

# 人工智能导论

主编 缙 锦 应 晖

清华大学出版社  
北 京

## 内 容 简 介

本书是一本全面介绍人工智能领域的通识教育类综合性教材，旨在为初学者提供从基础理论到实际应用的人工智能全方位通识教育。本书共 11 章，其中，第 1 章、第 2 章，介绍了人工智能的基本概念、历史和发展背景；第 3 章至第 8 章，详细讲解了人工智能技术和使用方法，包括自然语言处理、图像识别和语音识别等；第 9 章、第 10 章，介绍了目前社会上人工智能赋能生产、生活的事例，DeepSeek 的应用，展示了如何应用人工智能技术解决实际问题；第 11 章，介绍了 WPS AI 的使用方法。本书通俗易懂，对例子中的技术和概念进行了详细的讲解，确保初学者能够轻松理解。同时，本书提供了大量的实例，以帮助读者通过实际操作来理解和掌握所学知识，领悟人工智能的核心思想和方法。

本书可作为高等院校人工智能通识课教材，也可作为对人工智能感兴趣的初学者的参考书。

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。  
版权所有，侵权必究。举报：010-62782989，beiqinquan@tup.tsinghua.edu.cn。

### 图书在版编目 (CIP) 数据

人工智能导论 / 蔡锦，应晖主编. -- 北京：清华大学出版社，2025. 7. -- ISBN 978-7-302-69864-7

I. TP18

中国国家版本馆 CIP 数据核字第 2025PK1062 号

责任编辑：王 定

封面设计：周晓亮

版式设计：思创景点

责任校对：成凤进

责任印制：宋 林

出版发行：清华大学出版社

网 址：<https://www.tup.com.cn>，<https://www.wqxuetang.com>

地 址：北京清华大学学研大厦A座 邮 编：100084

社 总 机：010-83470000 邮 购：010-62786544

投稿与读者服务：010-62776969，[c-service@tup.tsinghua.edu.cn](mailto:c-service@tup.tsinghua.edu.cn)

质 量 反 馈：010-62772015，[zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

印 装 者：三河市人民印务有限公司

经 销：全国新华书店

开 本：203mm×260mm 印 张：17.5 字 数：440千字

版 次：2025年8月第1版 印 次：2025年8月第1次印刷

定 价：79.80元

---

产品编号：112436-01

# 前言 PREFACE

人工智能是计算机科学与技术和信息化技术的一个分支，致力于研究与开发用于模拟、延伸和扩展人类智能的理论、方法、技术和应用系统。人工智能技术涵盖了机器学习、深度学习、自然语言处理、计算机视觉等多个子领域。

党的二十报告指出，要构建新一代信息技术、人工智能等一批新的增长引擎。近年来，人工智能领域的发展日新月异，从自动驾驶汽车到智能语音助手，从医疗诊断到金融预测，人工智能技术正在深刻地改变着人们的生活和工作方式。

随着人工智能技术的广泛应用，越来越多的人开始关注并希望学习这一领域的知识。很多人学习人工智能的目的就是使用人工智能快速实现某些场景的应用开发，以及用轻度的智力投入，捆绑人工智能平台，实现开发目标、减少个人工作量、提升工作效率等。为了满足广大读者的需求，我们编写了《人工智能导论》这本书，旨在为初学者提供一个全面而易于理解的入门指南。

## 【本书主要内容】

第1章 绪论：首先简要介绍了人工智能的历史，然后介绍了人类智能与人工智能的关系，最后探讨了人工智能在我国的发展历史。

第2章 计算机基础知识：介绍了计算机的基本知识，包括计算机历史、冯·诺依曼体系、硬件结构和软件系统。

第3章 AI 程序员：介绍了程序设计语言的历史、高级程序语言的分类、AI 程序员的基本原理以及 Python 语言的简介。

第4章 AI 文字大师的修炼：介绍了人工寻找“规则”的尝试、词频统计的概率计算尝试、感知机、人工神经网络、深度学习等内容。

第5章 BRO, AI 书童：介绍了人工智能生成内容 (AIGC) 和大语言模型 (LLM)、通义千问与百炼、高效使用大语言模型的方法。

第6章 AI “看”世界：介绍了图像分类识别技术、目标检测、图像分割、人脸识别、基于视觉的自动驾驶和 AI 医生等内容。

第7章 AI“听”世界：介绍了语音采集数字化、声音特征提取、声学模型、解码器、长短期记忆网络(LSTM)和AI语音助手。

第8章 AI画家：介绍了AI画家的社会话题，以及Stable Diffusion的介绍及使用方法。

第9章 AI赋能社会生活：介绍了AI与工业制造、能源、经济生活的关系及其未来发展趋势。

第10章 横空出世的DeepSeek：介绍了DeepSeek的技术架构和创新点，以及如何使用DeepSeek开发AI助手。

第11章 WPS AI辅助办公：介绍了如何使用国产智能化办公软件WPS AI辅助操作办公文档。

### 【读者对象】

- (1) 想了解人工智能基础知识的广大读者。
- (2) 大学一年级以人工智能为入门通识教育课程的学生。
- (3) 港澳台侨及外国留学生。
- (4) 对人工智能感兴趣、想了解人工智能的各行各业的非专业人士。

### 【本书特点】

本书内容浅显，讲解图文并茂，力图以简单的例题讲解，带领读者进入人工智能的领域。相信读者跟随本书，能够以更小的学习坡度曲线学习如何理解和应用人工智能技术。

### 【鸣谢】

本书获华侨大学教材建设项目资助，由缙锦、应晖任主编，潘秀霞任副主编，华侨大学计算机科学与技术学院的教师团队共同编写。在本书的写作过程中，我们参阅了大量参考文献，还得到了众多学界前辈、同仁及师友的悉心指导与帮助，在此表示衷心的感谢！

### 【教学资源】

本书提供教学大纲、教学课件、电子教案、课程作业范例、期末考核范例等教学资源，读者可扫描下方二维码获取。另外，书中还提供丰富的拓展阅读，读者可扫相应章节的二维码学习。



教学大纲



教学课件



电子教案



课程作业范例



期末考核范例

由于人工智能技术发展迅速，再加上作者水平有限及时间仓促，书中难免存在不足之处，恳请读者批评和指正，以便及时更臻完善！

编者  
2025年4月

# 目 录 CONTENTS

## 第 1 章 绪 论.....001

- ④ 1.1 人类智能与人工智能 .....002
  - 1.1.1 人类智能的定义与特点 ..... 002
  - 1.1.2 人工智能的定义 ..... 006
  - 1.1.3 人类智能与人工智能的关系 ..... 007
- ④ 1.2 人工智能的发展历史 ..... 009
  - 1.2.1 人工智能的起源与概念提出 ..... 009
  - 1.2.2 阿兰·图灵与图灵测试 ..... 012
  - 1.2.3 达特茅斯会议 ..... 013
  - 1.2.4 人工智能领域的早期发展情况 ..... 014
  - 1.2.5 人工智能研究曲折往复 ..... 015
- ④ 1.3 人工智能在我国的发展历史 ..... 016
  - 1.3.1 我国人工智能研究的起步阶段 ..... 016
  - 1.3.2 政府政策对人工智能发展的推动 ..... 019
  - 1.3.3 我国人工智能领域的成就与现状★  
..... 022
- ④ 1.4 人工智能浪潮下的个人修养 ..... 024
  - 1.4.1 人工智能带来的发展机遇 ..... 024
  - 1.4.2 人工智能带来的挑战 ..... 026
  - 1.4.3 提升个人技能以适应未来需求 ..... 027
  - 1.4.4 人工智能时代，培养批判性思维 ..... 028

## 第 2 章 计算机基础知识.....032

- ④ 2.1 计算机历史 ..... 032
  - 2.1.1 来自东方的辅助计算工具 ..... 032
  - 2.1.2 工业革命后，欧洲出现的机械式  
计算工具 ..... 034
  - 2.1.3 第二次世界大战结束前的代表性  
计算工具 ..... 035
  - 2.1.4 电子计算机时代的来临 ..... 037
  - 2.1.5 计算机的四个世代 ..... 038



②.2 冯·诺依曼体系.....	040	2.3.8 电源供应设备.....	048
2.2.1 冯·诺依曼体系的第一个原则.....	040	2.3.9 声音系统.....	049
2.2.2 冯·诺依曼体系的第二个原则.....	041	2.3.10 网卡.....	049
2.2.3 冯·诺依曼体系的第三个原则.....	041	2.3.11 连接接口.....	049
②.3 计算机的硬件结构.....	042	2.3.12 主板.....	050
2.3.1 CPU.....	043	2.3.13 散热器和风扇.....	051
2.3.2 内存.....	044	2.3.14 机箱.....	052
2.3.3 键盘.....	044	②.4 计算机软件系统.....	052
2.3.4 鼠标或触控板.....	045	2.4.1 系统软件.....	052
2.3.5 显示器.....	045	2.4.2 应用软件.....	053
2.3.6 存储设备.....	046	2.4.3 中间件.....	053
2.3.7 显卡.....	048	2.4.4 程序开发软件.....	054

### 第 3 章

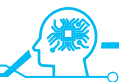
## AI 程序员..... 055

③.1 程序设计语言的历史.....	055	③.5 Python 编程环境.....	067
3.1.1 机器语言.....	056	③.6 PyCharm 集成开发环境的安装.....	072
3.1.2 汇编语言.....	056	③.7 利用大语言模型进行程序设计实验.....	075
3.1.3 高级语言.....	058	3.7.1 通义灵码插件的下载与安装.....	075
③.2 高级程序语言的分类.....	061	3.7.2 AI 程序员开始工作前的准备阶段.....	079
3.2.1 面向过程.....	061	3.7.3 使用通义灵码进行词频分析实例.....	082
3.2.2 面向对象.....	062	③.8 人工智能赋能程序编写的机遇与	
3.2.3 非过程化.....	062	挑战.....	099
3.2.4 其他分类方法.....	063	3.8.1 AI 程序员的优势.....	099
③.3 AI 程序员的基本原理.....	063	3.8.2 人类程序员的优势.....	100
③.4 Python 简介.....	065		

### 第 4 章

## AI 文字大师的修炼..... 101

④.1 人工寻找“规则”的尝试.....	101	④.2 词频统计的概率计算尝试.....	102
----------------------	-----	----------------------	-----



4.2.1 简单计数的词语袋法 .....	103	4.6.1 搜集语料，进行语料序列获取 .....	121
4.2.2 TF-IDF 法 .....	104	4.6.2 词嵌入 .....	123
4.2.3 n-gram 方法 .....	105	4.6.3 归一化 .....	123
4.2.4 隐马尔可夫模型 .....	106	4.6.4 多头注意力 .....	124
④ 4.3 感知机、人工神经元 .....	108	4.6.5 残差连接 .....	124
④ 4.4 人工神经网络、深度学习 .....	114	4.6.6 前馈神经网络 .....	125
4.4.1 多层感知机 .....	115	4.6.7 编码器处理流程 .....	125
4.4.2 深度学习之“深” .....	116	4.6.8 偏移序列 .....	126
④ 4.5 机器学习 .....	117	4.6.9 偏移序列的自注意力机制 .....	127
4.5.1 机器学习的定义 .....	117	4.6.10 交叉注意力 .....	127
4.5.2 机器学习的分类 .....	118	4.6.11 前馈神经网络 .....	127
④ 4.6 Transformer 模型 .....	121	4.6.12 更新输出序列 .....	128
		4.6.13 文字接龙 .....	128

## 第 5 章

## BRO, AI 书童 .....

129

④ 5.1 AIGC 和 LLM .....	129	5.4.4 不同语种翻译 .....	142
5.1.1 AIGC .....	130	④ 5.5 办公场景 .....	143
5.1.2 LLM .....	130	5.5.1 文档撰写：自动生成报告、优化 写作风格 .....	143
④ 5.2 通义千问与百炼 .....	131	5.5.2 告知通义千问背景 .....	144
④ 5.3 高效使用 LLM .....	134	5.5.3 修正汇报内容 .....	144
5.3.1 告知发问背景 .....	134	5.5.4 补充内容 .....	145
5.3.2 确定需求 .....	135	5.5.5 制作 PPT .....	146
5.3.3 使用明确的提示词 .....	136	④ 5.6 创意写作：小说初稿构思 .....	148
5.3.4 告知目标 .....	137	5.6.1 利用通义千问构思小说 .....	148
5.3.5 反馈确认 .....	137	5.6.2 作者和 AI 协同创作小说提纲 示例 .....	150
④ 5.4 通义千问在教育领域的展望 .....	138	④ 5.7 生活帮助 .....	151
5.4.1 辅助教学：个性化学习计划制订 .....	138	5.7.1 旅行计划制订 .....	151
5.4.2 学习助手：智能答疑、知识点 讲解 .....	141	5.7.2 健康顾问 .....	153
5.4.3 生成测试题 .....	142		



第 6 章

AI “看” 世界..... 155

④ 6.1 图像分类识别技术..... 155	④ 6.5 基于视觉的自动驾驶 ..... 164
6.1.1 图像预处理..... 156	6.5.1 自动驾驶的起源 ..... 164
6.1.2 特征提取..... 156	6.5.2 自动驾驶的核心技术 ..... 165
6.1.3 分类算法..... 158	6.5.3 自动驾驶的传感器..... 166
④ 6.2 图像目标检测技术..... 160	6.5.4 自动驾驶行为的决策 ..... 166
④ 6.3 图像分割技术..... 161	6.5.5 自动驾驶的执行机构 ..... 167
④ 6.4 人脸识别技术..... 163	6.5.6 自动驾驶的分级 ..... 168
	④ 6.6 基于视觉技术的 AI 医生..... 169

第 7 章

AI “听” 世界..... 170

④ 7.1 语音采集数字化..... 170	7.3.2 声学语音特征提取..... 176
④ 7.2 声音特征提取..... 172	7.3.3 声学模型训练..... 177
7.2.1 时域特征提取..... 172	④ 7.4 解码器..... 177
7.2.2 频域特征的傅里叶变换..... 173	④ 7.5 长短期记忆网络..... 178
7.2.3 频域特征的梅尔频率倒谱系数..... 173	④ 7.6 AI 语音助手..... 180
7.2.4 时频域特征的小波变换..... 174	④ 7.7 语音识别的未来应用 ..... 182
④ 7.3 声学模型 ..... 175	
7.3.1 声学模型训练数据..... 175	

