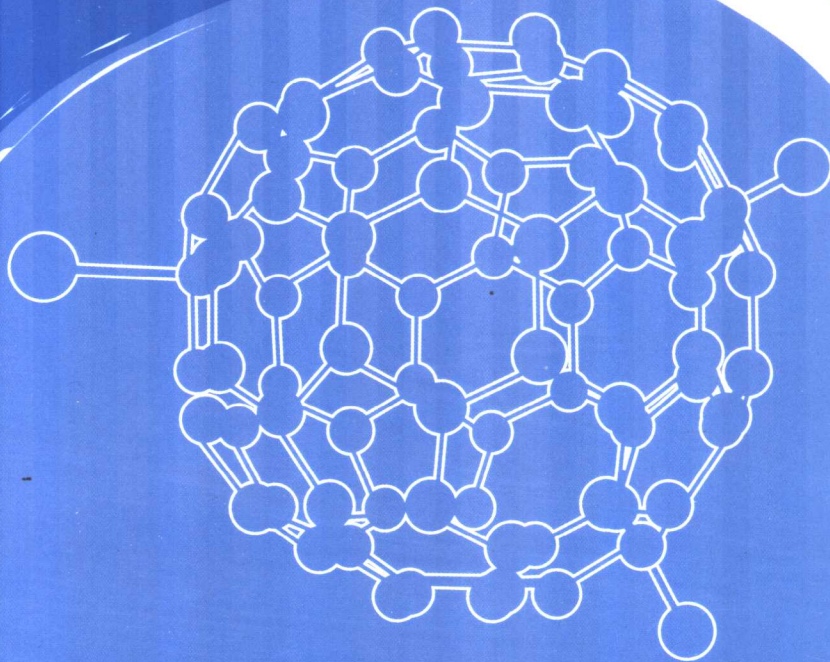


现代计算机科学技术精品教材



人工智能基础

朱福喜 朱三元 伍春香 编著

教程



光盘演示经典算法的
程序实现及推理系统



清华大学出版社

人工智能基础教程

朱福喜 朱三元 伍春香 编著

清华大学出版社

北京

内 容 简 介

本书系统地阐述了人工智能的基本原理、实现技术及其应用,全面地反映了国内外人工智能研究领域的最新进展和发展方向。

全书共 13 章,前 8 章系统地阐述了传统的人工智能原理和方法,内容包括状态空间和搜索技术、各种知识表示和处理技术、几种典型的推理技术、专家系统开发技术、机器学习、自然语言处理原理和方法。这些内容能够使读者对人工智能的基本概念和人工智能系统的构造方法有一个比较清楚的认识。第 9 章“Agent 技术”、第 10 章“知识获取的新技术”、第 11 章“遗传算法”、第 12 章“群集智能算法”是反映人工智能研究领域里新的进展,主要讨论分布式人工智能、数据挖掘技术、生物计算或仿生学计算等。第 13 章“次协调逻辑与自动推理”则主要讨论在不协调环境下的知识推理问题。

为增强学习过程的趣味性、可视性和可理解性,本书中的经典例子和算法特别用程序加以实现,附在光盘中可供演示。

本书强调具有先进性、实用性和可读性,可作为计算机、信息处理、自动化和电信等 IT 相关专业的高年级本科生学习人工智能的教材,也可供从事计算机科学研究、开发和应用的教学和科研人员参考。

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

本书防伪标签采用特殊防伪技术,用户可通过在图案表面涂抹清水,图案消失,水干后图案复现;或将表面膜揭下,放在白纸上用彩笔涂抹,图案在白纸上再现的方法识别真伪。

图书在版编目(CIP)数据

人工智能基础教程/朱福喜,朱三元,伍春香编著.—北京:

清华大学出版社,2006

ISBN 7-302-12577-5

I. 人... II. ①朱... ②朱... ③伍... III. 人工智能 IV. TP18

中国版本图书馆 CIP 数据核字(2006)第 012718 号

出 版 者:清华大学出版社

<http://www.tup.com.cn>

社 总 机:010-62770175

地 址:北京清华大学学研大厦

邮 编:100084

客 户 服 务:010-82896445

组稿编辑:夏非彼

文稿编辑:科海

封面设计:林陶

版式设计:科海

印 刷 者:北京科普瑞印刷有限责任公司

发 行 者:新华书店总店北京发行所

开 本:787×1092 1/16 印张:24.5 字数:596千字

版 次:2006年3月第1版 2006年3月第1次印刷

书 号:ISBN 7-302-12577-5/TP·8041

印 数:0 001~4 000

定 价:36.00元(1CD)

