

人工智能 人才培养系列

# 人工智能

## 原理与实践

### 基于 Python 语言和 TensorFlow

◎ 张明 何艳珊 杜永文 编著

深入浅出，  
原理与案例相结合，  
涵盖 TensorFlow 的  
主要内容，  
适合人工智能零基础  
初学者。

减少纯数学  
理论的研究探讨，  
对概念和理论知识  
简化处理。

从实际问题出发，  
在实践中介绍知识点  
和关联应用，  
突出 TensorFlow  
开发技能。

工信出版集团

人民邮电出版社  
POSTS & TELECOM PRESS

人工智能 人才培养系列

# 人工智能

## 原理与实践

基于 Python 语言和 TensorFlow

© 张明 何艳珊 杜永文 编著

人民邮电出版社  
北京

## 图书在版编目 (C I P) 数据

人工智能原理与实践：基于Python语言和  
TensorFlow / 张明, 何艳珊, 杜永文编著. — 北京：  
人民邮电出版社, 2019. 8  
ISBN 978-7-115-50929-1

I. ①人… II. ①张… ②何… ③杜… III. ①人工智  
能—算法②软件工具—程序设计 IV. ①TP18  
②TP311.561

中国版本图书馆CIP数据核字(2019)第040865号

## 内 容 提 要

TensorFlow 是谷歌公司开发的一款开源的人工智能应用框架, 用于帮助研究者研究和部署深度神经网络, 并为研究者提供一个加速深度学习的平台, 到目前为止, TensorFlow 已经成为世界上应用最广泛的机器学习平台。本书从人工智能的基础讲起, 深入到人工智能的框架原理、模型构建、源代码分析等各方面内容。全书分为基础篇和实战篇两部分, 共 11 章。基础篇部分包括第 1 章到第 4 章, 讲解人工智能的入门知识, 包括人工智能的概述、Python 编程语言的基本知识点、TensorFlow 的基础知识点和运行方式, 这 4 章是后续章节学习和实践的基础。实战篇部分包括第 5 章到第 11 章, 讲解如何用 TensorFlow 完成 MNIST 机器学习、卷积神经网络、字词嵌入、递归神经网络、Mandelbrot 集合、偏微分方程模拟仿真和人脸识别的应用实战。

本书深入浅出, 原理与案例相结合, 涵盖了 TensorFlow 的主要内容, 非常适合对人工智能和 TensorFlow 感兴趣的零基础初学者阅读学习。

- 
- ◆ 编 著 张 明 何艳珊 杜永文  
责任编辑 刘 博  
责任印制 陈 犇
  - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号  
邮编 100164 电子邮件 315@ptpress.com.cn  
网址 <http://www.ptpress.com.cn>  
三河市中晟雅豪印务有限公司印刷
  - ◆ 开本: 787×1092 1/16  
印张: 13.25 2019 年 8 月第 1 版  
字数: 331 千字 2019 年 8 月河北第 1 次印刷
- 

定价: 49.80 元

读者服务热线: (010) 81055256 印装质量热线: (010) 81055316

反盗版热线: (010) 81055315

广告经营许可证: 京东工商广登字 20170147 号

# 前言

2016年，在谷歌公司的AlphaGo击败韩国围棋选手李世石后，掀起了新一轮的人工智能学习热潮。经过近两三年的发展与普及，人工智能的应用遍地开花，逐渐渗透到人们生活的各个领域，并对这些领域的发展起到了不同程度的推动作用。

众所周知，计算机发展到高级阶段，人工智能是其重要发展方向，机器学习则是人工智能的一个工具，而深度学习又包含在机器学习中，是机器学习的一部分。不管是人工智能的学习还是机器学习或者是深度学习，它们皆是基于线性代数、矩阵、微积分、概率论等数学理论的，具有复杂的数学模型和算法。这些枯燥复杂的数学公式增加了人工智能入门的难度，影响了许多对人工智能领域感兴趣的读者继续学习下去的信心，甚至让他们经历了从有兴趣到放弃的过程。能够使对人工智能感兴趣的零基础初学者顺利走完由有兴趣到入门，由零基础到略懂的阶段，则是作者撰写本书的动力。希望本书能够帮助广大的人工智能初学者跨过初期学习的难关，享受到学习人工智能技术的乐趣。

深度学习的框架有许多，其中应用最广泛的是谷歌公司开发AlphaGo所用的TensorFlow平台框架，它集合了神经网络的各种算法和函数，并形成了一个工具箱。同时，TensorFlow是基于Python编程语言的，具有简单易学、能轻松上手的优势，因此一经开源便受到了各行各业的喜爱。本书从人工智能的基础原理讲起，逐一介绍Python语言、TensorFlow的基本概念和安装、TensorFlow的基础运作方式、使用TensorFlow实现MNIST机器学习、卷积神经网络、字词的向量表示、递归神经网络、人脸识别等内容。在介绍这些实战例子的同时，本书深入浅出地介绍了算法模型背后的理论内容。本书尽量减少对纯数学理论的研究探讨，对概念和一些理论知识也做了一定的简化处理；而且注重从实际问题出发，在章节设定上做到循序渐进，在实践中介绍知识点和关联应用，确保读者能够快速掌握各个知识点及相关技术，可以使用TensorFlow来做一些实际工作，并逐渐具备进一步自学的能力。