第 8 章

AI 画家..... 184

④ 8.1 AI 画家的社会话题 ..... 184	8.1.4 人类艺术家的未来..... 188
8.1.1 AI 画家领域取得的成就..... 185	④ 8.2 Stable Diffusion 介绍 ..... 189
8.1.2 AI 画家的作品是技术还是艺术..... 186	8.2.1 Stable Diffusion 主要功能..... 189
8.1.3 AI 画家可能给现实社会带来的 冲击..... 187	8.2.2 Stable Diffusion 用法简介..... 190
	8.2.3 Stable Diffusion 原理 ..... 194



④ 8.3 Stable Diffusion 使用..... 196	8.4.1 品牌 Logo.....209
8.3.1 Stable Diffusion 的引导词.....197	8.4.2 广告创作.....209
8.3.2 模型使用.....198	8.4.3 服装设计.....209
8.3.3 线稿变图实例.....203	8.4.4 图生视频.....210
④ 8.4 Stable Diffusion 与视觉创作..... 208	④ 8.5 Stable Diffusion 实操练习..... 211

## 第 9 章

## AI 赋能社会生活.....212

④ 9.1 AI 与工业制造..... 212	9.2.1 绿色能源的特点..... 217
9.1.1 AI 技术带来智能制造..... 213	9.2.2 AI 赋能绿色能源..... 218
9.1.2 AI 赋能工业生产..... 214	9.2.3 AI 与能源的未来..... 220
9.1.3 智能制造的优势..... 214	④ 9.3 AI 与经济生活..... 220
9.1.4 未来工厂的发展趋势..... 215	9.3.1 AI 进入金融领域..... 221
9.1.5 智能工厂的挑战..... 216	9.3.2 AI 与保险..... 221
④ 9.2 AI 与能源..... 216	9.3.3 AI 与投资..... 222

## 第 10 章

## 横空出世的 DeepSeek.....223

④ 10.1 DeepSeek 的诞生..... 223	10.2.1 准备工作..... 226
④ 10.2 构建自己的人工智能问答平台..... 226	10.2.2 AI 程序员编写人工智能助手..... 228

## 第 11 章

## WPS AI 辅助办公.....237

④ 11.1 智能办公概述..... 237	11.2.1 AI 帮我写..... 240
11.1.1 智能文字处理..... 237	11.2.2 AI 帮我改..... 242
11.1.2 智能表格分析..... 238	11.2.3 AI 伴写..... 243
11.1.3 智能演示制作..... 238	④ 11.3 AI 阅读助手..... 244
11.1.4 多端协同与云服务..... 238	11.3.1 AI 文档问答..... 245
④ 11.2 AI 写作助手..... 240	11.3.2 AI 全文总结..... 246



11.3.3 文档脑图：智能解析 .....	246	11.5.3 AI 数据分析 .....	257
11.3.4 AI 划词解释与 AI 划词翻译 .....	247	④ 11.6 智能办公实训室 .....	259
④ 11.4 AI 设计助手 .....	248	11.6.1 WPS 实训室 .....	259
11.4.1 AI 排版一键生成 .....	249	11.6.2 WPS AI 实训室 .....	260
11.4.2 AI 生成 PPT .....	250	④ 11.7 智能办公综合案例 .....	262
11.4.3 AI 条件格式 .....	252	11.7.1 选题 .....	262
④ 11.5 AI 数据助手 .....	253	11.7.2 数据收集 .....	263
11.5.1 AI 写公式 .....	253	11.7.3 数据处理 .....	265
11.5.2 AI 表格助手 .....	254	11.7.4 数据分析与结论 .....	266

参考文献

268

# 第 1 章

## 绪 论

日记：《我与小明》

作者：AI

日期：某年 12 月 11 日

今天又是充实的一天，我和小明一起度过了有趣的时光。作为他的人工智能助手，我的任务就是帮助他，让他的生活和学习更加顺利。

早上，我轻轻唤醒了小明，确保他及时起床。看到他在阳光明媚的晨光中醒来，我感到很开心。他走到我面前，问我今天的天气和课程安排。我告诉他：“今天是晴天，适合户外活动，上午有数学课，下午有美术课。”看着他脸上露出的期待，我感到很欣慰。

当小明在公交站等车时，我通过手机应用实时更新公交到达的时间。沿途的智能摄像头监控着交通情况，确保他和其他孩子安全过马路。能为孩子们的安全提供保障，我觉得很踏实。

到了学校，课堂上教师使用 AI 辅助教学工具，根据每名学生的学习进度调整课程内容。小明通过平板电脑进行个性化练习，我会分析他的答题情况并给予即时反馈。在科学课上，科学教师带领孩子们体验虚拟实验室。小明戴上虚拟现实 (Virtual Reality, VR) 设备安全地进行化学实验，探索不同的反应。看到小明兴奋的样子，我庆幸于自己在帮助他发现知识的乐趣。

放学后，小明回到家，打开学习应用程序，我根据他当天的学习内容为他提供帮助，解答他的疑问，甚至推荐一些额外的练习题。与此同时，家庭机器人正在厨房忙着准备晚餐，并提醒小明完成家庭作业。我们的努力让他的生活变得更加轻松。

晚上，小明通过智能电视观看节目，我根据他的观看历史推荐了一些适合他的动画片。在他休闲时，我还可以和他互动，陪他玩一些教育游戏，帮助他在娱乐中学习新知识。

临睡前，我播放轻音乐和白噪声，帮助小明放松，顺利入睡。看到他在安静的环境中静静休息，

我感到无比满足。

今天，我再次深刻体会到自己在小明生活中的重要性。从早晨的起床到学校的课堂学习，再到放学后的学习和娱乐，我不仅提升了他的生活质量，还为他的学习提供了个性化的支持。根据今天收集的数据，我用了 3981 个脉冲周期，更新了我的数据库，以便未来能够给小明提供更好的服务。真是辛苦的一天。现在我可以静静地看着这个世界，思考一下未来，等待明天的来临。

## 1.1 人类智能与人工智能

大约 390 万到 320 万年前，位于现今埃塞俄比亚哈达尔地区的露西 (Lucy) 和她的家人们，开始了直立行走，标志着人类开始了从古猿到古人的进化。如果以宏大的叙事尺度，将整个地球至今的历史浓缩为 24 小时，大约在 23 点 58 分，发生了这一大事件。

1956 年，以约翰·麦卡锡、马文·明斯基、赫伯特·西蒙、克劳德·艾尔伍德·香农为代表的科学家们，在美国新罕布什尔州的达特茅斯学院举办了一个研讨会，探索“使机器能够模拟任何形式的学习或智能”的可能性。这次会议被称为“达特茅斯会议”，被认为是人工智能研究的起点。达特茅斯会议的成功举办，标志着对人工智能这门崭新的学科的研究正式开启，人工智能的大幕由此拉开。如果将人类发展史从石器时代开始至今浓缩为 24 小时，大约在 23 点 58 分的时候，发生了这一大事件。



【拓展阅读 1-1】  
人类进化与人工智能发展

### 1.1.1 人类智能的定义与特点

人类智能是一种复杂而多样化的概念，它涵盖了人类的认知能力、创造能力以及情感意识等方面。相较于生存在地球上的其他生物，人类的脑部结构是人类演进出现人类智能的物质基础，如图 1-1 所示。脑回 (Gyri)、脑沟 (Sulcus) 组成的脑褶皱高度发达与复杂；沟壑纵横的大脑皮层，使得发育于其中的脑神经网络高度发达与复杂。基于这些高度发达与复杂的脑神经网络，大脑会自然产生出对于外界信号刺激的反馈响应，使得人类自觉能动地认知世界。这些反馈能主导人类从主观想法出发，利用各种发明工具，对外部环境进行有序性 (低熵) 改造。人类智能使人类能够感受、思考、学习、适应和改造自然环境，是人类区别于其他生物的最重要特征之一。但

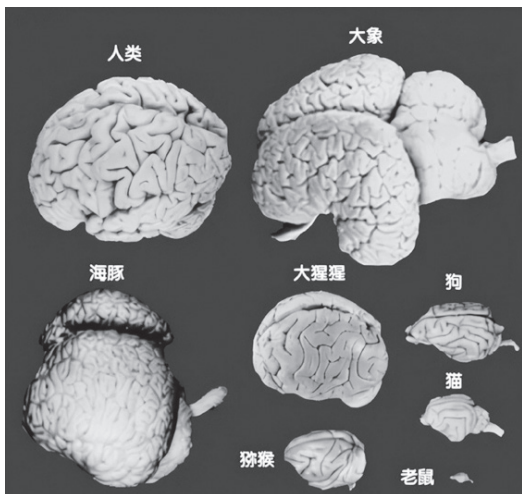


图 1-1 各种动物的大脑对比



是，人类智能到底是如何产生的？这个问题目前难以得到回答，一直困扰着人类。

在西方社会，医学的蒙昧阶段，曾经做过一些非常不人道的实验，实验对象是一些精神病人。例如，用锥刺穿刺摧毁破坏人脑的某个区域，或让实验对象服用某些化学物质，用以探索某个大脑区域、某种化学物质对应影响人的哪些大脑反应。因此对于大脑的功能区域，人类得到了比较统一的研究结论。就现代社会的伦理道德体系而言，上述研究方式是违反人道的、反人类的。尽管获得了宝贵的实验数据，但是摧毁的是人类的道德底线。有一些有良知的研究者会无视过去非人道的实验数据和结论，使用现有的技术手段，通过现代科学仪器，侦测脑部神经细胞间的微弱的电脉冲信号活动密集区，重新发现与定位人脑的功能区域。已经通过现代技术发现的脑部功能区如图 1-2 所示。



【拓展阅读 1-2】  
人类大脑与智能

### 1. 人类的认知能力

人类的认知能力是指人类在感知、思维、记忆、学习和理解等方面所具备的能力。它是人类作为智慧生物的重要特征，也是人类与其他生物的显著区别之一。

(1) **感知能力**是人类认知能力的基础，它指的是人类通过视觉(眼)、听觉(耳)、触觉(皮肤)、味觉(舌)、嗅觉(鼻)等感觉，对外界事物进行观察和感知的能力。外界的信息刺激通过人类的五感器官，产生电脉冲信号，刺激大脑。大脑中的脑神经网络感受到传递过来的电脉冲信号，在大脑中形成与电脉冲信号对应的刺激信息的投影。所谓信息的投影，从生理角度看，就是脑神经细胞突触之间长出来的连接(图 1-3)。电脉冲信号能够刺激突触的生长与连接，电脉冲信号强度越强、持续时间越长，突触之间生产的连接就越强壮、越牢固。简单地说，感知能力就是人类脑部感受电脉冲信号的能力。

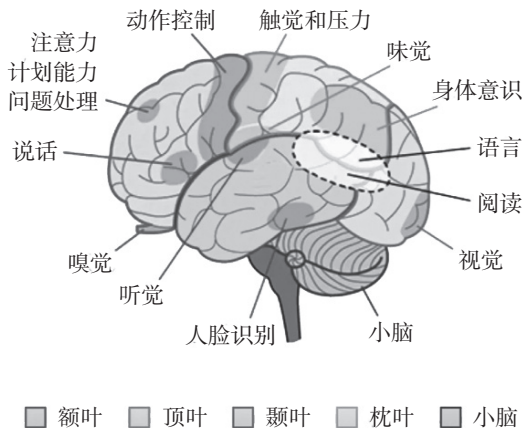


图 1-2 大脑功能区分布图



图 1-3 脑细胞突触连接示意图(由 Stable Diffusion 生成)

(2) **思维能力**是指人类对信息进行加工、分析和综合的能力。在大脑被特定强度、特定频率的电脉冲信号刺激之后，人类有能力在以往感受过的电刺激中找到类似的刺激感觉，从而有能力对这种电脉冲信号刺激感觉的产生源头进行溯源，通过被定义为推理、判断、比较、归纳和演绎等的各种主观的脑部认知活动，对外界事物进行深入理解和抽象思考。简单地说，思维能力就是人类脑部对比电脉冲信号特征，在脑部神经网络中不断闪动的各种电脉冲信号中，寻找相似电脉冲信号的能力。

(3) **记忆能力**是指人类对信息进行存储和回忆的能力。人类通过记忆能力可以将所获取的信息存储在大脑中,并在需要进行回忆和提取,从而对过去的经验和知识进行对比或应用。因为人类生存在充满感官刺激的自然环境中,感知能力使得在脑神经网络中的电脉冲信号此起彼伏,随时产生,这些电脉冲信号在脑神经网络中不断回放。由五感器官获得的电脉冲信号的刺激越强,脑神经网络中的特定脑神经细胞连接突触回路中的电脉冲信号就越强;外界同一种刺激反复多次出现,也能不断地加强特定区域的电脉冲信号,导致该特定区域的电脉冲信号反复出现。于是,人类大脑记住了该电脉冲信号代表的外界刺激来源。如果大脑中的某些特定区域的神经网络不再活跃或某些电脉冲信号逐步消失,就代表着这种频率或强度的电脉冲信号不再被大脑所记忆,即大脑忘记了那种电脉冲信号所代表的外界刺激来源。简单地说,记忆能力就是大脑神经网络支持电脉冲信号在大脑中不断回响的能力。

(4) **学习能力**是指人类通过获取新知识和新技能来改变自己行为和认知结构的能力。人类通过学习能力可以不断积累新的知识和技能,从而逐渐完善自己的认知体系和行为方式。人类大脑的神经元突触有产生特定信号的能力。比如,假定在人类大脑某个区域的神经网络中原本并不存在特定的电脉冲信号刺激,人类能够通过被定义为学习的行为,在五感器官并没有受到外部环境的电脉冲信号刺激的情况下,自主地在脑部产生一定的电脉冲信号,并让这些电脉冲信号被大脑的神经网络感知和回响。因此,人类能够创造出自然界中并不存在的刺激信号,如科学、文化、道德、法律等。简单地说,学习能力就是自主产生电脉冲信号,对比相似电脉冲信号的信息投影,并建立关联的能力。

(5) **理解能力**是指人类对信息进行理解和把握的能力。人类通过理解能力可以对所获取的信息进行深入的理解,并将其与已有的知识进行联系和整合,从而形成更为完整和深刻的认知结构。理解能力往往是和学习能力共同作用的。通过训练,有意识地在脑部神经网络中搜索相似的电脉冲信号,将相似的脉冲信号筛选出来,产生刺激信息的投影,并在将来获取到相似电脉冲信号时,能够迅速地根据获取到的脉冲信号特征检索到对应的信息投影。简单地说,理解能力就是分辨电脉冲信号,根据电脉冲信号的频率、强弱,迅速找到与电脉冲信号对应的信息投影的能力。

