2	事例的表示与组织	(175)
5.7.3	事例的检索	(176)
5.7.4	事例的改写	(176)
5.8	基于时序的推理	(176)
5.8.1	时间区间关系	(176)
5.8.2	时间区间和时间点的关系	(179)
5.8.3	时序的推理实例	(180)
5.9	归纳法推理	(181)
5.9.1	归纳法推理的研究现状	(182)
5.9.2	归纳法推理的理论基础	(183)
5.9.3	归纳法推理的基本概念	(185)
5.9.4	归纳法推理的定理证明器	(187)
5.9.5	归纳法推理研究中的主要难点	(188)
5.9.6	归纳法推理的理论局限	(189)
第六章	专家系统	(192)
6.1	专家系统的类型	(193)
6.1.1	专家系统的任务类型	(193)
6.1.2	分析型专家系统	(195)
6.1.3	设计型专家系统	(195)
6.1.4	组合型专家系统	(196)
6.2	知识获取的直接方法	(197)
6.2.1	概述	(197)
6.2.2	知识获取的直接方法	(197)
6.2.3	知识获取的新进展	(199)
6.3	专家系统的解释机制	(200)
6.3.1	预制文本解释法	(200)
6.3.2	路径跟踪解释法	(201)
6.3.3	自动程序员解释法	(201)
6.3.4	策略解释法	(202)
6.4	专家系统开发工具与环境	(202)
6.5	专家系统开发方法	(204)
6.5.1	专家系统开发的步骤	(204)
6.5.2	专家系统开发方法	(205)
6.6	专家系统开发实例	(205)

6.6.1	动物识别专家系统	(206)
6.6.2	MYCIN 专家系统	(207)
第七章	机器学习	(209)
7.1	概述	(209)
7.1.1	机器学习的定义和意义	(209)
7.1.2	机器学习的研究简史	(209)
7.1.3	机器学习方法的分类	(210)
7.1.4	机器学习中的推理方法	(212)
7.2	归纳学习	(213)
7.2.1	归纳概念学习的定义	(213)
7.2.2	归纳概念学习的形式描述	(215)
7.2.3	归纳概念学习算法的一般步骤	(216)
7.2.4	归纳概念学习的基本技术	(218)
7.3	解释学习	(225)
7.3.1	解释学习的基本原理	(225)
7.3.2	解释学习的一般框架	(226)
7.3.3	解释的学习过程	(227)
7.4	类比学习	(228)
7.4.1	类比学习的一般原理	(228)
7.4.2	类比学习的表示	(228)
7.4.3	类比学习的求解	(229)
7.4.4	逐步推理和监控的类比学习	(230)
7.5	神经网络学习	(232)
7.5.1	人工神经网络的发展简史	(232)
7.5.2	人工神经网络的基本原理	(233)
7.5.3	人工神经网络模拟的数学基础	(237)
7.5.4	人工神经网络的基本结构模式	(238)
7.5.5	人工神经网络互连结构	(239)
7.5.6	神经网络的学习算法	(242)
7.5.7	神经网络模型分类	(244)
第八章	机器人视觉	(247)
8.1	传感器和图像	(247)
8.1.1	数字图像	(247)
8.1.2	图像处理中的噪音	(249)
8.2	计算机视觉	(249)
8.2.1	图像理解	(249)
8.2.2	视觉和思考	(250)

8.3	视觉作为恢复过程	(251)
8.3.1	恢复什么	(252)
8.3.2	用几何的观点看图像构成	(253)
8.3.3	透视投影	(253)
8.3.4	正交投影	(254)
8.3.5	侧面透视投影	(255)
8.3.6	形状表示	(256)
8.3.7	表面的方向性及透视的形状	(256)
8.3.8	表面的方向性及在正交下的形状	(256)
8.3.9	立体画投影	(256)
8.3.10	透视投影的几何特性	(257)
8.3.11	透镜成像	(257)
8.3.12	图像构成的光度方式	(258)
8.4	恢复图像描述	(258)
8.4.1	边缘探测	(259)
8.4.2	基于微分法的方法	(259)
8.4.3	基于模型的方法	(262)
8.4.4	边缘聚合与 Hough 变换	(262)
8.4.5	图像分割	(263)
8.5	由轮廓得到形状	(264)
8.5.1	基于边缘标志的定性分析	(265)
8.5.2	基于斜对称的定量分析	(265)
8.6	从阴影到形状	(267)
8.6.1	反射率映射	(268)
8.6.2	解决病态问题	(270)
8.6.3	立体光度	(270)
8.7	从纹理到形状	(270)
8.7.1	纹理元素的密度	(271)
8.7.2	纹理反射率图	(272)
8.8	立体画	(273)
8.8.1	对应问题的描述	(273)
8.8.2	基于亮度的匹配	(274)
8.8.3	基于边缘的匹配	(275)
8.9	视觉运动的分析	(276)
8.9.1	运动域	(276)
8.9.2	运动域的估算	(277)
8.9.3	运动域的解释	(280)
8.10	主动视觉	(281)
8.11	应用	(282)