本书如存在文字不清、漏印以及缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:(010)82896445

前 言

使计算机具有智能、能够模仿人的思维和行为，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和困苦，但不畏艰难的从事思维和计算科学研究的研究人员并没有放弃对这个理想的追求。人工智能作为一门研究和模仿人的智能的学科，最早的萌芽可追溯到 20 世纪 20 年代图灵（Turing）等人的研究，他们在计算机出现之前，就预见将来会有一种会思维的机器。人工智能在 20 世纪 50 年代麦卡锡（McCarthy）提出“Artificial Intelligence”一词后逐步兴起，形成一门涉及心理学、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，计算机网络尤其是 Internet 的发展，为人工智能注入了新的内容和新的应用领域，加上多媒体技术、生物计算技术、分布式人工智能和多智能体（Multi-Agent）以及数据挖掘与知识发现等计算机技术的兴起，使得人工智能研究更加活跃，更加拓广了其研究和应用领域，人工智能正朝着健康和成熟的方向发展。而且一些曾经走入低谷的人工智能的一些分支和学派又重新活跃起来，如人工神经网络成为当前的研究热点之一，这也标志着生物学派的崛起。这些都表明，当今的人工智能比上个世纪八九十年代有了很大的发展，现在不再是符号主义一枝独秀的时代。因此，我们有必要对人工智能新的发展的概念和技术进行重新归纳、整理，形成人工智能教学新的核心内容。

本书主要阐述人工智能的基本原理和技术，包括问题空间与搜索技术、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。为反映人工智能研究的最新研究方向和成果，最后几章分别介绍了以前人工智能教材很少涉及的新技术：知识获取的新技术（其中包括数据挖掘、Web 挖掘、文本挖掘）；遗传算法；群集智能算法（其中包括人工蚂蚁、人工鱼等）。本书最后一章讨论次协调逻辑，旨在引起人们对不协调知识的处理的关注。本书参考了许多较新的国外同类教材和其他文献，力图保持新颖性和实用性，强调基本概念和基本观点，注重理论和实践相结合。

此外，为增强学习过程的趣味性、可视性和可理解性，还特别将本书中的经典例子和算法用程序加以实现，附在光盘中可供演示。其中，作者认为值得一提的是基于归结原理的推理系统和次协调推理系统，它们可使读者在学习有关知识的同时，自己动手体会推理过程，对于加深理解会有很大的帮助。

本书适合用作大学本科学习人工智能的教科书，虽然在一学期内无法讲完全部内容，

但可以介绍完问题求解的搜索算法（包括遗传算法）、知识表达、各种推理机制等基本方法与技术。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。

本书共 13 章，第 1、2、4、5、7、8、9、10、12 章由朱福喜执笔，第 3、6 章由伍春香、朱福喜、傅建明执笔，第 11、13 章由朱三元、朱福喜执笔，全书最后由朱福喜统稿。本教材的光盘演示实例主要由朱三元、朱福喜提供。