由于编写仓促，编写水平有限，书中难免出现疏漏和不妥之处，敬请各位读者多多批评指正。

本人与几位同事合作编写了本书，并在此过程中得到了他们的大量帮助，在此向何老师、杜老师表示衷心的感谢。

张明

2019年5月

# 目 录

## 基础篇

<b>第1章 绪论</b> .....	1
1.1 人工智能简介.....	1
1.1.1 人工智能的概念.....	1
1.1.2 现代人工智能的兴起.....	5
1.1.3 人工智能的学术流派.....	5
1.2 人工智能的发展历史.....	8
1.2.1 孕育期（1956年之前）.....	8
1.2.2 形成期（1956~1969年）.....	9
1.2.3 发展期（1970年之后）.....	11
1.3 人工智能技术的研究内容与应用领域.....	13
1.3.1 神经网络.....	14
1.3.2 机器学习.....	15
1.3.3 模式识别.....	15
1.3.4 自然语言理解.....	16
1.3.5 专家系统.....	17
1.3.6 博弈.....	17
1.3.7 智能控制.....	18
1.3.8 其他.....	18
1.4 人工智能与TensorFlow.....	18
1.4.1 机器学习与深度学习.....	18
1.4.2 TensorFlow概念.....	20
1.4.3 TensorFlow的应用.....	23
<b>第2章 Python基础应用</b> .....	25
2.1 引言.....	25
2.2 Python的安装.....	25
2.3 数据类型与数据结构.....	29
2.4 数字.....	29

2.5 变量及其命名规则.....	29
2.6 语句和表达式.....	30
2.7 字符串.....	31
2.8 容器.....	32
2.8.1 列表.....	32
2.8.2 元组.....	35
2.8.3 字典.....	35
2.8.4 复制.....	36
2.9 函数.....	38
2.9.1 常用内置函数及高阶函数.....	38
2.9.2 用户自定义函数.....	42
2.10 常用库.....	43
2.10.1 时间库.....	43
2.10.2 科学计算库（NumPy）.....	47
2.10.3 可视化绘图库（Matplotlib）.....	54
2.10.4 锁与线程.....	58
2.10.5 多线程编程.....	59
<b>第3章 TensorFlow基础</b> .....	62
3.1 TensorFlow的架构.....	62
3.2 TensorFlow的开发环境搭建.....	66
3.3 数据流图简介.....	77
3.3.1 数据流图基础.....	77
3.3.2 节点的依赖关系.....	80
3.4 TensorFlow中定义数据流图.....	83
3.4.1 构建一个TensorFlow数据流图.....	83
3.4.2 张量思维.....	87
3.4.3 张量的形状.....	90
3.4.4 TensorFlow的Op.....	91
3.4.5 TensorFlow的Graph对象.....	93
3.4.6 TensorFlow的Session.....	94
3.4.7 输入与占位符.....	97

3.4.8 Variable 对象 .....	98	6.2 卷积神经网络的模型架构 .....	142
3.5 通过名称作用域组织数据流图 .....	100	6.2.1 ImageNet-2010 网络结构 .....	142
3.6 构建数据流图 .....	105	6.2.2 DeepID 网络结构 .....	143
3.7 运行数据流图 .....	108	6.3 卷积运算 .....	144
<b>第 4 章 TensorFlow 运作方式</b> .....	114	6.3.1 输入和卷积核 .....	145
4.1 数据的准备和下载 .....	114	6.3.2 降维 .....	145
4.2 图表构建与推理 .....	115	6.3.3 填充 .....	145
4.2.1 图表构建 .....	115	6.3.4 数据格式 .....	145
4.2.2 推理 .....	116	6.4 卷积常见层 .....	146
4.3 损失与训练 .....	117	6.4.1 卷积层 .....	146
4.3.1 损失 .....	117	6.4.2 池化层 .....	149
4.3.2 训练 .....	117	6.4.3 归一化 .....	150
4.4 状态检查与可视化 .....	118	6.4.4 高级层 .....	151
4.4.1 状态检查 .....	118	6.5 TensorFlow 和图像 .....	152
4.4.2 状态可视化 .....	119	6.5.1 图像加载 .....	152
4.5 评估模型 .....	120	6.5.2 图像格式 .....	152
4.6 评估图表的构建与输出 .....	123	6.5.3 图像操作 .....	152
4.6.1 评估图表的构建 .....	123	6.5.4 颜色空间变换 .....	153
4.6.2 评估图表的输出 .....	123	6.6 模型训练 .....	153
<b>实 战 篇</b>		6.7 模型评估 .....	154
<b>第 5 章 MNIST 机器学习</b> .....	125	6.8 多 GPU 的模型训练 .....	154
5.1 MNIST 数据集简介 .....	125	<b>第 7 章 字词的向量表示</b> .....	155
5.2 MNIST 数据下载 .....	127	7.1 WordEmbedding 的基本概念和知识 .....	156
5.2.1 数据的准备 .....	129	7.2 Skip-Gram 模型 .....	158
5.2.2 数据重构 .....	130	7.2.1 数据集的准备 .....	160
5.2.3 数据集对象 .....	130	7.2.2 模型结构 .....	161
5.3 softmax 回归模型简介 .....	131	7.2.3 处理噪声对比 .....	162
5.4 模型的训练与评估 .....	132	7.2.4 模型训练 .....	163
5.5 TensorFlow 模型基本步骤 .....	135	7.3 嵌套学习可视化与评估 .....	164
5.6 构建 softmax 回归模型 .....	135	7.4 优化实现 .....	166
<b>第 6 章 卷积神经网络</b> .....	138	<b>第 8 章 递归神经网络</b> .....	168
6.1 卷积神经网络 .....	138	8.1 递归神经网络的架构 .....	169
		8.2 PTB 数据 .....	170
		8.3 模型及 LSTM .....	170

