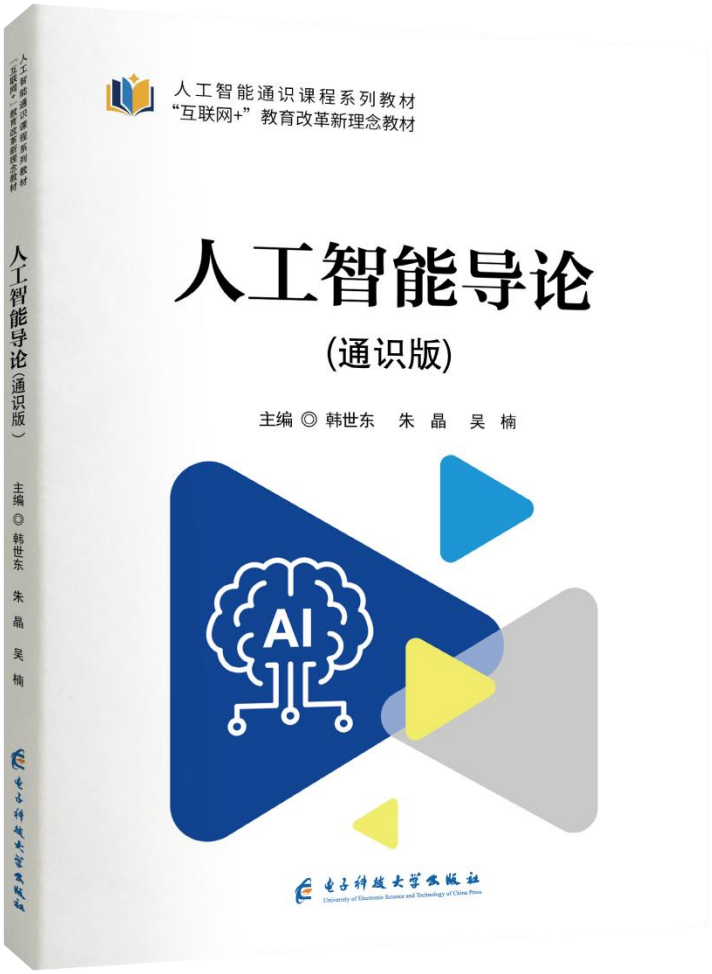
人工智能导论（通识版）



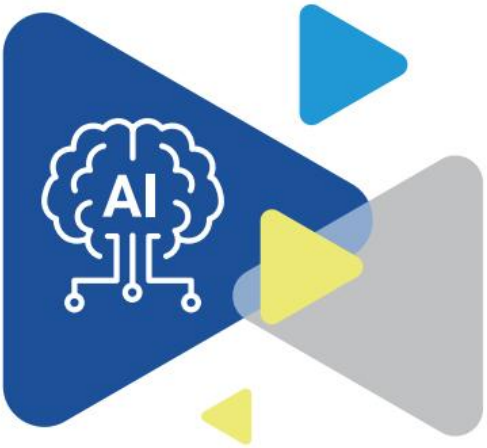
类目：计算机类
书名：人工智能导论（通识版）
主编：韩世东 朱 晶 吴 楠
出版社：电子科技大学出版社
开本：大 16 开
书号：978-7-5770-2409-7
使用层次：通用
出版时间：2026 年 6 月
定价：59.80 元
印刷方式：双色
是否有资源：有

人工智能通识课程系列教材
“互联网+”教育改革新理念教材

人工智能导论

(通识版)

主编◎韩世东　朱　晶　吴　楠



电子科技大学出版社
University of Electronic Science and Technology of China Press

人工智能通识课程系列教材
“互联网+”教育改革新理念教材

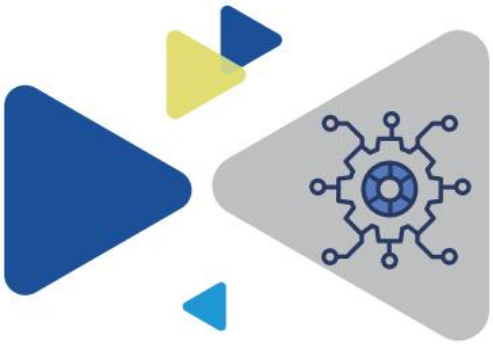
人工智能导论(通识版)

主编◎韩世东　朱　晶　吴　楠

电子科技大学出版社

人工智能导论

(通识版)



策划编辑：万晓桐
责任编辑：万晓桐
封面设计：鼎语书装

电子科技大学出版社

ISBN 978-7-1770-2409-7
定价：59.80元

人工智能通识课程系列教材
“互联网 +”教育改革新理念教材

人工智能导论

（通识版）

主　编 ◎ 韩世东　朱　晶　吴　楠
副主编 ◎ 马　军　黄　凯　尹光辉
吕尚征　陈超文　周　睿
何　毅　程洪丽
参　编 ◎ 郝又平



电子科技大学出版社
University of Electronic Science and Technology of China Press
·成都·

图书在版编目(CIP)数据

人工智能导论 ：通识版 / 韩世东，朱晶，吴楠主编.
成都 ：成都电子科大出版社，2026. 6. -- ISBN 978-7
-5770-2409-7

Ⅰ. TP18

中国国家版本馆 CIP 数据核字第 2026TQ1980 号

人工智能导论（通识版）

RENGONG ZHINENG DAOLUN (TONGSHI BAN)

韩世东　朱　晶　吴　楠　主编

策划编辑　万晓桐
责任编辑　万晓桐
责任校对　李雨纾
责任印制　梁　硕

出版发行　电子科技大学出版社
　　成都市一环路东一段 159 号电子信息产业大厦九楼　邮编　610051
主　　页　www. uestcp. com. cn
服务电话　028 - 83203399
邮购电话　028 - 83201495

印　　刷　三河市祥达印刷包装有限公司
成品尺寸　210 mm×285 mm
印　　张　18
字　　数　476 千字
版　　次　2026 年 6 月第 1 版
印　　次　2026 年 6 月第 1 次印刷
书　　号　ISBN 978-7-5770-2409-7
定　　价　59. 80 元

版权所有　侵权必究

前言 Preface

教育部于2024年启动了教育系统人工智能大模型示范应用行动，旨在打造人工智能通识课程体系，赋能理工医农文等各类人才培养。随着“人工智能＋”时代的来临，高校迫切需要构建一套完整的人工智能通识课程体系，为社会培养大量具有AI素养的优秀专业人才。

党的二十大报告提出：要实施科教兴国战略，强化现代化建设人才支撑。强调要深化教育领域综合改革，加强教材建设和管理。为了响应党中央的号召，深入贯彻全国教育大会和《教育强国建设规划纲要（2024—2035年）》精神，更好地构建适应社会需求的人工智能通识课程教育体系，我们在充分调研和论证的基础上，精心编写了这本《人工智能导论（通识版）》教材。

随着ChatGPT将人工智能技术呈现在人们的面前，人工智能技术展示了其颠覆性的力量，DeepSeek的异军突起更是引发了整个社会关注人工智能的热潮。身处人工智能时代，我们已经切身感受到了人工智能对人们思维模式和生活方式的改变，人工智能技术必将对人类的社会生产和生活产生重大而深远的影响。

与同类书相比，本书具备以下几个特点。

⁘循序渐进，由浅入深

本书为人工智能的入门教材，以人工智能的整个处理流程为主线，逐一介绍了人工智能所涉及的相关技术，帮助读者形成对人工智能知识体系及其应用领域的轮廓性认识，为读者在人工智能领域的继续深造学习奠定基础。

⁘说理透彻，注重实操

本书在重视基本理论的前提下，不忽视实际的可操作性，注重问题的解决。本书在每个项目均设有大量实训操作，以帮助读者巩固所学知识。

⁘技能与思政教育紧密结合

本书在讲解人工智能专业知识的同时，紧密结合思政教育主旋律，从专业知识角度触类旁通地引导学生提升相关的思政品质。

❖项目式教学，实操性强

本书采用项目式教学，把人工智能应用知识分解并融入一个个实践操作的训练项目中，以增强实用性。

由于编者能力有限，书中难免存在不妥之处，恳请读者提出宝贵意见，我们编写组人员不胜感激。

编　者

2026 年 2 月

CONTENTS 目录

项目一 初识庐山——人工智能概述 …… 1

项目导读 …… 1
知识目标 …… 1
技能目标 …… 1
素质目标 …… 2
1.1 人工智能概述 …… 2
1.2 人工智能的技术架构 …… 8
1.3 人工智能的三个阶段 …… 11
1.4 人工智能的行业应用场景 …… 14
1.5 人工智能的展望 …… 20
思考与练习 …… 23

项目二 深窥堂奥——人工智能中基本的方法 …… 25

项目导读 …… 25
知识目标 …… 25
技能目标 …… 25
素质目标 …… 26
2.1 搜索算法 …… 26
2.2 博弈 …… 32
2.3 逻辑 …… 35
2.4 产生式系统和专家系统 …… 38

2.5 进化计算 …… 41
思考与练习 …… 48

项目三 博采众长——机器学习与深度学习 …… 50

项目导读 …… 50
知识目标 …… 50
技能目标 …… 50
素质目标 …… 51
3.1 机器学习的基础 …… 51
3.2 监督学习 …… 56
3.3 无监督学习 …… 64
3.4 深度学习的基础 …… 70
3.5 深度学习的模型 …… 83
思考与练习 …… 90

项目四 明察秋毫——计算机视觉 …… 92

项目导读 …… 92
知识目标 …… 92
技能目标 …… 92
素质目标 …… 93
4.1 计算机视觉基础 …… 93
4.2 图像处理技术 …… 96
4.3 特征提取与描述 …… 102
4.4 目标检测与识别 …… 107
4.5 计算机的视觉应用 …… 111
思考与练习 …… 115

项目五 声声入耳——语音处理 …… 117

项目导读 …… 117
知识目标 …… 117
技能目标 …… 117
素质目标 …… 118
5.1 语音信号基础 …… 118

5.2　语音信号分析 …… 121
5.3　语音识别技术 …… 124
5.4　语音合成技术 …… 128
5.5　语音应用技术 …… 130
思考与练习 …… 134

项目六　知书达理——自然语言理解 …… 136

项目导读 …… 136
知识目标 …… 136
技能目标 …… 136
素质目标 …… 136
6.1　自然语言处理基础 …… 137
6.2　词法与句法分析 …… 139
6.3　语义理解 …… 144
6.4　文本生成与翻译 …… 147
6.5　情感分析与意图识别 …… 150
思考与练习 …… 156

项目七　身手不凡——智能机器人 …… 159

项目导读 …… 159
知识目标 …… 159
技能目标 …… 159
素质目标 …… 160
7.1　智能机器人基础 …… 160
7.2　机器人感知 …… 164
7.3　运动规划与控制 …… 169
7.4　人机交互 …… 175
7.5　智能机器人应用 …… 179
思考与练习 …… 187

项目八　无限创意——人工智能生成内容 …… 189

项目导读 …… 189
知识目标 …… 189

技能目标 …… 189
素质目标 …… 189
8.1 AIGC 概述 …… 190
8.2 图像生成技术 …… 193
8.3 文本生成技术 …… 213
8.4 音视频生成技术 …… 236
8.5 代码生成技术 …… 264
8.6 AIGC 的应用与思考 …… 274
思考与练习 …… 277

