

 普通高等教育电气工程与自动化(应用型)“十二五”规划教材

Artificial Intelligence and
Its Applications

人工智能 原理及应用

© 罗兵 李华嵩 李敬民 编著



 机械工业出版社
CHINA MACHINE PRESS

普通高等教育电气工程与自动化（应用型）“十二五”规划教材

人工智能原理及应用

罗 兵 李华嵩 李敬民 编著
章 云 主审



机械工业出版社

前 言

机器真的可以思考吗？机器可以具有智能吗？机器可以有感情吗？机器最终会像人一样甚至取代人吗？

人工智能的传说可以追溯到古埃及，好莱坞也有感人的科幻电影《人工智能》。但直到 1941 年电子计算机的出现和发展，机器具有智能才成为可能。“人工智能”（Artificial Intelligence）一词最初是在 1956 年 Dartmouth 学会上正式提出的：人工智能就是要让机器的行为看起来就像是人所表现出的智能行为一样。从那以后，研究者们发展了众多的理论和原理，人工智能的概念也随之扩展，甚至有人认为：可以设计创造出机器不仅可以像人的思维一样思考和推理，机器还可以产生和人完全不一样的知觉和意识，使用和人完全不一样的推理方式。

在它还不长的历史中，人工智能的发展比预想的要慢，但也取得了很大进展。从五十多年前出现到现在，“人工智能”已从传统的、基于符号智能的“狭义人工智能”，发展成为现代的，兼容符号智能、计算智能、集成智能和分布智能的“广义人工智能”。“广义人工智能”为智能科学技术的发展及应用提供了广阔的理论、方法和技术基础，也将是 21 世纪以信息技术为主导的网络和知识经济时代的核心关键技术之一，并对推动科学技术的进步和发展发挥更大的作用。目前，人工智能是以计算机为硬件平台，通过人的编程设计，由人和计算机来共创的智能。

本书介绍了人工智能的基本原理，通过技术方法结合实际应用讲解，帮助读者更好地理解人工智能的原理，并掌握其应用技术。全书共分 8 章，分别介绍人工智能的基本原理和技术应用。每章附有习题，书后附有五个相关实验的实验指导。其中第 1、8 章和实验指导由罗兵编写，第 2、3、4 章由李华嵩编写，第 5、6、7 章由李敬民编写，全书由罗兵统稿，章云教授审核全书并提出了宝贵的修改意见。