由于本书是作者在总结和吸取了由朱福喜、汤治群、傅建明等编著的《人工智能原理》教材及教学经验的基础上编写的，在此向汤治群、傅建明表示深深的感谢。

在本书的编写过程中，邓炜、丁圣超、王慧、李忠波、韦福如、查桂峰等帮助制作了部分演示实例，在此向他们表示衷心感谢。

同时，本书的顺利出版得到了科海夏非彼总编的大力支持，在此表示衷心感谢。

由于时间和水平所限，书中难免存在错误和不足之处，敬请广大读者批评指正，使本书得以改进和完善。

作 者

2006 年 2 月于武昌珞珈山

目 录

第 1 章 概述	1
1.1 什么是人工智能	1
1.2 AI 的产生及主要学派	3
1.3 人工智能、专家系统和知识工程	5
1.4 人工智能的技术特征	7
1.5 AI 模拟智能成功的标准	9
1.6 人工智能应用系统	9
第 2 章 问题求解与搜索方法	15
2.1 问题的状态和状态空间	15
2.1.1 如何定义状态空间及其搜索	15
2.1.2 问题特征分析	19
2.2 盲目的搜索方法	24
2.2.1 宽度优先搜索	24
2.2.2 深度优先搜索	24
2.2.3 分支有界搜索	24
2.2.4 迭代加深搜索	25
2.3 启发式搜索方法	25
2.3.1 启发式信息的表示	25
2.3.2 几种最基本的搜索策略	30
2.4 图搜索策略	35
2.4.1 一个通用的图搜索算法	35
2.4.2 A 算法与 A* 算法	39
2.5 问题归约与 AO* 算法	48
2.5.1 问题归约求解方法与与/或图	48
2.5.2 与/或图搜索	50
2.5.3 与/或图搜索的特点	51
2.5.4 与/或图搜索算法 AO*	53
2.5.5 对 AO* 算法的进一步观察	54
2.5.6 用 AO* 算法求解一个智力难题	55
2.6 博弈	59
2.6.1 概述	59
2.6.2 极小极大搜索过程	61

2.6.3 $\alpha - \beta$ 剪枝算法.....	64
习题 2.....	68
第 3 章 知识表示与处理方法.....	70
3.1 概述.....	70
3.1.1 知识和知识表示的含义.....	70
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.....	72
3.1.4 知识表示要注意的问题.....	73
3.2 逻辑表示法.....	74
3.3 产生式表示法.....	75
3.3.1 产生式系统的组成.....	75
3.3.2 产生式系统的知识表示.....	76
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.....	84
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.....	85
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.....	100
习题 3.....	103
第 4 章 谓词逻辑的归结原理及其应用.....	105
4.1 命题演算的归结方法.....	105
4.1.1 基本概念.....	105
4.1.2 命题演算的归结方法.....	106
4.2 谓词演算的归结.....	107
4.2.1 谓词演算的基本问题.....	107
4.2.2 将公式化成标准子句形式的步骤.....	107
4.2.3 合一算法.....	109
4.2.4 为什么要变量分离标准化.....	112

4.2.5 谓词演算的归结算法.....	113
4.3 归结原理.....	114
4.3.1 谓词演算基础.....	115
4.3.2 归结方法可靠性证明.....	116
4.3.3 归结方法的完备性.....	119
4.4 归结过程的控制策略.....	128
4.4.1 简化策略.....	128
4.4.2 支撑集策略.....	130
4.4.3 线性输入策略.....	131
4.5 几种归结方法及其应用实例.....	132
4.5.1 归结方法.....	132
4.5.2 归约及其应用.....	134
4.5.3 利用一般归结和等式归结的推理实例.....	137
习题 4.....	139
第 5 章 进一步的推理方法.....	141
5.1 非单调推理.....	141
5.1.1 单调推理与非单调推理的概念.....	141
5.1.2 默认逻辑.....	142
5.2 非单调推理系统 TMS.....	143
5.2.1 TMS 的依据.....	143
5.2.2 TMS 中信念的状态.....	143
5.2.3 TMS 中信念的表示方法.....	144
5.2.4 TMS 中的证实和推理.....	144
5.3 Dempster-Shafer (D-S) 证据理论.....	146
5.4 不确定性推理.....	153
5.4.1 不确定性.....	153
5.4.2 主观概率贝叶斯方法 (Bayes Approaches).....	154
5.5 MYCIN 系统的推理模型.....	157
5.5.1 理论和实际的背景.....	157
5.5.2 MYCIN 模型.....	159
5.5.3 MYCIN 模型分析.....	160
5.5.4 MYCIN 推理网络的基本模式.....	162
5.5.5 MYCIN 确定性因子的评价.....	164
5.6 模糊推理.....	164
5.6.1 模糊集论与模糊逻辑 (Fuzzy Sets and Fuzzy Logic).....	164
5.6.2 模糊聚类分析.....	167
5.7 基于案例的推理.....	174
5.7.1 基于案例推理的基本思想.....	175