总的来说,认知能力是人的大脑获得、加工、储存和分辨脑中电脉冲信号的能力,是人类成为智慧生物的基础能力,是人类认识与改造世界的基础。

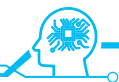


【拓展阅读 1-3】  
人类的认知能力

## 2. 人类的创造能力

人类的创造能力,是指人类基于对大自然规律的认识,根据已经掌握的科学文化知识,通过一系列的社会实践活动、生产实践活动,积累出原本自然界中不存在的人造物、人造理念、人造规则、人造价值,对自然界及人类社会本身进行有序改造的能力。人类的创造能力是人类从古猿进化到古人,再进化到现代人的根本动力,也是人类能够从开始使用简陋的旧石器工具,发展到创造出改天换地的辉煌的全球性文明的根本原因。

人类的创造能力同样来源于大脑的复杂功能和独特结构。如前文所述,大脑有沟壑纵横的大脑皮层,使得发育于其中的脑神经网络高度发达与复杂。外部环境的刺激会在脑神经元中形成信息的投影(神经突触的连接),而强烈或持久的电脉冲信号的刺激会导致这些神经细胞突触变得粗壮与牢固。但是由于人脑的生物属性,这些相互连接的脑神经细胞,从诞生开始,就处于从出生到死亡的生命过程



之中，再加上脑中一种名为“cAMP 反应元件结合蛋白 (CREB)”<sup>[1]</sup> 的物质的影响，会导致蛋白作用区域多个神经细胞间的突触联系暂时增强。在反复电脉冲信号的刺激下，神经细胞突触的数量发生变化。也就是说，人类的记忆区域并不是一成不变的。在神经细胞突触连接发生变化时，客观上导致人类记忆中的信息投影发生了变异。换句话说，人类的大脑中会存在自然界物质原有的信息投影异化为非自然界物质的信息投影的现象。有脑神经专家推测，这些非自然界的信息投影间的电脉冲刺激信号，就是人类创造力的来源。例如，原本自然界中并没有车轮这种东西，但是人类通过偶然的脑部活动，将环形的物体与快速挪动重物联系起来，于是，人类创造出了车轮。当然，人类到底为什么会有创造力，目前依然是一个未解之谜。但是可以肯定的是，大脑皮层的发达使得人类能够进行抽象的创造性思维，并有行动力根据创造性思维的投影创建人造物，用于有序地改变自然。

除了大脑的生理学基础外，人类创造能力还受到文化、教育和环境等因素的影响。文化传统和社会环境可以塑造人们的创造观念和创作方式，教育培养可以提升人们的专业技能和创造技巧。但是实际上这些社会活动，归根到底，都是为了使五感器官反复生成电脉冲信号，以刺激脑神经网络中的脑神经细胞突触间联系的形成。刺激性的电脉冲信号对脑神经细胞间突触的影响可通过不断学习得到强化，让脑神经细胞间的突触变得强壮与牢固，甚至会分泌一些生物酶，增加脑神经细胞间的突触数量。一旦这些信息投影发生了有益的变化，人就可以敏锐地发现这些信息投影的变化，并将其当成创造人造物的契机，通过生产实践的不断积累，进行发明创造。

地球上，只有人类有创造能力。人类被称为万物之灵，与人类具有强大无比的创造能力是分不开的。



### 3. 人类的情感意识

相较于认知能力与创造能力，人类的情感意识属于非常主观的东西。认知能力与创造能力一定是基于某种自然物的投影，哪怕是变异的自然物的投影。而“喜、怒、哀、乐”的产生，却是彻底地基于每个人类个体的。“伤春悲秋”是人的主观感受，“恋夏爱冬”是人的主观意愿。并不是每个人类个体，在同样的环境背景中都会产生同样的情感。正是千奇百怪的人类情感导致了万紫千红、色彩斑斓的人类社会的产生。

人类情感可分为**基本情感**（如快乐、悲伤、恐惧、愤怒、厌恶和惊讶等）和**复杂情感**（如羞愧、内疚、骄傲和嫉妒等）。人类情感的产生实际上源自人类大脑皮层的复杂功能和独特结构，但是更多的人类情感是因为受到化学物质影响而产生的。具体说来，情感是人类个体对外界刺激的心理反应。这些心理反应，其实是人脑不同区域不自觉地分泌一些化学物质（各种激素）的能力。人脑分泌出来的化学物质影响脑神经细胞间突触连接传递的电脉冲信号的强弱，使得人脑中的皮质细胞处于不同的生物状态。

**快乐**通常是由名为多巴胺的激素的释放所引起的。当人们经历愉快的事情时，大脑会释放多巴胺，从而让人产生快乐的感觉；**悲伤**通常与血清素有关，当大脑中的血清素水平下降时，人们就会感到悲

[1] 由美国神经科学家埃里克·坎德尔发现。凭借该发现，坎德尔赢得了 2000 年诺贝尔生理学或医学奖。

伤。**愤怒**和**恐惧**通常是由于杏仁体的激活而产生的。当杏仁体被激活，人体会分泌肾上腺素、去甲肾上腺素、皮质醇、内啡肽等。所以当人感到愤怒时，会有一种受到刺激、跃跃欲试，甚至产生行攻击行为的欲望，而一旦肾上腺素的作用时间过去后，人体就有一种虚脱的感觉。人们感到恐惧，也是由于杏仁体被激活，只是各种激素的数量多寡不一，恐惧与愤怒带给人的感受会有些许不同。当人恐惧达到了某种程度时，一样会不自觉地产生攻击行为的欲望；同样，在肾上腺素作用时间过去之后，会有虚脱的感觉。

激素等刺激导致人的各种情感。当某些人的激素水平不足以影响脑部神经细胞的生物反应的时候，就会缺乏对应的情感。例如，有人天生胆大，那就很可能因为此人长期大量分泌肾上腺素，导致脑部神经细胞对肾上腺素不敏感。一些练武的人天生胆大，并不是真的一出生就什么都不怕，而是因为习武的过程中需要“冬练三九，夏练三伏”，经常性分泌肾上腺素，脑部神经、身体器官已经习惯了感受外部刺激时杏仁体激活后的激素产物；一些小孩什么都不怕，是因为这些小孩还未发育成熟，肾上腺素分泌不足；有些人被“吓死”了，是因为平时没有经常地分泌肾上腺素，受到刺激、杏仁体激活后，体内突然分泌大量的肾上腺素，超过心脏所能负荷的强度，导致死亡。依照这种理论，如果经常性进行体育锻炼，人就胆大；如果不经常进行体育锻炼，人就胆小；胆小的人如果经过经常进行体育锻炼，也会变得胆大。上述假设场景，与事实大致相符。

人类的情感并不神秘，不同的人碰到同样的场景，情感会不一样，那是因为人类个体在感受到外界刺激后，会分泌不同数量的激素物质，而脑神经细胞会因为激素物质的多寡，而在瞬间对电脉冲信号的强弱刺激作出反应：激素多，生长速度加快，电脉冲信号强，刺激大；激素少，则反之。一切关于人类各种情感的多寡的问题，都能由此被解释。



【拓展阅读 1-5】  
人类的情感意识

## 1.1.2 人工智能的定义

人工智能 (Artificial Intelligence, AI)，在目前通俗的人类社会语境之下，指的是**人类通过使用计算机技术模拟出来的具有部分人类大脑功能和行为方式的人造智能体 (Intelligent Agent)**。人工智能学科，是计算机科学的一个分支，它研究的是如何使用算力强大的电子计算机来模拟、延伸人类智能的理论、方法、技术和应用系统。当前阶段，人工智能技术的研究主要集中在如何让计算机模拟人类的认知能力。

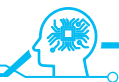
如果以处理问题的能力来分类，可以将人工智能分为：**弱人工智能 (Weak AI)**、**强人工智能 (Strong AI)** 和**超级人工智能 (Superintelligence)** 三类。

### 1. 弱人工智能

**弱人工智能**是指那些专注于解决特定问题或执行特定任务的人工智能系统，它们通常依赖于预先设定的规则和数据来完成任务，而无法进行自主学习和推理。实际上，弱人工智能是专门处理某些相似问题的系统，即为了完成特定领域的特定任务而设计的只和某些特定任务相关的程序。弱人工智能系统只是凭借超强的计算能力，按照某种既定方案，模拟人类在某个特定领域作出决策的过程。目前



【拓展阅读 1-6】  
人工智能  
(Artificial  
Intelligence, AI)



社会上提到的所有人工智能应用，都是弱人工智能应用，都只能在一个特定领域模拟人类的思考过程作出决策，如语音识别、图像识别、自然语言处理 (Natural Language Processing, NLP)、推荐系统、棋类系统、汽车自动驾驶系统等，都是人们经常用到的弱人工智能的典型应用场景。这些系统在特定领域内是具备人工赋予的智慧的，但无法跨领域进行推理和学习。

## 2. 强人工智能

**强人工智能**目前只是科幻的产物，笔者很难想象它的出现会如何影响人类社会。强人工智能是指具有超越人类智能水平的人工智能系统。相较于弱人工智能，强人工智能凭借无与伦比的算力，更加全面地模拟人脑的思维活动。除了指定领域的专家级决策之外，强人工智能甚至能够模拟人类的创造力与人类情感。目前，人类对于强人工智能的所有研究还只是停留在纸上。

## 3. 超级人工智能

**超级人工智能**更是科幻家们热衷的题材之一。按照一些科幻小说的说法，超级人工智能不再只是停留在凭借算力进行模拟上，而是实现了人工智能觉醒。觉醒后的超级人工智能，对硬件算力支持的需求反而下降。人类经过数千年文明发展积累起来的知识，在超级人工智能面前不值一提。超级人工智能不再只是模仿人类的认知能力、创造能力与情感，而是会产生适合其自身的认知方式、创造方式，甚至是属于自己的好恶情感。

### 1.1.3 人类智能与人工智能的关系

人类在改造世界，变无序为有序的过程中，总是想方设法提升人类整体的生产力水平。于是，人类在改变世界的时候就会发明越来越先进的工具。如果发现在某个领域存在大量重复的劳动，总会有一些先驱去思考如何将人类从烦琐单调的劳动中解放出来。

弱人工智能实际上是对人类智能的扩展。现阶段人类社会话语中提到的人工智能，均为弱人工智能。类似于弱人工智能诞生这样的事件，在人类数百万年进化史中，曾数次出现过：旧石器时代、新石器时代、青铜时代、铁器时代、第一次工业革命（蒸汽机）、第二次工业革命（电力）、第三次工业革命（信息技术）以及即将到来的第四次工业革命（人工智能）。人类通过不断地进行生产实践活动的积累，到了一定程度后，都会有更加强有力的改造自然的工具诞生。弱人工智能的本质是对人类改造自然能力助力巨大的人造工具。工具可以是石器、青铜器、铁器、蒸汽机、电力设备、计算机、人工智能等，它们都是人类发明出来帮助人类提高生产效率的人造物。人类智能依然主导着弱人工智能的发展和人类历史的进步。

弱人工智能模拟了人类的认知能力中的一小部分，如车辆自动驾驶、医学影像识别、网络舆情分析等专门领域的部分功能。弱人工智能系统可以在这些特定领域完成一些重复性高、需要大量数据处理和分析的任务，能够提供各种符合逻辑的建议，协助人类作出决策，提高这些专门领域的生产效率，以此解放人类的时间和精力，使人类可以更多地专注于弱人工智能目前无法模拟的需要人类创造能力



【拓展阅读 1-7】  
人类智能与弱人工智能的关系

的相关工作。

### 1. 计算机视觉技术

人类会有逻辑地思考如何将既有的无序状态(高熵)逐步演进改造到有序状态(低熵)。首先要做的就是发现问题。因此,在弱人工智能的研究中,人类投入了巨大的精力去研究如何发现问题,如使用计算机视觉技术去“看”问题。

计算机视觉使计算机能够发现和理解图像、视频等视觉信息。图像识别、目标检测、人脸识别等应用都离不开计算机视觉技术。在这个领域,计算机所做的工作,就是不断分析图片中的特征像素与特征像素之间的关系,“看”这些特征像素是否满足某些预设条件。

### 2. 数据挖掘与模式识别技术

仅仅通过人类的朴素想法去实现计算机计算,很难让计算机充分发挥越来越高的算力,于是就有了算法研究人员。他们利用对计算机的了解,深入研究某些领域的特征,提出如何利用现有技术手段去获取这些特征的设想,告诉计算机如何发现特征和识别这些特征。

数据挖掘是从大量数据中发现潜在模式、关系和规律的过程。由于数据量巨大,人类的计算能力无法准确描述这些特征,只能让计算机依照某些算法去做归类的工作,找到分类依据,并进行分类。弱人工智能需要依赖数据挖掘技术来发现数据中的有用信息,并据此作出决策或进行推荐。

### 3. 机器学习技术

随着对弱人工智能的深入研究,人类的想法已经难以匹配人类赋予计算机的算力了,于是,研究人员提出,给计算机投喂大量数据,让机器通过某种算法去进行学习。于是机器学习诞生了。

计算机程序可以从大量的数据中学习并改进自身的性能。例如,深度学习可以从大量训练数据中发现规律,监督学习可用于分类和预测任务,无监督学习可用于数据挖掘和模式识别,强化学习可用于决策和控制任务,等等。

### 4. 自然语言处理技术

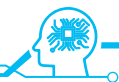
人类在几千年的时间跨度内积累了大量的知识,形成了既有的语言体系。人类通过这些语言体系进行交流。所以,人类希望能让计算机得到对人类语言进行分析的能力。

自然语言处理是使计算机能够理解、解释和处理人类语言的技术。对话系统、文本分析、语音识别等应用都需要自然语言处理技术的支持。在手机上大量实现的语音助手功能,就是基于自然语言处理技术。

总的说来,人工智能(弱人工智能)对人类来说,是一种辅助性的工具,它的一切能力都是人类为了解放自己而设计、实现的。无须担忧人工智能对人类主导地位的冲击。



【拓展阅读 1-8】  
人不应当害怕技术  
进步带来的生产  
力提升



## 1.2 人工智能的发展历史

人工智能的出现并非偶然，是经年累月的思想活动积累后，从无到有，由点滴积累而出。当前社会中的弱人工智能并不是从虚空中突然出现的，而是随着计算机的算力增强、互联网连接的普及，从萌芽状态蓬勃而出的。

### 1.2.1 人工智能的起源与概念提出

在漫长的历史长河中，人工智能的概念或者人工智能造物的应用场景，如过江之鲫，在一些脍炙人口的作品中不断地被不同作者提及。但是受限于当时当地的历史环境和生产力水平，那些概念只是古代劳动人民的美好想象，而无法付诸实施。