8.3.1 LSTM 的概念 .....	172	11.2 人脸识别的流程 .....	188
8.3.2 LSTM 的结构 .....	173	11.2.1 人脸图像的采集 .....	188
8.3.3 LSTM 的控制门 .....	173	11.2.2 人脸图像的检测 .....	189
8.4 反向传播的截断 .....	175	11.2.3 人脸图像的预处理 .....	189
8.5 输入与损失函数 .....	175	11.2.4 人脸图像的特征提取 .....	189
8.6 多个 LSTM 层堆叠 .....	175	11.2.5 人脸图像的匹配与识别 .....	190
8.7 代码的编译与运行 .....	176	11.2.6 活体鉴别 .....	190
<b>第 9 章 Mandelbrot 集合</b> .....	177	11.3 人脸识别种类 .....	190
9.1 库的导入 .....	178	11.3.1 人脸检测 .....	190
9.2 会话和变量初始化 .....	179	11.3.2 人脸关键点检测 .....	191
9.3 定义并运行计算 .....	179	11.3.3 人脸验证 .....	194
<b>第 10 章 偏微分方程模拟仿真</b> .....	180	11.4 人脸检测 .....	194
10.1 计算函数的定义 .....	180	11.4.1 LFW 数据集 .....	194
10.2 偏微分方程的定义 .....	182	11.4.2 数据预处理与检测 .....	195
10.3 仿真 .....	183	11.5 性别和年龄识别 .....	196
<b>第 11 章 人脸识别</b> .....	185	11.5.1 数据预处理 .....	198
11.1 人脸识别概念 .....	185	11.5.2 模型构建 .....	198
		11.5.3 模型训练 .....	203
		11.5.4 模型验证 .....	204

# 基础篇

## 01

## 第1章 绪论

### 1.1 人工智能简介

人工智能 (Artificial Intelligence, AI) 与空间技术、能源技术并称为世界三大尖端技术, 也被称为继三次工业革命后的又一次革命, 它是在计算机科学、控制论、信息论、神经生理学、哲学、语言学等多种学科研究的基础上发展起来的, 是一门新思想、新观念、新理论、新技术不断涌现的前沿性学科和迅速发展的综合性学科。目前, 人工智能在很多学科领域都得到了广泛应用, 并取得了丰硕成果, 无论在理论还是实践上都已自成系统。

2017年12月, 人工智能入选“2017年度中国媒体十大流行语”。经过多年的演进, 人工智能的发展进入了新阶段。AlphaGo的胜利, 无人驾驶的成功, 模式识别的突破性进展, 这些人工智能的发展成果一次又一次地牵动着我们的神经。

#### 1.1.1 人工智能的概念

什么是智能? 什么是人工智能? 人工智能和人的智能、动物的智能有什么区别和联系? 这些是每个人工智能初学者都会问到的问题, 也是学术界长期争论却没有定论的问题。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能进行计算、推理和其他思维活动的智能机器的必然结果; 从理论基础上讲, 它是控制论、信息论、系统论、计算机科学、神经生理学、心理学、数学和哲学等多种学科相互渗透的结果; 从物质基础上讲, 它是电子计算机的出现和广泛应用的结果。

为了更好地理解人工智能的内涵, 本节首先介绍一些与之相关的基本概念。

##### 1. 智能的概念

智能的拉丁文表示是 *Legere*, 意思是收集、汇集。但究竟智能是什么? 智能的本质是什么? 智能是如何产生的? 尽管相关的学者和研究人员一直在努力探究, 但这些问题仍然没有完全解决, 依然困扰着人类。

虽然近年来, 神经生理学家、心理学家等对人脑的结构和功能有了一些初步认识, 但对整个神经系统的内部结构和作用机制, 特别是人脑的功能原理还没有完全研究清楚, 因此对智能做出一个精确、可被公认的定义显然是不可能的, 研究人员只能基于自己的研究领域, 从不同角度对智能进

行侧面描述，如思维理论的观点、知识阈值理论的观点和进化理论的观点。对这些观点的了解可以帮助我们勾勒出智能的内涵和特征。

思维理论认为智能的核心是思维，人的一切智慧或智能都来自大脑的思维活动，人类的一切知识都是人们思维的产物，因而通过对思维规律与方法的研究可以揭示智能的本质。思维理论来源于认知科学，认知科学是研究人们认识客观世界的规律和方法的一门学科。

知识阈值理论认为，智能行为取决于知识的数量和知识的一般化程度，系统的智能来自它运用知识的能力，智能就是在巨大的搜索空间中迅速找到一个满意解的能力。知识阈值理论强调知识在智能中的重要意义和作用，推动了专家系统、知识工程等领域的发展。

进化理论认为，人的本质能力是在动态环境中的行走能力、对外界事物的感知能力、维持生命和繁衍生殖的能力，这些本质能力为智能发展提供了基础，因此智能是某种复杂系统所呈现的性质，是许多部件交互作用的结果。智能仅仅由系统总的行为及行为与环境的联系所决定，它可以在没有明显可操作的内部表达的情况下产生，也可以在没有明显的推理系统出现的情况下产生。进化理论是由美国麻省理工学院（MIT）的布鲁克斯（Brooks）教授提出的，他是人工智能进化主义学派的代表人物。

综合上述各种观点，可以认为智能是知识与智力结合的产物。知识是智能行为的基础，智力是获取知识并运用知识求解问题的能力。智能具有以下特征。

### （1）感知能力

感知能力是指人类通过诸如视觉、听觉、触觉、味觉、嗅觉等感觉器官感知外部世界的能力。感知是人类最基本的生理、心理现象，也是获取外部信息的基本途径。人类通过感知能力获得关于世界的相关信息，然后经大脑加工成为知识，感知是智能活动产生的前提和基础。

事实上，人类通常对感知到的外界信息有两种不同的处理方式：在紧急或简单情形下，不经大脑思索，直接由底层智能机构做出反应；在复杂情形下，通过大脑思维，做出反应。

### （2）记忆与思维能力

记忆与思维都是人脑的重要特征，记忆存储感觉器官感知到的外部信息和思维产生的知识；思维则对记忆的信息进行处理，动态地利用已有知识对信息进行分析、计算、比较、判断、推理、联想、决策等，是获取知识、运用知识并最终求解问题的根本途径。