参考文献 …… 279

项目一

初识庐山——人工智能概述

项目导读

人工智能的出现，不仅是人类在智能科技领域的一次重大突破，更是开启了人类与机器协同发展的全新篇章。从简单的自动化任务到复杂的决策支持，从虚拟助手的贴心服务到自动驾驶汽车的智能导航，从医疗影像的精准诊断到艺术创作的无限想象，人工智能正以惊人的速度渗透我们生活的每一个角落，深刻改变着我们的工作方式、生活方式，甚至思维方式。

知识目标

1. 了解人工智能的定义和分类。
2. 了解我国人工智能的发展现状及大学生应具备的人工智能素养。
3. 掌握人工智能的技术架构及其主要组成部分。
4. 熟悉人工智能的三个阶段。
5. 认识人工智能在各行业的应用场景。
6. 认识人工智能伦理治理的核心问题与基本原则。
7. 展望人工智能的未来发展方向。

技能目标

1. 能够结合具体案例，分析不同 AI(artificial intelligence，AI)技术的特点及适用场景，初步判断技术选型的合理性。

2. 具备运用基础 AI 工具解决简单学习或实践问题的能力。例如，利用 AI 进行数据处理、文献梳理等。

3. 能够通过专业渠道跟踪 AI 技术前沿动态，初步分析技术发展对所学专业的影响。

4. 初步具备识别 AI 应用中伦理风险的能力，在实践中遵循数据安全与隐私保护规范。

素质目标

1. 激发对人工智能技术的探索热情，培养科学精神与创新意识。
2. 增强社会责任感与伦理意识，树立“科技向善”的技术价值观。
3. 感受我国在人工智能领域的发展成就，提升国家意识与民族自豪感。
4. 培养跨学科思维，促进计算机、数学、人文等多领域知识的融合应用。

1.1 人工智能概述

人工智能的应用在我们日常生活中无处不在，从智能语音助手到自动驾驶汽车，再到智能推荐系统，都在使用人工智能技术。随着计算能力的提升和算法的发展，人工智能正在逐渐改变各行各业的工作方式，推动着科技、经济和社会的发展。

1.1.1 人工智能的定义

人工智能是计算机科学的核心分支，旨在让机器模拟、延伸和扩展人类智能，实现自主感知、推理、学习、决策等智能行为。其研究范畴涵盖知识表示、逻辑推理、机器学习、计算机视觉、自然语言处理等多个领域，核心目标是让机器具备类人的思维能力，能够独立应对复杂问题并自主优化解决方案。

从本质上看，人工智能并非复制人类智能，而是通过技术手段实现“类人智能效果”。它可能以不同于人类思维的方式完成任务，甚至在特定领域表现出超越人类的效率与精度。自1956 年达特茅斯会议正式提出“人工智能” 概念以来，其理论体系不断完善，技术手段持续革新，应用边界也从最初的学术探索拓展到社会生产生活的各个角落。

实现人工智能的技术路径主要分为以下两种。

第一种是工程学方法(engineering approach)。这种方法采用传统编程技术，旨在让系统呈现智能行为，而不关注其实现方式是否与人类或生物体的智能机制相同。该方法的优势在于已经取得了诸多实际应用成果，如文字识别、计算机对弈等领域。然而，这种方法需要开发者完全依靠人工编写详细的程序逻辑，不仅编程工作量大、容易出错，而且在发现问题后必须修改源代码，重新进行编译、调试，最终通过版本更新或补丁来修复问题，整个流程相当烦琐。

第二种是模拟法(modeling approach)。与工程学方法不同，这种方法不仅追求智能效果，更强调实现机制要与人类或生物体的智能原理相似。开发者需要为每个智能体设计专门的智能模块。虽然这种方法要求开发者具备生物学思维，入门门槛较高，但一旦掌握就能实现更广泛的应用。由于不需要预先详细定义所有行为规则，在处理复杂问题时，这种方法往往比工程学方法更加高效。



当前，人工智能已展现出诸多显著特征，包括：自主决策能力、环境适应性、智能交互特性、海量数据处理、持续学习进化、实时响应机制、系统高度集成、精准模式识别、容错恢复能力和并行计算优势等。随着技术持续突破，人工智能的特性和能力边界还将不断扩展和深化。

1.1.2　人工智能的分类

人工智能可以根据其功能和工作原理分为多种不同的类别，这些分类有助于理解人工智能的不同应用场景和技术特点。

1. 按用途进行划分

(1)决策式人工智能

决策式人工智能专注于分析情况并做出决策。它通过评估多种选项和可能的结果，帮助用户或系统选择最佳的行动方案。例如，自动驾驶中的车道选择、车速调节等决策，金融领域的信贷风险评估，校园中的智能排课系统等，都是通过决策式 AI 实现的高效规划。

(2)生成式人工智能

生成式人工智能专注于创造全新内容。它可以根据学习到的数据自动生成文本、图像、音乐等内容。例如，将多篇专业论文输入生成式 AI，可快速生成囊括核心观点的文献综述；通过文字描述即可生成对应的绘画作品，或辅助编写课程报告框架等。

2. 按智能水平进行划分

(1)弱人工智能(weak AI)

弱人工智能也称为“狭义人工智能(narrow AI)”，是当前最主流的应用形态。其专注于单一领域或特定任务，仅具备处理该领域问题的能力，缺乏跨领域适应力和自主学习进化的能力。例如，手机语音助手(仅处理语音交互)、人脸识别门禁(仅完成身份核验)、论文查重系统(仅比对文本相似度)等，都属于弱人工智能。

(2)强人工智能(strong AI)

强人工智能也称为“广义人工智能(General AI)”，旨在具备“全面类人智能”，能够像人类一样在多个领域自主感知、理解、推理和决策，具备自我意识和情感感知能力，无须专门训练即可适应全新任务。例如，科幻作品中能够独立思考、自主交流的智能机器人，就属于强人工智能的范畴。目前，强人工智能仍处于理论研究阶段，其实现需要突破计算资源、算法逻辑、认知科学等多重瓶颈。

(3)超人工智能(superintelligent AI)

超人工智能指在所有领域的智能水平都超越人类的终极形态，能够通过自我学习持续进化，具备高度创造性和自主性，甚至能独立开展科学研究、推动技术革命。目前，这一概念仅存在于理论探讨中，其实现不仅面临技术挑战，还涉及伦理、安全等重大社会问题，是学界和业界共同关注的长期议题。

1.1.3　人工智能伦理治理

随着 AI 技术的飞速发展，AI 已经深刻改变了我们的社会、经济和生活方式。然而，AI

的广泛应用也带来了许多伦理和社会问题，如隐私泄露、算法歧视、责任归属等。这些问题需要我们进行深思熟虑的伦理治理，以确保人工智能在给人类带来利益的同时，能够遵循道德和法律的规范，避免对社会产生负面影响。

人工智能伦理治理是指在 AI 的开发、应用和监管过程中，确保其符合伦理原则、社会价值和法律规范的一系列政策和措施。伦理治理不仅包括对 AI 技术本身的监管，还涉及如何制定合理的政策、法规来引导 AI 技术的健康发展。

1. 人工智能伦理的核心问题

(1)隐私与数据保护

随着 AI 技术对大数据和个人信息的依赖，如何保护用户隐私成为一个重要问题。AI 在获取、存储和使用数据时，可能侵犯用户的个人隐私，导致敏感数据泄露。例如，智能语音助手、面部识别技术、健康监测设备等，都涉及大量个人数据的收集和处理。如果没有合理的保护措施，这些数据就可能被滥用，造成对用户隐私的侵害。

(2)算法偏见与歧视

AI 系统的决策往往依赖于大数据和训练算法，而这些数据和算法可能存在偏见。例如，在招聘、司法判决、信贷评分等领域，如果 AI 系统的训练数据存在性别、种族或社会阶层的偏见，那么 AI 做出的决策可能加剧社会的不公平。这种算法歧视可能导致某些群体受到不公平待遇，影响社会公正。

(3)自动化对就业的影响

随着自动化和 AI 技术的普及，许多传统的工作岗位被机器人或自动化系统取代。这种变化虽然提高了生产效率，但也可能导致大量失业，尤其是在低技能劳动密集型行业。如何平衡 AI 带来的经济效益与对劳动力市场的影响，成为社会亟待解决的问题。

(4)责任归属问题

当 AI 系统出现错误或造成损害时，责任应由谁承担？例如，自动驾驶汽车发生事故时，是否应该由 AI 开发者、汽车制造商、用户或其他相关方负责？这种责任归属问题在法律和道德上都尚未得到明确的规定，需要通过完善的法律框架来解决。

(5)AI 的自主性与控制问题

随着 AI 技术的不断发展，尤其是在深度学习和强化学习领域，机器可能会变得越来越“自主”。这引发了对 AI 失控的担忧。例如，如果一个 AI 系统在没有人工干预的情况下做出错误的决策，或者在未经授权的情况下执行任务，后果可能是无法预测和控制的。因此，如何保持 AI 系统的透明性、可控性和可解释性，避免 AI 技术对社会造成不可逆的伤害，是一个重要的伦理问题。

2. 人工智能伦理治理的核心原则

为了确保 AI 技术在发展过程中能够遵循伦理和法律规范，各国和各地区在人工智能伦理治理方面提出了若干核心原则。这些原则旨在为 AI 的研发和应用提供道德指导，帮助实现技术与社会责任的平衡。

(1)公平性与公正性

AI 系统应避免算法偏见，确保其决策和行为公正、不偏不倚，避免对特定群体的歧视。



在 AI 应用中应考虑到不同人群的需求，避免加剧社会不平等。例如，招聘 AI 应屏蔽性别、年龄等无关因素，聚焦能力评估。

(2)隐私保护与数据安全

保护个人隐私和数据安全是 AI 伦理治理的核心原则之一。在数据采集、存储和使用过程中，必须严格遵守隐私保护规定，确保数据的匿名化、加密处理，防止信息泄露或滥用。

(3)透明性与可解释性

AI 系统的决策过程应具有透明性，并能够向用户或相关方提供足够的信息，使得决策过程可解释。特别是在一些高风险领域(如医疗、司法等)，AI 决策的可解释性至关重要，要确保决策的可信度和公正性。

(4)责任与问责制

在 AI 应用中，必须明确责任归属，确保当 AI 系统出现问题时，能够追溯责任。无论是在自动驾驶、医疗诊断还是金融服务中，都应确保开发者、制造商、使用者和监管者的责任明确，以便在出现错误时进行有效的问责。

(5)人类中心主义

人类中心主义强调 AI 技术的应用应始终服务于人类的福祉。AI 应当补充和增强人类能力，而不是替代或威胁人类的尊严与生存。AI 系统的设计和使用应始终关注人类的利益，确保其应用有利于社会的可持续发展。

3. 保护个人隐私政策

为落实伦理治理原则，各国纷纷出台隐私保护法规，为 AI 发展划定“红线”。

①2018 年，欧盟颁布《通用数据保护条例》(GDPR)，对全球数据保护立法产生了深远影响。它赋予个人对其数据的广泛控制权，要求企业在处理欧盟公民的数据时必须遵守严格的法律框架，否则将面临高额罚款。

②2020 年，美国颁布《加州消费者隐私法案》(CCPA)，为加州居民提供了对其个人信息的更多控制权。CCPA 要求企业披露数据收集和销售情况，并赋予消费者拒绝出售其个人信息的权利，推动了美国其他地区的隐私立法进程。