8.11.1	自动机动车驾驶	(282)
8.11.2	物体识别	(284)
第九章	自然语言处理	(289)
9.1	语言的组分	(290)
9.1.1	实义词和虚词	(290)
9.1.2	短语结构	(290)
9.2	上下文无关语法	(291)
9.2.1	重写规则	(291)
9.2.2	语法分析	(292)
9.3	上下文无关语法分析	(293)
9.3.1	产生后继状态的算法	(294)
9.3.2	利用词典	(295)
9.3.3	建立语法分析树	(295)
9.4	包含特殊的语法	(298)
9.4.1	引进特征	(298)
9.4.2	特征匹配	(299)
9.5	利用图表的高效语法分析	(302)
9.5.1	Chart 数据结构	(302)
9.5.2	有多种解释的句子	(303)
9.6	语义解释	(305)
9.6.1	词的意思	(305)
9.6.2	利用特征的语义解释	(307)
9.6.3	消除词的多义	(309)
9.7	生成自然语言	(310)
9.8	在上下文中的自然语言	(311)
9.8.1	言语的行为	(311)
9.8.2	创建参考	(312)
9.8.3	处理数据库的断言和问题	(313)
9.9	量词的辖域	(316)
第十章	规 划	(319)
10.1	概述	(319)
10.1.1	规划的定义	(319)
10.1.2	规划的分类	(320)
10.2	规划问题的表达——积木世界	(321)
10.3	规划系统的组成	(323)
10.3.1	挑选可应用的规则	(323)
10.3.2	应用规则	(323)

10.3.3	框架公理	(324)
10.3.4	检测解	(325)
10.3.5	检测死胡同	(326)
10.3.6	修正近似正确的解	(326)
10.4	三角形表格	(327)
10.5	用目标堆栈的简单规划	(328)
10.6	使用目标集的非线性规划	(333)
10.7	分层规划	(336)
10.7.1	长度优先搜索	(336)
10.7.2	NOAH 规划	(337)
10.8	多目标规划	(343)
第十一章	遗传算法	(349)
11.1	遗传算法的基本概念	(349)
11.2	遗传算法的形式描述	(351)
11.3	设计 GA 基本步骤	(353)
11.4	遗传算法的基本理论	(356)
11.5	遗传算法简例	(358)
第十二章	次协调逻辑及其自动推理	(361)
12.1	什么是次协调逻辑	(361)
12.2	专家系统与不协调知识库	(362)
12.3	广义 Horn 子句逻辑程序和语法	(362)
12.4	广义 Horn 子句逻辑程序的语义	(365)
12.5	前不动点的语义	(369)
12.6	合经典的广义 Horn 子句程序	(372)
12.7	广义 Horn 子句程序的操作语义	(376)
12.8	对合经典 GHP 的线性定序非 Horn 子句归结	(380)
12.9	不协调的医学专家系统及投资决策系统	(383)
12.10	次协调逻辑自动定理证明的理论	(385)
12.11	次协调逻辑线性归结原理	(388)
12.12	注解逻辑下的推理实例	(394)
12.13	实现策略	(397)

第一章 序 论

1.1 什么是人工智能?

人类的自然智能伴随着人类活动时处处都存在。人类的许多活动,如解题、下棋、猜谜、讨论、编制计划和编程,甚至驾车、骑车等,都需要智能。如果机器能够完成这些任务,那么就可以认为机器已经具有某种程度的“人工智能”。

什么是智能?什么是人工智能?智能与人工智能有什么区别和联系?这些都是广大科技工作者十分感兴趣和值得深入探讨的问题。人工智能的出现不是偶然的。从思想基础上讲,它是人们长期以来探索研制能够进行计算、推理和其他思维活动的智能机器的必然结果;从理论上讲,它是信息论、控制论、系统工程论、计算机科学、心理学、神经学、认知科学、数学和哲学等多学科相互渗透的结果;从物质和技术基础上讲,它是电子计算机和电子技术得到广泛应用的结果。为了熟悉人工智能,先必须掌握一些与它有关的概念。