思维包括逻辑思维、形象思维和灵感思维等，其中，逻辑思维与形象思维是最基本的两类思维方式；灵感思维是指人在潜意识的激发下获得灵感而“忽然开窍”的思维活动，也称顿悟思维。神经生理学家研究发现，逻辑思维与左半脑的活动有关，形象思维与右半脑的活动有关。

#### ① 逻辑思维

逻辑思维也被称为抽象思维，是根据逻辑规则对信息理性处理的过程，反映了人们以抽象、间接、概括的方式认识客观世界的过程。推理、证明、思考等活动都是典型的逻辑思维过程。逻辑思维具有以下特征。

- 思维过程是串行的、线性的过程。
- 容易形式化，可以用符号串表示思维过程。
- 思维过程严密、可靠，可用于从逻辑上合理预测事物的发展，加深人们对事物的认识。

#### ② 形象思维

形象思维以客观现象为思维对象，以感性形象认识为思维材料，以意象为主要思维工具，以指

导创造物化形象的实践为主要目的,也称直感思维,如图像识别、视觉信息加工等,形象思维具有以下特征。

- 思维过程是并行协同式的、非线性的过程。
- 较难形式化,对象、场合不同,形象的联系规则也不同,没有统一的形象联系规则。
- 信息变形或缺少时,仍然有可能得到比较满意的结果。

### ③ 灵感思维

灵感思维是显意识与潜意识相互作用的思维方式。思维活动中经常遇到的“茅塞顿开”“恍然大悟”等情况,都是灵感思维的典型例子,在这样的过程中除了能明显感觉到的显意识外,感觉不到的潜意识也发挥了作用。灵感思维具有以下特点。

- 具有不定期的突发性。
- 具有非线性的独创性及模糊性。
- 穿插于形象思维与逻辑思维之中,有突破、创新、升华的作用。
- 灵感思维过程更复杂,至今无法描述其产生和实现的原理。

### (3) 学习与自适应能力

学习是人类的本能,这种学习可能是自觉的、有意识的,也可能是不自觉的、无意识的;可能是有教师指导的学习,也可能是通过自身实践的学习。人类通过学习,不断地适应环境、积累知识。

### (4) 行为能力

行为能力是指人们通过语言、表情、眼神或者形体动作对外界刺激做出反应的能力,也被称为表达能力。外界的刺激可以通过感知直接获得的信息,也可以是经过思维活动得到的信息。

## 2. 人工智能的定义

考虑到人工智能学科本身相对较短的发展历史及学科所涉及领域的多样性,人工智能的概念是一个至今仍存在争议的问题,目前还没有一个被绝对公认的定义。在其发展过程中,不同学术流派、具有不同学科背景的人工智能学者对其有着不同的理解,提出了一些不同的观点。下面是人工智能领域一些比较权威的科学家所给出的关于人工智能的定义。

人工智能之父、达特蒙斯会议的倡导者之一、1971年图灵奖的获得者麦卡锡(McCarthy)教授认为,人工智能是使一部机器的反应方式就像是一个人在行动时所依据的智能。

人工智能逻辑学派的奠基人、美国斯坦福大学人工智能研究中心的尼尔森(Nilsson)教授认为,人工智能是关于知识的科学,即怎样表示知识、获取知识和使用知识的科学。

美国人工智能协会前主席、麻省理工学院(MIT)的温斯顿(Winston)教授认为,人工智能就是研究如何使计算机去做过去只有人才能做的智能工作。

人工智能之父、达特蒙斯会议的倡导者之一、首位图灵奖的获得者明斯基(Minsky)认为,人工智能是让机器去做需要人的智能才能做到的事情的一门学科。

知识工程的提出者、大型人工智能系统的开拓者、图灵奖的获得者费根鲍姆(Feigen Baum)认为,人工智能是一个知识信息处理系统。

综合各种人工智能观点,我们可以从“能力”和“学科”两方面对人工智能进行理解。人工智能从本质上讲,是指用人工的方法在机器上实现智能,是一门研究如何构造智能机器或智能系统,使之能够模拟人类智能活动的的能力,以延伸人类智能的学科。

### 3. 其他相关的概念

随着计算机科学技术、控制理论和技术、信息理论和技术、神经生理学、心理学、语言学等相关学科的飞速发展，人工智能领域又出现了广义人工智能、狭义人工智能、计算智能、感知智能及认知智能等概念，这里对它们给出一个简要的介绍。

#### (1) 广义人工智能和狭义人工智能

2001年，中国人工智能学会第9届全国人工智能学术会议在北京举行，中国人工智能学会理事长涂序彦在题为“广义人工智能”的大会报告中，提出了广义人工智能（Generalized Artificial Intelligence, GAI）的概念，给出了广义人工智能的学科体系，认为人工智能这个学科已经从学派分歧的、不同层次的、传统的“狭义人工智能”转变为多学派兼容的、多层次结合的“广义人工智能”。广义人工智能的含义如下。

① 它是多学派兼容的，能模拟、延伸与扩展“人的智能”及“其他动物的智能”，既研究机器智能，也开发智能机器。

② 它是多层次结合的，如自推理、自联想、自学习、自寻优、自协调、自规划、自决策、自感知、自识别、自辨识、自诊断、自预测、自聚焦、自融合、自适应、自组织、自整定、自校正、自稳定、自修复、自繁衍、自进化等，不仅研究专家系统、人工神经网络，而且研究模式识别、智能机器人等。

③ 它是多智体协同的，不仅研究个体的、单机的、集中式人工智能，而且研究群体的、网络的、多智体（Multi-Agent）分布式人工智能（Distributed Artificial Intelligence, DAI），模拟、延伸与扩展人的群体智能或其他动物的群体智能。