③2021 年，我国颁布《中华人民共和国数据安全法》，强调数据在国家安全中的重要性。它规定了数据处理者的安全义务，包括建立数据分类分级保护制度、进行风险评估和报告数据安全事件等，以保障国家数据主权和个人信息安全。

④2021 年，我国《中华人民共和国个人信息保护法》正式实施，旨在规范个人信息处理活动，保护个人隐私权。该法律明确规定了收集、存储、使用和传输个人信息的合法条件和程序，对违规行为设定了严厉的处罚措施。

1.1.4　我国人工智能发展现状

在全球科技变革与社会经济结构调整的交汇点，我国已将人工智能定位为突破发展制约、引领产业结构优化的战略性工具。面对劳动力比较优势减弱、人口老龄化趋势加剧以及对高端专业人才需求日增的现实，发展人工智能及智能装备不仅是实现生产模式由“人”向“机”转变、促进产业升级的必要路径，亦为应对当前经济社会挑战、激发创新潜能提供了重要契机。

“人工智能＋X”正逐步演化为促进社会整体创新与创业活动的新范式及核心推动力。

1. 历史溯源与基础奠定

我国对人工智能的系统性探索可追溯至 20 世纪 50 年代末，以中国科学院计算技术研究所的成立为标志。作为国内最早的计算机科研机构之一，该所在推动我国早期计算机科技进步的同时，也为我国在人工智能领域的后续研究奠定了初步基础。此后，伴随计算机硬件与软件技术的快速迭代，我国逐步积累了坚实的技术储备，为人工智能领域的持续突破与发展提供了有力支撑。

至 20 世纪 80 年代，我国开始规划性地发展人工智能技术，并在语音识别、机器翻译等特定领域取得了初步进展。此阶段，我国的计算机科学家与工程师积极探索人工智能的理论框架与核心技术，开展了一系列具有开创性的研究工作，如提出了基于模型的机器翻译方法及神经网络方法等概念。国防科技工业委员会于 1984 年主持召开了全国智能计算机及其系统学术讨论会，1985 年召开了全国首届第五代智能计算机学术研讨会。自 1986 年起，智能计算机系统、智能机器人和智能信息处理等关键项目被正式纳入国家高技术研究发展计划(“863 计划”)。1987 年 7 月，《人工智能及其应用》一书由清华大学出版社出版，成为国内首部拥有自主知识产权的人工智能系统性专著。1993 年开始，智能控制与智能自动化等项目亦被纳入国家科技攀登计划。

2. 从数字化向智能化的演进

在 2024 年之前，尽管“人工智能”这一术语已存在，相关产业(如半导体、芯片、算力、光刻设备及移动终端等)亦有相应发展，但社会整体的关注焦点更多地集中于“数字化”转型与“数字经济”的构建。“信息化”“互联网＋”是彼时更为通行的概念。相较而言，“智能化”因其蕴含的趋近甚至可能超越人类认知与执行能力的特性，在当时的语境下仍被视为一种更为高端和前沿的技术形态。此阶段，我国在人工智能领域的研究虽持续进行，但总体上呈现出对国际(特别是美国)发展趋势的跟随态势，如对 OpenAI 的 ChatGPT 及大语言模型的关注。

3. 我国人工智能的加速发展

2024 年成为我国人工智能发展历程中的一个关键转捩点。随着“新质生产力”概念的提出以及内生性科技进步的驱动，我国人工智能领域开始显现出强劲的自主发展动能，国产化芯片与软件产品的相继推出是其突出表征。至 2024 年年底，人工智能相关产业在国内呈现快速发展态势。国际观察甚至将 2024 年视为“人工智能元年”的重要延续，部分原因在于我国在年内密集推出了自主研发的大模型，如早期的豆包、百度的文心一言、阿里巴巴的通义千问等。值得注意的是，早期涌现的部分国产大模型后被证实存在对底层开源模型的“封装”现象，真正具备完全自主芯片架构与核心算法的模型数量有限，其中，豆包、腾讯的混元大模型、阿里巴巴的通义千问及万相是较为典型的代表。然而，这一系列动态无疑标志着我国开始在人工智能领域加速追赶，力求与国际先进水平接轨。

4. 迈向国际前沿

进入 2025 年，DeepSeek 大模型的出现，在特定层面进一步缩短了中美在尖端人工智能技术上的差距。DeepSeek 等国产大模型的广泛关注，极大地激发了全社会对人工智能应用的



探索热情与创新活力。此后，我国在人形机器人、脑机接口、智能体(agent)、具身智能等前沿方向亦相继取得了阶段性突破。

5. 核心成就与政策导向

①当前，我国在人工智能领域取得的成就主要体现在以下三个层面。

a. 大模型技术：尽管与国际顶尖水平仍存在差距，但已具备一定的可比性，并处于持续快速迭代的进程中。

b. 核心硬件与软件：在芯片、光刻技术、算法等关键环节取得了一定进展，逐步形成了部分自主研发与生产能力。

c. 标志性技术突破：以 DeepSeek 等为代表的高水平模型，体现了我国在人工智能领域研发能力与技术水平的提升。

②国内学者在问题求解、不确定性原理、泛逻辑理论、拓扑学在人工智能中的应用、模式识别(如文字、语音、指纹、人脸、虹膜及步态识别等)、图像处理、机器学习、专家系统、智能计算及智能控制等多个分支领域均有重要建树，取得了一批具有国际先进水平的原创性成果。这些研究成果广泛应用于生物医学、卫星遥感、机器人视觉、货物检测、目标跟踪、自主导航、安全防卫、金融服务、交通管理、国防军事、电子商务及多媒体网络通信等关键行业与领域。

③国家持续强化顶层设计与战略支持。2026 年 3 月 5 日，国务院总理李强在第十四届全国人民代表大会第四次会议上所作的《政府工作报告》中明确提出，建立未来产业投入增长机制，培育生物制造、量子科技、具身智能、6G 等未来产业。《政府工作报告》亦强调，持续推进“人工智能＋”行动，将数字技术与制造优势、市场优势更好结合起来，支持大模型广泛应用，并大力发展智能网联新能源汽车、人工智能手机和电脑、智能机器人等新一代智能终端以及智能制造装备。

6. 挑战与展望

我国人工智能产业正处于高速发展阶段，政策支持持续加强，应用场景日趋丰富。然而，核心技术瓶颈的持续攻坚、能源消耗的优化管理以及数据安全与伦理规范的构建等问题，仍是未来发展中亟待解决的关键挑战。我国人工智能正以坚定的步伐和前所未有的发展态势，向更高技术水平和更广泛应用领域迈进。其深度融合与创新应用，必将为国家经济社会的高质量发展注入强大而持久的智能动力。

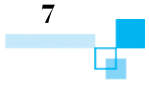
1.1.5　大学生应具备的人工智能素养

在当今人工智能快速发展的时代，大学生作为未来的科技从业者、创新者和社会建设者，需要具备“知识－能力－价值－伦理”四位一体的 AI 素养，才能适应时代发展需求。

1. 体系化知识

认知是智能的重要表现，体系化知识是其基石。大学生需要对人工智能进行整体性理解和系统化分类，从而更清晰地认识到人工智能的内涵、边界和外延。

人工智能具有广泛的交叉渗透特点，涉及计算机科学、数学、逻辑学、心理学等多个学



科。例如，机器学习算法需要掌握数学中的概率论和线性代数等知识，而自然语言处理则需要掌握语言学和心理学原理等知识。通过掌握体系化知识，学生可以更好地理解这些学科之间的联系，为后续的学习和应用打下坚实基础。

2. 构建式能力

构建式能力是指大学生在掌握体系化知识的基础上，能够自主构建知识体系，形成自己的思维框架和解决问题的能力。这种能力的培养需要学生具备较强的逻辑思维能力和创新能力。

在人工智能领域，构建式能力尤为重要。因为人工智能技术的发展日新月异，只有具备自主构建知识体系的能力，才能跟上时代的步伐，不断更新自己的知识和技能。

3. 创造性价值

创造性价值是指大学生在掌握体系化知识和构建式能力的基础上，能够运用所学知识解决实际问题，创造出有价值的成果。这种价值的创造需要学生具备创新意识和实践能力。

在人工智能领域，创造性价值体现在能够运用人工智能技术解决实际问题，提高工作效率，降低成本，甚至改变人们的生活方式。因此，培养学生的创造性价值对于推动人工智能技术的发展具有重要意义。

4. 人本型伦理

人本型伦理是指在人工智能的发展和应用过程中，始终以人为本，尊重人的尊严和权利，保障人的安全和利益。这种伦理观要求我们在发展人工智能的同时，关注其可能带来的社会影响和潜在风险。

在人工智能领域，人本型伦理尤为重要。因为人工智能技术的发展可能会对就业、隐私、安全等方面产生影响，甚至可能导致道德和法律问题。因此，培养学生的人本型伦理观，使他们在未来的工作中能够遵循伦理原则，为社会的和谐发展作出贡献。

1.2 人工智能的技术架构

人工智能的技术架构通常由基础层、技术层和应用层构成，如图 1-1 所示。每一层都有其特定的功能和任务，共同构成了人工智能系统的整体架构。

1.2.1 基础层

基础层是人工智能技术架构的最底层，它为 AI 的学习、决策和执行提供了坚实的支持。这个层次主要涉及数据的收集与处理、平台系统的支持和硬件设施的提供。基础层的作用是为技术层和应用层提供必要的资源和工具，确保 AI 系统能够高效、稳定地运行。

1. 数据

数据是人工智能的核心资源之一，AI 系统的学习和决策过程通常依赖于大量的高质量数



图 1-1　人工智能的技术构架

据。数据可以分为以下几种类型：

①图像数据：用于计算机视觉任务，如物体识别、人脸识别、图像分类等。图像数据通常以像素的形式呈现，可以通过图像处理和深度学习算法进行分析。

②视频数据：用于视频分析、行为识别、视频监控等任务。视频数据是一系列连续的图像，通过对视频数据的帧间变化进行分析，AI 能够理解动态场景中的信息。

③语音数据：语音数据用于语音识别、语音生成、语音翻译等任务。AI 通过将语音信号转化为文本或指令进行处理，为语音助手、智能客服等应用提供支持。

④文字数据：文字数据广泛应用于自然语言处理(Natural Languag Procossin，NLP)领域，包括文本分类、情感分析、机器翻译、问答系统等。文本数据可以从书籍、文章、网页等方面获取，经过处理后帮助 AI 理解和生成语言。

⑤结构化数据：这类数据通常以表格或数据库的形式存在，适用于传统的数据分析、预测建模等任务。结构化数据通常用于监督学习模型的训练，通过标注数据进行学习和决策。

这些数据为人工智能提供了学习的基础和决策的依据，通过数据的分析与处理，AI 能够识别模式、做出预测，并完成复杂的任务。

2. 平台系统

平台系统为基础层提供了数据的存储、处理和计算资源，确保人工智能能够高效地运行。以下是基础层中几个重要的平台系统。

①云计算平台：云计算为人工智能提供了强大的计算资源支持。AI 应用需要大量的计算能力，尤其是深度学习模型的训练，通常需要高性能的计算资源。通过云计算平台，AI 可以动态获取计算资源，按需进行大规模数据处理和存储。常见的云计算平台有 Amazon Web Services(AWS)、Google CloudPlatform(GCP)、Microsoft Azure、阿里云、华为云等。