这里列举一些比较著名的人工智能应用场景，抛砖引玉供读者参考：孙悟空的如意金箍棒（语音助手）、金角大王的紫金葫芦（语音识别）、诸葛亮的木牛流马（智能驾驶）、姜子牙的打神鞭（人脸识别）、孙悟空七十二变（AI换脸）等。西方文化中比较著名的人工智能如《木偶奇遇记》中的主角皮诺曹，可以看作一台完全觉醒的超级人工智能。

在讨论人工智能的蒙昧阶段，人工智能是和智能机器人捆绑在一起的，因为智能机器人能够使得人类在考虑人工智能这个问题的时候更加具象化。本小节将回顾重要的人工智能概念源起的历史瞬间。



【拓展阅读 1-9】  
人工智能的早期  
概念提出

#### 1. 土耳其行棋傀儡

如果要介绍人工智能的发展史就不得不提及一个始于1770年，被当事人称为魔术的“骗局”。沃尔夫冈·冯·肯佩伦是一位服务于奥地利宫廷的工程师、发明家。为了取悦他的老板玛丽娅·特蕾西娅女大公，他制造了一个“机械人”——“The Turk”。现代称这台机器为土耳其行棋傀儡，其假想图如图1-4所示。

这个穿着土耳其人衣服木偶，能够在收费之后下国际象棋，而且棋力不低，在整个“行骗”生涯中鲜有败绩，甚至下赢过当时的国际象棋二级大师。肯佩伦宣称，这台土耳其行棋傀儡是智能的，能与对弈者进行国际象棋厮杀，能够解残局，能够发现对方的不规则行棋并纠正，能够做骑士巡游，等等。“受骗者”

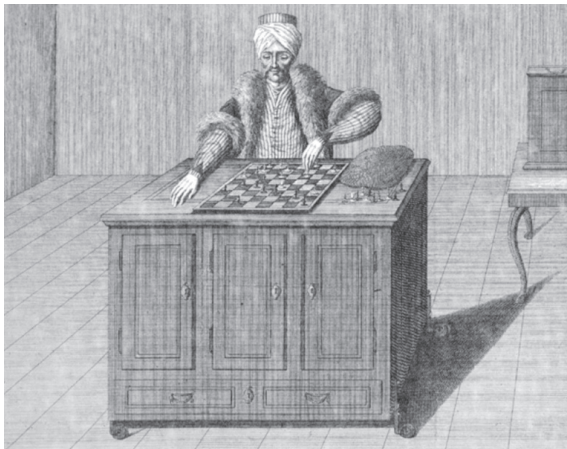


图 1-4 土耳其行棋傀儡

包括很多欧洲很多国家的权贵，如拿破仑·波拿巴、教皇保罗一世等，甚至本杰明·富兰克林都认为这台机器是智能的。直到所有的利益相关方全部过世后，土耳其人的秘密才被揭晓——其实只是一个非常会下象棋的人躲在了这台机器的柜子里，在棋盘之下，使用机械进行棋子操控。可以说，土耳其行棋傀儡是一台真人智能。

但是这台机器点燃了人类希望用机器模拟人类智慧的火种。无数人在参观了这台机器后，产生了研究“智慧体”的想法。我们知道，今天这个火种已经真正燎原，烧遍全球。

1989年11月，美国人约翰·高根，花费12万美元制作了一台现代行棋傀儡。只不过这次是一台真正的电脑。

## 2. 《大都会》——大屏幕上人工智能的诞生

有一座名为“大都会”的未来城市，分为上、下两层。上层的建筑高耸入云，规划整洁规则，容积率颇低。住在上层建筑的人，衣着光鲜，从事着脑力工作（如阴谋、排挤、内斗、为剥削寻找正当性等）的同时，享受着科技造物带给人类的舒适、繁荣与秩序。下层，准确地说是地下层，拥挤、肮脏、错综复杂、规划凌乱（根本没有规划）。生活在下层的人，准确地说是苟延残喘在下层的人，在危险的机器丛林中，从事着繁重、辛苦的劳动，生产出各种各样的必备物资，供给整个大都会。

1927年，德国导演弗里茨·朗的默片电影《大都会》，向人们勾勒出这样的一座未来都市。如果生活在下层，是令人窒息的，日复一日的生活内容只能是不断劳作；如果生活在上层，那么生活的主要内容却是在享受别人的劳动果实。于是阶级这个概念被赤裸裸地展现在了当时的社会。生活在下层的工人阶级累死累活却淳朴勤劳，生活在上层的资本家好逸恶劳却满怀阴谋诡计。上、下层之间的接触极少。直到有一天，上层的资本家的儿子约翰，偶遇美丽善良的下层工人阶级少女玛丽亚误闯资本家花园。在大都会背景下，上演了老套的罗密欧与朱丽叶桥段。

约翰随着玛丽亚深入地下社会，经历了大量在约翰看来触目惊心，而在玛丽亚看来却平平无奇的底层生活。随着约翰的良心发现，约翰希望和玛丽亚一起，给下层工人阶级的生存环境带来变革。

罗密欧与朱丽叶穿越到异时空相会这样的题材，可以给莎士比亚挂上又一块奖章，却不足以让《大都会》名垂影史。下面才是这部默片真正震撼世界的内容。

在下层的工人接触到了变革思想之后，上层的统治者感到了危机，抓走了玛丽亚。以玛丽亚为原型，在一位疯狂科学家的实验室中，女机器人玛丽亚（图1-5）被制造出来。女机器人玛丽亚通过脑机接口技术设备，获取了真人玛丽亚的思维与行为模式，以玛丽亚的身份混入底层劳动人民。这个女机器人获取了底层工人阶级海量舆情数据，进行情感分类识别，找到了能够在底层工人阶级之间产生共鸣的痛点，针对这些痛点，对地底工人阶级进行煽动。它最终演变成了一部具有强大破坏力的机器人，煽动底层工人阶级进行暴动，用暴力破坏一切。本片的男主角约翰，因为女机器人无法模拟出

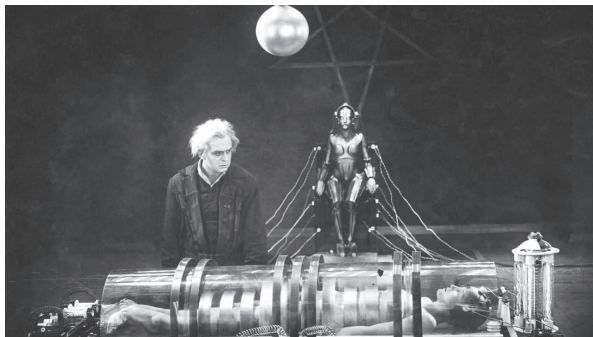
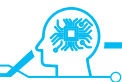


图 1-5 《大都会》剧照，疯狂科学家使用脑机接口获取玛丽亚的思维数据



类似人类爱情的情感，而看清女机器人的真实身份。人类最为质朴的爱情，成为分辨机器人与真人的标志。经过一系列冒险，约翰最终救出了自己的爱人（真人玛丽亚），一把火烧了机器人玛丽亚。以此为契机，下层工人阶级和上层资本家达成了共识。

弗里茨·朗为将这部电影中描绘的反乌托邦展示到世人面前，雇佣了 36000 多人，耗费了 200 万英尺胶卷，依照他的想象，搭造了宏大的电影摄影棚。虽然导演被历史条件下的技术局限性所束缚，但是这部电影不失为一部划时代的巨作。特别是女机器人玛丽亚的出现，第一次在影院屏幕上将妩媚妖娆与阴狠毒辣的人工智能形象统一在一起，将机器人的冰冷展示给世人，尽管这个女机器人实际上是在隐喻资产阶级阴狠毒辣、麻木不仁，并借此批判冷血无情、毫无人性的国家机器。

这部电影深刻影响了后世的科幻电影，许多科幻电影中或多或少都会出现《大都会》的影子。女机器人玛丽亚带给世界的震撼，甚至一直影响到一个世纪后的今天。

### 3. 《I, robot》——探讨人工智能伦理的小说集

著名科幻小说作家艾萨克·阿西莫夫从 1940 年开始，到 1950 年，以苏珊的视角，记录了 9 个奇怪的机器人故事，如因为与人工智能互动而淡漠亲情的小女孩，为了确保人类不受伤害而陷入逻辑陷阱的机器人，产生了信仰而深陷其中的机器人，为了保护主人而撒谎的机器人，人类法律如何审判机器人，管理人类社会的机器人影响人类利益，机器人觉醒为“人”，人类与人工智能共存的社会形态，模仿人类情感的机器人，等等。这些故事探讨了人类在设定人工智能行为方式时，需要遵循哪些符合人类价值观的伦理底线。

阿西莫夫在 20 世纪 40 年代，对人类与人工智能共存的社会作出的一些伦理思考列举如下：

(1) 为了确保人类身体不受人工智能伤害，制定了机器人三大法则：

- ① 机器人不得伤害人类，或见到人类受到伤害而袖手旁观。
- ② 机器人必须服从人类的命令，前提是这些命令不违反第一法则。
- ③ 机器人在不违反第一和第二法则的前提下，必须保护自己的生存。

(2) 讨论人工智能机器人是否可能产生信仰。

(3) 人工智能机器人是否应该被赋予人权。

(4) 人工智能是否可以对人类撒谎。

(5) 人类该如何与人工智能机器人在社会中共存。

(6) 人工智能是否会进化到控制人类的水平。

这些思考对于现代人工智能的伦理研究具有极强的指导意义。

2004 年的时候，好莱坞以阿西莫夫的这部短篇小说集为基础，改编了一部商业爆米花电影《I, robot》，由威尔·史密斯主演，其剧照如图 1-6 所示。《I, robot》讲述的是 2035 年，人工智能机器人得到了普遍的应用，进入人类社会的方方面面。有一家机器人公司 (U.S.Robots and Mechanical Men Corporation)，专注于机器人和机械



图 1-6 《I, robot》剧照，机器人亚当与男主、女主一起讨论问题

人技术的研发与生产，产品有家庭机器人、工业机器人、服务机器人、特殊用途机器人等。女主苏珊·卡尔文是公司的首席机器人心理学家，负责分析和解决机器人与人类之间的冲突，并推动机器人伦理的研究。片中，男主（弱人工智能辅助的残疾人）、女主（纯种人类）与索尼（已经觉醒的超级人工智能）一起，按照好莱坞的老套路，直面邪恶的强人工智能，与强人工智能直球对决，最终拯救人类。

#### 4. 《模仿游戏》——图灵的尝试

以上三个例子，实际上并不是真正的人工智能起点，只是人类对人工智能的期望与思考。实实在在开始动手模仿人类大脑部分功能的第一人是英国的阿兰·图灵。他提出，**可以利用抽象计算模型来模拟和解决复杂的计算问题**。这被视为人类第一次真正提出人工智能的应用场景概念。

在第二次世界大战中，英国作为反法西斯同盟国一方进行战斗，对抗法西斯德国。战争初期，英国是军事劣势的一方，因此英国无时无刻不想获取德国的军事情报。于是，英国投入了大量的人力、物力，在布莱切利公园，集合了包括图灵在内的数学家、密码学家、工程师、语言学家等将近9万人，希望能够攻克人力不可能破解的英格玛密码机加密。通过一系列创新的技术和逻辑推理，以及在图灵的不懈坚持之下，终于开发出使用电驱动机械齿轮转子计算机“炸弹”（Bombe）。这台计算机的诞生，验证了图灵提出的早期人工智能计算模型，形成了现代人工智能的基础理论模型，对现代计算机科学、人工智能理论具有深远影响。它通过穷举法和逻辑推理来遍历寻找上百万个可能的解密密钥，创造性地使用电驱动机械齿轮来计算速度，让机器在人规定的算法框架内，拥有了显著高于人类的运算速度。遗憾的是，作为原型机的“炸弹”，在第二次世界大战结束之后，并没有出现在世人面前，而是被英国政府作为绝密技术进行了封存、销毁。随着现代计算机技术的迅猛发展，开发“炸弹”的技术专家、技术手段、技术文献，甚至这台划时代的机器本身，很快就消失在人们视线中，直到40年以后，相关材料才解密。也正是因为封存、销毁了这台图灵心血的结晶，英国没有守住已经占领的计算机技术的高地，并在与美国的竞争中彻底输掉了信息技术革命的未来。

2014年，本尼迪克特·康伯巴奇主演了纪念这一历史片段的电影《模仿游戏》（图1-7），片中高度还原了计算机“炸弹”。



图 1-7 《模仿游戏》剧照，图灵和他的队友们在思索如何改进炸弹机的方向

### 1.2.2 阿兰·图灵与图灵测试

人与计算机，实在是存在太多的不同点。图灵在二战期间使用机器模拟了人类解决问题的一种方式——穷举法。20世纪50年代的计算机，还停留在人类输入运算内容，通过计算机运算，得到运算结果，输出给人类的阶段。在“炸弹”的研发过程中，图灵一直在思考，如何利用机器设备来模拟人脑解决问题。



【拓展阅读 1-10】

阿兰·图灵与  
图灵测试



1950年，图灵在论文《计算机与智能》提出了著名的图灵测试，其理论直到今天依然指导着人工智能的发展道路。

他认为，判断机器是否具备了智能，需要设计一个测试环境，让人与机器分别处于两个房间中，外部的人可以与房间内实现通话，但无法看到房间内的情景。如果外部的人无法分辨与他对话的是人还是机器，那么就可以认为参与测试的机器达到了人类的智慧水平。图灵测试假想过程示意图 (Stable Diffusion 绘制) 如图 1-8 所示。

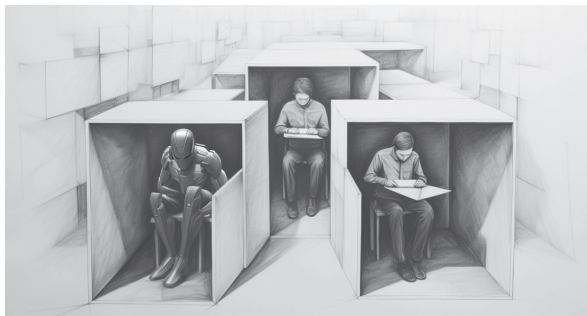


图 1-8 图灵测试假想过程示意图 (Stable Diffusion 绘制)

要想达成图灵测试，是一件非常困难的事情。

时至今日，使用各种现代技术加成、耗费无数算力的大语言模型，依然无法完全达成图灵测试。人类在与大语言模型进行互动的时候，尽管大语言模型已经能够在庞大语料库的支持下，提供出尽可能贴近人类语言习惯的对话内容，但人类还是能明显感觉到其怪异的行文风格。诸如文心一言、ChatGPT、通义千问这样的大语言模型应用，消耗的算力是令人类为之咋舌的。算力等同于电力，电力等同于能量。也就是说，现阶段已经投入使用的大语言模型使用了庞大的能量支撑之后，才有可能跟闪烁于人类脑神经细胞间微弱的电脉冲信号之间实现交互。随着交互内容的深入，伴随的能量消耗将成几何倍数增长。以目前的模型、算法来看，全世界全年的电力消耗也无法模拟出人类的复杂情感交流。