首先我们看看什么是信息。我们知道,信息与物质及能量构成宇宙。信息是物质和能量运动的形式,是以物质和能量为载体的客观存在。它们之间的关系是:

(1)能量和物质可以相互转换,但信息与能量、信息与物质不会直接转化。

(2)物质和能量表现为信息,或者产生信息。

(3)信息能够控制

{	能量与能量之间的转换;
	物质与物质之间的转换;
	能量转换为物质;
	物质转换为能量。

人们不能直接认识物质和能量,而是通过它们的信息来认识物质和能量。

人的认识过程为:信息经过感觉输入到神经系统,再经过大脑思维变为认识。

那么什么是认识呢?认识就是用符号去整理研究对象,并确定其间的联系。

从认识我们可以继续探讨什么是知识?什么是智力?知识是用人们对于可重复信息之间的联系的认识,所以知识是被认识了的信息和信息之间的联系,它是信息经过加工整理、解释、挑选和改造而形成的。

关于智力,科学家们有不同的定义,以下是几种对智力的定义:

Wisterw: 智力是指个体有意识地以思维活动来适应新情况的一种潜力,也就是个体对生活中新问题和新条件在心理上的一般适应能力。

Terman: 智力是抽象思维的能力。

Buckingham: 智力是学习的能力。

Storrdard: 智力是从事艰难、复杂、抽象、敏捷和创造性的活动以及集中精力和保持情

绪稳定的能力。

Piaget: 智力的本质就是适应,使个体与环境取得平衡。

Guilford: 智力是对信息进行处理的能力。

总而言之,智力可看做个体的各种认识能力的综合,特别强调解决新问题的能力,包括抽象思维、学习能力、对环境的适应能力。

有了知识和智力的定义后,一般将智能定义为:智能=知识集+智力。所以智能主要指运用知识解决问题的能力,推理、学习和联想是其重要因素。

至于人工智能,像许多新兴学科一样,至今尚无统一的定义。下面是几位著名的人工智能方面的科学家分别在不同的年代对人工智能给出的定义:

1978 年 P. Winston

“人工智能是研究使计算机更灵活有用,了解使智能实现成为可能的原理。因此,人工智能研究结果不仅是使计算机能模拟智能,而且是了解如何帮助人们变得更有智能。”

1981 年 A. Barr 和 E. Feigenbum

“人工智能是计算机科学的一个分支,它关心的是设计智能计算机系统,该系统具有与人的行为相联系的智能特征,如了解语言、学习、推理、问题求解等。”

1983 年 Elaine Rich

“人工智能是研究怎样让电脑模拟人脑从事推理、规划、设计、思考、学习等思维活动,解决至今认为需要由专家才能处理的复杂问题。”

1987 年 Michael R. Genesereth 和 Nils J. Nilsson

“人工智能是研究智能行为的科学。它的最终目的是建立关于自然智能实体行为的理论和指导创造具有智能行为的人工制品,这样一来,人工智能有两个分支,一个为科学人工智能,一个为工程人工智能。”

Michael 和 Nilsson 关于人工智能的定义引出了科学人工智能和工程人工智能的概念。

关于科学人工智能,它的目的是发展概念和词汇以帮助我们了解人和其他动物的智能行为。

关于工程人工智能,它主要研究建立智能机器的概念、理论和实践。例如:

- ◆ 专家系统:在专门的领域内的咨询(医疗、探矿、财务等)服务系统。
- ◆ 自然语言处理:在有限范围内的问题回答系统。
- ◆ 程序验证系统:通过定理证明途径验证程序的正确性。
- ◆ 智能机器人:人工智能研究计算机视觉和智能机。

以上是人工智能的一些比较权威的定义。人工智能似乎还有一个比较模糊的定义,那就是“如果某个问题在计算机上没有解决,那么这个问题就是人工智能问题”,因为一旦解决了某个问题,也就有了解决算法,因而也就不那么“智能”了。从这个意义上讲,人工智能永远是一个深奥而无止境的追求目标。

1.2 AI 的产生及主要学派

丘奇(Church)、图灵(Turing)和其他一些学者关于计算本质的思想,将形式推理概念与当时即将发明的计算机之间建立起了联系。他们在这方面的重要工作是关于计算和符号

处理的理论。早在计算机产生之前,丘奇和图灵就已发现,数值计算并不是计算的主要方面,它们仅仅是解释机器内部状态的一种方法。被称为“人工智能之父”的图灵,不仅仅创造了一个简单的非数字计算模型,而且直接证明了计算机可能以某种被认为是智能的方式进行工作,这就是人工智能思想的萌芽。