5.7.2 案例的表示与组织.....	176
5.7.3 案例的检索.....	177
5.7.4 案例的改写.....	177
5.8 归纳法推理.....	177
5.8.1 归纳法推理的理论基础.....	178
5.8.2 归纳法推理的基本概念.....	180
5.8.3 归纳法推理研究中的主要难点.....	183
5.8.4 归纳法推理的研究成果.....	184
习题 5.....	185
第 6 章 专家系统.....	188
6.1 概述.....	188
6.1.1 什么是专家系统.....	188
6.1.2 专家系统的结构.....	188
6.1.3 专家系统的特点.....	189
6.1.4 专家系统的类型.....	189
6.1.5 成功专家系统的实例.....	191
6.2 知识获取的直接方法.....	192
6.2.1 概述.....	192
6.2.2 知识获取的直接方法.....	194
6.2.3 知识获取的新进展.....	195
6.3 专家系统的解释机制.....	196
6.3.1 预制文本解释法.....	196
6.3.2 路径跟踪解释法.....	197
6.3.3 自动程序员解释法.....	197
6.3.4 策略解释法.....	198
6.4 专家系统开发工具与环境.....	199
6.5 专家系统开发方法.....	200
6.5.1 专家系统开发步骤.....	200
6.5.2 专家系统开发方法.....	202
6.6 专家系统开发实例.....	202
6.6.1 动物识别专家系统.....	203
6.6.2 MYCIN 专家系统.....	204
习题 6.....	205
第 7 章 机器学习.....	206
7.1 概述.....	206
7.1.1 机器学习的定义和意义.....	206
7.1.2 机器学习的研究简史.....	206

7.1.3 机器学习方法的分类.....	207
7.1.4 机器学习中的推理方法.....	209
7.2 归纳概念学习.....	210
7.2.1 归纳概念学习的定义.....	211
7.2.2 归纳概念学习的形式描述.....	213
7.2.3 归纳概念学习算法的一般步骤.....	214
7.2.4 归纳概念学习的复杂度和精确度.....	215
7.2.5 归纳概念学习的基本技术.....	215
7.3 基于解释的学习.....	224
7.3.1 基于解释学习的基本原理.....	224
7.3.2 基于解释学习的一般框架.....	224
7.3.3 基于解释学习的过程.....	225
7.4 基于类比的学习.....	226
7.4.1 类比学习的一般原理.....	226
7.4.2 类比学习的表示.....	227
7.4.3 类比学习的求解.....	228
7.4.4 逐步推理和监控的类比学习.....	228
7.5 人工神经网络学习.....	230
7.5.1 人工神经网络的发展简史.....	230
7.5.2 人工神经网络的基本原理.....	232
7.5.3 人工神经网络模拟的数学基础.....	235
7.5.4 人工神经网络的基本结构模式.....	237
7.5.5 人工神经网络的学习算法.....	238
7.5.6 神经网络模型分类.....	240
习题 7.....	243
第 8 章 自然语言处理.....	244
8.1 语言的组成部分.....	245
8.1.1 实词和虚词.....	245
8.1.2 短语结构.....	245
8.2 上下文无关语法.....	246
8.2.1 重写规则.....	246
8.2.2 语法分析.....	247
8.3 上下文无关语法分析.....	248
8.3.1 产生后继状态的算法.....	249
8.3.2 利用词典.....	251
8.3.3 建立语法分析树.....	251
8.4 特殊语法的分析.....	254
8.4.1 引进特征.....	255

8.4.2 特征匹配.....	256
8.5 利用图表的高效语法分析.....	259
8.5.1 Chart 数据结构.....	259
8.5.2 有多种解释的句子.....	259
8.6 语义解释.....	261
8.6.1 词的意思 (Sense).....	262
8.6.2 利用特征的语义解释.....	263
8.6.3 消除词的多义意思.....	265
8.7 生成自然语言.....	266
8.8 在上下文中的自然语言.....	267
8.8.1 言语的行为 (Speech Acts).....	268
8.8.2 创建引用.....	268
8.8.3 处理数据库的断言和问题.....	269
习题 8.....	273
第 9 章 Agent 技术.....	275
9.1 Agent 的定义.....	275
9.1.1 Agent 的弱定义.....	276
9.1.2 Agent 的强定义.....	276
9.2 Agent 的分类.....	276
9.2.1 按功能划分.....	277
9.2.2 按属性划分.....	278
9.2.3 其他的特殊 Agent.....	281
9.3 移动 Agent.....	281
9.3.1 移动 Agent 系统的一般结构.....	282
9.3.2 移动 Agent 的分类.....	284
9.3.3 移动 Agent 的优点.....	286
9.3.4 移动 Agent 的技术难点.....	287
9.3.5 移动 Agent 技术的标准化.....	288
9.4 移动 Agent 平台的介绍.....	289
9.4.1 General Magic 公司的 Odyssey.....	290
9.4.2 IBM 公司的 Aglet.....	290
第 10 章 知识获取的新技术.....	292
10.1 数据挖掘.....	292
10.1.1 数据挖掘的产生和最新发展.....	292
10.1.2 数据挖掘研究的主要内容.....	293
10.1.3 数据挖掘的特点.....	294
10.1.4 数据挖掘的分类.....	294