本书配有电子课件，欢迎选用本书的老师索取，索取邮箱：ai_courseware@163.com，或登录 www.cmpedu.com 注册下载。

限于作者水平，书中难免有不妥和错误之处，恳请读者批评指正。

作 者

目 录

前言

第 1 章 绪论 1

1.1 人工智能的概念和研究意义 1

1.1.1 人工智能的概念 1

1.1.2 人工智能的研究意义 3

1.1.3 人工智能的研究目标和特点 4

1.2 人工智能的起源和发展 5

1.2.1 萌芽期 5

1.2.2 形成期 7

1.2.3 发展期 7

1.2.4 人工智能的最新研究进展和 争论 9

1.3 人工智能的研究内容和应用领域 10

1.3.1 人工智能研究的课题 10

1.3.2 人工智能的应用领域 11

习题一 14

第 2 章 知识的表示 15

2.1 知识及其表示概述 15

2.1.1 知识的含义与结构 15

2.1.2 知识的特征、分类和表示 16

2.1.3 人工智能系统所关心的知识 18

2.1.4 陈述性知识与过程性知识 19

2.2 状态空间表示 19

2.2.1 状态 19

2.2.2 操作 20

2.2.3 状态空间 20

2.2.4 问题的解 21

2.2.5 状态空间表示法求解步骤 21

2.3 谓词逻辑表示 22

2.3.1 谓词逻辑的基本内容 22

2.3.2 个体词、谓词与量词 23

2.3.3 谓词公式 24

2.3.4 谓词公式的解释 26

2.3.5 谓词逻辑表示知识 28

2.4 问题归约表示 30

2.4.1 问题归约表示的过程 31

2.4.2 问题归约法的与/或图表示 31

2.4.3 问题归约法的节点定义 31

2.4.4 问题归约表示的可解性 32

2.5 语义网络表示 35

2.5.1 语义网络表示法 35

2.5.2 语义网络的网络结构 35

2.5.3 语义网络的语义表示 36

2.5.4 连词和量化的表示 37

2.5.5 语义网络推理的性质继承及 匹配 40

2.5.6 语义网络表示法的特点 42

2.6 框架表示 42

2.6.1 框架理论 43

2.6.2 框架结构 43

2.6.3 附加过程 45

2.6.4 框架系统中的知识组织 46

2.6.5 框架系统的推理机制 47

2.6.6 框架表示法的评价 48

2.7 过程表示 49

2.7.1 过程知识表示的概念 49

2.7.2 示例：使用过程表示法求解九宫 问题 49

2.7.3 过程表示的特点 51

2.8 面向对象的表示 52

2.8.1 面向对象的概念与特性 52

2.8.2 面向对象的原则 54

2.8.3 面向对象的要素 54

2.8.4 类与类继承 57

2.8.5 面向对象的知识表示的内容 57

2.8.6 面向对象表示的实例 60

2.8.7 面向对象知识表示的特点 63

2.9 Prolog 语言概述 64

2.9.1 Prolog 语言的背景	64	4.2.2 概率推理方法	126
2.9.2 Prolog 语言的逻辑思想示例	64	4.3 主观贝叶斯方法	128
2.9.3 Prolog 语言的逻辑程序定义	67	4.3.1 知识不确定性的表示	128
2.9.4 Prolog 语言的数据结构和递归	67	4.3.2 证据不确定性的表示	131
习题二	72	4.3.3 主观贝叶斯方法的推理过程	132
第3章 确定性推理方法	73	4.3.4 主观贝叶斯方法应用举例	135
3.1 推理概述	73	4.4 可信度方法	139
3.1.1 推理的概念	73	4.4.1 可信度的概念	139
3.1.2 推理的方法	74	4.4.2 基于可信度的不确定性表示 ——CF 模型	139
3.1.3 推理的控制策略	77	4.4.3 可信度方法的推理算法	142
3.1.4 推理中的冲突	81	4.4.4 可信度方法应用举例	143
3.2 确定性推理的逻辑基础	82	4.5 证据理论 (D-S Theory)	146
3.2.1 命题公式的解释	82	4.5.1 证据理论的形式化描述	146
3.2.2 等价式	83	4.5.2 证据理论的不确定性推理 模型	152
3.2.3 永真蕴含式	84	4.5.3 证据理论应用举例	153
3.2.4 前束范式与 Skolem 范式	84	4.6 模糊推理	157
3.2.5 置换与合一	85	4.6.1 模糊集理论与模糊逻辑	157
3.3 演绎推理方法	86	4.6.2 模糊知识的表示和模糊概念的 匹配	163
3.3.1 演绎推理的概念	86	4.6.3 模糊规则推理方法	166
3.3.2 演绎推理的特点	92	习题四	168
3.4 归结推理方法	94	第5章 搜索策略	170
3.4.1 子句集及其化简	94	5.1 搜索概述	170
3.4.2 Herbrand (海伯伦) 定理	98	5.1.1 搜索的概念及类型	170
3.4.3 Robinson (鲁宾逊) 归结 原理	102	5.1.2 状态空间的搜索	171
3.4.4 利用归结推理进行定理证明	104	5.1.3 与/或树的搜索	173
3.4.5 应用归结原理进行问题求解	107	5.2 状态空间的盲目搜索策略	176
3.5 归结过程中的控制策略	109	5.2.1 状态空间图的一般搜索过程	176
3.5.1 引入控制策略的原因	109	5.2.2 广度优先搜索和深度优先 搜索	178
3.5.2 归结控制策略	110	5.2.3 代价树的搜索	180
习题三	114	5.3 状态空间的启发式搜索策略	183
第4章 不确定性推理方法	117	5.3.1 启发信息与估价函数	183
4.1 不确定推理概述	117	5.3.2 最佳优先搜索	184
4.1.1 不确定推理的概念	117	5.3.3 A* 算法	187
4.1.2 不确定推理的基本问题和方法 分类	118	5.3.4 A* 算法应用举例	188
4.1.3 不确定性推理与产生式表示	121	5.4 与/或树的盲目搜索策略	189
4.2 概率推理	123		
4.2.1 概率的基本性质和计算公式	123		

5.4.1 与/或树的搜索过程	189	习题六	228
5.4.2 与/或树的广度优先搜索	190	第7章 计算智能	230
5.4.3 与/或树的深度优先搜索	191	7.1 计算智能概述	230
5.5 与/或树的启发式搜索策略	192	7.1.1 计算智能的概念	230
5.5.1 解树的代价与希望树	192	7.1.2 计算智能的研究发展过程	230
5.5.2 与/或树的启发式搜索过程	194	7.1.3 计算智能与人工智能的关系	231
5.6 博弈树的启发式搜索	195	7.2 神经计算	231
5.6.1 博弈树的搜索过程	195	7.2.1 生物神经元	232
5.6.2 极大极小分析法	196	7.2.2 人工神经元	232
5.6.3 α - β 剪枝	198	7.2.3 神经网络的互连结构	233
习题五	200	7.2.4 神经网络的典型模型	235
第6章 机器学习	201	7.3 模糊计算	240
6.1 机器学习概述	201	7.3.1 模糊集及其运算	240
6.1.1 机器学习的概念	201	7.3.2 模糊关系及其运算	242
6.1.2 研究机器学习的意义	202	7.4 遗传算法	244
6.1.3 机器学习的发展过程	202	7.4.1 遗传算法的基本概念	244
6.1.4 机器学习的主要策略	203	7.4.2 遗传算法的基本原理	245
6.2 机械学习	204	7.4.3 遗传算法的应用	249
6.2.1 机械学习的过程	204	7.5 蚁群算法	251
6.2.2 机械学习系统要考虑的问题	205	习题七	253
6.3 归纳学习	206	第8章 人工智能应用研究	255
6.3.1 示例学习	206	8.1 专家系统	255
6.3.2 观察与发现学习	209	8.1.1 专家系统的起源和发展	255
6.4 解释学习	210	8.1.2 专家系统的结构	258
6.4.1 解释学习的基本原理	210	8.1.3 专家系统的特点	262
6.4.2 解释学习过程和算法	210	8.1.4 专家系统的建立	263
6.4.3 解释学习举例	211	8.1.5 专家系统应用实例: 混凝土成品 料温专家控制系统	264
6.5 类比学习	212	8.2 自然语言理解	268
6.5.1 类比学习的概念	212	8.2.1 概述	268
6.5.2 类比学习的表示	213	8.2.2 句法分析	271
6.5.3 类比学习的求解	214	8.2.3 词性标注	283
6.6 决策树学习	214	8.3 数据挖掘与知识发现	286
6.6.1 ID3 算法	215	8.3.1 概述	286
6.6.2 实例计算	216	8.3.2 关联规则挖掘	290
6.7 神经网络学习	219	8.3.3 分类	292
6.7.1 神经网络学习的概念	219	8.3.4 聚类分析	298
6.7.2 感知器学习	221	8.4 自动规划	301
6.7.3 BP 网络学习	224	8.4.1 概述	301
6.7.4 Hopfield 网络学习	226		