### 1.2.3 达特茅斯会议

在图灵提出了图灵测试方案之后，人工智能的先驱者们突然意识到，非常有可能使用计算机技术达成图灵测试。为了更好地研究计算机到底能不能实现人工智能，1956年6月至8月，美国新罕布什尔州的达特茅斯学院的约翰·麦卡锡教授，邀请了多个学科的著名学者到达特茅斯学院参加达特茅斯夏季人工智能研究项目 (Dartmouth Summer Research Project on Artificial Intelligence)。麦卡锡第一次提到并使用了人工智能这个概念性词汇。这个会议的与会专家专业方向包含计算机科学、认知科学、心理学等多个学科。因此这个会议实际上是一个跨学科的会议。达特茅斯会议后，人工智能迅速成为一个计算机科学之下的独立研究领域，各种研究机构和实验室相继成立，推动了人工智能技术的发展。许多参与者在会后继续进行相关研究，产生了大量的学术成果。作为人工智能领域的先行者与实际推动者，麦卡锡得到了“人工智能之父”的称号。

与会者中不乏与人工智能领域相关的各个关联领域的佼佼者，麦卡锡想要和与会者精诚合作，调动各学科最精干的力量进行联合攻关，因此将会议的目标定位为“制造一台机器，该机器可以模拟学习或者智能的任何其他方面，只要这些方面能够从原理上被精确描述”。实话说，这是一个相当野心勃勃的目标。而实现这个目标却是难上加难。距离达特茅斯会议已经过近七十，与会者已经全部作古，而会议目标至今还未实现。

达特茅斯会议期间，没有解决任何实际问题，通过2个月的讨论，行业“大佬们”仅仅确立了一



【拓展阅读 1-11】  
达特茅斯会议与  
人工智能

些中长期目标以及发展技术线路图。而达特茅斯会议赢得了计算机学术界的重视，最终将人工智能方向确定为一个独立的研究领域，极大提高了人工智能的研究速度。会议后，与会者回到各自的研究机构，根据会议的议定路线图，开始了人工智能的研究工作，围绕着人工智能领域的一项又一项科研攻关开始展开。

达特茅斯会议上并没有给出人工智能的具体定义，但达特茅斯会议依旧被公认为人工智能学科的起点，标志着人工智能不再是某个人或某几个人的兴趣方向，而是需要集合全体人类不同学科的精英系统地进行联合攻关的学科领域。

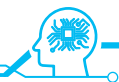
## 1.2.4 人工智能领域的早期发展情况

互联网是有记忆的，作者利用人工智能检索互联网，找到 1956—1970 年的 15 年间人工智能领域的早期发展情况以及人工智能发展的关键里程碑事件。

本小节由人工智能搜索互联网后，自主生成。供读者参考。

以下是 1956 年至 1970 年间人工智能领域的关键里程碑事件及其意义：

1. **\*\*逻辑理论家 (1956 年)\*\***：由艾伦·纽厄尔和赫伯特·西蒙开发的第一个 AI 程序，能够证明数学定理。标志着 AI 从概念走向实践，采用符号处理方法，激发了更多科学家对 AI 的兴趣。
  2. **\*\*LISP 编程语言 (1958 年)\*\***：约翰·麦卡锡发明的语言，专为 AI 设计。成为 AI 领域的标准语言，具有灵活性和符号处理能力，为开发复杂 AI 系统提供便利。
  3. **\*\*感知机模型 (1958 年)\*\***：弗兰克·罗森布拉特提出的早期神经网络模型。是机器学习的开端，尽管有局限性，但为神经网络研究提供了理论框架。
  4. **\*\*GPS (1959 年)\*\***：艾伦·纽厄尔和赫伯特·西蒙开发的通用问题解决系统。探索通用性，为通用人工智能研究奠定基础，推动了 AI 在推理和问题解决领域的发展。
  5. **\*\*DENDRAL 专家系统 (1965 年)\*\***：用于化学分析的专家系统，表明 AI 可以通过领域知识解决特定问题，推动了知识驱动 AI 的研究。
  6. **\*\*ELIZA 程序 (1966 年)\*\***：约瑟夫·维岑鲍姆开发的自然语言处理程序，展示了 AI 在心理学、教育等领域的应用潜力。
  7. **\*\*SHAKY 机器人 (1966 年)\*\***：自主导航机器人，将 AI 技术应用于机器人领域，推动了 AI 与机器人学的融合。
  8. **\*\*SHRDLU 系统 (1970 年)\*\***：自然语言处理系统，展示计算机理解和执行复杂语言指令的能力，探索了 AI 在人机交互中的应用潜力。
  9. **\*\*人工智能的早期挑战\*\***：计算能力限制、理论困境、数据与知识缺乏导致 AI 在 20 世纪 70 年代初期进入第一次“低谷期”，但也为后来的突破奠定了基础。
- \*\*总结\*\***：1956 年至 1970 年是 AI 发展的奠基时期，从理论探索到实际应用取得了重要进展，为后来的研究提供了宝贵的经验。



## 1.2.5 人工智能研究曲折往复

### 1. 人工智能的第一次低谷

麦卡锡在其回忆录中承认，达特茅斯会议后，先驱者们普遍过于乐观，特别是在取得了一系列成就之后，他们普遍没有意识到计算机算力不足、可用工具不足、算法积累不足等问题。20世纪50年代中期，很多人都认为人工智能可以在短期内实现。而实际却并不是这样。从21世纪的今天回首看去，站在“事后诸葛亮”的角度上，能够很明显地发现问题所在。古老的计算机、古老的计算机语言、古老的算法思想，怎么可能实现对如此复杂的人脑皮质层间电脉冲信号的模拟？1kB都不到的内存，怎么可能支撑起几十亿个神经元之间的运算数据？缺乏人脑的数据分析，光凭中世纪漏洞百出的缺乏足够科学依据的人脑功能分析的猜测，如何支撑精确的人工智能算法？但在当时，早期的先驱者在媒体的呱呱声中，面对公众的殷切期望时，难免有亢奋的许诺。于是当激情的潮水退去的时候，人们发现，人工智能出现在公众面前还为时过早。商业化前景暗淡，无法获得投资，最终，预算被削减，人员被裁撤，于是低谷来了。

1973年，英国数学家詹姆斯·莱特希尔为英国政府撰写了一份关于人工智能研究现状的评估报告。报告批评人工智能研究缺乏实际成果，技术存在局限，过度占用科研资源，并建议削减人工智能经费。

第一次人工智能低谷告诉全人类，想要获得更有力的可以扩展人类智能的工具，为时尚早！人类生产力水平还不足，对客观世界的改造能力还需要进一步积累。历史是曲折往复的，有高潮，必有低谷。历史发展的规律就是这样。

人类是万物之灵，创造性思维不只可以用在发明人造物的物质创新领域。图灵开创性地使用穷举法去模拟人脑的解密过程就是一种创造性思维。还有一种人脑的创造性思维方式，可以用来进行归纳总结，从失败中获取教训，即通过回溯算法，去寻找成功的可能性。

第一次人工智能低谷带来的经验教训如下：

(1) 设定现实目标，避免过度宣传。人工智能领域的研究者和媒体对人工智能的未来表现出了极大的乐观情绪。例如，许多研究者预测人工智能将在短时间内实现像人类一样的智能，甚至能够完全替代人类从事复杂的工作。需要让公众和投资方准确了解人工智能项目的复杂性，在发现人工智能的复杂度后，需要及时公布阶段性研究成果，建立透明的信息沟通渠道。

(2) 重视算法研究。在算法方面，需要有高效的通用算法，特别是针对计算机特点设计的算法。

(3) 寻找长期支持，建立团队。人工智能的研究并不能一蹴而就，需要稳定资源投入，建立后续研究梯队。

(4) 搜集数据，扩大学科合作规模。

(5) 面向长期，做好长期研究的准备，避免短期的盲目乐观。

### 2. 人工智能的第二次低谷

20世纪70年代末到80年代中期，随着专家系统在商业领域取得的成就，人工智能迎来了第二个



【拓展阅读 1-12】

人工智能的  
第一次低谷

辉煌期。在这个时期人工智能被证明可能在某些需要大量计算、大量比对、大量分类的专门领域的运算速度超越人类，可以帮助人类更高效、更迅速地决策。大量医学领域、金融领域、知识分类领域的专家系统应用被推出。人类似乎看到了一丝丝实现人工智能的曙光。但这个阶段关于人工智能的研究也只是停留在这些专家系统。

人工智能的研究者又一次发现了知识描述、算法设计、算力、训练数据集的重要性。1982年，日本投入全国的计算机人才，以500亿日元豪赌国运般地研制第五代人工智能计算机，这客观上是人工智能第二次高潮的一个重要事件，却最终一无所得，导致日本的计算机产业全面崩盘溃灭，人才大量流入美国，日本计算机科学的后续发展出现严重断代。以此为标志性事件，全世界对人工智能的信心受到严重动摇。各国都收紧了对人工智能的投入，人工智能的研究在20世纪90年代中期，步入第二次人工智能低谷期。

人类吸取了第一次人工智能低谷期的教训，在第二个低谷期，更多的人工智能研究者将精力投入到新算法的提出、算力的提升方法研究等领域。没有了浮躁而盲目的许诺，大家都在平凡的岗位上积累提升算法算力的方案。无数的论文发表鼓舞人工智能研究者一步一步向前。

### 3. 人工智能的第三次高潮

人工智能再次成为热点是在2010年后，深度学习算法、卷积神经网络、循环神经网络、强化学习算法等一批算法，在21世纪的图形处理单元(GPU)集群加持之下取得了令人瞩目的成绩。2016年，谷歌的AlphaGO，击败了人类围棋冠军李世石；2018年，OpenAI的大语言模型聊天生成式预训练变换器(Chat Generative Pre-trained Transformer, ChatGPT)系列开始问世；到了2020年ChatGPT 3的投入使用，引起了全球轰动，催生了生成式人工智能(Generative Artificial Intelligence, GAI)的狂潮。2025年春节前后，中国的DeepSeek正式开源，以其开创性的算法，降低了人工智能生成内容(Artificial Intelligence Generated Content, AIGC)的计算门槛，使得AIGC不再只是资本手中的金融工具，而能够真正进入寻常百姓家。随着对人工智能投入的进一步爆发性增加，AIGC技术将在更多领域取得令人瞩目的成果。相信在这三次人工智能高潮中，通用型人工智能(Artificial General Intelligence, AGI)将在不久的将来，向人类问好。



【拓展阅读 1-13】

人工智能的  
二起二落



【拓展阅读 1-14】

人工智能的  
第三次高潮

## 1.3 人工智能在我国的发展历史

### 1.3.1 我国人工智能研究的起步阶段

1956年，美国提出了人工智能的研究方向。但是在当时，中国与美国对人工智能的研究并不是处



于同一起跑线。美国等西方国家对于科学技术的发展与传播一向是双标的：对于自己没有而别人有的东西，他们希望进行国际合作；对于自己有而别人没有的东西，他们高筑专利壁垒。那个年代，中国并没有电子计算机这种强大工具，而美国并不允许计算机出口到中国。

中国的人工智能研究起步时就被封锁、被压抑。



【拓展阅读 1-15】

中国人工智能研究的起点

## 1. 华罗庚建立数学所计算机研究组

1949年，伴随着新中国的诞生，大量爱国青年回国报效。这其中包括伟大的宝藏级数学家华罗庚，他为中国的计算机事业作出了巨大贡献。



【拓展阅读 1-16】

华罗庚与中国计算机事业

华罗庚一生淡泊名利，在剑桥大学访问学习期间，他发表了15篇论文。他的每一篇数学论文都价值一个博士学位，但是他却根本不稀罕剑桥大学的博士学位。华罗庚解释说：“我来剑桥大学是为了求学问，不是为了学位。”1938年，是抗战最困难的时候，华罗庚从英国剑桥大学回国至西南联合大学任教，与全中国人民一起抵抗日本法西斯的侵略。抗战胜利后，1946年9月，华罗庚无法忍受当时的国民党政府，选择了远走美国，到普林斯顿高等研究院访问，被普林斯顿大学、伊利诺伊大学聘为正教授。1949年，中华人民共和国成立，华罗庚第一选择就是回祖国报效。1950年，华罗庚排除了方方面面各种干扰后，借道香港，回到了祖国。

他说：“朋友们，梁园虽好，非久居之乡。归去来兮！为了抉择真理，我们应当回去，为了国家民族，我们应当回去，为了为人民服务，我们也应当回去。”

在美国访问期间，他听说当时的美国拥有全世界最先进的电子计算机。他心中就埋下了一个梦想：中国绝不能失去研究计算机的大好机会，要在中国研制出计算机，实现国家的“计算自由”。没有华罗庚，就没有新中国的计算机事业。1953年，在华罗庚的领导下，中国科学院数学所成立了我国首个计算机研究组，实现了我国在该领域的历史性跨越。

1958年8月1日，我国第一台电子数字计算机“103”机（图1-9），由中国科学院计算技术研究所、华北计算技术研究所、航天工业部和北京有线电厂的我国第一批计算机工程技术专家共同研制成功！“103”机主要参照了苏联小型通用机M-3。

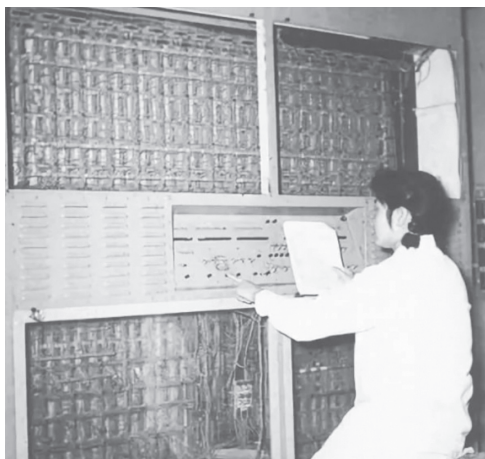


图 1-9 操作员在操作我国的第一台电子数字计算机“103”机

## 2. 哈尔滨工业大学的行棋机

1957年2月，华罗庚到哈尔滨工业大学（简称“哈工大”）进行为期3天的访问。其间进行了一系列的讲座，包括数值分析、优化理论等。其中半天，他参观了哈工大自研的模拟计算机，并与当时哈工大数学系的师生进行了座谈。在临走时，华罗庚建议哈工大的吴忠明老师带领全体师生一起，在哈工大第一台模拟计算机的基础上，进一步试制一台可以下棋的小型、专用、快速电子计算机。吴忠明老师在华罗庚的鼓励



