

人工智能通识入门微课 - Pilot

Neo Lee

<2025-05-14 Wed>

目录

1	AI 是什么?	2
1.1	什么是智能 <i>Intelligence</i>	2
1.1.1	思考与练习	3
1.2	什么是人工智能 <i>Artificial Intelligence</i>	3
1.2.1	思考与练习	4
1.3	弱 AI 与强 AI	4
1.3.1	思考与练习	5

1. 什么是 AI? AI 与传统数字化工具的区别;
2. 阿里“通义”基础功能演示 (文本生成、改写、思维导图等);
3. AI 的潜在局限与风险 (偏见、错误信息、隐私泄露等);
4. 初步渗透 AI 伦理与社会影响;
5. 上机操作, 亲身体验通义简单指令使用。

1 AI 是什么?

人工智能 (*Artificial Intelligence, AI*), 从诞生以来就是一个令人兴奋的有趣概念, 不仅是计算机科学的重要领域, 还无数次出现在科幻题材的小说、电影、动漫和游戏中, 引发无数灵感和经久不衰的热烈讨论, 其中最吸引人的可能是关于人类与 AI 的深层互动乃至人类本质的探讨。

但长期以来, 精确定义“什么是 AI”仍然是一个难题。在 AI 大牛 Stuart Russell 和 Peter Norvig 的经典教材《人工智能: 现代方法 (*Artificial Intelligence A Modern Approach*)》中, 用了一整个小节 4 页纸来说明这个问题, 希望更准确更系统了解的同学可以去读一读, 会非常有教益, 我们这里则尝试用更简单的方式来帮助大家理解“AI 是什么”。

1.1 什么是智能 *Intelligence*

智能 *Intelligence* 通常指在未知具体解决方案的情况下完成复杂任务的能力, 尤其是那些需要运用特定知识 *knowledge* 和技能 *skill* 才能完成的任务。我们把拥有智能的对象称为“智能体”。

这里第一个要点在于对“未知”问题的探索, 如果已知解决方案、按部就班操作就能完成的任务, 通常不能充分体现出“智能”的作用。

另一个要点则是“复杂”, 简单的任务并不体现智能性, 比如走路、喝水、穿衣等, 而“智能型”任务包含对任务的分析、分解、规划、尝试与组合, 比如制订一份旅游计划就是一个典型的例子。

1.1.1 思考与练习

举出更多“智能型”任务的例子，尝试思考和总结自己解决这类任务通常的思维模式与方法。

1.2 什么是人工智能 *Artificial Intelligence*

所有非生物的智能体即为人工智能 *Artificial Intelligence*。人工智能通常是通过某种计算机程序来实现的。

这里的关键问题是，计算机程序有很多，什么样的程序才算“人工智能”？实际上这个问题没有精确的答案，而且会随着技术的发展而不断变化。比如在三十年前，使用 *A** 等启发式图算法被视为 *AI*，专家系统被视为 *AI*，但在今天，这些都是普通的计算机程序/系统了；十多年前，手写字符或者人脸的识别被视为 *AI*，今天是很日常应用的标配；今天我们耳熟能详的各种生成式人工智能，什么时候会被更先进的人工智能所取代呢？

诚然，所有这些技术及应用都是人工智能技术发展过程中的产物，某种意义上都是人工智能，但它们之间的差别和差距是巨大的，本着与时俱进的精神和持续发展的眼光，我们今天重点关注最新的、以机器学习相关技术为基础的、自回归的大语言模型 (*auto regressive language model*) 为代表的人工智能技术及应用，是合理且务实的。

在这一前提下，通常认为具备以下至少一个特征的程序，就属于“人工智能”的范畴：

- 学习：通过数据驱动改进和加强自身能力（如机器学习、深度学习），例如：推荐系统根据用户行为调整推荐内容。
- 推理与决策：能处理不确定性问题（如概率推理、模糊逻辑），例如：医疗诊断系统综合症状推断疾病。
- 感知与交互：识别图像、语音、文本等非结构化数据，示例：自动驾驶汽车解析摄像头和雷达信号。
- 自主：在动态环境中无需人工干预（如工业机器人调整动作路径）。

- 泛化：将经验迁移到新场景（如预训练大模型适应不同任务）。

相反的，非 AI 程序的典型特征包括：

- 固定规则：输出完全由预设逻辑决定（如计算器）。
- 无适应性：无法从数据或经验中自我优化。
- 任务单一性：仅能处理结构化输入（如数据库查询）。

当然，这样的定义无法回答更深刻的、涉及伦理的一些判断，比如是否要具备“意识”或“创造力”才算 AI（当前技术尚未达到）？这是一个开放性的问题，交给你去继续探究。

1.2.1 思考与练习

其实人工智能已经渗透入一些传统应用的特定功能点，比如很多软件都有的翻译功能，大部分使用的是较老的不那么“AI”的实现，但也有使用全新一代大语言模型技术实现的翻译功能，比如 2025 年初小红书上线的自动翻译功能，就是这样的例子，你可以实际尝试一下，体会一下它与前代翻译技术的显著差异。

想想你最希望拥有的人工智能产品功能是什么，与你身边的人分享，并尝试讨论一下它的可实现性。

1.3 弱 AI 与强 AI

弱人工智能 Narrow Intelligence，是指只能完成特定类型任务的人工智能。

英文名字更能体现出实质：所谓弱人工智能其实可能很强，它只是能力专一，比如人工智能围棋程序 AlphaGo 可以轻松击败顶级人类棋手，但它只会下围棋，不会国际象棋也不会画画更不会编程。

强人工智能 Artificial General Intelligence，指能以类似人类的水平完成几乎任何任务的人工智能，包括学习。

同样的，更好的说法是“通用人工智能”，或者简称 *AGI*。通用人工智能需要程序能够自行解决全新的问题，包含强大的学习、探索、试错与自我优化能力，这是人工智能领域的终极目标，目前尚无法预测它如何或者合适能够实现，但我们知道，一旦这一目标实现，人工智能将快速超越人类的智能水平，从 *AGI* 进化到 *ASI*（超级人工智能）。

1.3.1 思考与练习

为什么我们说一旦实现了 *AGI*，很快就会进化到 *ASI*？