8.4.2 自动规划的原理	302	实验一 Visual Prolog 知识表示与推理	
8.4.3 自动规划技术	304	实验	323
8.4.4 自动规划技术的最新发展	307	实验二 状态空间搜索——八数码	
8.5 分布式智能与 Agent	311	问题	327
8.5.1 分布式智能概述	311	实验三 基于 MATLAB 的模糊推理系统	
8.5.2 Agent 的概念和模型	312	设计	328
8.5.3 多 Agent 的概念及其建模	317	实验四 基于 MATLAB 的神经网络	
8.5.4 Agent 之间的通信	320	设计	330
8.5.5 Agent 的体系结构	321	实验五 应用遗传算法求解优化问题	332
习题八	322	参考文献	334
附录 实验指导	323		

第 1 章 绪 论

人工智能 (Artificial Intelligence) 是计算机科学、控制论、信息论、神经生理学、心理学和语言学等多种学科互相渗透而发展起来的一门综合性新学科, 其诞生可追溯到 20 世纪 50 年代中期。1956 年夏季, 在美国达特茅斯大学, 由青年数学助教麦卡锡 (J. McCarthy) 和他的三位朋友明斯基 (M. Minsky, 哈佛大学青年数学和神经学家, 现 MIT 教授)、罗切斯特 (N. Lochester, IBM 公司信息研究中心负责人) 和香农 (C. E. Shannon, 贝尔实验室信息部数学研究员) 共同发起, 邀请 IBM 公司的莫尔 (T. More) 和塞缪尔 (A. Samuel)、MIT 的塞福里奇 (O. Selfridge) 和所罗门夫 (R. Solomonoff) 以及 RAND 公司和卡内基理工大学的诺维尔 (A. Newell) 和西蒙 (H. A. Simon, 均为 CMU 教授) 等人参加夏季学术讨论班, 历时两个月。这十位学者都是在数学、神经生理学、心理学和计算机科学等领域中从事教学和研究工作的, 他们在会上第一次正式使用了人工智能 (AI) 这一术语, 从而开创了人工智能这一研究方向。

1.1 人工智能的概念和研究意义

1.1.1 人工智能的概念

人工智能这个词看起来似乎一目了然——人制造的智能, 但是要给人工智能这个科学名词下一个准确的定义却很困难, 至今尚无统一的定义。

人类对于自身智能的思考和智能拓延的渴望古而有之, 比如传说中, 中国古代的能工巧匠们制造过指南车、铜人、木牛流马等。到了近现代, 随着社会生产的发展和变革, 以及社会分工日益细化、专业化, 出现了大量单一烦琐的复杂工作, 让机器来代替人从事这些单调重复但又需要较高专业技能的劳动的需求开始增多。与此同时, 科学技术的发展使得人们对客观世界的观察与理解达到了前所未有的高度, 人们开始更加关注对人类自身奥秘的探索。而人最本质也是最独特之处就是人的智能现象, 因此对于智能的研究开始成为新兴的热门领域。计算机技术的出现和飞速进步, 为人们研究自身的智能提供了崭新的途径。人工智能科学与技术, 就是在这种大背景下开始蓬勃发展的。对人工智能科学的研究, 既能满足人类智能的延伸并改进、改善甚至改变人类社会生产方式的需要, 又能促进对人类智能的深入认识, 从而服务于人类自身的发展。

人类对智能理解的渴望可以追溯到几百甚至上千年前。我国的《西游记》就是一个最好的例子。孙悟空的汗毛所变成的小猴子们, 不就是一个理想化的智能机器人吗?

人类的自然智能伴随着人类活动处处时时存在。人类的很多活动, 如解算题、猜谜语、进行讨论、编制计划和编写计算机程序, 甚至驾驶汽车和骑自行车等, 都体现了人类才有的智能。如果机器能够执行这些任务, 就可以认为机器已具有智能了吗?

从狭义的概念上来讲, 人工智能是计算机科学中涉及研究、设计和应用智能机器的一个

分支，是对智能计算机系统的研究。

它的近期主要目标在于研究用机器来模仿和执行人脑的某些智力功能，并开发相关理论和技术。其中智能机器是指能够在各类环境中自主地或交互地执行各种拟人任务（anthropomorphic tasks），与人智力相当或相近的机器，具体地说，是能够对人类语言进行理解、学习和推理的机器。

从广义上来讲，人工智能是指人类智能行为规律、智能理论方面的研究。

狭义方面已经做了一些工作，如专家系统的研究与开发。广义的理解至今还没有得出令人兴奋的结果。

对于人工智能虽然难以给出严格的定义，但可从不同侧面进行描述和理解。

人工智能是计算机科学的一个分支，是研究使计算机来完成能表现出人类智能的任务的学科，主要包括计算机实现智能的原理、制造类似于人脑的智能计算机以及使计算机更巧妙地实现高层次的应用。它涉及计算机科学、心理学、哲学和语言学等学科，总的目标是增强人的智能。

还可以说，人工智能是研究怎样使计算机来模仿人脑所从事的推理、学习、思考和规划等思维活动，来解决需人类专家才能处理的复杂问题，如医疗诊断、石油测井解释、气象预报、运输调度和管理决策等课题。从实用观点来看，人工智能是一门知识工程学：以知识为对象，研究知识的获取、知识的表示方法和知识的使用。

与传统的计算机程序相比较，人工智能首先研究的是以符号表示的知识而不是以数值数据为研究对象；其次，采用的是启发式推理方法而不是常规的算法；另外，人工智能的控制结构与领域知识是分离的，还允许出现不正确的解答。