②云操作系统：云操作系统为云计算平台提供操作系统级别的管理和控制功能。它负责管理云平台上的虚拟化资源、计算资源、存储资源以及网络连接。通过云操作系统，人工智

能应用可以在云端灵活部署，确保数据处理的高效性和稳定性。

③大数据系统：大数据系统处理和分析海量的、复杂的非结构化或半结构化数据，为人工智能提供丰富的训练数据和实时分析能力。通过大数据系统，AI 可以快速地处理来自不同来源的数据，进行数据挖掘和知识发现。常见的大数据处理框架有 Hadoop、Spark 等。

④计算资源调度系统：计算资源调度系统负责高效地分配和管理计算资源，确保 AI 任务在高效、合理的计算环境下进行。尤其在进行深度学习训练时，计算资源调度系统能够根据任务的优先级和需求动态调整资源分配，从而提高系统整体的运行效率。

3. 硬件设施

硬件设施为人工智能提供了必要的物理基础。它包括处理器、存储设备、传感器等，确保数据能够高效地传输和计算。以下是基础层中常见的硬件设施。

①芯片：AI 芯片是专为加速 AI 运算而设计的硬件。它们优化了 AI 计算中的特定任务，如深度学习中的矩阵计算。

②云计算服务器：云计算服务器提供了大量的计算资源和存储空间，支持大规模数据的处理和模型训练。通过云计算服务器，AI 能够利用大量的计算力进行数据分析和训练，而无须依赖本地硬件。

③传感器：传感器用于感知环境并收集数据，是 AI 系统获取外部信息的“眼睛”和“耳朵”。常见的传感器包括图像传感器(摄像头)、激光雷达、温度传感器、压力传感器等，这些传感器提供的数据可以被 AI 系统用于实时决策和控制。

④GPU/CPU 和 RAM：GPU 和 CPU 是计算机系统的核心处理单元，GPU 具有强大的并行处理能力，适合处理深度学习和大数据分析任务；CPU 则适合处理一般的计算任务。RAM(随机存取内存)则用于存储数据和程序，为 AI 应用提供快速的存取能力。

这些硬件设施共同构成了人工智能基础层的物理基础，它们为数据处理、模型训练和推理提供了强有力的支持。

1.2.2 技术层

技术层是人工智能系统的核心部分，涉及 AI 算法、技术和底层框架的实现。该层主要包括通用技术、算法和底层框架，它们共同为人工智能系统提供强大的智能功能。通过这些技术和算法，AI 能够处理各种复杂的任务，如语音识别、图像处理、决策制定等。

1. 通用技术

通用技术是人工智能实现智能功能的基础能力，它们使得 AI 能够处理各种类型的数据并进行复杂的决策，如自然语言处理、计算机视觉和智能语音等。

2. 算法部分

人工智能的核心能力来自各类算法。不同的 AI 任务通常需要不同的算法来进行处理和解决。

算法部分包括深度学习、人工神经网络、强化学习、集成学习、回归模型、决策树模型、支持向量机和聚类分析等，这些算法是实现人工智能功能的关键。



3. 底层框架

为了实现上述算法，人工智能需要依赖底层框架提供支持。底层框架为 AI 开发人员提供了算法实现平台、高效的计算工具和数据处理能力。常见的人工智能底层框架包括以下几种。

(1)TensorFlow

TensorFlow 是一个由 Google 开发的开源机器学习框架，广泛适用于深度学习和神经网络的研究与应用。它支持多种平台和设备，能够高效处理大规模数据和复杂计算，特别适合进行深度学习模型的训练和推理。

(2)PyTorch

PyTorch 是一个由 Facebook 开发的开源深度学习框架，因其灵活性和易用性受到广泛欢迎。它支持动态图计算，方便进行实验和调试。PyTorch 在学术界和工业界都有广泛应用，特别是在计算机视觉和自然语言处理等领域。

(3)Caffe

Caffe 是一个深度学习框架，特别适用于计算机视觉领域。它的设计注重速度和模块化，能够快速处理图像和视频数据。Caffe 常用于图像分类、目标检测、语义分割等任务。

(4)MXNet

MXNet 是一个由 Apache 开发的深度学习框架，支持多种语言(如 Python、Scala 等)，并提供了高效的分布式计算能力。MXNet 特别适用于大规模的深度学习模型训练和推理等。

1.2.3　应用层

应用层是人工智能技术架构中的最上层，它是 AI 技术在实际行业和领域中的具体实现。应用层将技术层的算法和模型转化为可操作的解决方案，广泛地应用于各个行业和领域，包括智能工业、智能交通、智慧城市、智能医疗、自动驾驶、智能农业、智能政务、智慧财务和智能零售。通过人工智能的感知、学习和决策能力，应用层解决了许多现实世界中的问题，提高了工作效率，推动了各行业的智能化发展。

1.3　人工智能的三个阶段



人工智能的发展历程是一个不断演进的过程，从最初的简单计算到如今的复杂认知，经历了多个重要阶段。通常，我们将人工智能的发展划分为三个阶段："计算智能"阶段、"感知智能"阶段和"认知智能"阶段。每个阶段都有其独特的特点和技术突破，推动着人工智能技术不断向前发展。

1.3.1　人工智能的"计算智能"阶段

计算智能(computational intelligence，CI)是人工智能发展的初期阶段，侧重于模拟和实现计算机系统通过数据处理、模式识别和计算推理等方法来解决实际问题的能力。这个阶段

的人工智能系统主要依赖于大量的数据输入和算法模型，以模仿人类的计算思维过程，但它们并不具备复杂的认知或自我意识。

“计算智能”阶段的人工智能通常包括诸如规则系统、进化算法、模糊逻辑、人工神经网络等技术。这些技术基于逻辑推理、经验积累和启发式搜索，帮助计算机系统做出决策和预测。这一阶段的人工智能虽然无法具备类似人类的感知、推理和创造能力，但它通过有效地分析、处理大量数据，在特定任务上表现出了显著的优势。

计算智能的核心理念在于模拟人类的计算和逻辑推理能力，并通过算法让机器能够执行从简单到复杂的任务。其关注点不仅仅是通过规则和公式解决问题，还包括如何在没有明确规则的情况下，通过学习和优化来提升任务执行效率。

虽然计算智能在多个领域取得了显著成绩，但它仍然存在一定的局限性。首先，计算智能系统通常依赖于大量的先验知识和经验，在没有充分数据支持的情况下，系统的推理能力可能受到限制。其次，尽管进化算法、模糊逻辑和神经网络等技术可以模拟某些复杂的决策过程，但这些系统缺乏灵活性和适应能力，难以应对不断变化的环境和未知的挑战。计算智能阶段的系统通常是依赖于数据训练或规则编写的固定模型，难以具备自我意识和深度推理能力。

随着人工智能技术的不断发展，计算智能也在向更高层次的智能系统进化。未来，计算智能将可能融入更多自我学习和自我优化的能力。例如，深度学习和强化学习等新型算法的出现，使得计算智能系统不仅能够从数据中进行学习，还能够通过不断的试错和优化提升系统的性能。此外，随着大数据、云计算和物联网等技术的发展，计算智能将进一步拓宽应用领域，成为智能城市、智慧医疗、智能制造等领域的核心支撑。

1.3.2 人工智能的“感知智能”阶段

感知智能(perceptual intelligence)是人工智能发展的第二个阶段，主要集中于模拟人类感知能力的实现，旨在通过计算机系统使机器能够“感知”外部环境。这一阶段的人工智能系统侧重于从外界获取信息，并对这些信息进行处理和理解，从而做出响应。

感知智能的关键技术包括计算机视觉、语音识别、传感器技术等。这些技术使得人工智能能够像人类一样通过视觉、听觉等感官感知世界。这一阶段的人工智能突破了计算智能的局限，不再仅仅依赖于规则和算法进行推理，而是通过对环境的感知进行自我调整和决策，从而推动人工智能向更为智能化的方向发展。

在“感知智能”阶段，人工智能的任务不再仅仅是对已有知识的处理和计算，而是通过数据输入(如图像、声音、传感器数据等)来理解和分析周围的世界。通过模拟人类的感知过程，感知智能不仅可以识别环境中的物体和情境，还能根据这些信息做出相应的行为决策。

尽管感知智能取得了显著进展，但仍面临一些挑战。首先，感知智能系统在处理复杂、动态的环境时仍存在局限性。例如，自动驾驶系统可能在复杂的天气条件下表现不佳(如雨雪天气中的物体识别)。其次，感知智能的准确性往往受到传感器质量、数据质量和训练数据质量的限制。例如，语音识别系统可能受到噪声环境的影响，导致识别准确度下降。

未来，感知智能将朝着更加智能、更加自主的方向发展。随着深度学习、传感器技术和



数据处理能力的不断进步，感知智能将更加精准地理解和响应周围环境的变化。特别是在自动驾驶、医疗影像分析、智能家居等领域，感知智能将发挥越来越重要的作用。随着人工智能技术的不断突破，感知智能将更加融入我们的日常生活，并进一步改变人类与机器的交互方式。

自动驾驶汽车是感知智能阶段的一个重要应用案例。首先，它通过多种传感器，如摄像头、激光雷达、毫米波雷达等，感知周围的交通环境，包括道路状况、车辆位置、行人等信息。然后，它通过深度学习算法对这些信息进行处理和分析，从而实现自动驾驶功能。例如，华为问界的自动驾驶系统能够自动识别交通标志、车道线，并根据路况自动调整车速和行驶方向。自动驾驶汽车的发展不仅展示了感知智能阶段的强大能力，也为未来智能交通的发展提供了新的思路。

1.3.3　人工智能的“认知智能”阶段

认知智能(cognitive intelligence)是人工智能发展的第三个阶段，旨在模拟和实现人类的高级认知能力，如理解、推理、学习和决策。与“计算智能”阶段和“感知智能”阶段侧重于任务的执行和环境的感知不同，“认知智能”则更多关注如何让人工智能系统具备类似人类的思维能力，能够在复杂和不确定的环境中做出合理的判断和决策。

认知智能的核心目标是让人工智能能够理解并处理复杂的信息，从而执行更高级的任务，如计划、推理、创造性思维和自我优化。在这一阶段，人工智能不再只是依赖规则和算法来解决问题，而是通过推理、学习、理解上下文以及解决具有高度不确定性的问题来进行智能决策。这一阶段的技术突破使得人工智能从单纯的执行者转变为更加智能、具有自我调节能力的系统。

认知智能的研究不仅仅依赖于技术上的创新，还涉及认知科学、心理学、哲学等多个学科。通过借鉴人类大脑的工作原理，认知智能试图为机器赋予更为复杂的思维能力，使其能够在动态的环境中灵活应对。

智能教育系统是“认知智能”阶段的一个重要应用案例。它通过自然语言处理技术理解学生的问题和需求，结合教育心理学和教学法，为学生提供个性化的学习方案。例如，智能辅导系统可以根据学生的学习进度和知识掌握情况，自动调整教学的内容和难度，提供有针对性的练习和讲解。同时，通过情感分析技术，系统还能够感知学生的情绪状态，及时给予鼓励和支持，提高学生的学习效果和积极性。

1. 认知智能的挑战与未来

(1)推理与决策的复杂性

尽管认知智能系统在处理复杂问题时取得了显著进展，但其推理和决策过程仍然面临挑战。复杂的现实世界包含大量的动态变化和不确定因素，认知智能系统需要在有限的信息和不断变化的环境中做出最佳决策。这要求系统具备更强的推理能力和更高的灵活性。