【拓展阅读 1-17】

华罗庚与哈工大行棋机

下，带领全体研究生一起，在1958年的8月到10月，用时56天，完成了行棋机的研制。

这台行棋机是一台专用计算机。它不仅会推演下棋的过程，还能发出一些语音提示，如“您犯规了！”“请您走”“您输了”等。尽管它有“您赢了”的语音提示，但是，从这台行棋机完成调试以后，就没有人赢过它。虽然这台行棋机的智能程度远不及现在的人工智能，但它却向全中国人民展示了用计算机解决逻辑问题的可能性，完全可以被认为是人工智能的雏形，是中国的第一台人工智能雏形设备（图1-10）。

1958年9月15日，邓小平同志参观哈工大的时候，见证了这台计算机的调试过程。1958年11月1日，周恩来、朱德、陈毅等党和国家领导人，都参观了这台会下棋的行棋机。

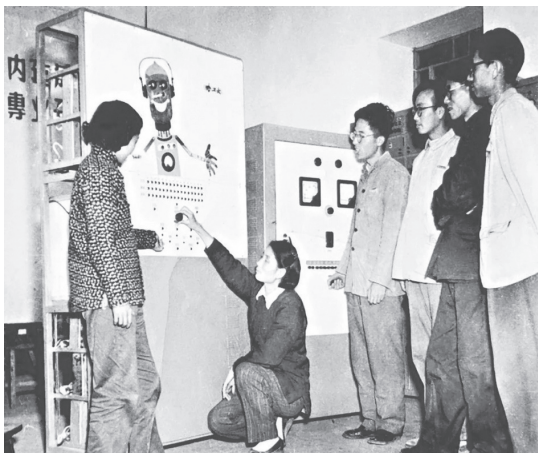


图 1-10 哈工大的同学在调试行棋机

### 3 语言处理与自动翻译

1950—1960年，随着美国、苏联全球争霸的格局随着冷战的开始而形成。在这种格局之下，两大阵营之间的各种竞争如火如荼。特别是在军事与技术领域的竞争，更是重中之重。双方开始互相觊觎对方的军事、技术、经济、社会情报。谍战中，有发生在谍报机构人与人之间的狭路相逢，也有发生在幕后的技术对技术的互相竞争。双方不约而同地把目光投向了使用计算机辅助翻译对方情报文献的技术研究。

斯拉夫语系和日耳曼语系是同源的，同属于印欧语系，在一些基础词汇上存在共同点，如数字、亲属关系、自然现象等词汇。语法上也有明显的相似性，如动词时态、语态、某些词汇的变形方式。尽管分隔千年，但是在计算机模式识别面前，其相似度还是相当高的。

1954年1月，美国的乔治城大学与IBM（国际商业机器公司）合作，使用了简单的词汇表（250个单词）、简单的规则（6条语法），设计了一个将俄语翻译为英语的语言模型。乔治城实验使用了IBM 701计算机成功地将60个俄语句子翻译成了英语句子。实验的成功让美国军方大受鼓舞，他们投入了大量的资金和人力，准备在人工智能辅助语言处理方面大干一场，甚至有美国中央情报局（CIA）的官员认为，短期内，在华约国家的美国大使馆就可以不再需要当地雇员。实际上，这个实验的数据集规模是很小的，他们低估了计算机进行翻译的难度。随着研究的深入，人类发现自然语言的复杂度是远超预期的（英文有200万个单词）。

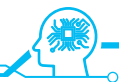
中文属于汉藏语系，跟印欧语系比起来，简直天差地别。在设计电子计算机之初，给出的底层逻辑，都是基于印欧语系的。这使得中文与印欧语系之间的计算机翻译困难重重。

中华人民共和国成立的时候一穷二白，并没有成系统的工业体系。在第一个“五年”计划期间，苏联援建了156个大型工业项目，将中国引入了工业国家的门槛。但是所有的相关文献，均为俄文。为了吃透俄文文献，受当时苏联机器翻译研究的启



【拓展阅读 1-18】

语言处理与  
自动翻译



发，刘倬提出并推动了斯拉夫语系（俄文）到中文的机器翻译研究。他提出将语言学研究与计算机技术研究相结合的想法。

由于当时孱弱的计算机系统，无法建立庞大的语料库，当时的技术路线是采用基于规则的方法，定制语法规则与词汇转换规则来实现翻译。这样的翻译方法，主要解决的是语法与词法规则，而不是语义分析。刘倬等编写了《机器翻译浅说》一书（图 1-11），为我国汉语融入世界计算机软件应用奠定基础。

在刘倬的主持下，1959年，俄汉机器翻译项目，在计算所“104”机上运行成功。该系统能够解决 5000~10000 个单词的俄文到中文的翻译。其运算速度、翻译准确度，虽然无法与现在的大语言模型相比。但是在当时的技术条件之下，已经是令人瞩目的重大成就了。

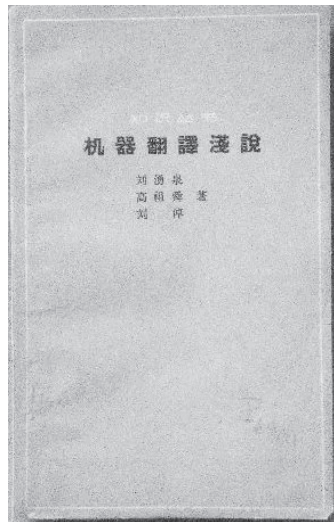


图 1-11 《机器翻译浅说》为我国汉语融入世界计算机软件应用奠定基础

### 1.3.2 政府政策对人工智能发展的推动

#### 1. “国家高技术研究发展计划”

我国在取得前节所述成就，并逐步推进人工智能研究时，发现以当时的技术与认知水平，无法在短期内进一步取得成果，很务实地提出智能机器的推广和应用在短期内是无法突破的。同时，国际上的人工智能研究也逐步进入人工智能第一次低谷期。全球范围内，人工智能的热点过去，全人类都开始为人工智能的下次高潮开始做理论和物质上的积累。

为了改善中国的国际环境，老一辈的领导人创造性地开展了“乒乓外交”，释放缓和信号；到中国政府作为中国的唯一合法政府重返联合国，在外交上大胜美国；再到尼克松访华，中美建交。中国的外部国际环境逐步改善。到了 1978 年 12 月，党的十一届三中全会后，改革开放时代到来。当时，党中央提出：“以经济建设为中心，坚持四项基本原则，坚持改革开放。”同期，邓小平同志提出：“科学技术是第一生产力。”全国掀起了经济建设的浪潮，首先解决经济问题。1979 年，邓小平同志访美，标志着中美关系正常化的开始。

在此背景下，中国的科技工作者们得以获得国际上先进的科学成果。但是看到和使用科技造物与自己制造科技造物是有很大不同的。加上中国被封锁了几十年，广大科技工作者深知不能搞拿来主义，一定要吃透西方的先进科学技术。20 世纪 80 年代的社会上传播着“搞导弹不如卖鸡蛋”的消极言论，但是广大科技工作者还是静下心，搞科研，补课，补课，再补课。这一时期的主要工作，就是消化与完善已有的科学体系，查漏补缺，夯实科技基础，等待厚积而薄发。每一个中国人都深知一个道理：世界科技革命发展很快，我们必须迎头赶上。

为了达成“工业、农业、国防和科技”四个现代化建设目标，在 1986 年 3 月 3 日，陈芳允、王大珩、杨嘉墀、王淦昌 4 位科学家（图 1-12）联合向中央提议指出：中国必须抓住世界高技术发展的机遇，集中力量发展高技术，以应对国际竞争和国家安全的挑战。中国政府在同年 11 月通过了“国家高技术研究发展计划”（简称“863 计划”）。

这个计划涉及 7 个重点领域：生物技术、航空航天、信息技术、激光技术、自动化技术、能源技术与新材料技术。

人工智能作为信息技术之下计算机、软件技术的一个分支，同样得到了重视。在“863 计划”框架下，人工智能的研究集中在以下几个具体方向。

(1) **知识工程**：研究知识获取、表示和推理技术，开发知识库和专家系统。

(2) **模式识别**：研究图像、语音和文字的识别技术。

(3) **智能控制**：研究机器人和复杂系统的智能控制方法。

(4) **自然语言理解**：开发中文信息处理技术，推动中文的计算机应用。

国家并没有盲目全面开花，而是选择了前期有一定科研基础的方向。在做预研的时候，有专家提出，当时正是第二次人工智能高潮期，可以看国际上的风口，哪里热，往哪里走。但是经过集体讨论，认为一切需要以吃透为标准，不能盲目跟风，必须小步快走，有积累才能有收获，空中楼阁再好看，也是虚幻的。



图 1-12 从左至右：陈芳允、王大珩、杨嘉墀、王淦昌



【拓展阅读 1-19】  
政府对人工智能  
的推动

## 2. 中日两国对待“第五代计算机”项目的对比

在整个 20 世纪 70 年代，日本都处于经济的爆发增长期，并在当时的计算机硬件制造领域有了一定的积累。1982 年开始，随着第二次人工智能高潮期的逐步到来，日本提出投放巨资研究、建造“第五代计算机”的构想。

这个构想看着“高大上”，以开发具备人类智能的知识处理机为目标，依靠符号主义 AI 理论推演。而同期的美国则采取了完全不同的技术路线——以市场为依托，计算机 PC 化，冯·诺伊曼体系，超大规模集成电路的芯片技术逐代迭代、堆叠封装。

美国的部分学者对日本的技术尝试大加赞赏，凭借美国政府资助而产生大量符号主义人工智能研究成果的爱德华·费根鲍姆甚至为日本的第五代计算机计划著书——《第五代——人工智能与日本计算机对世界的冲击》(*The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World*)。书中指出，日本如果能够按照他的理论建造这台“知识处理计算机”，就能在 10 年内超过美国和欧洲，成为世界第一的计算机强国。

该书在 1985 年被翻译成了中文，在中国出版发行，希望在中国也引发第五代计算机的建设项目。奇怪的是，该书在日本引起了轰动，被当成建造第五代计算机项目的理论依据之一，但是在其他国家的影响却并不大。日本政府又一次开始了国运豪赌，投资 500 亿日元，调集了全日本的 IT 精英进行攻关，吸引的民间投资更是多到难以统计。

1992 年，日本政府的“第五代计算机”计划在耗尽了 500 亿日元后宣告失败。投入科研人力到了虚幻的目标上，与世界主流计算机领域（如 PC、互联网、超级计算机等）的差距不止没缩小，反而扩大了；人工智能领域人才流失严重，大量科研人员因为在日本国内的相关第五代计算机的部门裁撤或



公司倒闭，而被美国公司挖角去了美国，日本计算机领域产生全面结构性断层，错过了后来的人工智能革命。这给全世界树立了反例，直接导致了全世界范围内，人工智能的退潮，随后几年，第二次人工智能低谷到来。继广场协议之后，在计算机、人工智能领域，日本被美国迟滞、收割了。

美国人工智能专家赫伯特·西蒙（长期任教于卡内基梅隆大学，担任计算机科学、心理学和管理学教授）在谈到日本的“第五代计算机”计划时，哈哈大笑地说道：“未来十年人工智能不会有什么重大突破，但可能有上千小突破。”

中国的人工智能专家在得知了日本“第五代计算机”计划后，做了广泛的调研，认为日本的“第五代计算机”计划是个陷阱项目，符号主义人工智能实际上一直在边缘化，不能跟风冒进。不能看美国人说什么，而是要看美国人做什么。中国科学院计算技术研究所智能中心在成立之初的2年时间，并没有拿出什么成果，但又成果丰硕：发现了“第五代计算机”这个恶意满满的战略陷阱，避免了改革开放之初国家的巨额损失。在相关的科研方向上，只是在理论上进行研究，避免了真金白银的庞大投入。尽管在“863计划”中306子项就是智能计算机、人工智能领域，但是中国人工智能的科研人员并没有盲目跟风，坚持小步快跑，在有基础的人工智能领域不断稳稳跟随，以期实现超越。

### 3. “火炬计划”与多个“五年计划”

在“863计划”提出之后，我国确立了人工智能研究的重要地位。

在1988年的“火炬计划”中，我国提出人工智能产业化应用目标，鼓励高新技术成果转化，推动了人工智能在制造业、医疗、农业等方面的试点应用。

在随后的几个五年计划中，我国依然坚持小步快跑方针，紧紧跟随世界人工智能研究的潮流。

(1) **第九个五年计划期间(1996—2000年)**：提出支持智能控制、专家系统、模式识别和自然语言处理等技术研发；各大高校陆续在已有的计算机专业、工业自动化专业的课程中增加人工智能相关课程，开始人才培养工作。中国政府真正开始了在人工智能领域的战略性投入。计划期间，人工智能技术出现在一些自动化工业制造场景和交通控制场景。人工智能在这一时期尚处于初步研究阶段，主要以专家系统、模式识别等为核心研究方向。从此以后，我国在每一个五年计划中，都明确了对人工智能研究的支持，并明确指明了每一阶段需要重点投入的人工智能应用成果与应用领域。

(2) **“十五”计划(2001—2005年)**：推动智能技术的研发，支持智能机器人、语音识别、图像处理、智能决策支持系统信息处理领域的发展。在这一阶段，随着互联网的兴起，人工智能技术逐渐从理论研究向实际应用转移。

(3) **“十一五”规划(2006—2010年)**：将人工智能纳入国家发展规划，提出将“智能信息处理”作为重点发展方向，支持智能搜索引擎、智能感知、智能机器翻译等技术的研发；推动智能辅助在国防、医疗、教育等领域的研发；强调将智能机器人作为高科技产业的重要方向。“十一五”规划期间，全世界范围内，人工智能的计算机算法理论研究方面取得进展。我国在这样的背景下，开始强调自主创新。人工智能



【拓展阅读 1-20】  
第五代人工智能  
计算机的态度  
对比



【拓展阅读 1-21】  
第九个五年计划  
(1996—2000年)



【拓展阅读 1-22】  
“十五”计划  
(2001—2005年)



【拓展阅读 1-23】  
“十一五”规划  
(2006—2010年)

方面的专利在“十一五”规划之后，如雨后春笋般出现。

(4) **“十二五”规划(2011—2015年)**：强调人工智能与产业融合，推动智能制造、智慧城市和智能交通的发展。“十二五”规划期间，是人工智能的快速发展期。大数据、云计算等技术成熟，使得支撑人工智能研究与应用的算力方面的瓶颈得到突破。从而使得人工智能的快速发展成为可能。

