

同濟大學

TONGJI UNIVERSITY

毕业设计（论文）

课题名称

人工智能导论报告

副标题

浅谈强化学习的发展背景与技术现状

学院

软件学院

专业

软件工程

学生姓名

刘淑仪

学号

2251730

指导教师

曾进

日期

2024 年 3 月 19 日

人工智能导论报告

摘要

强化学习作为当今人工智能领域的相关研究方向之一，在其技术状态方面取得了显著突破。本文总结了强化学习在人工智能领域的发展背景、技术现状、技术发展趋势和应用场景。起源于心理学领域的强化学习概念，经过几十年的发展，已成为人工智能领域的重要研究方向之一。文章首先介绍了强化学习的基本概念和发展历程，然后概述了强化学习的关键技术，包括价值函数、策略搜索、深度 Q 网络等。接着，分析了强化学习技术的发展趋势。最后，总结了强化学习在智能游戏、自动驾驶、金融交易等多个领域的应用场景，并展望了未来的发展前景。本文旨在为读者提供关于强化学习的全面理解，并激发对强化学习技术应用和研究的更多思考。

关键词：人工智能，强化学习，机器学习

1 人工智能总体发展历史

1.1 基本介绍

1.1.1 研究目的

人工智能（Artificial Intelligence, AI）研究目的是通过探索智慧的实质，扩展人类智能——促使智能主体会听（语音识别、机器翻译等）、会看（图像识别、文字识别等）、会说（语音合成、人机对话等）、会思考（人机对弈、专家系统等）、会学习（知识表示，机器学习等）、会行动（机器人、自动驾驶汽车等）。

1.1.2 发展史

在 1943 年-20 世纪 60 年代，人工智能开始起步发展。美国神经科学家麦卡洛克和逻辑学家皮茨提出了神经元数学模型，现代人工智能学科奠基石就此诞生。而 56 年达特茅斯学院人工智能夏季研讨会上正式使用了人工智能（artificial intelligence, AI）这一术语。这是人类历史上第一次人工智能研讨，标志着人工智能学科的诞生。

在 20 世纪 70 年代，人工智能迎来了反思发展期。虽然人们此时对人工智能抱有较大的期望，但计算力和理论知识的匮乏使得其发展目标不免落空，人工智能的发展走向低谷。80 年代时，人工智能走入应用发展的新高潮。在 90 年代到 2010 年，AI 平稳发展。人工智能的研究重心从基于知识系统转向了机器学习方向。

2011 年至今，科学与应用之间的技术鸿沟被大幅度跨越，诸如图像分类、语音识别、知识问答、人机对弈、无人驾驶等人工智能技术实现了重大的技术突破，迎来爆发式增长的新高潮。

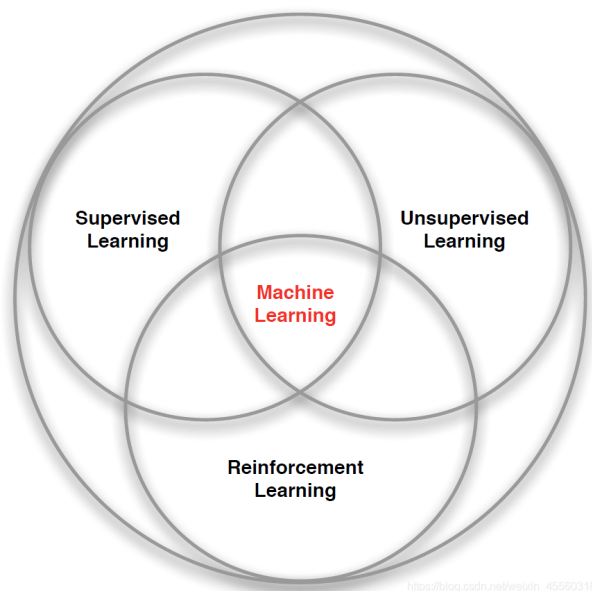


图 1.1 内容交叉示意图

1.2 强化学习发展背景

强化学习最早可以追溯到巴甫洛夫的条件反射实验，它从动物行为研究和优化控制两个领域独立发展，最终经 Bellman 之手将其抽象为马尔可夫决策过程 (Markov Decision Process, MDP)；在 1956 年 Bellman 提出动态规划方法时，强化学习的发展历史拉开序幕。