智能是什么？如何理解智能？至今人们对智能的认识还很肤浅，仅介绍几种观点。

图灵（A. M. Turing）提出的智能实验，参加者是计算机、被实验的人以及主持实验的人。由主持人提出问题，计算机和被实验的人来回答，被实验者在回答问题时尽可能地向主持人表示他是“真正的”人，计算机也尽可能逼真地模仿人的思维。如果主持人通过听取对问题的回答分辨不出哪个是人的回答时，便可认为被试验的计算机是有智能的了。有人对这样设计的实验提出了疑问：认为这种实验只反映了结果的比较而没有涉及思维的过程，而且也没明确被实验的人是个孩子还是有良好素质的成年人。

当一个计算机系统能给出有关问题的正确答案或有用的建议，而解决问题所用的概念和推理与人相当、还能解释推理过程时，便可以说这样的计算机系统是有智能的了。

费根鲍姆（Feigenbaum）认为只要告诉机器做什么，而不告诉怎么做，机器便能完成任务时，就可以说机器有智能了。

渡边慧认为，人类智能主要体现在演绎能力和归纳能力，机器具有这种能力就是有智能了。

近年来在人类智能的理解上形成了两种不同的观点。一种观点称做符号主义，它认为人类的智能基本单元是符号，认知过程就是符号表示下的符号运算，从而思维就是符号计算，这种观点基于“物理符号系统假设”。一些成功的专家系统和自然语言理解系统等都是基于这种观点的。另一种观点称做联接主义，即人工神经网络，它认为智能的基本单元是神经元，认知过程是由神经网络构成的，是并行分布的，而不是符号的运算。这种观点可避免知识表示问题，带来的新问题是各神经元联接权值和网络结构设定的困难。一般说来，人工智

能系统可分两类：一类是以物理符号的途径实现的；另一类是用联接机理的途径实现的。两种途径各有优缺点，如能用其所长避其所短或把两者结合起来就更合适些。

现在，衡量一个程序是否是一个智能程序，常常从其自然语言理解能力、自动推理能力以及机器学习能力来分辨。从这种观点看，多数程序不是智能程序，少数程序可视做智能程序的雏形，原因是人们对机器学习以及自然语言理解的机理的认识还是很初步的。

对于计算机会不会有智能或者说计算机能不能思维这个有争论的问题，人们可这样来理解：“人的思维过程不是神，人脑是物质的就应该可以认识，一旦掌握了规律便可由机器来实现，这样的看法是正确的。然而人脑在开发，机器又是要人来制造的，从这点看不能说可以完全由机器来代替人脑，但是部分地代替是没有问题的。”

人们还可以问人工智能学科的基础理论是什么？有的学者认为，人工智能学科不需要基础理论的指导，说在没有空气动力学的情況下就造出了飞机。这是实干派的观点，有人称他们为“鸟飞派”。多数学者认为人工智能学科也要有理论，针对“鸟飞派”说应该回到空气动力学。他们认为：人工智能学科要有理论也要有实践，认知科学（或思维科学）是人工智能的理论基础；可把人工智能列入认知科学内的一种工程技术学科；从认知观点看，现今人工智能仅限于逻辑思维范围是不够的，要考虑形象思维、灵感思维方能促进人工智能的重大发展。数学常被认为是多种学科的基础科学，它在机械运动中、复杂的无组织的如分子运动中都有重要贡献。数学还应进入语言和思维领域，人工智能学科不借用数学工具是不可想象的。数学在人工智能中远没有发挥作用，仅限于标准逻辑、模糊数学以及一些搜索方法的使用。数学应进入人工智能学科，人工智能学科应促进数学的发展，两者将相互促进。

1.1.2 人工智能的研究意义

计算机是迄今为止最有效的信息处理工具，以至于人们称它为“电脑”。但是计算机系统的智能却相当低下，例如缺乏自适应、自学习和自优化等能力，也缺乏社会常识或专业知识等，而只能是被动地按照人们为它事先安排好的工作步骤进行工作，因而它的功能和作用就受到很大的限制，难以满足越来越复杂和广泛的社会需求。

既然计算机和人脑一样都可进行信息处理，那么是否能让计算机同人脑一样也具有智能呢？这正是人们研究人工智能的初衷。事实上，如果计算机自身也具有一定智能的话，那么，它的功效将会发生质的飞跃，成为名副其实的电“脑”。这样的电脑将是人脑更为有效的扩大和延伸，也是人类智能的扩大和延伸，其作用将是不可估量的。例如，用这样的电脑武装起来的机器人就是智能机器人。

智能机器人的出现，标志着人类社会进入一个新的时代。研究人工智能也是当前信息化社会的迫切要求。人类社会现在已经进入了信息化时代。信息化的进一步发展，就必须有智能技术的支持。例如，当前迅速发展着的互联网（Internet）、万维网（WWW）和网格计算（Grid Computation）就强烈地需要智能技术的支持。也就是说，人工智能技术在互联网、万维网和网格计算上将发挥重要作用。

智能化也是自动化发展的必然趋势。自动化发展到一定水平，再向前发展就必然是智能化。事实上，智能化将是继机械化、自动化之后，人类生产和生活中的又一个技术特征。

另外，研究人工智能，对探索人类自身智能的奥秘也可提供有益的帮助。因为人们可以通过电脑对人脑进行模拟，从而揭示人脑的工作原理，发现自然智能的渊源。事实上，现在

有一门称为“计算神经科学”的学科正迅速崛起，它从整体水平、细胞水平和分子水平对大脑进行模拟研究，以揭示其智能活动的机理和规律。

1.1.3 人工智能的研究目标和特点

虽然人工智能的定义无法准确统一，但就其本质而言，它是研究如何制造出人造的智能机器或智能系统，来模拟人类智能活动的的能力，以延伸人类智能的科学。至于人类智能活动的的能力是什么含义，人们也是有共同认识的。一般地说，它是人类在认识世界和改造世界的活动中，由脑力劳动表现出来的能力，更具体一些可概括为以下几个方面。