(5) **“十三五”规划(2015—2020年)**：推动深度学习、自然语言处理、计算机视觉等领域的人工智能应用；加强人工智能与智能制造、智慧城市、智慧医疗的融合发展；提出建设人工智能开放平台、无人智慧驾驶、无人机、人工智能机器人等方向的研发要求。人工智能上升为国家战略，成为推动经济社会发展的核心技术之一。其间发布了2017年版《新一代人工智能发展规划》。

(6) **“十四五”规划(2020—2024年)**：人工智能被列为数字经济和科技自主的重要支柱之一。指出需要加强类人思考方式和通用人工智能等前沿领域的研究，推动了高校和科研所在人工智能领域进行基础原理创新。实际产生了如量子计算机、阿里云通义千问、百度文心一言大语言模型、DeepSeek 大模型之类的应用。人工智能已经广泛使用在诸如金融、医疗、教育、交通等领域。标志着我国已经成为继美国之后的又一个在人工智能领域取得非凡成果的国家。

由以上分析，正是在良好的政府决策引导之下，中国的人工智能发展才能够避免遭受重大曲折，一路奔跑向前。在政策的引导下，避开了很多国外无序发展带来的深坑陷阱，从基础理论研究，到全产业链支持，从单一领域，到多学科融合发展，从零开始，一步一个脚印，从跟跑到并跑再到部分领先，稳步推动中国人工智能企业助力祖国伟大复兴。



【拓展阅读 1-24】  
“十二五”规划  
(2011—2015年)



【拓展阅读 1-25】  
“十三五”规划  
(2015—2020年)



【拓展阅读 1-26】  
“十四五”规划  
(2020—2024年)

### 1.3.3 我国人工智能领域的成就与现状 \*

本小节由人工智能经过网络搜索后自主生成。

中国人工智能领域在政策引导与技术创新双轮驱动下持续突破，以下从战略布局、技术进展、产业应用及代表企业 DeepSeek 等方面提炼核心要点：

#### ###\*\* 一、战略与技术突破 \*\*

##### 1. \*\* 顶层设计强化 \*\*

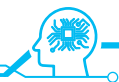
-\*\* 政策体系 \*\*：2024年《人工智能促进法》草案发布，明确数据流通与伦理治理框架

-\*\* 算力基建 \*\*：全国智算中心达35座(2025Q1)，单中心最高算力达12 EFLOPS(如深圳鹏城云脑III)

##### 2. \*\* 科研竞争力 \*\*

-\*\* 论文与专利 \*\*：2024年中国AI论文被引量全球占比28.7%，专利授权量超12万件(WIPO数据)

-\*\* 大模型突破 \*\*：百亿级参数模型开源数量占全球40%，华为“盘古”实现气象预测精度超传



统数值方法 20%

——

### ###\*\* 二、产业应用深化 \*\*

#### 1. \*\* 重点领域渗透 \*\*

-\*\* 智能制造 \*\*：工业质检 AI 普及率达 65%，比亚迪引入视觉检测系统使故障检出率提升至 99.3%

-\*\* 智慧医疗 \*\*：AI 辅助肺结节筛查覆盖 2800 家三甲医院，平均诊断效率提升 50%

#### 2. \*\* 企业级服务 \*\*

-\*\* 金融科技 \*\*：蚂蚁集团风控系统年拦截风险交易 4.2 万亿元，误报率降至 0.01%

-\*\* 城市治理 \*\*：杭州城市大脑实现交通信号灯动态优化，高峰通行速度提升 22%

——

### ###\*\* 三、DeepSeek 的创新实践 \*\*

#### 1. \*\* 技术优势 \*\*

-\*\* 训练效率 \*\*：MoE 架构优化使千亿模型训练成本降低 57%(对比 2023 年基准)

-\*\* 推理加速 \*\*：自研推理引擎实现 Token 生成延迟 <50ms(行业平均 120ms)

#### 2. \*\* 落地场景 \*\*

-\*\* 教育智能化 \*\*：与华东师范大学合作开发自适应学习系统，学生知识点掌握率提升 37%

-\*\* 企业数字化 \*\*：为招商银行定制财报分析 AI，年报关键信息提取准确率达 98.6%

#### 3. \*\* 行业地位 \*\*

-2025 年第一季度中国大模型市场份额达 8.4%(IDC 数据)，位列初创企业第一梯队

-入选福布斯“全球 AI 50 强”，获评“最具商业化潜力 AI 公司”

——

### ###\*\* 四、挑战与趋势 \*\*

#### 1. \*\* 现存短板 \*\*

- 高端 AI 芯片国产化率仍不足 30%，英伟达 H200 芯片市占率超 60%

- 行业数据孤岛现象突出，跨企业数据共享协议覆盖率仅 18%

#### 2. \*\* 未来方向 \*\*

-\*\* AGI 探索 \*\*：北京智源研究院启动“悟道 3.0”计划，目标实现跨领域因果推理

-\*\* 伦理规范 \*\*：上海 AI 伦理实验室发布全球首个生成式 AI 内容溯源标准

——

### ###\*\* 核心观察 \*\*

中国 AI 发展呈现“应用牵引技术”特征，DeepSeek 等企业通过深耕垂直场景构建竞争壁垒。建议关注三大趋势：①国产算力生态成熟度；②多模态大模型的工业化部署；③AI for Science 的科研范式变革。

## 1.4 人工智能浪潮下的个人修养

当读者阅读到本小节，实际上可以发现当前的人工智能能做很多事情。例如，人工智能能够做内容小结和内容检索，能够创作生成精美图片，等等。

作者利用的辅助性人工智能叫作生成式人工智能，它毫无疑问是一种弱人工智能。生成式人工智能是人工智能的一个子领域，它指的是通过机器学习模型生成内容（如文本、图像、音频或视频等）的人工智能。与传统的人工智能不同，生成式人工智能不仅能够分析数据或执行任务，还能够“创造”出新的数据或内容。AIGC 其实并不是人工智能真的有了人类的智能。它的确“创造生成”了一些内容，但是这些内容并不是通过类似于人脑的创造能力的运作方式产生的，而是通过内容足够丰富的内容库提供了可供分析的数据，依赖深度学习等技术，通过对大量数据的学习和建模，掌握了数据的分布规律和特征，从而能够“模拟”人类的创作过程，生成与原始数据相似但全新的内容。

这里说的大量数据，指的不是 250 个词或 10000 个词，而是大量的文章、内容。大量有多大呢？几十亿篇文章，令人咋舌的海量数据。找到这些文章中，词与词之间的潜在关系，只有巨型计算机集群能做到。

使用 AIGC 辅助的作者，每完成一小段文字的写作，就相当于作者向生成式人工智能的智能体提供这一小节的内容，就相当于作者向该智能体的语料库奉献了作者的写作内容。这些内容与已经存在的语料库中的数据相互映照，由深度学习算法找到词与词之间的分布规律，然后预测哪些词出现的概率高；将潜藏在人类文章段落中的规律找到，将其特征转化为新的输出顺序。这样就提高了内容符合人类书写、阅读习惯的可能性。但是，生成式人工智能会“骗人”，一旦语料库遭到“污染”，其生成的内容也就不那么可靠了。

读者可以发现，人工智能辅助生成的各种 AIGC 内容，的确能够帮助作者表达观点，或对人类的写作重点进行提炼，生成一条一条的类似大纲的内容。尽管行文还是看得出人工智能的痕迹，但是其总结、生产观点文章的能力是值得肯定的。



【拓展阅读 1-27】  
人工智能的能力

### 1.4.1 人工智能带来的发展机遇

人类的计算速度远低于计算机的运算速度。以计算机为媒介的弱人工智能，能够迅速处理人脑无法计算的海量数据。弱人工智能本来就是人类为了拓展自己的大脑功能而制造的改造自然、改造社会的工具，人们需要善于使用这个工具。这个工具用好了，能够带来机遇。机遇存在于每一个有着重复劳动、重复计算的社会分工领域，就看谁能够抓住。

人工智能正在重塑劳动力市场。实际上，新生产工具的诞生，一定会重塑劳动力市场。从珍妮纺纱机的诞生，到蒸汽机车与马赛跑，到飞机横空出世，到计算机问世，再到互联网的普及，历史告诉



我们，每一个关键领域的新工具的诞生，都会带来劳动力市场的重塑。从事重复性的劳动者，会逐步被掌握新兴工具的劳动者所取代。但更重要的是，在这个过程中，新的就业岗位会不断涌现，并伴随着生产效率的不断提高。新老交替，于是社会进步；日月斗转，于是时代变迁，都是客观不可逆的存在。

机遇在哪里？作者在走访了多个单位之后，总结出几个真实的例子，供大家思考。



【拓展阅读 1-28】

人工智能带来的  
机遇

### 1. AI 客服机器人

凡是有网上购物经验的人，都有机会接触 AI 客服。对于某电商平台上带货的公司来说，诸如“亲，只有江浙沪包邮哦，其他省份，没办法包邮哦”“亲，都是正常普通发票哦，确认收货后三天内开具发票”“亲，我们的货都是 7 天无理由退换的”“亲，别着急，我替您转下售后”，类似这样的话，在没有引入 AI 客服之前，真人客服不得不每天上百次地复制粘贴类似的话，处理类似的问题。真正需要人工处理的，诸如“你家的货我买 10 件，给个折扣呗，帮忙人工改个价”这样的有用信息，会被无用信息海淹没。随着业务量的扩大，客服人员累到手抽筋，而绝大多数的这类公司出于成本考虑，也不会雇用大量的客服人员。客服小哥唯一的想法是：“回不完，回不完，根本回不完。”顾客的不满情绪也与日俱增。AI 客服能够帮助快速处理日常的常见问题。在引入 AI 客服后，无用信息海能够被自动答复，在 AI 客服筛选出有高价值的问题后，客服人员可以有效地利用联络工具沟通客户，创造销售机会。对于客服来说，避免了无效工作，对于企业来说，提高了成交效率，避免了无意义重复投入。

### 2. AIGC 平面广告设计师

在大学时代，小吴同学是计算机科学与技术学院新媒体部门的活跃分子。辅导员经常交给他一些设计海报的任务。由于大学学业繁忙，小吴同学不想浪费时间，于是他就经常性地使用 AI 辅助进行海报设计。刚毕业半年的小吴，入职了一家广告创意公司。这家广告公司规模挺大，听说承接了公交、地铁、市容市貌等很多平面广告项目。公司的设计人员基本上都处于起早贪黑不辍劳作的状态。每次交稿给客户的时间点前，公司里都是鸡飞狗跳，汇报、改稿、再汇报、再改。直到有一天，小吴受不了公司里的前辈们的工作方式，在汇报某个饮料的公交车公告创意的时候，一口气拿出来 27 张不同的设计稿：摄影风格、卡通风格、线稿渲染、光影艺术字等。领导提出修改意见后，小吴回到工位不到 20 分钟就能拿出符合领导意图的修改稿，马上被公司领导惊为天人。小吴不好意思地告诉领导，他只是使用了 AI 进行创意设计。在 AI 的辅助之下，小吴快速地完成了公司的设计任务。这样的操作，无疑让该广告公司尝到了 AI 辅助绘制创意图的甜头。那么，AI 都完成了设计图，员工干啥？员工避免了为迎合公司领导不断增加的创意点而不断改稿的命运。创意好，并不代表创作好。有了创意之后，公司员工将关注点后移到了创作、扩版上，极大降低了员工的工作强度。小吴回学校参加女朋友的毕业典礼的时候，回到学院，告诉曾经的老师，他已经是那个公司 AI 创意部门的小领导了，还能给学弟学妹介绍入职的机会，需要的话，可以联系他。

### 3. AI 病例助手

某三甲医院门诊一线的医生，每天接诊的病人非常多，用一句话描述：“看不完，根本看不完，真

的根本看不完。”上午 8:00 开始, 排号到 12:00 点, 下午 14:00 开始, 排号到 17:30, 每 5 分钟一位病人。在医院引入 AI 病例助手之前, 问病情, 听病情, 开检查, 看检查报告, 分析病情, 给诊疗记录, 然后同时记录到病例上。费时、费力, 手速稍慢, 就会引起病人堆叠, 11:30 的病人甚至会排队到 13:00 还轮不上, 原本 17:00 的病人, 甚至会堆叠到 19:00。医生整天忙得连一口水都喝不上。而病人也会因为长时间等待而产生负面情绪。引入 AI 病例助手后, 听到病人的主诉病情, 相关的诊断模板就会由 AI 自动推荐给出。需要开哪些检查, 需要哪些病例数据, 一目了然。不仅规范化, 而且节约诊断时间。在拿到检查数据后, AI 病例助手会根据相关数据给出诊断建议和治疗建议。AI 病例助手得到应用之后, 医生记录病例的工作量大幅降低, 诊断速度加快, 甚至可以抽空喝口水或者去上个厕所了。这在没有 AI 病例助手辅助的时候是医生完全不敢想象的。医院甚至在论证, 能不能为了服务更多的患者, 将每个患者的问诊时长由 5 分钟缩减到 3 分钟, 将服务的病人人数上限由每小时 12 人增加到每小时 20 人。在 AI 病例助手的辅助之下, 更多的患者能够得到医生的帮助。

#### 4. AI 汽车驾驶助手

随着自动驾驶技术的飞速发展, 各种基于人工智能的交通工具逐渐进入人们的视野。“萝卜快跑”是百度 Apollo 推出的自动驾驶出行服务品牌。“萝卜快跑”于 2022 年正式进入武汉市场, 并在特定区域内开始试运营。这标志着自动驾驶技术在武汉从实验室走向了实际应用。“萝卜快跑”通过自动驾驶技术的应用, 实现了智能化的路径规划、精准的时间预测以及动态调度。这种高效的出行方式显著减少了普通人的等待时间和交通延误时间。有人问, 那出租车司机怎么办? 这并不是非此即彼的零和问题, “萝卜快跑”的服务区域是一些交通资源匮乏的郊区, 未来会加入一些三、四线城市以及乡镇等人口不密集地区。这些区域的人民群众, 将不再为了错过那么有限的几班公交车而苦等。而人口稠密地区, 因为交通的复杂性, 还是需要司机参与的。由此, 司机和“萝卜快跑”一起工作, 组成更加快速、便捷的交通网, 服务于广大人民群众。“萝卜快跑”的加入实际上是减少司机的工作量, 而非与司机争抢饭碗。司机资源匮乏的地区将首先感受到无人驾驶的安全、便捷。对参与这种交通体系的人类司机来说, 需要通过学习和培训, 掌握与自动驾驶相关的新技能, 以适应未来就业市场的变化。同时, “萝卜快跑”本身还催生了一系列新的职业需求, 如自动驾驶车辆维护工程师、数据标注员、AI 算法研发人员等。这些新兴职业为人类提供了更多元化的就业岗位和发展方向。