#### (2) 计算智能

1992年，美国学者贝兹德克（Bezedek）在 *Approximate Reasoning* 学报上首次提出了计算智能的观点。1994年，IEEE 在美国佛罗里达州奥兰多市举办的首届国际计算智能大会（World Congress on Computational Intelligence, WCCI）上，第1次将神经网络、进化计算和模糊系统3个领域合并，形成“计算智能”这个统一的学科范畴。

然而，计算智能（Computational Intelligence, CI）目前还没有一个统一的形式化定义，使用较多的是贝兹德克的观点，即如果一个系统仅处理底层的数值数据，含有模式识别部件，没有使用人工智能意义上的知识，且具有计算适应性、计算容错力、接近于人的计算速度和近似于人的误差率这4个特征，则它是计算智能的。

计算智能是信息科学、生命科学及认知科学等不同学科共同发展的产物，它借鉴仿生学的思想，基于生物系统的结构、进化和认知，以模型（计算模型、数学模型）为基础，以分布、并行、仿生计算为特征去模拟生物体和人类的智能，其研究领域包括神经计算（网络）、模糊计算（系统）、进化计算、群体智能、模拟退火、免疫计算、DNA 计算和人工生命等。

关于计算智能与人工智能的关系，目前存在两种不同的观点。第1种观点的代表人物是贝兹德克，该观点认为生物智能（Biological Intelligence, BI）包含了人工智能，人工智能又包含了计算智能，计算智能是人工智能的一个子集。第2种观点的代表人物是艾伯哈特（Eberhart），该观点认为人工智能与计算智能之间虽然有重合，但计算智能是一个全新的学科领域，无论是生物智能还是机器智能，计算智能都是其核心部分，人工智能则是外层。

### 1.1.2 现代人工智能的兴起

尽管人工智能的历史背景可以追溯到遥远的过去，因为人类很早就有制造机器来帮助人类的幻想，但一般认为人工智能这门学科诞生于1956年的达特蒙斯（Dartmouth）学院。

1946年世界上第1台通用电子计算机ENIAC诞生于美国，最初被用于军方弹道表的计算。计算机科学技术经过大约10年的发展，人们逐渐意识到除了单纯的数字计算外，计算机应该还可以帮助人们完成更多的事情。1950年，艾伦·图灵（Alan Turing）在他的论文《计算机器与智能》中提出了著名的图灵测试（The Turing Test）。在图灵测试中，测试员通过文字与隔离开的房间中的一台机器和一个人进行自由对话。如果测试员无法分辨出与他对话的是机器还是人，那么将判定参与测试的机器通过测试。虽然到目前为止，图灵测试的科学性始终受到质疑，但是它在过去的几十年间一直被认为是测试机器智能的一个重要标准，图灵测试对人工智能的发展产生了极为深远的影响。1951年，普林斯顿大学数学系的一位24岁的研究生马文·闵斯基（Marvin Minsky）建立了世界上第1个神经网络机器，这个神经网络机器拥有40个神经元，通过这个神经网络机器，人们第1次模拟了神经信号的传递。这项开创性的工作为人工智能奠定了深远的基础。闵斯基也由于在人工智能领域所做的一系列奠基性的贡献，在1969年获得了计算机科学领域的最高奖项——图灵奖（Turing Award）。

1956年夏季，在美国达特蒙斯学院，达特蒙斯学院年轻的数学助教麦卡锡（McCarthy，后为斯坦福大学教授）、哈佛大学数学与神经学初级研究员明斯基（Minsky，后为MIT教授）、IBM公司信息研究中心负责人罗切斯特（Lochester）和贝尔实验室信息部数学研究员香农（C. Shannon，信息论的创始人）共同发起了一个持续两个月的夏季学术讨论会，并邀请了IBM公司的莫尔（More）和塞缪尔（Samuel）、MIT的塞尔弗里奇（Selfridge）和索罗门夫（Solomonoff）、兰德（RAND）公司和卡内基梅隆大学（CMU）的纽厄尔（Newell）和西蒙（Simon）等人参加，会议的主题涉及自动计算机、为计算机编程使其能够使用语言、神经网络、计算规模理论、自我改造、抽象、随机性与创造性等几个方面，在会上他们提出了“学习和智能的每一个方面都能被精确地描述，使得人们可以制造一台机器来模拟它”，这次会议给这个致力于通过机器来模拟人类智能的新的科学领域起了个名字——人工智能（Artificial Intelligence, AI），也就是在这个讨论会上，人工智能这一术语被第1次正式使用，它开创了人工智能的研究方向，标志着人工智能作为一门新兴学科的正式诞生。

### 1.1.3 人工智能的学术流派

随着人工神经网络的再度兴起和布鲁克斯机器虫的出现，20世纪80年代到21世纪初人工智能的研究形成了相对独立的三大学派，即符号主义学派、连接主义学派和行为主义学派。当然，从其他角度来看，人工智能学派还有另外的划分方法，如强人工智能学派与弱人工智能学派、简约派与粗陋派、传统人工智能学派与现场人工智能学派等，本小节将对这些学派进行详细说明。

#### 1. 符号主义学派、连接主义学派与行为主义学派

人工智能的研究途径是指研究人工智能的观点与方法，从一般的观点来看，根据人工智能研究途径的不同，人工智能的学者被分为以下三大学术流派。

##### （1）符号主义学派

符号主义学派也称为心理学派、逻辑学派，这一学派的学者主要基于心理模拟和符号推演的方法进行人工智能研究。早期的代表人物有纽厄尔、肖、西蒙等，后来还有费根鲍姆、尼尔森等，其