1) 通过视觉、听觉、触觉等感官活动，接受并理解文字、图像、声音、语言等各种外部的“自然信息”，这就是认识和理解世界环境的能力。

2) 通过人脑的生理与心理活动以及有关的信息处理过程，将感性知识抽象为理性知识，并能对事物运动的规律进行分析、判断和推理，这就是提出概念、建立方法，进行演绎和归纳推理、做出决策的能力。

3) 通过教育、训练和学习过程，日益丰富自身的知识和技能，这就是学习的能力。

4) 对变化多端的外界环境条件，如干扰、刺激等作用能灵活地做出反应，这就是自我适应的能力。

总之，人类智能是涉及信息描述和信息处理的复杂过程，因而实现人工智能是一项艰巨的任务。尽管如此，这门学科还是引起了许多科学和技术工作者的浓厚兴趣，特别是在计算机科学和技术飞速发展和计算机应用日益普及的情况下，许多学者认为实现人工智能的手段已经具备，人工智能已经开始走向实践阶段。

目前研究人工智能主要有两条途径。一条是心理学家、生理学家们认为大脑是智能活动的物质基础，要揭示人类智能的奥秘，就必须弄清大脑的结构，也就是要从大脑的神经元模型着手研究，弄清大脑信息处理过程的机理，这样人工智能的实现就可迎刃而解。显然，由于人脑有上百亿神经元，而且现阶段要进行人脑的物理模拟实验还很困难，因而完成这个任务极其艰巨。但可以看出，这一学派是企图创立“信息处理的智能理论”作为实现人工智能的长远研究目标的，这个观点是值得重视的。另一条是计算机科学家们提出的从模拟人脑功能的角度来实现人工智能，也就是把通过计算机程序的运行，从效果上达到和人们智能行为活动过程相类似作为研究目标，因而这派学者只是局限于“建造工程性智能机器或系统”，解决相关原理和技术作为实现人工智能的近期目标，这个观点比较实际，目前引起较多人的注意。

总之，不论从什么角度来研究人工智能，都是通过计算机来实现，因此可以说，人工智能的中心目标是要搞清楚实现人工智能的有关原理，使计算机有智慧、更聪明、更有用。

人工智能的研究有以下几个特点：

1) 人工智能是一门关于知识和应用的科学。以知识为对象，研究知识的获取、表示和使用；以应用知识解决问题为研究目标。

2) 人工智能的系统过程是：数据处理→知识处理，数据→符号。符号表示的是知识而不是数值、数据。

3) 问题求解过程有启发，有推导。

4) 人工智能是引起争论最多的科学之一。

当前人工智能的研究还存在着很多困惑：是应该以人类的普遍思维规律为主，还是以特定知识的处理和运用为主？智能的本质是什么？机器能达到人的水平吗？总体看来，人工智能研究是非常困难的。

麦卡锡认为：人工智能的所有问题都是难解的。明斯基认为：人工智能是有史以来最难的科学之一，其困难在于实现智能需要大量的知识，而最难对付的知识是常识而不是专业知识。德里夫斯（Dreyfus）也认为：常识问题是实现人工智能的最大障碍。

总之，万能的逻辑推理体系至今没有创造出来，并不是因为人工智能专家的本领不够，而是因为这种万能的体系从根本上就是不可能有的。人工智能系统的最大弱点就是缺乏知识，缺乏人类在几千年的文明史上积累起来的知识，在实际生活中，人是根据知识行事的，而不是根据在抽象原则上的推理行事的。

即使就推理体系来说，它的主要技术是状态空间搜索，而在执行中遇到的主要困难就是“组合爆炸”，事实表明，单靠一些思维原则是解决不了组合爆炸问题的，要摆脱困境，只有大量使用理性的知识。

所以人工智能这门学科既有研究意义和应用价值，在研究中又存在很大困难。

1.2 人工智能的起源和发展

人工智能是和众多学科紧密相关的，其起源与发展也和其他众多学科及科学家紧密相关。其发展历程大致可以分为3个阶段：萌芽期（1956年以前）、形成期（1956~1961年）和发展期（1961年以后）。

1.2.1 萌芽期

自古以来，人类就力图根据认识水平和当时的技术条件，用机器来代替人的部分脑力劳动，以提高征服自然的能力。公元前850年，古希腊就有制造机器人帮助人们劳动的神话传说。在我国，公元前900多年，也有歌舞机器人传说的记载，这说明古代人就有人工智能的幻想。

随着历史的发展，到12世纪末至13世纪初，西班牙的神学家和逻辑学家路易（Romen Luee）试图制造能解决各种问题的通用逻辑机。17世纪，法国物理学家和数学家帕斯考（B. Pascal）制成了世界上第一台会演算的机械加法器并获得实际应用。随后德国数学家和哲学家莱布尼茨（G. W. Leibniz）在这台加法器的基础上发展并制成了进行全部四则运算的计算器。他还提出了逻辑机的设计思想，即通过符号体系，对对象的特征进行推理，这种“万能符号”和“推理计算”的思想是现代化“思考”机器的萌芽，因而他曾被后人誉为数理逻辑的奠基人。19世纪，英国数学和力学家巴贝基（C. Babbage）致力于差分机和分析机的研究，虽因条件限制未能完全实现，但其设计思想不愧为当时人工智能的最高成就。