(2)知识表示的局限性

尽管知识表示技术取得了长足进展，但如何有效地表示和处理世界知识仍然是认知智能的一个难题。复杂的知识体系和隐性知识的表示、推理仍然面临技术和理论的挑战。

(3)伦理与社会影响

认知智能的普及可能引发伦理和社会问题。例如，如何确保认知智能系统做出的决策符合人类价值观？如何处理人工智能在医疗、司法等领域的决策透明性和责任归属问题？这些问题需要在发展认知智能的同时加以关注和解决。

2. 未来展望

随着技术的不断进步，认知智能将逐步进入更加智能和自主的阶段。未来的认知智能系统将能够在更广泛的领域内执行复杂的任务，并且通过不断学习和自我优化，提高解决问题的效率。随着人工智能在各行各业中的渗透，认知智能将成为智能决策、智慧城市、自动化服务等领域的核心驱动力。

1.4 人工智能的行业应用场景

AI 作为一种具有广泛适应性的技术，已经深入渗透到各个行业并推动了多个领域的创新与发展。从智能安防到智慧城市，AI 的应用场景不断扩展，并不断提升各个行业的效率、精度和智能化水平。

1.4.1 智能安防

智能安防是 AI 应用得最成熟的领域之一，核心是通过 AI 技术提升安全防护的效率和准确性，广泛应用于校园、社区、企业等场景。

1. 视频监控与人脸识别

视频监控系统是智能安防的核心组成部分，随着人工智能技术的发展，传统的视频监控设备逐渐向智能化方向转型。通过集成计算机视觉与深度学习算法，智能视频监控能够实现对实时视频流的自动分析，自动识别和跟踪异常事件或人物。例如，在企业办公大楼的入口处安装人脸识别门禁系统，员工无须携带门禁卡，只需刷脸即可进入，同时也能防止未授权人员进入，如图 1-2 所示。



图 1-2　人脸识别门禁系统

2. 智能警报与预警系统

传统的警报系统依赖人工触发或预设触发条件，然而，智能警报系统通过集成传感器、视频监控和 AI 分析，能够实现实时自动监控并根据情境智能响应。当发生入侵、火灾、异常噪声等事件时，AI 系统能够通过数据分析即时触发警报，减少误报并提高响应速度。同时，智能预警系统还可以根据大数据和机器学习模型对潜在威胁进行预测，提前做好防范措施。

1.4.2　智能驾驶

智能驾驶是 AI 与交通领域的深度融合，核心是让车辆具备自主感知、决策和控制能力，提升驾驶的安全性和道路通行的效率。

1. 自动驾驶技术与车辆感知技术

自动驾驶技术使得汽车能够在无须人工干预的情况下自主完成驾驶任务。该技术主要依赖于 AI 算法，通过集成传感器(如摄像头、雷达、激光雷达等)来感知周围环境并做出决策。自动驾驶系统能够识别道路标志、行人、车辆等交通元素(如图 1-3 所示)，通过深度学习和计算机视觉技术优化路径规划和避障操作，从而实现安全、智能的驾驶。例如，华为问界搭载的 ADS 2.0 及以上高阶智驾系统，可以在不依赖高精地图的情况下，依靠多传感器融合与端到端大模型，自主完成城市道路的跟车、自动变道、障碍物避让、红绿灯启停及自动泊车/远程召唤等。

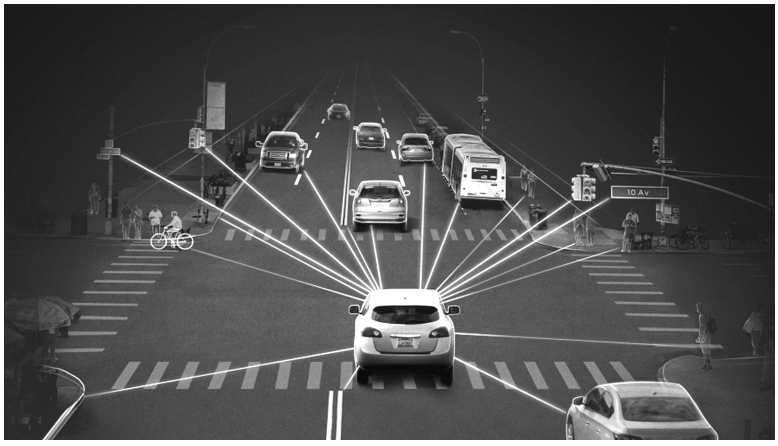


图 1-3　识别交通元素

2. 路况预测与智能导航

智能驾驶不仅仅是车辆的自动化控制，还涉及交通流量的预测与导航优化。AI 技术通过分析实时交通数据、天气状况、历史交通流量等信息，能够为驾驶员提供实时路况预测，帮助驾驶员避免拥堵，选择最佳行驶路线。智能导航系统能够实时更新地图信息，预测潜在的交通问题，并在行驶过程中提供准确的引导，确保驾驶的安全与高效，如图 1-4 所示。

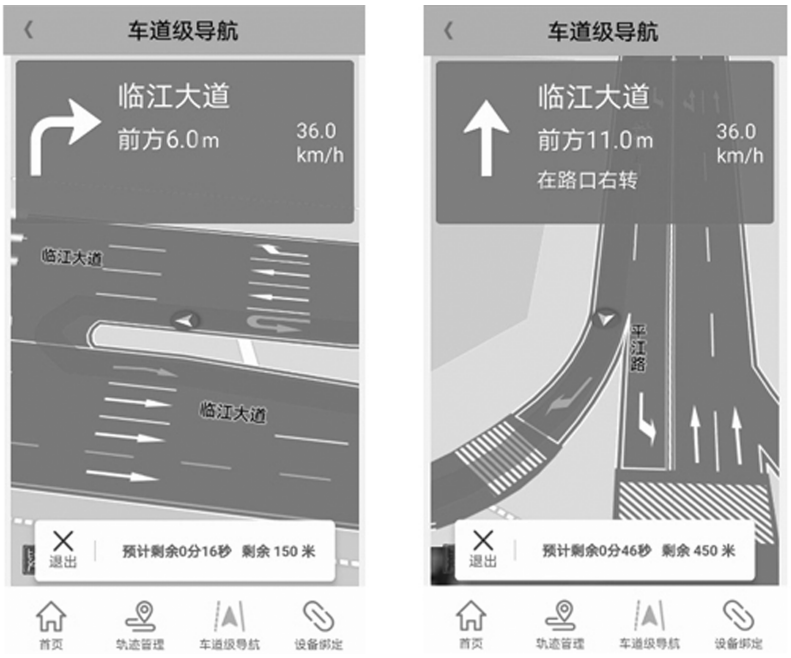


图 1-4　智能导航

1.4.3　智能医疗

智能医疗是 AI 与医疗领域的结合，核心是提升诊断准确性、优化医疗服务流程，缓解医疗资源紧张的问题。

1. AI 在医学诊断中的应用

人工智能在医学领域的应用已经改变了传统的诊断方式。通过深度学习、图像识别、自然语言处理等技术，AI 能够辅助医生诊断疾病，尤其是在医学影像分析中。AI 系统能够高效处理 X 光片、CT 扫描、MRI 等影像数据，快速识别潜在的疾病，如肿瘤、肺结核、骨折等。通过机器学习模型，AI 还能帮助医生从病历数据中提取关键特征，提供个性化的治疗建议。

2. 智能化医疗设备与远程医疗相结合

智能医疗设备通过集成 AI 算法，能够实时监测和分析患者的健康数据。例如，智能手表和可穿戴设备能够跟踪心率、血氧、血压等生理数据(如图 1-5 所示)，AI 算法则分析这些数据并提供健康报告或警示。远程医疗通过视频会议、AI 诊断等手段实现医患之间的远程互动，为偏远地区或行动不便的患者提供便捷的医疗服务。



图 1-5　智能手表跟踪血压

1.4.4　智能服务

智能服务涵盖生活、学习、工作中的各类服务场景，核

心是通过 AI 技术提升服务效率、优化用户体验。

1. 智能客服与个性化推荐相融合

智能客服系统通过自然语言处理(NLP)技术，能够理解用户的查询意图并提供相应的答案。与传统人工客服相比，智能客服不仅可以高效处理大量客户的需求，还能够提供 24h 不间断服务。通过对用户行为的分析，AI 可以根据用户的购物历史或浏览记录提供个性化推荐，如推荐相关产品或服务。

2. 服务自动化与机器人应用

服务自动化利用 AI 技术取代人工服务，提升效率并减少错误。例如，在餐厅、酒店等服务行业中，智能机器人能够进行自助点餐、服务员引导、餐品配送(如图 1-6 所示)等工作。AI 驱动的机器人还能够通过语音识别与自然语言处理技术与客户进行互动，提升客户体验，同时降低人工成本。



图 1-6　智能机器人餐品配送

1.4.5　智能家居

智能家居是 AI 在家庭场景的应用，核心是实现家居设备的自动化控制和智能交互，提升居住的便捷性和舒适性。

1. 家庭自动化与设备控制

智能家居系统通过 AI 技术实现家居设备的自动化管理与控制。智能家居设备，如智能灯泡、智能空调、智能音响等，通过手机、语音助手等接口进行远程控制，为用户提供更加便捷和舒适的居住环境。AI 技术使得这些设备能够根据家居的环境变化或用户习惯进行自我调节，如自动调节室内温度、光照强度，提升居住舒适度，如图 1-7 所示。



图 1-7　智能家居系统设备控制

2. 智能家居的安全性与舒适性

智能家居不仅能够提供便捷的自动化控制，还能够提升家庭的安全性。智能监控摄像头通过集成人脸识别和运动检测技术，可以识别家庭成员或陌生人，实时监控家庭安全。结合智能警报系统，家居安全系统能够在有异常发生时自动报警。此外，智能家居还通过环境监测技术，实时调整室内的空气质量、温湿度等，确保舒适的居住环境。

1.4.6 智慧农业

智慧农业是 AI 与农业的结合，核心是实现精准种植、高效管理，提升农业的产量和质量，减少资源浪费。

1. AI 助力农业生产与管理

AI 技术在农业生产中的应用正在改变传统农业的管理方式。通过大数据分析、机器学习和传感器技术，AI 可以实时监测作物的生长状况、土壤质量和天气变化，从而提供精准的农业管理方案。AI 还能够通过分析农田数据来预测农作物的生长趋势，帮助农民进行科学规划，提高产量和质量。

2. 精准农业与智能灌溉

精准农业利用 AI 与物联网技术，实现对农业生产的精确管理。通过环境监测设备和传感器，AI 能够分析土壤湿度、温度等因素，优化灌溉过程，确保水资源的高效利用，减少浪费。此外，智能灌溉系统能够根据天气预报和土壤湿度数据自动调节灌溉量，保证作物生长所需的最佳水分供应，如图 1-8 所示。



图 1-8 智能灌溉系统

1.4.7 智慧交通

智慧交通是 AI 在交通领域的综合应用，核心是优化交通管理、缓解拥堵、提升通行效率。

1. 交通流量预测与管理

智慧交通系统通过大数据与 AI 算法对交通流量进行分析与预测，从而优化交通管理。例如，通过实时收集和分析交通摄像头、传感器和 GPS 数据，AI 能够预测未来的交通流量，提前采取调节措施，如通过调整交通信号灯、引导车辆绕行等来缓解交通拥堵，提高道路通行的效率，如图 1-9 所示。