代表性的理念是“物理符号系统假设”，认为人对客观世界的认知基元是符号，认知过程就是符号处理的过程。

“心理模拟，符号推演”是从人脑的宏观心理层面入手，以智能行为的心理模型为依据，将问题或知识表示成某种逻辑网络，采用符号推演的方法，模拟人脑的逻辑思维过程，实现人工智能。采用这一途径与方法的原因如下。

① 人脑可意识到的思维活动是在心理层面上进行的，如记忆、联想、推理、计算、思考等思维过程都是一些心理活动，心理层面上的思维过程是可以语言符号显式表达的，因而人的智能行为可以用逻辑来建模。

② 心理学、逻辑学、语言学等实际上也是建立在心理层面上的，因此这些学科的一些现成理论和方法可供人工智能参考或直接使用。

③ 数字计算机可以方便地实现语言符号型知识的表示和处理。

④ 可以直接运用人类已有的显式知识（包括理论知识和经验知识）建立基于知识的智能系统。

符号推演法是人工智能研究中最先使用的方法之一，采用这种方法取得了人工智能的许多重要成果，如自动推理、定理证明、问题求解、机器博弈、专家系统等。由于这种方法是模拟人脑的逻辑思维，利用显式的知识和推理来解决问题，因此它擅长实现人脑的高级认知功能，如推理、决策等抽象思维。

## （2）连接主义学派

连接主义学派也被称为生理学派，主要采用生理模拟和神经计算的方法进行人工智能研究，其代表人物有麦卡洛克、皮茨、罗森布拉特、科厚南、霍普菲尔德、鲁梅尔哈特等。连接主义学派早在 20 世纪 40 年代就已出现，但由于种种原因发展缓慢，甚至一度出现低潮，直到 20 世纪 80 年代中期才重新崛起，现已成为人工智能研究中不可或缺的重要途径与方法，每年国际国内都有很多关于人工神经网络的专门会议召开，用于相关领域工作的交流。

“生理模拟，神经计算”是从人脑的生理层面，即微观结构和工作机理入手，以智能行为的生理模型为依据，采用数值计算的方法模拟脑神经网络的工作过程来实现人工智能。具体来讲，就是用人工神经网络作为信息和知识的载体，用称为神经计算的数值计算方法来实现网络的学习、记忆、联想、识别和推理等功能。

神经网络具有高度的并行分布性、很强的健壮性和容错性，它擅长模拟人脑的形象思维，便于实现人脑的低级感知功能，如图像、声音信息的识别和处理。

由于人脑的生理结构是由 10~10<sup>4</sup> 个神经细胞组成的神经网络，它是一个动态、开放、高度复杂的巨系统，人们至今对它的生理结构和工作机理还未完全弄清楚，因此要想实现对人脑的真正和完全模拟，一时还难以办到，目前的生理模拟只是局部的或近似的。

## （3）行为主义学派

行为主义学派也称进化主义、控制论学派，是基于控制论“感知-动作”控制系统的人工智能学派，其代表人物是 MIT 的布鲁克斯教授。行为主义认为人工智能起源于控制论，人工智能可以像人类智能一样逐步进化，智能取决于感知和行为，取决于对外界复杂环境的适应，而不是表示和推理。这种方法通过模拟人和动物在与环境交互、控制过程中的智能活动和行为特性（如反应、适应、学习、寻优等）研究和实现人工智能。

行为主义的典型工作是布鲁克斯教授研制的六足智能机器虫，这个机器虫可以被看作是新一代的“控制论动物”，它虽然不具备人那样的推理、规划能力，但应对复杂环境的能力却大大超过了原有的机器人，在自然环境下具有灵活的防碰撞和漫游行为。

## 2. 传统人工智能学派与现场人工智能学派

事实上，由于行为主义的人工智能观点与已有的传统人工智能的看法完全不同，有人把人工智能的研究分为传统人工智能与现场人工智能两大方向。以卡内基梅隆大学为代表的传统人工智能观点认为，智能是表现在对环境的深刻理解及深思熟虑的推理决策上的，因此智能系统需要有强有力的传感和计算设备来支持复杂环境建模和寻找正确答案的决策方案，他们采用的是“环境建模规划-控制”的纵向体系结构。现场人工智能强调的是智能体与环境的交互，为了实现这种交互，智能体一方面要从环境获取信息，另一方面要通过自己的动作对环境施加影响，而且这种影响行为不是深思熟虑的，而是一种反射行为，采用的是“感知-动作”的横向体系结构。

## 3. 弱人工智能学派与强人工智能学派

在人工智能研究中，根据对“程序化的计算机的作用”所持有的不同观点，人工智能分为强人工智能学派和弱人工智能学派。弱人工智能学派持工具主义的态度，认为程序是用来解释理论的工具，而强人工智能学派持实在论的态度，认为程序本身就是解释对象——心灵。

弱人工智能学派认为，不可能制造出能真正推理和解决问题的智能机器，机器只能执行人的指令，而所谓的智能是被设定的，机器只能通过运行设定好的计算机程序对外界刺激做出对应的反应，它并不真正拥有智能，也不会有自主意识。

弱人工智能所表现出的行为方式是人类预先设定好的，其目的是生命以外的，程序性的，其知识是有限的，运算过程也是处在一个相对封闭的环境中。即使它的能力在今后的发展中得到非常大的提升，这样的提升也仅仅是知识量的增加、运算速度的提高和运算范围的扩大。人类依靠弱人工智能产品极大地减轻了人脑的工作量，使自己从简单的重复劳动中解放出来，并将有限的精力投入到更重要的研究中。这样的人工智能完全是受人类掌控的，所以它作为人类手中的工具对人类来说相对安全。