10.1.5	数据挖掘中常用的技术.....	295
10.1.6	数据挖掘过程.....	296
10.1.7	数据挖掘的研究方向及面临的困难.....	296
10.1.8	数据挖掘中的关联规则挖掘.....	297
10.2	Web 挖掘.....	303
10.2.1	Web 挖掘概述.....	303
10.2.2	Web 挖掘分类.....	304
10.2.3	Web 数据挖掘的技术难点.....	310
10.3	文本挖掘.....	310
10.3.1	文本挖掘的概念.....	310
10.3.2	文本挖掘预处理.....	311
10.3.3	文本挖掘的关键技术.....	312
10.3.4	文本挖掘系统的评价标准.....	313
第 11 章	遗传算法.....	314
11.1	概述.....	314
11.1.1	遗传算法的基本概念.....	314
11.1.2	遗传算法的基本流程.....	315
11.2	遗传编码.....	316
11.2.1	二进制编码.....	317
11.2.2	Gray 编码.....	318
11.2.3	实数编码.....	318
11.2.4	有序编码.....	318
11.2.5	结构式编码.....	319
11.3	适应值函数.....	319
11.4	遗传操作.....	320
11.4.1	选择操作 (Selection).....	320
11.4.2	交叉操作 (Crossover).....	323
11.4.3	变异操作 (Mutation).....	325
11.5	初始化群体.....	326
11.6	控制参数的选取.....	326
11.7	算法的终止准则.....	327
11.8	遗传算法的基本理论.....	327
11.8.1	模式定理.....	327
11.8.2	隐含并行性.....	329
11.8.3	构造块假设.....	329
11.8.4	收敛性.....	329
11.9	遗传算法简例.....	329
11.10	遗传算法的应用领域.....	332

习题 11	334
第 12 章 群集智能算法	335
12.1 群集智能算法的研究背景	335
12.2 群集智能基本算法的介绍	335
12.2.1 蚁群算法	335
12.2.2 Flock 算法	340
12.3 集智系统介绍	342
12.3.1 “人工鱼”	342
12.3.2 Terrarium 世界	346
12.4 群集智能的优缺点	352
第 13 章 次协调逻辑与自动推理	353
13.1 次协调逻辑的概述	353
13.1.1 传统的人工智能与经典逻辑	353
13.1.2 人工智能中不协调的数据和知识库	353
13.1.3 次协调逻辑的含义	354
13.2 注解谓词演算	354
13.2.1 多真值格	354
13.2.2 注解逻辑	356
13.2.3 注解谓词公式的语义	356
13.2.4 APC 中的不协调、非、蕴含	359
13.3 基于 APC 的 SLDa-推导和 SLDa-反驳	361
13.3.1 SLDa-推导和 SLDa-反驳	361
13.3.2 注解逻辑推理方法	362
13.3.3 注解逻辑推理举例	362
13.4 注解逻辑的归结原理	364
13.5 应用实例	368
13.6 控制策略	375
习题 13	376
参考文献	377

第1章 概 述

1.1 什么是人工智能

人类的自然智能伴随着人类活动无时不在、无处不在。人类的许多活动，如解题、下棋、猜谜、写作、编制计划和编程，甚至骑车、驾车等，都需要智能。如果机器能够完成这些任务的一部分，那么就可以认为机器已经具有某种程度的“人工智能”。

什么是人的智能？什么是人工智能？人的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，而且值得深入探讨的问题。人工智能的出现不是偶然的。从思维基础上讲，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。