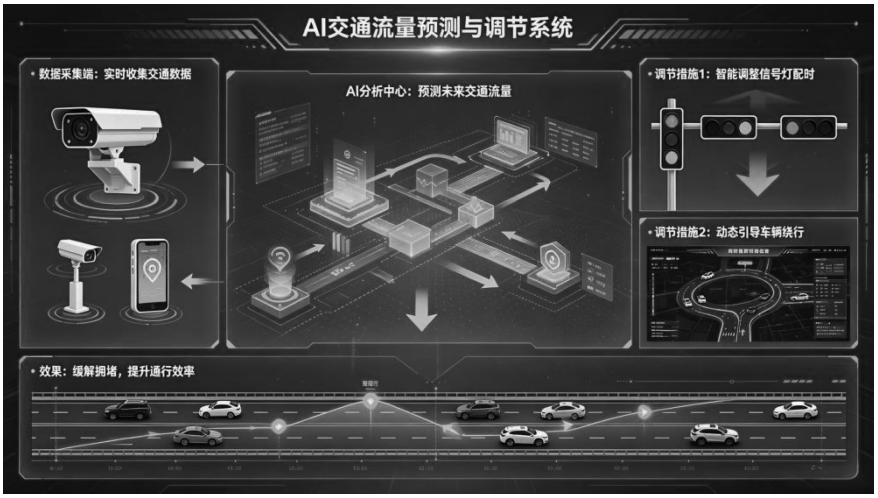


图 1-9　交通流量测量

2. 智能停车与无人配送

智能停车系统通过 AI 技术帮助驾驶员快速找到空闲车位，减少因寻找停车位而造成的拥堵，如图 1-10 所示。在无人配送领域，AI 驱动的自动驾驶车辆能够通过优化算法选择最优路径，进行高效配送。无人配送系统通过集成传感器、图像识别和 AI 决策，使得配送过程变得自动、安全、高效。



图 1-10　智能停车

1.4.8 智慧城市

智慧城市是 AI 在城市管理中的综合应用，核心是整合城市资源、优化管理流程，以提升城市运行效率和居民生活质量。

1. 城市管理与智能化服务

智慧城市通过整合 AI、大数据和物联网技术，提升城市的管理效率与服务质量。通过智能城市管理系统，AI 能够优化能源分配、垃圾处理、交通控制等各类城市服务。例如，智能路灯系统可以根据交通流量和环境亮度自动调节光照强度，节省能源；智能垃圾桶系统能够监控垃圾填充情况，自动通知清理。

2. 环境监测与资源优化

智慧城市的环境监测系统通过 AI 技术实时监测空气质量、水质、噪声等环境因素，并根据分析结果提供优化建议。例如，智能环境监测系统能够在空气污染指数达到警戒线时自动调节城市公共交通和工业排放，减少污染源。此外，AI 还能够在资源管理(如水、电等)方面进行优化，提高资源的使用效率，降低浪费。

1.5 人工智能的展望

随着人工智能技术的快速发展，它不仅推动了多个行业的创新，也为社会的智能化转型提供了源源不断的动力。未来，人工智能将变得更加易于使用，更加安全、高效，同时广泛应用于各个领域。

1.5.1 更易用的人工智能开发框架

未来的 AI 开发将打破“专业壁垒”，让非技术背景的用户也能轻松使用 AI 工具，实现人人皆可使用 AI。

1. 无须深入编程的 AI 开发工具

随着人工智能技术的普及，越来越多的开发者和非技术人员都希望能够使用 AI 来解决实际问题。因此，开发者不再需要掌握复杂的编程知识或深入了解机器学习的数学原理。一些易于使用的 AI 开发工具正在逐步普及，这些工具通过可视化界面和简化的操作流程，使得普通用户也能进行 AI 应用开发。例如，Google 的 AutoML、Microsoft 的 Azure AI，以及 IBM 的 Watson Studio 等平台提供了自动化机器学习工具，用户只需上传数据，选择模型类型，平台就会自动完成模型的训练和优化。这种“无代码”的 AI 开发方式，使得更多领域的人能够参与 AI 的应用，推动 AI 的普及和商业化。

2. 开放源代码与社区支持

随着开源技术的蓬勃发展，越来越多的人工智能框架和工具逐渐开放源代码，开发者可



以自由地访问、修改和分享代码。这种开放源代码的方式不仅提高了技术的透明度，还促进了全球开发者社区的协作和创新。例如，TensorFlow、PyTorch、Keras、魔搭社区、open I 启智社区等开源深度学习框架，为全球研究人员和开发者提供了强大的支持，并推动了 AI 技术的不断进步。

此外，强大的社区支持为 AI 开发者提供了丰富的资源，如教程、文档、在线讨论和问题解答等，降低了学习和应用 AI 技术的门槛。这种开放与共享的模式，进一步加速了 AI 技术的推广与普及。

1.5.2　更完善的人工智能数据服务

数据是人工智能的基础，而数据的质量和可获得性将直接影响到人工智能技术的表现。未来的 AI 数据服务将在“数据质量、数据安全、数据共享”三个方面实现突破，为 AI 发展提供更坚实的支撑。

1. 数据收集与共享的创新方法

数据是人工智能发展的核心资源。随着物联网(IoT)、智能设备和各种传感器的普及，未来将产生更多的数据，这为 AI 的发展提供了宝贵的原材料。然而，如何有效地收集、整合和共享这些海量数据，仍然是一个重大挑战。

未来，数据收集与共享将通过更加智能和高效的方式实现。例如，基于区块链技术的数据共享平台能够确保数据的透明性和可追溯性，避免数据的滥用与篡改。同时，数据提供者和使用者之间的隐私保护与利益共享也将得到更好的平衡。此外，智能化的数据清洗和预处理技术将进一步提高数据的质量和可用性，为 AI 模型提供更加准确和全面的数据支持。

2. 数据质量与数据安全保障

数据质量直接影响人工智能模型的性能和准确性。未来的 AI 系统将更加注重数据的质量控制，通过多层次的数据清洗、标注、验证和标准化来确保数据的高质量。同时，随着数据量的不断增长，如何确保数据的安全性和隐私保护也变得尤为重要。

随着 GDPR(通用数据保护条例)等隐私保护法规的实施，AI 数据服务将需要满足更严格的合规性要求。未来，数据加密、访问控制、匿名化等技术将在数据共享和使用过程中得到更广泛的应用，以保障个人和企业的数据安全。

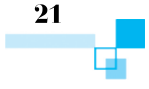
1.5.3　更安全的人工智能数据共享

随着人工智能技术的不断发展，数据的隐私和安全性问题也日益突出。未来将通过技术创新和制度建设，实现“数据可用不可见”，既发挥数据价值，又保护个人和企业隐私。

1. 数据隐私与加密技术

在数据驱动的人工智能系统中，数据隐私问题日益突出。个人数据、企业敏感信息以及政府数据的保护成为 AI 发展中的一项挑战。因此，数据隐私保护和加密技术将成为未来 AI 发展的重要方向。

未来，AI 技术将通过采用高效的加密方法，如同态加密、差分隐私等技术，确保数据在



使用过程中的安全性和隐私性。例如，同态加密技术使得数据在加密状态下进行处理，避免了数据泄露的风险。而差分隐私技术则允许数据在不暴露个体信息的情况下进行统计分析，从而保护用户的隐私。

2. 合规性与监管挑战

随着 AI 应用的扩大，数据使用的合规性和监管问题将成为发展的瓶颈。不同国家和地区对数据隐私和安全有不同的法规，如何在全球范围内既实现数据的安全共享，又符合当地的法律规定，将是一个巨大的挑战。

为了应对这一挑战，AI 企业和政府需要加强合作，制定全球统一的合规标准与政策，确保数据使用的透明性和公平性。同时，AI 技术的发展也需要考虑伦理问题，确保人工智能的使用不对社会公平性和个人自由造成负面影响。

1.5.4 更优化的人工智能算法模型

人工智能算法模型是 AI 发展的核心，未来的人工智能将越来越依赖于更优化的算法模型。算法模型的优化将包括提升计算效率、降低算法能耗、提高模型准确性和泛化能力等方面。

1. 高效地进行模型训练与优化

随着人工智能的普及，越来越多的 AI 应用需要处理大量的数据和复杂的算法模型。在这一过程中，如何提高 AI 模型的训练效率和优化性能将成为研究的重点。

未来，AI 算法的训练将通过更多的创新技术得到提升。例如，迁移学习技术允许模型将一个领域的学习成果迁移到另一个领域，减少数据需求和训练时间。此外，量子计算的引入可能会大幅提高 AI 模型的计算能力，使得超大规模的数据集和复杂模型得以在短时间内完成训练。

2. 泛化能力与可解释性

AI 模型的泛化能力和可解释性将是未来发展的关键。泛化能力指的是模型在处理未见过的数据时，仍能保持较好的预测能力；而可解释性则指的是模型能够清晰地说明其决策过程和依据，避免“黑箱”问题。

为了提升 AI 模型的可解释性，未来的 AI 系统将更加注重开发可解释的算法和工具。例如，利用模型可视化、决策树生成等方法，使得人工智能的决策过程透明化，从而提高其在医疗、金融等关键领域的可接受性和信任度。

1.5.5 “端-边-云”全面发展的人工智能算力体系

1. 分布式计算与智能终端的结合

人工智能计算的需求不断提升，而传统的集中式云计算可能无法满足低延迟、高响应等要求。因此，“端-边-云”计算架构应运而生，成为人工智能发展的新趋势。

①端(edge)：是指智能终端设备，如智能手机、可穿戴设备等，它们通过集成 AI 芯片进行本地数据处理。



②边(edge)：是指边缘计算节点，如路由器、服务器等，在本地进行数据预处理和计算。

③云(cloud)：云端则负责处理更为复杂的大规模计算任务，同时提供强大的计算资源和存储能力。

通过“端-边-云”三者的结合，AI 可以更高效地处理数据，同时保证低延迟和高计算能力。

2. 边缘计算与云计算的协同

边缘计算和云计算的协同工作将成为未来 AI 算力发展的重要趋势。边缘计算将负责实时数据的采集和初步处理，而云计算则承担大规模数据分析、深度学习模型训练等计算任务。这种分工合作可以大大提高 AI 应用的响应速度，减少延迟，并提高资源使用效率。

例如，自动驾驶汽车通过边缘计算实时处理来自传感器的数据，而云计算则提供实时更新的地图和交通信息，帮助车辆进行决策。这样的协同模式提高了系统的效率和可靠性。

1.5.6　更广泛的人工智能服务

1. AI 服务的普及与商业化应用

人工智能的普及和商业化应用将进一步推动其在各行业的渗透。从智能客服到个性化推荐，从智能家居到工业自动化，AI 服务将在未来成为各行各业的基础设施。随着技术的成熟，AI 服务将更加普及，成为各企业提升效率和竞争力的重要手段。

例如，越来越多的企业已经将人工智能集成到客户服务中，利用自然语言处理技术提供 24/7 不间断的智能客服服务。与此同时，电商平台也通过 AI 算法分析用户行为，提供个性化推荐，提升用户购物体验和平台销量。

2. 行业内不同领域的应用扩展

未来，人工智能的应用将在各个行业进一步扩展。在医疗领域，AI 将帮助医生进行精准诊断，提升治疗效率；在金融领域，AI 将用于风险控制和投资预测；在教育领域，AI 将通过个性化学习方案帮助学生提升学习效果。此外，AI 还将在农业、能源、物流等行业发挥重要作用，推动各行业的智能化转型。