强人工智能观点认为，有可能制造出真正能推理和解决问题的智能机器，并且这样的机器被认为是具有知觉的，有自我意识的。玛格丽特·A.博登 (Margaret A. Boden) 编写的《人工智能哲学》一书中曾这样描述：带有正确程序的计算机确实可被认为具有理解和其他认知状态，在这个意义上讲，能恰当编程的计算机其实就是一个心灵。

强人工智能有类人的人工智能和非类人的人工智能之分，前者认为机器的思考和推理与人的思维相似，而后者则认为机器具有和人完全不一样的知觉和意识，其推理方式也和人完全不同。

## 4. 简约派与粗陋派

简约 (Neat) 和粗陋 (Scruffy) 的区分最初是由夏克 (Schank) 于 20 世纪 70 年代中期提出的，用于表征他本人在自然语言处理方面的工作同麦卡锡、纽厄尔、西蒙、科瓦斯基 (Kowalski) 等人的工作不同。

简约派认为问题的解决应该是简洁的、可证明正确的、注重形式推理和优美的数学解集，把智能看成自上而下、以逻辑与知识为基础的推理行为。粗陋派则认为智能是非常复杂的，难以用某种简约的系统解决，不主张从上到下逻辑地构建智能系统，而是自下至上地通过与环境的交互，不断

地学习、试验，最后形成适当的响应。通过这个思想建立的智能系统有很强的自主能力，对未知的环境有很高的适应能力，系统的健壮性也很强。对于简约派来说，粗陋的方法看起来有些杂乱，成功案例只是偶然的，不可能真正对智能如何工作这个问题有实质性的解释。对粗陋派来说，简约的方法看起来有些形式主义，在应用到实际系统时太慢、太脆弱。

还有一些学者认为，简约和粗陋两个学派的争论还有些地理位置和文化的原因。粗陋派与明斯基在 20 世纪 60 年代领导的 MIT 的人工智能研究密切相关，MIT 人工智能实验室由于其研究人员“随心所欲”的工作方式而闻名，他们会花大量时间调试程序直到其达到要求。在 MIT 开发的比较重要的、有影响的“粗陋”系统包括威森鲍姆的 ELIZA、维诺格拉德的 SHRDLU 等。然而，由于这种做法缺乏整体设计，很难维持程序的更大版本，太杂乱以至于无法被扩展。MIT 人工智能实验室和其他实验室研究方法的对比也被描述为“过程和声明的差别”。类似于 SHRDLU 的程序被设计成能实施行动的主体，这些主体可以执行过程；其他实验室的程序则被设计成推理机，推理机能操控形式语句（或声明），并能把操控转换成行为。

简约派和粗陋派的分歧在 20 世纪 80 年代达到顶峰。1983 年，尼尔森在国际人工智能学会的演说中曾讨论了这种分歧，但他认为人工智能既需要简约派，也需要粗陋派。事实上，人工智能领域大多数成功的实例来源于简约方法与粗陋方法的结合。

粗陋方法在 20 世纪 80 年代被布鲁克斯应用于机器人学，他致力于设计快速、低价并脱离控制（Out of Control，是 1989 年布鲁克斯与弗林（Flynn）合著论文的标题）的机器人，这样的机器人与早期的机器人不同，它们没有基于机器学习算法对视觉信息进行分析，从而建立对世界的表示，也没有通过基于逻辑的形式化描述规划行为，只是单纯地对传感器信息做出反应。它们的目标只是生存或移动。

20 世纪 90 年代，人们开始用统计学和数学的方法进行人工智能研究，如贝叶斯网络和数学优化这些高度形式化的方法，这种趋势被诺维格（Norvig）和罗素（Russell）描述为“简约的成功”。简约派的问题解决方式在 21 世纪获得了很大的成功，也在整个技术工业中得到应用，然而这种解决问题的方式只是用特定的解法解决特定的问题，对于一般的智能问题仍然无能为力。

尽管简约和粗陋这两个学派之间的争论仍然没有结论，但简约和粗陋这两个术语在 21 世纪已经很少被人工智能的研究人员使用。诸如机器学习和计算机视觉这类问题的简约解决方式已经成为整个技术工业中不可或缺的部分，而特定的、详细的、杂乱无章的粗陋解决方案在机器人和常识方面的研究中仍然占主导地位。

## 1.2 人工智能的发展历史

从 1956 年在达特蒙斯会议上“人工智能”作为一门新兴学科被正式提出至今，人工智能走过了一条坎坷曲折的发展道路，也取得了惊人的成就，并得到了迅速发展。回顾其发展历史，可以归纳为孕育、形成和发展 3 个主要阶段。

### 1.2.1 孕育期（1956 年之前）

尽管现代人工智能的兴起一般被认为开始于 1956 年达特蒙斯的夏季讨论会，但实际上自古以来，

人类就在一直尝试用各种机器来代替人的部分劳动,以提高征服自然的能力。例如,中国道家的重要典籍《列子》中有“偃师造人”一节,描述了能工巧匠偃师研制歌舞机器人的传说;春秋后期,据《墨经》记载,鲁班曾造过一只木鸟,能在空中飞行“三日不下”;古希腊也有制造机器人帮助人们从事劳动的神话传说。当然,除了文学作品中关于人工智能的记载之外,还有很多科学家为人工智能这个学科的最终诞生付出了艰辛的劳动和不懈的努力。

古希腊著名的哲学家亚里士多德(Aristotle)曾在他的著作《工具论》中提出了形式逻辑的一些主要定律,其中的三段论至今仍然是演绎推理的基本依据,亚里士多德本人也被称为形式逻辑的奠基人。

提出“知识就是力量”这一警句的英国哲学家培根(Bacon)系统地提出了归纳法,对人工智能转向以知识为中心的研究产生了重要影响。

德国数学家和哲学家莱布尼茨(Leibniz)在法国物理学家和数学家帕斯卡(Pascal)所设计的机械加法器的基础上,发展并制成了能进行四则运算的计算器,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。