进入20世纪后，人工智能相继出现若干开创性的工作。1936年，年仅24岁的英国数学家图灵在他的一篇“理想计算机”的论文中，就提出了著名的图灵机模型，1945年他进一步论述了电子数字计算机设计思想，1950年他又在《计算机能思维吗？》一文中提出了机器能够思维的论述，可以说这些都是图灵为人工智能所作的杰出贡献。1938年德国青年工程师祖斯（Zuse）研制成了第一台累计数字计算机Z-1，后来又进行了改进，到1945年他

又发明了 Plankal 程序语言。此外，1946 年美国科学家茅克利 (J. W. Mauchly) 等人制成了世界上第一台电子数字计算机 ENIAC。还有同一时代美国数学家维纳 (N. Wiener) 控制论的创立，美国数学家香农信息论的创立，英国生物学家阿斯比 (W. R. Ashby) 所设计的脑模型等，这一切都为人工智能学科的诞生在理论和实验工具上做出了巨大贡献。

回顾人工智能的萌芽期，相关的重要科学家及其贡献有：

1) 古希腊伟大的哲学家、思想家亚里士多德 (Aristotle) (公元前 384 ~ 322)，他的主要贡献是为形式逻辑奠定了基础。形式逻辑是一切推理活动的最基本的出发点。

在他的代表作《工具论》中，就给出了形式逻辑的一些基本规律，如矛盾律、排中律，并且实际上已经提到了同一律和充足理由律。此外，亚里士多德还研究了概念、判断问题，以及概念的分类和概念之间的关系，判断问题的分类和它们之间的关系。他最著名的创造就是提出人人熟知的三段论。

2) 英国的哲学家、自然科学家培根 (Bacon) (1561 ~ 1626)，他的主要贡献是系统地给出了归纳法，成为和亚里士多德的演绎法相辅相成的思维法则。培根的另一个功绩是强调了知识的作用。培根的著名警句是“知识就是力量”。

3) 德国数学家、哲学家莱布尼茨提出了关于数理逻辑的思想，把形式逻辑符号化，从而能对人的思维进行运算和推理。他做出了能进行四则运算的手摇计算机。

4) 英国数学家、逻辑学家布尔 (Boole) (1815 ~ 1864)，他初步实现了布莱尼茨的思维符号化和数学化的思想，提出了一种崭新的代数系统——布尔代数。

5) 美籍奥地利数理逻辑学家哥德尔 (Godel) (1906 ~ 1978)，他证明了一阶谓词的完备性定理：任何包含初等数论的形式系统，如果它是无矛盾的，那么一定是不完备的。

此定理的意义在于，人的思维形式化和机械化的某种极限，在理论上证明了有些事是做不到的。

6) 英国数学家图灵 1936 年提出了一种理想计算机的数学模型——图灵机，1950 年提出了图灵试验，发表了“计算机与智能”的论文。当今世界上计算机科学最高荣誉奖励为“图灵奖”。

著名的图灵试验是：当一个人与一个封闭房间里的人或者机器交谈时，如果他不能分辨自己问题的回答是计算机还是人给出时，则称该机器是具有智能的。以往该试验几乎是衡量机器人人工智能的唯一标准，但是从 20 世纪 90 年代开始，现代人工智能领域的科学家开始对此试验提出异议：反对封闭式的，机器完全自主的智能；提出与外界交流的，人机交互的智能。

7) 美国科学家茅克利等人于 1946 年发明了电子数字计算机 ENIAC。

8) 美国神经生理学家麦克库洛 (McCulloch)，建立了第一个神经网络数学模型。

从某种意义上可以说近代人工智能的发展，首先是从人工神经网络研究开始的。但是由于某种原因，神经网络的研究一度进入低潮。

9) 美国数学家香农 1948 年发表了《通讯的数学理论》，标志着“信息论”的诞生。

10) 美国数学家、计算机科学家麦卡锡，人工智能的早期研究者。1956 年，他和其他一些学者联合发起召开了世界上第一次人工智能学术大会，在他的提议下，会上正式决定使用“人工智能”一词来概括这个研究方向。

参加大会的有明斯基、罗切斯特、香农、莫尔、塞缪尔、塞福里奇、所罗门夫、西蒙和

纽维尔等数学家、心理学家、神经生理学家、计算机科学家。麦卡锡也被尊为“人工智能之父”。

1.2.2 形成期

1956年,在美国的达特茅斯大学的一次历史性的聚会被认为是人工智能学科正式诞生的标志,从此在美国开始形成了以人工智能为研究目标的几个研究组:如纽维尔和西蒙的Carnegie-RAND协作组;塞缪尔和基内特(Gelernter)的IBM公司工程课题研究组;明斯基和麦卡锡的MIT研究组等,这一时期人工智能的研究工作主要在下述几个方面。

1957年,纽维尔、肖(J. Shaw)和西蒙等人的心理学小组编制出一个称为逻辑理论机LT(The Logic Theory Machine)的数学定理证明程序,当时该程序证明了努塞尔(B. A. W. Russell)和怀特海德(A. N. Whitehead)的《数学原理》一书第二章中的38个定理(1963年修订的程序在大机器上终于证明完了该章中全部52个定理)。后来他们又揭示了人在解题时的思维过程大致可归结为3个阶段:

- (1) 先想出大致的解题计划。
- (2) 根据记忆中的公理、定理和推理规则组织解题过程。
- (3) 进行方法和目的分析,修正解题计划。

这种思维活动不仅用于解数学题,也大致适用于解决其他问题。基于这一思想,他们于1960年又编制了能解10种类型不同课题的通用问题求解程序(General Problem Solving, GPS)。另外他们还发明了编程的表处理技术和NSS国际象棋机。和这些工作有联系的纽维尔关于自适应象棋机的论文和西蒙关于问题求解和决策过程中合理选择和环境影响的行为理论的论文,也是当时信息处理研究方面的巨大成就。后来他们的学生还做了许多工作,如人的口语学习和记忆的EPAM模型(1959年)、早期自然语言理解程序SAD-SAM等。此外他们还还对启发式求解方法进行了探讨。