一、选择题

1. 人工智能的定义是（　　）。

A. 让机器完全替代人类进行工作

B. 使机器能够模拟和执行人类智能行为

C. 仅用于处理大数据的技术

D. 一种用于自动化生产的工具

2. 以下不属于人工智能分类的是(　　)。

A. 决策式人工智能　　B. 生成式人工智能

C. 感知式人工智能　　D. 弱人工智能

3. 以下不属于人工智能基础层内容的是(　　)。

A. 数据收集与处理　　B. 云计算平台

C. 深度学习算法　　D. 硬件设施

4. 人工智能的“感知智能”阶段主要侧重于(　　)。

A. 模拟人类的计算能力　　B. 模拟人类的感知能力

C. 模拟人类的决策能力　　D. 模拟人类的创造力

5. 以下不属于人工智能应用层的领域是(　　)。

A. 智能医疗　　B. 智能驾驶

C. 智能客服　　D. 智能硬件开发

6. 以下不属于人工智能技术层内容的是(　　)。

A. 自然语言处理　　B. 计算机视觉

C. 云计算服务器　　D. 深度学习框架

二、填空题

1. 人工智能的定义是计算机科学的一个分支，旨在使机器能够_____和执行人类智能行为。

2. 在人工智能的分类中，_____人工智能专注于特定任务或领域，而_____人工智能旨在模仿人类的全面智能。

3. 人工智能伦理治理的核心原则包括公平性与公正性、_____与数据安全、透明性与可解释性等。

4. 人工智能的技术架构通常由基础层、_____和应用层构成。

5. 人工智能的“认知智能”阶段旨在模拟人类的_____能力，如理解、推理、学习和决策等。

三、简答题

1. 简述人工智能的“感知智能”阶段的特点。

2. 列举人工智能在智能交通领域的两个应用场景，并简要说明其作用。

3. 简述大学生应具备的人工智能素养。

项目二 ▶▶▷

深窥堂奥——人工智能中基本的方法

项目导读

人工智能不仅仅局限于理论模型，它的许多经典算法和系统已经在实践中得到广泛应用。无论是在搜索问题的解决，还是在博弈策略的选择，甚至在复杂的逻辑推理和知识表示方面，人工智能技术都展现出了强大的潜力。

知识目标

1. 理解搜索算法的基本原理和应用，包括智能系统中的搜索、盲目搜索、知情搜索以及博弈中的搜索。

2. 掌握博弈论的基本概念和策略，包括二人博弈、囚徒困境和高级计算机博弈。

3. 学习逻辑在人工智能中的应用，包括逻辑和表示、模糊逻辑。

4. 了解产生式系统和专家系统的工作原理，包括产生式系统的构建和专家系统的开发。

5. 理解人工智能技术解决实际问题的方式，特别是在自动化、决策支持和知识推理等领域的应用。

6. 掌握进化计算的基本原理和方法：包括模拟退火算法、遗传算法和遗传规划。

技能目标

1. 能够应用基本的人工智能算法，如广度优先搜索(BFS)、深度优先搜索(DFS)、A*算法等。

2. 能够分析并实现简单的博弈算法，如极大极小算法和 Alpha-Beta 剪枝算法，应用于模拟博弈决策。

3. 能够设计和实现简单的产生式系统和专家系统。

4. 能够使用模拟退火、遗传算法等进化计算方法解决实际优化问题。

素质目标

1. 培养战略思维与决策意识，学会在不确定环境中权衡利弊、选择最优方案。
2. 提升逻辑推理与系统思维能力，能够将复杂问题拆解为可解决的步骤。
3. 强化实践创新与知识应用能力，敢于用 AI 方法解决本专业的实际问题。
4. 培养严谨的科学态度，在算法实现与优化中注重逻辑闭环与结果验证。

2.1 搜索算法

搜索算法是人工智能领域中的核心技术之一，它广泛应用于各种智能系统中，包括机器人导航、游戏智能、优化问题等。搜索算法的目标是在状态空间中寻找一个满足目标条件的路径或解。不同的搜索方法适用于不同的问题，选择合适的算法是解决问题的关键。

2.1.1 智能系统中的搜索

搜索是人工智能系统中至关重要的一部分，尤其是在解决问题时，智能系统需要找到最优的路径或解。搜索算法适用于在给定的状态空间中寻找从初始状态到目标状态的路径。智能系统中的搜索不仅涉及简单的查找任务，还包括决策、规划和博弈等应用。

例如，在自动驾驶系统中，搜索算法适用于规划从起点到目的地的路径，避开障碍物，确保安全且高效的行驶；在游戏 AI 中，搜索算法帮助游戏代理(如电脑对手)做出决策，在有限的时间内选择最优策略；在机器人的任务规划中，搜索算法适用于找到解决特定任务的最佳路径，如清理房间的路径规划。

搜索算法的精度和效率直接影响智能系统的性能。为了提高效率，许多智能系统会采用启发式信息，以便更快速地找到最优解。

2.1.2 盲目搜索

盲目搜索是最基础的搜索方法之一，在这种搜索方法中，算法没有任何关于目标的额外信息，完全依赖于遍历状态空间，逐步寻找解决方案。可以理解为“盲目”地在一个大迷宫中寻找出口，不依赖任何指示，只是按照一定规则一步一步前进。

盲目搜索的核心思想是从初始状态开始的，按照某种顺序逐步探索所有可能的状态，直到找到目标状态。它不利用目标信息，因此搜索的效率较低，但这种方法在某些情况下仍然是有效的。

盲目搜索方法的关键是搜索顺序的选择。常见的盲目搜索方法包括广度优先搜索(breadth-fist search，BFS)和深度优先搜索(depth-fist search，DFS)。

1. 广度优先搜索

广度优先搜索并不考虑结果的可能位置，而是系统地展开并检查图中的所有节点，以找寻结果。广度优先搜索不使用经验法则算法。在广度优先搜索过程中，所有因为展开节点而得到的子节点都会被加进一个先进先出的队列中。广度优先搜索从初始节点开始，应用运算符生成第一层节点，检查目标节点是否在这些节点的后继节点中，若没有，再用产生式规则将所有的第一层节点逐一扩展，得到第二层节点，并逐一检查第二层节点是否包含目标节点。若没有，再用产生式规则逐一扩展第二层所有节点，重复上述操作，直到发现目标节点为止，遍历顺序如图 2-1 所示。广度优先搜索保证找到从起点到目标的最短路径，适用于图中的无权图。例如，假设你在一个城市的地图上寻找从 A 点到 B 点的最短路径，BFS 会逐步探索从 A 点出发的所有可能路径，直到找到最短路径。

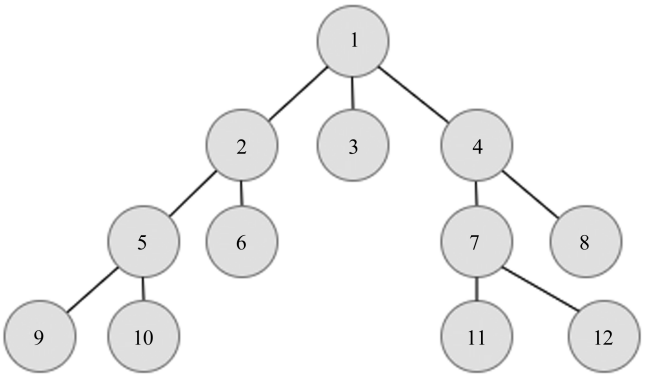


图 2-1　广度优先搜索遍历顺序

2. 深度优先搜索

深度优先搜索算法是把最近刚产生的节点优先扩展，直到达到一定的深度限制。若未找到目标或无法再扩展时，再回溯到另一个节点继续扩展。从初始状态开始，利用规则生成搜索树下一层任一个节点，检查是否出现目标状态，若未出现，以此状态利用规则生成再下一层任一个节点，检查是否出现目标状态，若未出现，则继续重复上面的操作过程，一直进行到叶节点(即不能再生成新状态节点)。当没有发现目标状态时，回溯到上一层结果，取另一可能扩展搜索的分支，生成新状态节点。若仍不是目标状态，就按该分支一直扩展到叶节点；若仍没有发现目标状态，采用相同的回溯方法回退到上层节点，扩展可能的分支生成新状态。如此进行下去，直到找到目标状态为止，遍历顺序如图 2-2 所示。例如，假设你在寻找一个迷宫的出口，DFS 可能会沿某条路径走到底，直到无法继续，再回溯尝试其他路径。

广度优先搜索算法通常需要存储产生的所有节点，占用的存储空间较大。所以当搜索树的节点较多，用其他方法易产生内存溢出时，通常使用深度优先搜索。深度优先搜索不会保留全部节点，而是将扩展完的节点从数据库中弹出并删去，因此通常在数据库中存储的节点数就是深度值，因此深度优先搜索占用空间较少。

盲目搜索方法的优点是简单、容易实现，并且在某些特定问题中能有效找到解。然而，它们的缺点在于效率低下，尤其是在状态空间非常大的情况下，搜索可能会非常慢，甚至需要遍历整个状态空间。

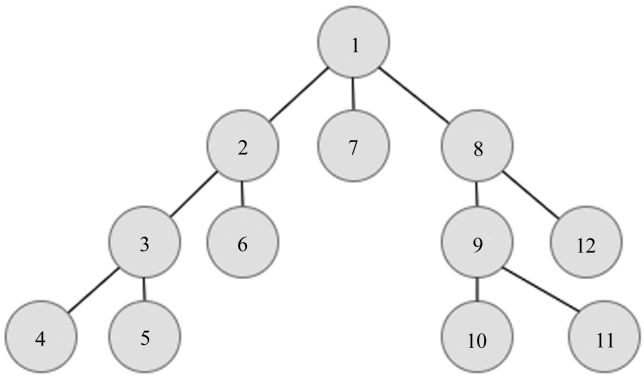


图 2-2　深度优先搜索的遍历顺序

虽然盲目搜索算法不如知情搜索高效，但在某些问题中，盲目搜索仍然具有其独特的应用价值。例如，当问题的规模较小，状态空间不复杂时，盲目搜索可能是最简单且有效的解决方案。在没有启发信息的情况下，盲目搜索提供了一个简单且直接的求解方式。

2.1.3　知情搜索

知情搜索(informed search)是人工智能中一种通过利用额外的信息或启发式函数来优化搜索过程的技术。与传统的盲目搜索(如深度优先搜索和广度优先搜索)不同，知情搜索通过引入启发式信息来帮助算法选择最有可能达到目标的路径，从而提高搜索效率。

1. 启发式函数

知情搜索的关键在于引入启发式函数，启发式函数(heuristic function)用于评估当前状态到目标状态的估计成本或距离，从而引导搜索过程更有效地进行。启发式搜索通常比盲目搜索更高效，尤其是在处理大规模或复杂问题时。

启发式函数[$h(n)$]是一个评估函数，它给出了从当前状态 n 到目标状态的估计成本。在知情搜索中，启发式函数的作用是帮助算法判断哪条路径可能是最优的。启发式函数的质量直接影响搜索效率和搜索结果的最优性。常见的启发式函数有下面几种。