为了解人工智能，先熟悉一下与它有关的一些概念，这些概念涉及到信息、认识、知识、智力、智能。不难看出，这些概念在逐步贴近人工智能。

我们首先看看什么是信息。信息、物质及能量构成整个宇宙。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。人们不能直接认识物质和能量，而是通过物质和能量的信息来认识它们。

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，并确定其联系。

由认识可以继续探讨什么是知识、什么是智力。

知识是人们对于可重复信息之间的联系的认识，是被认识了的信息和信息之间的联系，是信息经过加工整理、解释、挑选和改造而形成的。

人们接受和建立知识的能力往往被看作是智力。关于智力，科学家们有不同的定义。以下是几位科学家对智力的定义。

Wister: 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，是个体对生活中新问题和新条件的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。

Storddard: 智力是从事艰难、复杂、抽象、敏捷和创造性的活动，以及集中能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。

简言之，智力被看作个体的各种认识能力的综合，特别强调解决新问题的能力、抽象

思维、学习能力、对环境适应能力。

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、学习和联想是智能的重要因素。

至于人工智能，其英文是 Artificial Intelligence，简称 AI。字面上的意义是智能的人工制品。它是研究如何将人的智能转化为机器智能，或者用机器来模拟或实现人的智能。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在不同的年代对人工智能给出的定义：

“人工智能是研究使计算机更灵活有用、了解使智能的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模拟智能，而且是了解如何帮助人们变得更有智能。”

——1978 年 Winston P.

“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，该系统具有我们通常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。”

——1981 年 A.Barr 和 Feigenbaum E.

“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。”

——1983 年 Elaine Rich

“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。”

——1987 年 Michael R.Genesereth 和 Nils J.Nilsson

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，它研究的是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（医疗、探矿、财务等领域）内的咨询服务系统。
- 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，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，也就有了解决这个问题的模型或算法，因而也就被划分到某个学科或某个学科的分支中。因此，从某种意义上讲，人工智能永远是一个深奥而永无止境的追求目标。

1.2 AI 的产生及主要学派

AI 的起源可以追溯到丘奇 (Church)、图灵 (Turing) 和其他一些学者关于计算本质的思想萌芽。早在 20 世纪 30 年代, 他们就开始探索形式推理概念与即将发明的计算机之间的联系, 建立起了关于计算和符号处理的理论。而且, 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。

人工智能作为一门学科出现的突出标志是: 1956 年夏, 在美国达特茅斯 (Dartmouth) 大学由当时美国年轻的数学家麦卡锡 (John-McCarthy) 和他的朋友明斯基 (Minsky)、纽厄尔 (Newell)、西蒙 (Simon)、香农 (Shannon)、塞缪尔 (Samuel)、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, 举办了一个长达两个月的研讨会。会上麦卡锡提出了“Artificial Intelligence”一词, 尔后纽厄尔和西蒙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 3 个学派。

1. 符号主义学派

符号主义 (Symbolicism) 又称为逻辑主义 (Logicism)、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, 其代表性成果为启发式程序 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。

符号主义的主要理论基础是物理符号系统假设。符号主义将符号系统定义为如下 3 部分:

- 一组符号, 对应于客观世界的某些物理模型。
- 一组结构, 由与某种方式相关联的符号的实例所构成。
- 一组过程, 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。

在物理符号的假设下, 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, 即可用计算机的符号操作来模拟人的认知过程。这实质上就是认为, 人的思维是可操作的。

符号主义的基本信念是: 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。

2. 联结主义学派

联结主义 (Connectionism) 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，其主要理论基础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，用于否定基于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，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 (Artificial Neural Network, ANN) 研究的兴起，ANN 可以看作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，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 ANN 理论和应用研究。联结主义具有代表性的工作有：

- 霍普菲尔特 (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。
- 鲁梅尔哈特 (J.D.Rumelhart) 教授在 1986 年提出多层网络中的反向传播 (BP) 算法。

3. 行为主义学派

行为主义 (Actionism) 又称为进化主义 (Evolutionism) 或控制论学派 (Cyberneticsism)，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)。
- 智能行为只能通过现实世界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 专家布鲁克斯 (Brooks)。在 1991 年 8 月悉尼召开的第 12 届国际人工智能联合会议 (IJCAI) 上，布鲁克斯作为大会“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，并以他在 MIT 多年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”一文，对传统的 AI 提出了批评和挑战。