1956年塞缪尔研究的具有自学习、自组织、自适应能力的西洋跳棋程序是IBM小组有影响的工作,这个程序可以像一个优秀棋手那样,向前看几步来下棋。它还能学习棋谱,在分析大约175000副不同棋局后,可猜测出书上所有推荐的走步,准确度达48%,这是机器模拟人类学习过程卓有成就的探索。1959年这个程序曾战胜设计者本人,1962年还击败了美国一个州的跳棋大师。

在MIT小组,1959年麦卡锡发明的表(符号)处理语言LISP,成为人工智能程序设计的主要语言,至今仍被广泛采用。1958年麦卡锡建立的行动计划咨询系统以及1960年明斯基的论文《走向人工智能的步骤》,对人工智能的发展都起了积极的作用。

此外,1956年乔姆斯基(N. Chomsky)的文法体系,1958年塞福里奇等人的模式识别系统程序等,都对人工智能的研究产生有益的影响。这些早期成果,充分表明人工智能作为一门新兴学科正在茁壮成长。

1.2.3 发展期

20世纪60年代以来,人工智能的研究活动越来越受到重视。为了揭示智能的有关原理,研究者们相继对问题求解、博弈、定理证明、程序设计、机器视觉和自然语言理解等领域的课题进行了深入的研究。几十年来,不仅使研究课题有所扩展和深入,而且还逐渐理清

了这些课题共同的基本核心问题以及它们和其他学科间的相互关系。1974年尼尔森(N. J. Nilsson)对发展时期的一些工作写过一篇综述论文,他把人工智能的研究归纳为4个核心课题和8个应用课题,并分别对它们进行论述。

这一时期中某些课题曾出现一些较有代表性的工作,1965年洛宾森(J. A. Robinson)提出了归结(消解)原理,推动了自动定理证明这一课题的发展。20世纪70年代初,威洛格(T. Winograd)、锡安克(R. C. Schank)和西蒙等人在自然语言理解方面做了许多发展工作,较重要的成就是威洛格提出的积木世界中理解自然语言的程序。关于知识表示技术有格林(C. Green)1966年的一阶谓词演算语句,奎林(M. R. Quillian)1966年的语义记忆的网络结构,西蒙等人1973年的语义网结构,锡安克1972年的概念网结构,明斯基1974年的框架系统的分层组织结构等。关于专家系统自1965年研制DENDRAL系统以来,一直受到人们的重视,这是人工智能走向实际应用最引人注目的课题。1977年,E. A. 费根鲍姆提出了知识工程(Knowledge Engineering)的研究方向,导致了专家系统和知识库系统更深入的研究和开发工作。此外智能机器人、自然语言理解和自动程序设计等课题,也是这一时期较集中的研究课题,也取得不少成果。

从20世纪80年代中期开始,经历了10多年的低潮之后,有关人工神经网络的研究取得了突破性的进展。1982年生物物理学家霍普菲尔德(Hopfield)提出了一种新的全互联的神经网络模型,被称为Hopfield模型。利用该模型的能量单调下降特性,可用于求解优化问题的近似计算。1985年霍普菲尔德利用这种模型成功地求解了“旅行商(TSP)”问题。1986年罗姆哈特(Rumelhart)提出了反向传播(Back Propagation - BP)学习算法,解决了多层人工神经网络的学习问题,成为广泛应用的神经网络学习算法。从此,掀起了新的人工神经网络的研究热潮,提出了很多新的神经网络模型,并被广泛地应用于模式识别、故障诊断、预测和智能控制等多个领域。

1997年5月,IBM公司研制的“深蓝”计算机,以3.5:2.5的比分,首次在正式比赛中战胜了人类国际象棋世界冠军卡斯帕罗夫,在世界范围内引起了轰动。这标志着在某些领域,经过努力,人工智能系统可以达到人类的最高水平。

这一时期学术交流的发展对人工智能的研究有很大推动作用。1969年国际人工智能联合会成立,并举行第一次学术会议IJCAI-69(International Joint Conference on Artificial Intelligence),以后每两年召开一次。随着人工智能研究的发展,1974年又成立了欧洲人工智能学会,并召开第一次会议ECAI(European Conference on Artificial Intelligence),随后也是相隔两年召开一次。此外许多国家也都有本国的人工智能学术团体。在人工智能刊物方面,1970年创办了《Artificial Intelligence》国际性期刊,爱丁堡大学还不定期出版《Machine Intelligence》杂志,还有IJCAI会议文集,ECAI会议文集等。此外,ACM、AFIPS和IEEE等刊物也刊载人工智能的论著。

美国是人工智能的发源地,随着人工智能的发展,世界各国有关学者也都相继加入这一行列,英国在20世纪60年代就起步人工智能的研究,到70年代,在爱丁堡大学还成立了“人工智能”系。日本和西欧一些国家虽起步较晚,但发展都较快,前苏联对人工智能研究也开始予以重视。我国从1978年才开始人工智能课题的研究,主要在定理证明、汉语自然语言理解、机器人及专家系统方面设立课题,并取得一些初步成果。我国也先后成立了中国人工智能学会、中国计算机学会人工智能和模式识别专业委员会和中国自动化学会模式识别

与机器智能专业委员会等学术团体,开展这方面的学术交流。此外,国家已建立了若干个与人工智能研究有关的国家重点实验室,这些都将促进我国人工智能的研究,为这一学科的发展做出贡献。