英国逻辑学家布尔(Boole)创立了布尔代数,首次用符号语言描述了思维活动的基本推理法则。

19世纪末期,德国逻辑学家弗雷治(Frege)提出用机械推理的符号表示系统,从而发明了大家现在熟知的谓词演算。

1936年,英国数学家图灵提出了一种理想的计算机数学模型,即图灵机,这为后来电子计算机的问世奠定了理论基础。他还在1950年提出了著名的“图灵测试”,给智能的标准提供了明确的依据。

1943年,美国神经生理学家麦卡洛克(Maculloch)和数理逻辑学家匹茨(Pitts)提出了第1个神经元的数学模型,即M-P模型,开创了神经科学研究的新时代。

1945年,美籍匈牙利数学家诺依曼(Neumann)提出了以二进制和程序存储控制为核心的通用电子数字计算机体系结构原理,奠定了现代电子计算机体系结构的基础。

1946年,美国数学家莫克利(Mauchly)和埃柯特(Eckert)制造出了世界上第1台通用电子计算机ENIAC。这项重要的研究成果为人工智能的研究提供了物质基础,对全人类的生活影响至今。

此外,美国著名数学家维纳(Wiener)创立的控制论、贝尔实验室主攻信息研究的数学家香农创立的信息论等,都为日后人工智能这一学科的诞生铺平了道路。

至此,人工智能的基本雏形已初步形成,人工智能诞生的客观条件也基本具备。这一时期被称为人工智能的孕育期。

## 1.2.2 形成期(1956~1969年)

达特蒙斯讨论会之后,在美国开始形成了以智能为研究目标的几个研究组,他们分别是纽厄尔和西蒙的Carnegie-RAND协作组(也称为心理学组)、塞缪尔和格伦特尔(Gelernter)的IBM公司工程课题研究组以及明斯基和麦卡锡的MIT研究组。这3个小组在后续的十多年中,分别在定理证明、问题求解、博弈等领域取得了重大突破,人们把这一时期称为人工智能基础技术的研究和形成时期。鉴于这一阶段人工智能的飞速发展,也有人称其为人工智能的高潮时期。这一时期,人工智能研究工作主要集中在以下几个方面。

### 1. Carnegie-RAND 协作组

1957年，纽厄尔、肖（Shaw）和西蒙等人编制出了一个称为逻辑理论机（Logic Theory Machine）的数学定理证明程序，该程序能模拟人类用数理逻辑证明定理时的思维规律，证明了怀特黑德（Whitehead）和罗素（W.Russel）的经典著作《数学原理》中第2章的38个定理，后来又在一部较大的计算机上完成了该章中全部52条定理的证明。1960年，他们编制了十多种不同类型课题的通用问题求解（General Problem-Solving, GPS）程序，在当时就可以解决如不定积分、三角函数、代数方程、猴子摘香蕉、汉诺塔、人羊过河等11种不同类型的问题，它和逻辑理论机都是首次在计算机上运行的启发式程序。

此外，该组还发明了编程的表处理技术和NSS国际象棋机，纽厄尔关于自适应象棋机的论文以及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相关的研究工作。例如，1959年，人的口语学习和记忆的EPAM（Elementary Perceiving And Memory，初级知觉和记忆）程序模型，成功地模拟了高水平记忆者的学习过程与实际成绩；1963年，林德赛（R. Lindsay）用IPL-V表处理语言设计的自然语言理解程序SAD-SAM，回答了关于亲属关系方面的提问，等等。

### 2. IBM 公司工程课题研究组

1956年，塞缪尔在IBM 704计算机上研制成功一个具有自学习、自组织和自适应能力的西洋跳棋程序，该程序可以像人类棋手那样多看几步后再走棋，可以学习人的下棋经验或自己积累经验，还可以学习棋谱。1959年这个程序战胜了设计者本人，1962年击败了美国某一州的跳棋大师，他们的工作为发现启发式搜索在智能行为中的基本机制作用做出了贡献。

### 3. MIT 研究组

1958年，麦卡锡进行课题Advice Taker的研究，试图使程序能接受劝告而改善自身的性能，Advice Taker被称为世界上第1个体现知识获取工具思想的系统。1959年，麦卡锡发明了函数式的符号处理语言LISP，成为人工智能程序设计的主要语言，至今仍被广泛采用。1960年明斯基撰写了《走向人工智能的步骤》的论文。这些工作都对人工智能的发展起到了积极的作用。1963年，美国高等研究计划局给MIT投入了200万美金，用来作为开启新项目Project MAC（The Project on Mathematics and Computation）的研究经费，并在机器视觉和语言理解等研究领域激发了新的推动力。可以毫不夸张地说，Project MAC项目培养了早期的一大批计算机科学和人工智能的人才，并对这些领域的发展起到了极其深远的影响。随着该项目的不断进行，也就形成了现在MIT赫赫有名的计算机科学与人工智能实验室（MIT CSAIL），一系列的优秀研究成果也在此期间不断涌现。例如MIT的约瑟夫·维森鲍姆（Joseph Weizenbaum）教授创建了世界上第1款自然语言对话程序ELIZA，这款程序可以通过极其简单的模式匹配和对话规则与人进行文字交流。虽然以我们现在的技术水平来看，当时的文字对话程序显得比较简单，但是在当时，它的问世确实取得了让人惊叹的效果。

### 4. 其他

1957年，罗森布拉特（F. Rosenblatt）提出了感知器（Perceptron），用于简单的文字、图像和声音识别，推动了人工神经网络的发展。

1965年，鲁滨孙（Robinson）提出了归结原理（消解原理），这种与传统演绎推理完全不同的方法成为自动定理证明的主要技术。