

现代人工智能的本质、技术和途径

人工智能是指系统或智能体试图模拟或拥有人类的行为、思维 and 智能，虽然有不同途径期望通向人工智能，但这里我们只关注数学和工程相结合的技术路线。简单地说，现代人工智能技术就是通过机器学习及由其驱动而发展起来的计算机视觉、自然语言处理和语音识别技术来实现多模态数据融合的现实交互。人工智能的本质是处理匹配、决策和生成等任务，而这又和机器学习的三大学习范式——有监督学习、强化学习和无监督学习相呼应。数学上，现代人工智能技术则可以理解为利用数据驱动的方法求解具有组合结构的复杂高维问题。这也意味着计算机科学和统计学是现代人工智能的两个重要学科支柱。

“什么是智能”是人工智能的根本性问题。回答这个问题，需要考虑“人与理性”以及“思想与行为”两个维度的四种组合：类人行为、理性行为、类人思考、理性思考。类人思考和理性思考意味着智能体具有思考和自省能力，这超越了人工智能作为技术的范畴。类人行为和理性行为分别代表了当今人工智能的两种主要分类：弱人工智能和强人工智能。“图灵测试”被广泛用于测试系统是否具有智能特性，然而这只是对智能的一种定性描述。迄今为止还没有一种对智能的定量数学定义。找到一个类似于熵作为度量不确定性的这样严格定义可能很困难，但考虑类似 p -值作为描述给定数据对假设检验证据度量的定义或许是一种途径。另一方面，定性描述也使得智能边界弗远。

人工智能经历了热情高涨和期望无限的早期(1952-1969)，通用搜索机制所导致的回落期(1966-1973)，以专家系统为代表的基于规则学习的崛起期(1969-1986)，神经网络联结主义的回归期(1986-1995)，统计机器学习的复兴期(1995-至今)，以及大数据驱动的深度学习突破期(2006-至今)。因此，可以把现代人工智能理解为数据科学的高级形式。

人工智能的发展历程本质上就是试图寻找有效求解具有组合结构问题的方法的过程。在早期阶段，由于考虑的问题规模较小，且关注的任务相对单一，通常可以采用暴力搜索方法。然而随着问题规模变得越来越大以及任务变得越来越复杂，使用通用搜索机制求解问题变得不可行。因此，考虑利用数据或者规则的内在结构设计高级算法来求解。然而任何算法寻找精确解的适用性总是有其边界的，这迫使我们在效率和性能之间寻求折衷的方法——寻找近似解或次优解。引入不确定性和“探索和使用”机制产生了求解问题的突破性思路，数据驱动的统一

计方法由此被广泛应用。一方面数据的积累规模越来越大，且利用预训练和蒸馏等技术使得获取所需数据变得更加容易，另一方面从数据中挖掘信息和推理结论的算法技术也在不断改进和提升。人工智能的求解任务从匹配到决策以至生成，使得序贯动态决策和多模态数据生成都成为可能。

机器学习起源于计算机科学，但它和统计学也一脉相承，都是试图从数据或经验中学习以提升系统的能力或性能。机器学习是现代人工智能获得成功的一个最有效途径。特别地，深度学习和强化学习构成现代人工智能技术的两翼。深度学习利用层级化、局部化和正则化等思想提供了一种强大的数据表示方式或者函数逼近模型，且它的计算便于并行执行。深度学习给计算机视觉、自然语言处理、语音识别、棋牌游戏以及自然科学一些领域带来了颠覆性的突破，并为实现多模态数据生成提供了可行方法。强化学习则利用反馈、对抗和试错等机制提供了一种求解问题的方法论或者算法设计的数学表示，为我们实现在线决策提供了潜在思路。深度学习和强化学习相得益彰，融合两者的“深度强化学习”可以有效地克服求解具有组合结构的高维问题所面临的“维数诅咒”挑战，为我们实现现代人工智能提供了强大的技术方法。因果学习是实现人工智能的另一种潜在途径，它提供了表示对象之间逻辑关系的概率方法，因此在可解释性方面扮演着重要的角色，但是它面临着大规模数据的可扩展性问题。在深度学习和强化学习基础上，结合因果学习和统计物理模型，为设计新一代大语言模型赋予更多遐想。

人工智能发展方兴未艾，算法、应用和基础构成其研究的三个层面。首先，提出和开发新模型、新技术、新算法。其次，寻找人工智能技术更广泛的应用，定制具体领域场景的人工智能系统。第三，关注人工智能模型所蕴含的理论基础，旨在设计尽可能高效地利用资源、信息和计算能力的算法，或者判断什么情况下设计出信息和计算有效的算法是不可能的。我们通常从样本有效性和计算有效性两个方面来研究上述问题。样本有效推理是统计学的一个经典主题，而计算有效算法是计算机科学研究的核心课题。正如我们现在看到的，两者的交汇产生了有趣和令人惊奇的效果。如果孤立地研究它们，将无法了解这些效果。因此，这种更现代的观点对于理解和分析人工智能基础至关重要，同时也可能会给我们带来新的数学问题。

10年前深度学习在计算机视觉中以及近期强化学习在大语言模型中取得的成功都是算法和工程完美结合的巅峰之作。一个好的算法需要好的工程实现才能发挥其最大的效果。大模型的搭建、超参数的选取、数据的增广、以及精调、蒸馏和预训练等既是工程技术也是艺术，是算法的必要补充，也是驱动新算法提出的源泉。计算图和自动微分技术的提出让人工智能的开发和研究变得相对容易。

机器学习系统已成为了一门重要的领域。

人工智能和数据科学都是概率论在数据中的应用，前者更注重人机交互的智能属性，而后者则更注重问题解释的科学属性。人工智能是计算机科学的应用，数据科学是统计学的拓展。机器学习是计算机科学和统计学的交汇，是探索人工智能和数据科学的核心方法。因此，发展人工智能归根结底取决于计算机科学和统计学的学科发展和人才培养的底蕴。

发展人工智能技术需要高度的想象力、创造力和执行力，需要务实、理性、严谨的求是态度。人工智能的未来属于青年。因此，我们需要创造一个良好的环境和机制，潜心培养一批批有才华的年轻人，并创造机会让他们能脱颖而出。特别地，以做项目完成指标为导向和以做论文为研究驱动的博士生培养模式已经不能很好地服务科技的创新，我们同样需要有以培养下一代人才为纯粹要务的模式——即帮助学生构建较为完整、先进的学科知识体系和储备丰富的研究工具和手段，树立他们平等、独立和合作的科学精神，鼓励他们去自由探索最有挑战的领域和项目。唯有这样，我国人工智能或科技未来可期。

后记：我 2018 年 4 月在北大申请设立机器学习博士生项目作为计算机科学和统计学之间的纽带，未被认同，没能成功，颇为遗憾。本文写作的动机是鼓励我的学生们对现代人工智能技术应抱有极大的学习热情，但更重要是告诉他们统计学发展正当其时、大有作为，当务之要则是构建前沿的知识体系，掌握尽可能多的数学工具，要么研究人工智能无法逾越的基础问题，要么研究设计和分析人工智能技术的方法。

致谢：感谢周舒畅博士和我之前的学生王树森、赵申剑，由于和他们在大语言模型研发方面的交流让我对业界需求有比较清晰了解。感谢 John Hopcroft 教授，每次和他交谈总是受益良多，他关于人才培养的理念让我振聋发聩、诚惶诚恐。

张志华

2023.03.31