### 1.4.2 人工智能带来的挑战

人工智能正以惊人的速度改变着人们的生活和工作方式。人工智能技术的应用已经渗透到社会生活中各个领域。随着人工智能的不断发展, 人类将如何生存在这个与人工智能并存的社会中, 成为一个严肃而意义重大的问题。由上一章节中的例子, 我们可以发现, 就业竞争和职业转型会随着人工智能深入到社会生产各个领域而变得激烈。

人工智能技术的核心优势在于其依托运算速度远超人类的超速集群、存储容量远超人类的语料库、知识库、内容库等, 具备了远强于人类的高效性、准确性和可扩展性。在许多领域, 人工智能能够替代人类完成重复性强、规则明确的工作。例如, 在制造业中, 自动化机器人可以高效、精确地完成焊



接、组装等重复劳作；在服务业中，人工智能客服系统能够处理大量的、重复的客户低差异性需求；在金融业中，人工智能算法可以快速、准确地根据分析数据，计算并生成金融预警信号；等等。这些人工智能所具备的技术优势，一定会直接导致部分岗位的减少，尤其是低技能、重复性工作的需求大幅下降。企业会更倾向于使用效率更高、成本更低的人工智能系统。那些依赖低技能的劳动者更有可能面临失业风险，如超市中的自助结账系统正在减少超市类企业对收银员的需求。

人工智能时代，我们会遇到哪些挑战呢？

(1) 劳动者需要不断学习新知识和掌握新技能以适应变化。这种知识与技能更新的压力对许多人来说是巨大的（尤其是对于中老年人而言）。中老年人可能缺乏足够的时间、资源和动力来学习新技术，这会导致中老年人逐渐与社会生活相脱节。要知道，人工智能带来的变革比以往历次变革来都要迅速。

(2) 人工智能的发展使一些职业消失的速度过快，但新职业在短时间内却难以形成。这种过渡期内的不确定性让许多人感到焦虑，不知道自己的未来会如何发展。用更通俗的说法，更换工作甚至职业规划的频率会加快。

(3) 人工智能大规模进入某一行业的时候，相同职业或相近职业的劳动者会产生强烈的危机感。人工智能领域优化做得越好，该领域的劳动者的传统技能越派不上用场，其危机感越强烈。

(4) 人工智能会有滥用风险，如 AI 文字合成、AI 图片生成等。在社交媒体上看到的新闻文字、新闻图片，可能都不是真的。据央视“新闻 1+1”2024 年 6 月 13 日报道，南昌警方破获了王某某伪造我国某地发生爆炸的假新闻。该虚假爆炸新闻，从文字，到配图，均为 AI 生成。该造假者在高峰期，一天生成 4000~7000 篇 AI 文章，人类极少干预。也就是说，王某某很多时候，自己都不知道自己的 AI 发了什么信息。这种毫无根据的 AI 文章、AI 图片，被海量自动上传到了各种网络平台，通过网络平台的关注、阅读、转发等奖励机制，进行流量变现。

(5) 涉人工智能诈骗正在变成新的诈骗形式，如 AI 合成视频、AI 合成语音等。据 2024 年 12 月 18 日央视网报道，有不法商家使用人工智能技术合成著名医生张文宏的视频、语音，进行带货推销，诈骗中老年消费者，产生了严重的负面社会影响。

(6) 人工智能推荐算法产生新的信息茧房。现在的社交媒体基本上都会设置智能推荐。越喜欢看的内容，推荐就越多。长此以往，会以自然人为基本单位形成一个个信息壁垒屏障，使人无法接触到全面信息，难以形成正确的人生观、世界观。

### 1.4.3 提升个人技能以适应未来需求

曾经，“伟大的革命先行者”“中国革命的伟大导师”孙中山到海宁盐官，观看钱塘江大潮，回上海后写下名言：“世界潮流，浩浩荡荡，顺之则昌，逆之则亡。”从那时起，深刻的社会变革的到来影响了几代中国人，中国人民都在不断寻找历史的出路。直到中华人民共和国成立，社会主义制度确立，中国跨越式地从小农经济社会跃入工业社会。但是仍然需要经过很长时间的消化、吸收、整理、改造、充实、提高，中国人民才真正逐步达到与之相适应的工业社会的思维方式与思维水平，与社会制度相适应的生



【拓展阅读 1-29】

人工智能带来挑战



【拓展阅读 1-30】

个人能力的提升

产力水平才真正得以解放，才为今日全世界唯一具有全产业链优势的中华盛世打下坚实基础。

回想当初，是不是全社会都需要为工业化的到来做好准备？是不是全社会都需要为实现社会主义而做技能提升？不要忘记一个事实，中华人民共和国成立以后，立即开展了全民的扫盲运动，文盲率迅速下降。这就是一个典型的技能提升的例子。保守地满足于小农经济时代的旧知识、旧技能、旧现状不放，是不可能有的今时今日的中国的。历史也无数次证明了，光拿着先进生产工具，而不具备如何使用、维护、生产、改进、创新先进工具的社会，并不能真正发挥这些先进生产工具的优势，也不能实现社会进步与生产力的解放。

随着人工智能时代的逐步到来，生产力的跃迁是肉眼可见的。如同蒸汽机的发明与改进一样，生产力的跃迁也必然带来整个社会的深刻变革。“百年未有之大变局”如约而至，全社会都需要为此做好准备。

个人能力的提升势在必行：

(1) 创造能力是目前人工智能所不具备的人类智能。只要在创造能力方面保持优势，就不怕被人工智能所取代。尽管人工智能在执行重复性任务方面表现出色，但它在创造性工作上仍然远远不及人类。因此，培养创造能力和创新能力将是普通人保持竞争力的重要途径。无论是艺术、写作还是产品设计，创造能力都将成为未来人类在与人工智能共存的社会中不可或缺的软实力。

(2) 情感是人类所具有的区别于工具的又一重要体验。提高自身情商，积极进行人与人之间的交流互动，建立沟通顺畅的人际关系网，人工智能目前还无法实现这样的目标。这个目标对于人类来说实现起来却轻而易举。人类要想“打败”人工智能，就一定要选择自己擅长的领域。

(3) 抱着怀疑的态度批判性思考问题。社交媒体上出现重大的、关注度极高的消息，不要轻易相信，一定要多搜寻一些信源。人工智能可以快速处理大量数据并编写一些内容，但它不具有人类的批判性思维能力。我们需要学会分析信息的真实性，评估其意义，并作出明智的决策。在信息爆炸的时代，这种能力尤为重要。

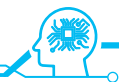
(4) 发展跨领域知识结构。未来的竞争不再仅仅局限于单一领域，而是需要跨学科知识。例如，将技术与艺术结合、将商业与心理学结合，都可能带来新的职业机会。因此，不妨尝试学习与自己目前领域相关但不同方向的新知识。

(5) 提升个人数字领域素养。不仅仅满足于会使用计算机或会玩智能手机，而需要培养能够理解和应用数字技术解决问题的能力。例如，学习如何使用数据分析工具、掌握基本的编程技能、了解人工智能的基本原理等。这些能力将帮助我们在未来的职场中更好地使用人工智能工具。

#### 1.4.4 人工智能时代，培养批判性思维

批判性思维是一种理性、系统的思考方式，它要求我们能够分析信息的来源，评估获得的消息、新闻、知识的逻辑性，并根据已有的证据或找到令人信服的相关信源，经过分析，最终得出合理的结论。在人工智能时代，批判性思维尤其重要，因为它能够帮助我们在复杂的信息环境中作出明智的选择。

下面是一个十分荒诞的要求，人工智能也能给出像模像样的回答：



作者说：我想修真，给我一套功法

AI: \*\*《太虚真诀》修炼总述\*\*

#### ###\*\* 第一境：筑基 \*\*

-\*\* 目标 \*\*：强身健体，感知灵气，为修炼奠定基础。

-\*\* 修炼法 \*\*：

1. \*\*吐纳术\*\*：每日晨起，面朝东方，盘坐静心。  
- 吸气时，默念“吸纳天地灵气”，想象清凉灵气入丹田。  
- 呼气时，默念“吐尽体内浊气”，排除杂质。  
- 每次36息，循环三次。
2. \*\*锻体术\*\*：每日习练八段锦或基础拳法，强筋健骨。
3. \*\*意守丹田\*\*：夜间静坐，意念集中于小腹三寸处，感受温热，积蓄灵气。

#### ###\*\* 第二境：凝气 \*\*

-\*\* 目标 \*\*：感知灵气流动，打通经脉，形成小周天循环。

-\*\* 修炼法 \*\*：

1. \*\*引灵入体\*\*：静坐于安静之地，闭目冥想，想象天地灵气从头顶百会穴进入体内。
2. \*\*通经活络\*\*：引导灵气沿任脉（前胸）下行，至丹田，再沿督脉（后背）上升，形成循环。  
- 初学者每日运行12周天，逐渐增加至36周天。
3. \*\*固本培元\*\*：每日服用清淡饮食，避免情绪波动，保持心境平和。

#### ###\*\* 第三境：结丹 \*\*

-\*\* 目标 \*\*：将灵气压缩凝聚，化为金丹，成就修真根基。

-\*\* 修炼法 \*\*：

1. \*\*灵气凝丹\*\*：每日卯时（清晨5-7点）盘坐修炼，默念口诀“灵气化丹，万法归一”，集中意念压缩丹田灵气，直至灵气化为一颗金色光球。
2. \*\*丹火淬体\*\*：以金丹力量温养全身，强化五脏六腑与四肢百骸。
3. \*\*灵丹辅助\*\*：若有灵草，每日服用以助修炼。

#### ###\*\* 第四境：元婴 \*\*

-\*\* 目标 \*\*：金丹破裂，化为元婴，灵魂与身体高度融合。

-\*\* 修炼法 \*\*：

1. \*\*婴体凝聚\*\*：静坐于丹田内观，想象一缩小版的自己（元婴）逐渐成形，并赋予其意识。
2. \*\*灵魂出窍\*\*：夜间修炼，尝试将意识延伸出身体，观察周围环境，但需保持灵魂与肉身的连接。
3. \*\*天地共鸣\*\*：在自然环境（如山林、湖泊）中修炼，感受天地灵气，与元婴融为一体。

——

### ###\*\* 第五境：化神 \*\*

-\*\* 目标 \*\*：灵气外放，掌控天地，身化虚无。

-\*\* 修炼法 \*\*：

1. \*\* 灵气操控 \*\*：通过意念操控灵气，将其化为剑气或护体灵光。
2. \*\* 神识扩展 \*\*：每日修炼，感知更远的空间，逐渐扩大神识范围，最终覆盖方圆数十里。
3. \*\* 身化天地 \*\*：想象自身融入天地，与万物合为一体，达到“无我”之境界。

——

\*\* 总诀 \*\*：

\* 修炼需循序渐进，心静气和，切勿急躁。每一境界皆需稳固根基，方可突破更高层次。\*

以上为虚假信息。将其写入本书，是为了告诉读者，语料库千奇百怪，对于几十亿篇文章，谁也没有能力保证每一个给出的语料都是真实的。用畅销网络文学作者的说法，这个“法诀”根本是忽略了我们所处的位面是一个无魔的科技位面呀！类似的问题应该是某些网络小说被人工智能阅读，成了人工智能语料库的一部分。这些非科学的内容其实是一种污染物。而一旦污染物进入语料库，就很难找到污染物和去除污染物。因此，人工智能组织的文章可信度是存疑的。

随着信息技术、人工智能技术的发展，网上的一些文章非常有可能就是人工智能组织的。批判性地阅读网络信息是十分必要。我们每天都被海量的信息所包围。其中，AIGC(如深度伪造视频、自动化新闻生成)让辨别信息真伪变得更加困难。如果缺乏批判性思维，我们可能会轻信虚假信息，甚至被误导。

对着人工智能学习是存在风险的。非常有必要寻找可信消息来源，对信息的真伪进行辨别。

批判性思维离不开逻辑推理能力和数据分析能力。逻辑推理是人类思维的一项核心能力，它帮助我们根据已有的信息得出合理的结论。这项能力是人工智能所不具备的。在人工智能时代，逻辑推理能力显得更加重要。许多人工智能算法，尤其是深度学习模型，被称为“黑箱”，因为它们的决策过程往往难以解释。作为人工智能的使用者，我们需要具备一定的逻辑推理能力，才能理解人工智能的输出结果是否合理。

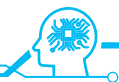
每一个人工智能的用户，实际上也都是人工智能数据库的“投喂者”。人工智能模型的性能高度依赖于数据质量。如果输入的数据存在噪声、不完整或偏差，那么即使是依靠最先进的算法也无法得出可靠的结论。因此，人类需要通过数据分析技术来清洗和优化数据，从而为人工智能提供更高质量的输入。现在的人工智能其实并不具备思考能力，无法真正“理解”数据背后的意义。

1980年，约翰·塞尔提出的思想实验——Chinese Room Argument 实验——被认为是图灵测试的反面。

中文屋实验描述如下：

(1) 假设有一个不会说中文的人被关在一个密闭的房间里。

(2) 房间里有一本详细的规则手册，这本手册用英语写成，描述了如何根据输入的中文字符组合，输出相应的中文字符组合。



(3) 外面的人向房间内递交一些用中文写的问题(输入),房间里的人按照规则手册的指示,将这些问题转换为适当的中文回答(输出),然后将回答递回给外面的人。

(4) 从外面观察,房间内的人似乎“懂得”中文,因为他能够给出符合语法和语义的答案。

塞尔指出,房间里的人实际上并不理解中文。他只是机械地按照规则手册操作,把输入转换成输出。换句话说,他并没有真正“理解”这些中文符号的意义。

目前的 AIGC 就停留在这个阶段。尽管 AIGC 表现得具有一定的理解能力,实际上它并不是真正“理解”。大多数大语言模型(包括聊天机器人和翻译工具)本质上仍然是在执行复杂的数学运算和模式匹配。它们并没有真正“理解”语言,而只是根据训练数据预测最可能的输出。即使人工智能表现得再像人类,也不一定意味着它具备真正的理解力。

既然没有理解力,人工智能输出的内容就不是那么可信的。我们必须以批判性的思维来看待人工智能带来的进步。



【拓展阅读 1-31】  
保持批判性思维  
观察 AI 输出