人工智能学科产生的突出标志是:1956年夏,在美国达特茅斯(Dartmouth)大学由当时美国年轻的数学家 John-McCarthy 和他的朋友明斯基(Minsky)、纽维尔(Newell)、西蒙(Simon)、香农(Shannon)、塞缪尔(Sauml)、莫尔(More)等数学、心理学、神经学、信息论、计算机科学方面的学者,举办了一个长达2个月的讨论会。会上 McCarthy 提出了“Artificial Intelligence”一词,尔后 Newell 和 Simon 提出了物理符号系统假设,从而创建了 AI 这一学科。主张系统符号假设的学派形成了 AI 研究的主要学派,即符号主义学派。目前,人工智能主要有以下三个学派。

1. 符号主义(Symbolicism)

符号主义又称为逻辑主义(Logicis)、心理学派(Psychologism)或计算机学派(Computerism)。该学派认为人工智能源于数理逻辑。数理逻辑在19世纪获得迅速发展,到20世纪30年代开始用于描述智能行为。计算机产生以后,又在计算机上实现了逻辑演绎系统,其代表的成果为启发式程序 LT(逻辑理论家),人们使用它证明了38个数学定理,从而表明了人类可利用计算机模拟人类的智能活动。

前面提到符号主义来源于物理符号系统假设,那么什么是符号系统呢?符号主义将符号系统定义为如下三部分组成:

- ◆ 一组符号:对应于客观世界的某些物理模型。
- ◆ 一组结构:它是由以某种方式相关联的符号的实例所构成。
- ◆ 一组过程:它作用于符号结构上而产生另一些符号结构,这些作用包括:创建、修改、消除等。

在这个定义下,一个物理符号系统就是能够逐步生成一组符号的产生器。在该假设下,符号主义认为,人的认知是符号,人的认知过程是符号操作过程。符号主义还认为,人就是一个物理符号系统,计算机也是一个物理符号系统,因此,我们就能够用计算机来模拟人的智能行为,即用计算机的符号操作来模拟人的认知过程。这实质就是认为,人的思维是可操作的。

符号主义还认为:知识是信息的一种形式,是构成智能的基础,AI 的核心问题是知识表示、知识推理和知识运用。知识可用符号表示,也可用符号进行推理。符号主义就是在这种假设之下建立起基于知识的人类智能和机器智能的核心理论体系。

符号主义曾长期一枝独秀,经历了从启发式算法到专家系统,再到知识工程理论与技术的发展道路,为 AI 作出了重要的贡献。至今符号主义仍是 AI 的主流学派。

2. 联结主义(Connectionism)

联结主义又称仿生学派(Bionicsism)或生理学派(Physiologism),是基于生物进化论的 AI 学派,其主要原理为神经网络及神经网络间的连接机制与学习算法。联结主义认为 AI 源于仿生学,特别是通过对人脑模型的研究,认为人的思维基元是神经元,而不是符号处理过程,人脑不同于电脑,并提出联结主义的大脑工作模式,用于取代符号操作的电脑工作模式。

如果说符号主义是从宏观上模拟人的思维过程的话,那么联结主义则试图从微观上解

决人类的认知功能,以探索认知过程的微观结构。联结主义从人脑模式出发,建议在网络层次上模拟人的认知过程。所以,联结主义本质上是用人脑的并行分布处理模式来表现认知过程。

联结主义的兴起标志着神经生理学和非线性科学向 AI 渗透,这主要表现为人工神经网络 ANN(Artificial Neural Network)研究的兴起,ANN 可以看为是一种具有学习和自组织能力的智能机器或系统。ANN 作为模拟人的智能和形象思维能力的一条重要途径,对 AI 研究工作者有着极大的吸引力。近年来,由于出现了一些新型的 ANN 模型和一些强有力的学习算法,大大地推动了有关 ANN 理论和应用研究。联结主义具有代表性的工作有:

- ◆ Hopfield 教授在 1982 年和 1984 年的两篇论文中提出用硬件模拟神经网络;
- ◆ Rumthart 教授在 1986 年提出多层网络中的反向传播(BP)算法。

3. 行为主义(Actionism)

行为主义又称为进化主义(Evolutionism)或控制论学派(Cyberneticism),其原理为控制论及“感知-动作”型控制系统。行为主义提出了智能行为的“感知-动作”模式,认为:

- ◆ 智能取决于感知和行动;
- ◆ 人工智能可以像人类智能一样逐步进化(所以称为进化主义);
- ◆ 智能行为只能在现实世界中与周围环境交互作用而表现出来。

行为主义是控制论向 AI 领域的渗透,它的理论基础是控制论,它把神经系统的工作原理与信息理论联系起来,着重研究模拟人在控制过程中的智能行为和作用,如自寻优、自适应、自校正、自镇定、自学习和自组织等控制论系统,并进行控制论动物的研究。这一学派的代表首推美国 AI 专家 Brooks。1991 年 8 月在悉尼召开的十二届国际人工智能联合会议上, Brooks 作为大会“计算机与思维”奖的得主,通过讨论 AI、计算机、控制论、机器人等问题的发展情况,并以他多年在 MIT 进行人造动物机器的研究与实践和他所提出的“假设计算机体系结构”研究为基础,发表了“没有推理的智能”的论文,对传统的 AI 提出了批评和挑战。

Brooks 的行为主义学派否定智能行为来源于逻辑推理及其启发式的思想,认为对 AI 的研究不应把精力放在知识表示和编制推理规则上,而应着重研究在复杂环境下对行为的控制。这种思想对 AI 主流派传统的符号主义思想是一次冲击和挑战。行为主义学派的代表作首推 Brooks 等人研制的六足行走机器人,它是一个基于“感知-动作”模式的模拟昆虫行为的控制系统。

1.3 人工智能、专家系统和知识工程

20 世纪 60 年代,人工智能的研究者试图通过找到通用问题求解方法来模拟复杂的思维过程。然而,这种策略虽取得了一些进展,但没有多大突破。单个程序能处理的问题越多,它处理一般问题的能力越差。于是,人工智能工作者认为一定存在着一种使计算机程序具有“智能”的方法。

既然开发通用问题求解程序太难,他们就把研究重心转移到开发一系列较特殊的程序。因此,在 70 年代,他们致力于问题表示技术——如何将问题求解形式化,使之易于求解;搜索技术——如何有效地控制解的搜索过程,使之不要浪费太多的时间和空间。使用这两种技术虽也取得了一定的进展,但仍未有所突破。

70年代后期,人工智能工作者才认识到:一个程序求解问题的能力来自它所具有的知识,而不仅仅是它所采用的形式化方法和推理策略。这个概念上的突破可以简单地叙述为:要使一个程序具有智能,就要给它提供许多关于某一问题领域特定的知识。

这一认识导致了特定问题求解的计算机程序的发展,这类程序就是人们所熟悉的专家系统。关于专家系统的定义,1985年Michie认为:“一个专家系统以某种形式将专家的基于知识和技能的部分嵌入到计算机中,使得该系统能产生智能行为和建议,并且当用户提出要求时就能证明其推理过程。”

一般认为,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问题提供具有专家水平的解答。

早期的专家系统通常用高级程序设计语言编写,尤其是LISP或PROLOG语言,它们常被选为实现语言。然而,在用高级编程语言作为专家系统的建造工具时,人们常常要花大量注意力在与被模型化的问题领域毫无关系的系统实现上。而且,领域专家知识和运用这些知识的算法紧密交织在一起,不易分开,致使系统一旦建成,便不易改变。然而专家知识和经验却总在改变。由于对以上特性的分析,研究者们清醒地认识到在开发专家系统中应该把算法与知识分开,从而使现今专家系统的基本模式成为:

专家系统=知识+推理

因此,一个专家系统主要由以下两个部分组成:

- (1)知识库,存放关于特定领域的知识;
- (2)推理机,包括操纵知识库中所表示的知识的算法。

现在,专家系统很少直接用高级编程语言编写,取而代之的是专家系统构造工具。在专家系统构造工具中,预先规定了知识表示形式并提供了相应的推理机,开发一个实际专家系统仅需要提供特定领域的知识,并以工具所要求的知识表示形式表示出来。知识库的开发独立于推理机的好处之一是知识库可以逐步开发与求精,在不对程序进行大量修改的情况下,来纠正错误和不足;另一个好处就是一个知识库可以被另一知识库所代替,从而形成完全不同领域的专家系统。

知识工程产生于20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,但建造一个成功的专家系统工程量巨大,其花费是以多少年来计算的,所以迫切需把专家系统的建造提高到工程的高度来认识。为此,美国的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。1977年著名的人工智能专家E. Feigenbun在第五届国际人工智能会议(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。从此,知识工程这个术语为全世界广泛使用,Feigenbun因此被誉为“知识工程之父”。

目前,知识工程的主要研究内容有:

1. 基础研究

基础研究的主要内容有:知识的本质、分类、结构和作用,知识的表示方法和语言,知识的获取和学习方法,推理和控制机制,解释和接口模型,认知模型等。

2. 实用知识系统的开发研究

实用知识研究主要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,如实用知识获取技术,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