1.2.4 人工智能的最新研究进展和争论

在物理符号系统假设和基于知识的启发式求解方法指导下,专家系统、机器翻译、机器视觉、问题求解等方面的研究已有实际应用。20世纪80年代以来,对知识表示、常识推理、机器学习、分布式人工智能以及智能机器体系结构等基础性研究方面也取得了可喜的进展。区别于物理符号假设的连接主义,神经网络的研究方面也有明显的进展,它避开了知识表示带来的困难,而神经元间联结的结构和权值选取又成了难题。人工智能研究目前处于什么状态说法不一,认为失败的毕竟是少数,多数认为近年来没有重大突破,甚至出现了某种程度的踏步不前,这值得人工智能学界反思。明斯基,布鲁克斯(Brooks),柯西(Kirsh)等提出了不同的观点和人工智能的基本问题。

1. 人工智能基础的5个基本问题

1987年5月在麻省理工学院(MIT)召开了人工智能专题讨论会,几位代表人物各自阐明了对人工智能基础的认识,而且评价了有关基础性工作。

1987~1988年,《Computational Intelligence》杂志先后发表了德莫特(Mc Dermott)写的《纯理性批判》、奇斯曼(Peter Cheesman)写的《计算机理解质疑》。1987年以来的人工智能学术会议上都有以非单调推理为背景的常识表示和常识推理的争论。1990年出版了帕特里奇(Partridge)和维尔克斯(Wilks)的《人工智能基础》和保顿(Magaret Boden)的《人工智能哲学》等专著。这些讨论和论述形成了人工智能基础辩论的局面。

1991年有代表性的杂志《Artificial Intelligence》第47卷发表了人工智能基础专辑,指出了人工智能研究的趋势,并就人工智能有关方法的基础性假设进行了辩论。柯西在专辑中提出了人工智能的5个基本问题:

- (1) 知识与概念化是否是人工智能的核心?
- (2) 认知能力能否与载体分开来研究?
- (3) 认知的轨迹是否可用类自然语言来描述?
- (4) 学习能否与认知分开来研究?
- (5) 所有的认知是否有一种统一的结构?

尼尔森代表的逻辑主义,对(1)~(4)给予肯定的回答,对(5)持中立观点。纽维尔(Newell)代表的认知主义,对(1)、(3)、(5)给予肯定的回答。而布鲁克斯代表的进化主义,对(1)~(5)完全否定。

2. 思维的社会

明斯基认为人工智能研究已进入死胡同,原因是人们追求某种统一的知识表示和理论基础。他认为人的智能中,根本不存在统一的理论,那里有几百个理论在起作用,彼此既有合作又有矛盾,但根本没有什么一致性、完备性。他把人的智能称为“思维的社会”(Society of Mind),意指由大量具有某种思维能力的单元组成的复杂社会。这种观点与麦卡锡主张的以非单调逻辑为中心的常识推理和与尼尔森主张的逻辑主义观点是完全不同的。

3. 无需表示、无需理性的智能

人工智能的研究至今过于集中在一些经过抽象的、过分简单化的现实世界模型（如积木）之上，研究一些玩具式的小例子（如八数码）。布鲁克斯认为应走出这种抽象模型的象牙塔，而以复杂的现实世界为背景，让人工智能理论和技术先经受解决实际问题的考验，并在这种考验中成长。

要求机器人像人一样的思维太困难了，在做像样的机器人之前，不如先做一个像样的“机器虫”，布鲁克斯果真实现了。机器虫没有像人那样的推理、规划能力，可是应付复杂环境的能力大大超过了现有的机器人。

布鲁克斯进而提出了无需表示、无需理性（推理）的智能。其基本观点是：到现场去、物理实现、初级智能、行为产生智能。

以这些思想为基础，布鲁克斯设计了一种控制结构，是一个松散的集合体，是一些相对独立的功能单元组成的分层异步分布式网络，把机器虫所需的一些基本功能如避让、前进、平衡等分离出来，分别发挥作用。布鲁克斯的设想获得了一定程度的成功，有些机器虫已成商品。

布鲁克斯的观点反对者大有人在，他们提出：今天的虫明天能变成人吗？

4. 逻辑的作用

尼尔森是主张用逻辑来研究人工智能的代表。他指出人工智能的逻辑学派基于 3 个论题：

- (1) 智能机器必须有关于自身环境的知识。
- (2) 通用智能机器要能陈述性地表达关于自身环境的大部分知识。
- (3) 通用智能机器表达陈述性知识的语言至少要有一阶逻辑的表达能力。

逻辑学派还对知识是什么、世界是什么、机器与世界如何交互以及设计智能机器的过程的任务等进行了研究，强调的是知识、概念化、模型论语义。

反对派也提出不同的观点和问题，责问逻辑学派究竟干了些什么，人类的思维大部分是非逻辑的，基于逻辑的机器不能达到人类的认识能力。把任何事物都用统一的逻辑框架来表达是不现实的。逻辑学派强调可靠的演绎推理，忽略了归纳、类比、概率推理。以模型论语义作为知识表示的核心也值得怀疑。

人工智能研究中的争论并未妨碍其快速发展和各国对其重视。我国把人工智能列为 863 高技术计划重点研究技术之一。随着计算机技术特别是互联网的普及和应用，对人工智能的需求变得越来越迫切，同时也给人工智能的研究提供了新的广阔的舞台。

1.3 人工智能的研究内容和应用领域

1.3.1 人工智能研究的课题

人工智能的知识领域浩繁，很难面面俱到，但是各个领域的思想和方法上有许多可以互相借鉴的地方。从基础理论的角度出发，其研究基本内容包括：

(1) 启发式搜索理论：搜索的方法很多，如回溯、图搜索、启发式等，主要是给定一些经验做指导提高搜索效率。该方面的研究已经有了比较成熟的技术。