①曼哈顿距离(manhattan distance)：在网格状的空间中，计算水平和垂直方向上两个点之间的距离。

②欧几里得距离(euclidean distance)：计算两个点之间的直线距离，适用于连续空间。

③不在位置的块数(misplaced tiles)：在拼图问题中，计算当前状态下每个块与目标状态中对应块的错位数量。

2. 知情搜索的常见算法

(1)A * 算法(A-star)

A * 算法是一种非常经典的知情搜索算法。它综合了 BFS 的全面性和 DFS 的高效性，利用启发式函数计算当前状态到目标的估计成本，并选择估计成本最小的路径进行扩展。在设计 A * 算法时，启发式函数的选择至关重要。启发式函数能够根据当前状态到目标状态的距离进行合理估计，从而引导搜索过程，减少搜索空间，提高搜索效率。启发式函数的质量直接影响 A * 算法的搜索效率和最终解的质量。



例如，在地图导航中，A * 算法不仅要考虑当前路径的代价，还要考虑到从当前位置到目的地的估算代价，从而选择最短的路径。

A * 算法的基本公式如下：

$$f(n)=g(n)+h(n)$$

其中，

$f(n)$：表示从初始状态到当前状态 n 的预估总代价(即选择的评估标准)。

$g(n)$：表示从初始状态到当前状态 n 的实际代价。

$h(n)$：表示从当前状态 n 到目标状态的启发式估算代价(即启发式函数)。

A * 算法的步骤如下。

①初始化。

• 创建一个开放列表，存储所有待扩展的节点。

• 创建一个封闭列表，存储已经扩展的节点。

• 将起点加入开放列表，并计算起点的 $f(n)$ 值。

②搜索过程。

• 从开放列表中选择 $f(n)$值最小的节点进行扩展，扩展的节点为当前节点的邻居节点。

• 对每个邻居节点，计算其 $g(n)$和 $h(n)$ 值，并更新其 $f(n)$。

• 如果邻居节点没有在开放列表中，可将其加入开放列表；如果邻居节点已存在于开放列表中，则检查是否需要更新其 $f(n)$值。

• 将当前节点移至封闭列表。

③终止条件。

• 如果开放列表为空，表示没有找到路径。

• 如果从开放列表中选择的节点为目标节点，则结束搜索，路径找到。

④路径回溯：从目标节点回溯，找出最优路径。

(2)贪心算法(greedy search)

贪心算法每次都选择当前最优的行动，而不考虑后续可能的影响。它适用于需要快速找到近似最优解的情况。

贪心算法的基本步骤如下：

①选择步骤：在每一步中，贪心算法都选择当前最优的解决方案，即做出局部最优的选择。

②可行性检查：确保当前选择是可行的，不会导致违反问题的约束。

③更新步骤：将当前的选择结果添加到已做的部分解中，更新剩余的子问题。

④终止条件：在满足终止条件时，返回最终的解决方案。

活动选择问题是贪心算法的经典例子之一。给定多个活动，每个活动都有开始和结束时间，目的是选择尽可能多的活动，使得它们之间没有时间冲突。其步骤如下：

①排序：按照活动的结束时间进行排序。

②选择活动：从第一个活动开始，选择结束时间最早的活动，如果该活动与已选择的活动不冲突，则将其加入解中。

③继续选择：重复上一步，直到无法选择更多的活动。

3. 知情搜索的应用场景

知情搜索算法比盲目搜索更高效，能够在较短的时间内找到最优解或近似最优解。在大规模复杂问题中，知情搜索尤其重要，因为它减少了搜索的范围，避免了不必要的状态扩展。

知情搜索算法因其高效性，被广泛应用于各种实际问题中。常见的应用场景包括以下几种。

①路径规划：在地图上寻找从起点到终点的最短路径。例如，Google 地图中的路线规划就是使用了知情搜索算法。

②游戏 AI：在棋类游戏、战斗游戏等中，AI 需要通过知情搜索算法计算最佳策略，A * 算法常用于游戏中的路径搜索。

③机器人导航：自主机器人在环境中导航时，常常需要根据环境特征和目标位置使用启发式搜索算法找到最佳行进路线。

2.1.4 博弈中的搜索

博弈是人工智能中的一个重要领域，它涉及多个智能体(如人类、计算机或机器人等)在面对冲突或合作时所做出的决策。博弈中的搜索问题与其他类型的搜索不同，因为它不仅要考虑如何找到目标状态，还需要预测其他参与者的行为，以及他们的可能反应。在这种情况下，搜索算法要能够处理复杂的动态决策过程。

博弈问题通常是指两个或更多参与者在有限的资源、规则和目标下进行决策的过程。每个参与者的决策会影响到其他参与者的利益和策略，因此博弈问题的求解不仅涉及单一决策者的选择，还需要预测和反应其他决策者的行为。

例如，在下棋游戏中，每一步棋的选择不仅依赖于当前棋盘状态，还必须考虑对方棋手的可能走法。在这种情况下，搜索不仅要找到自己有利的走法，还要考虑对方的回应和反应。

常见的博弈搜索算法有极大极小算法和 Alpha-Beta 剪枝。

1. 极大极小算法

极大极小算法通过递归分析所有可能的游戏状态，选择一个最有利于自己的策略。每个玩家假设对方也会选择最优策略，然后倒推选择最佳行动。其基本思想包括下面几部分。

①先设定博弈的双方中一方为 A，另一方为 B；然后为其中的一方(如 A)寻找一个最优行动方案。

②为了找到当前的最优行动方案，需要对各种可能的方案所产生的后果进行比较。具体来说，就是要考虑每一方案实施后对方可能采取的所有行动，并计算可能的得分。

③为计算得分，需要根据问题的特性信息定义一个估价函数，用来估算当前博弈树末端节点的得分。此时估算出来的得分称为静态估值。

④当端节点的估值计算出来后，再推算出父节点的得分，推算的方法是：对于“或”节点，选其子节点中一个最大的得分作为父节点的得分，这是为了使自己在可供选择的方案中选择一个对自己最有利的方案；对于“与”节点，选其子节点中一个最小的得分作为父节点的得分，这是为了立足于最坏的情况。这样计算出的父节点的得分称为倒推值。

⑤如果一个行动方案能获得较大的倒推值，则它就是当前最好的行动方案。

极大极小算法示意图如图 2-3 所示。

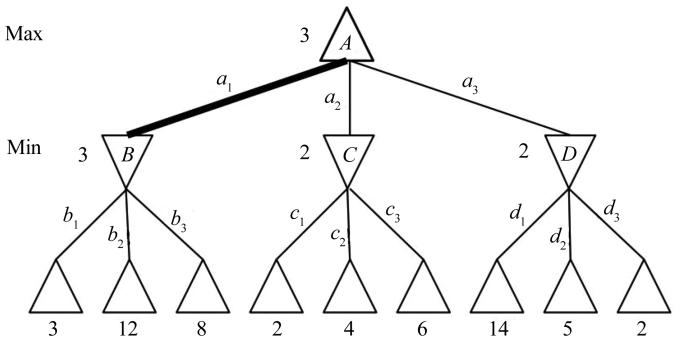


图 2-3　极大极小算法示意图

在博弈问题中，每一种格局可供选择的行动方案都有很多，因此会生成十分庞大的博弈树。据统计，西洋跳棋的博弈树非常庞大，试图利用完整的博弈树来进行极小极大分析是非常困难的。可行的办法是只生成一定深度的博弈树，然后进行极小极大分析，找出当前最好的行动方案。在此之后，再在已选定的分支上扩展一定深度，最终选择最好的行动方案。如此进行下去，直到取得胜败的结果为止。至于每次生成博弈树的深度，当然是越大越好，但由于受到计算机存储空间的限制，只能根据实际情况而定。

2. α-β 剪枝

α-β 剪枝是对极大极小算法的优化。它通过提前剪除掉一些不可能影响最终结果的分支，从而提高搜索效率，具体的剪枝方法包括下面几种。

①对于一个与节点 Min，若能估计出其倒推值的上确界 β，并且这个 β 值不大于 Min 的父节点(一定是或节点)的估计倒推值的下确界 α，即 $\alpha \geqslant \beta$，则就不必再扩展该 Min 节点的其余子节点了(因为这些节点的估值对 Min 父节点的倒推值已无任何影响了)。这一过程称为 a 剪枝。

②对于一个或节点 Max，若能估计出其倒推值的下确界 α，并且这个 α 值不小于 Max 的父节点(一定是与节点)的估计倒推值的上确界 β，即 $\alpha \geqslant \beta$，则就不必再扩展该 Max 节点的其余子节点了(因为这些节点的估值对 Max 父节点的倒推值已无任何影响了)。这一过程称为 β 剪枝。

在计算博弈树时，Alpha-Beta 剪枝可以提前终止一些无效的分支，避免在博弈树中浪费时间进行无意义的计算。通过这种优化，α-β 剪枝在大型博弈问题中能够有效提高搜索效率，使得计算机能够在有限的时间内做出较为精准的决策。例如，在围棋或国际象棋的计算机对弈中，α-β 剪枝能够显著减少需要探索的棋局数量，从而在更短的时间内做出决策。

博弈中的搜索不仅应用于棋类游戏(如国际象棋、围棋等)，还广泛应用于机器人对抗、自动化决策等领域。

2.2 博　　弈

博弈是多智能体之间的“互动决策”，每个参与者的选择都会影响他人的收益，而每个人都在追求自身利益最大化。博弈论不仅是 AI 的重要应用领域，更是理解人类社会互动、商业竞争、团队协作的核心工具。

2.2.1　二人博弈

二人博弈是博弈论中最基础的类型之一，通常涉及两个理性参与者，在一个规则明确的环境下，相互竞争或合作，做出决策以优化自身的利益。

1. 二人博弈的基本概念

在二人博弈中，两个参与者通过相互选择策略来实现自己的目标。每个参与者在做出决策时，必须考虑到对方的可能反应。因此，二人博弈不仅要预测对方的行为，还需要根据对方的选择调整自己的策略。

二人博弈包括下面几种主要特征。

①两个参与者：博弈中有两个决策者，每个决策者都在根据自己的目标选择策略。

②策略集合：每个参与者有一组可供选择的策略。策略的选择会影响最终的博弈结果。

③收益函数：每个参与者都有自己的收益函数，用来评估每种策略组合下的结果。收益函数根据参与者的策略和对方的策略来确定最终结果。

④博弈的结果：博弈的结果是基于两位玩家的策略选择，通常以支付矩阵的形式表示，其中显示了每种策略组合下参与者的收益。

2. 支付矩阵

支付矩阵(payoff matrix)是二人博弈中常用的一种工具，它帮助我们清晰地展示不同策略组合下参与者的收益。通过支付矩阵，玩家可以看到在每一种情况下，自己的收益是多少，同时也能看到对方选择某个策略时自己可能得到的结果。

以下是一个典型的二人博弈的支付矩阵示例，见表 2-1 所列。

表 2-1　二人博弈的支付矩阵

玩家 B/玩家 A	合作 (C)	背叛 (B)
合作 (C)	(3，3)	(0，5)
背叛 (B)	(5，0)	(1，1)

在这个支付矩阵中，玩家 A 和玩家 B 都有两个选择：合作(C)或者背叛(B)。每个单元格中的数字表示在对应策略组合下，玩家 A 和玩家 B 的收益。例如，如果两位玩家都选择合作，那么玩家 A 和玩家 B 的收益分别为 3。如果玩家 A 选择合作而玩家 B 选择背叛，那么玩