强化学习的目标是使得回报最大化。强化学习和非监督学习的关键部分就是回报的选择。强化学习是学习状态和行为之间的映射关系，以使得数值回报达到最大化。换句话说，在未知采取何种行为的情况下，学习者必须通过不断尝试才能发现采取哪种行为能够产生最大回报。

在 2015 年之前，人们对强化学习的研究还处于理论阶段。而 2015 年，DeepMind 公司提出了著名的 Deep-Q-Network 算法，将强化学习与深度学习进行了融合，设计出了 AlphaGo，并于 2016、2017 年相继战胜世界围棋冠军李世石和柯洁，自此强化学习名声大噪。

2 强化学习发展趋势

2.1 核心发展趋势

随着算力的飞速提升，深度学习 (Deep Learning, DL) 发展势头迅猛，掀起了新一轮人工智能发展热潮。在深度学习的带动下，强化学习 (Reinforcement Learning, RL) 重新步入了人们的视野。

如何赋予机器自主学习的能力，一直是人工智能领域的研究热点。强化学习是机器学习领域之

一，受到行为心理学的启发，主要关注智能体如何在环境中采取行动，以最大限度地提高累积奖励。

随着强化学习在协作和竞争环境中的应用增多，多智能体强化学习成为了一个重要的研究方向。不仅如此，迁移强化学习、可解释性强化学习和基于模型的强化学习也是最近的发展热门。

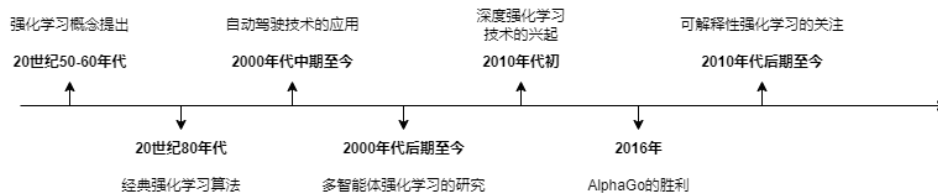


图 2.1 发展趋势图

2.2 结合发展

由于近些年来的技术突破，和深度学习 (Deep Learning) 的整合，使得强化学习有了进一步的运用。

人工智能在模拟大自然的两个方面，神经网络和强化学习，神经网络模拟的是人类和哺乳动物的神经系统神经元，而强化学习它模拟进化过程生物与环境的交互，所以一个微观一个宏观，它都显示了人类在探索人工智能道路上的向大自然学习的两个方面。

而把强化学习和深度学习结合，就会发现其更接近于通用的人工智能。

3 强化学习应用背景

强化学习的应用非常广泛，例如在机器人控制、游戏玩法、自然语言处理等领域中都有广泛的应用。常用的强化学习算法包括 Q-learning、Deep Q-Network (DQN)、Policy Gradient 等。其主要分类如下：

A. 智能游戏：

强化学习在智能游戏领域有着广泛的应用。此外，强化学习还被用于电子游戏中的角色控制、敌人行为设计等方面。

B. 自动驾驶：

自动驾驶汽车需要在复杂的交通环境中做出决策，并学习适应不同的驾驶情境。强化学习被用于自动驾驶系统中，以优化车辆的行为策略，确保安全、高效的驾驶。

C. 机器人控制：

强化学习可以用于机器人的路径规划、物体抓取、动作控制等方面。通过在仿真环境中进行训练，机器人能够学习执行各种复杂任务。

D. 金融交易：

在金融领域，强化学习被应用于股票交易、量化交易等方面。智能体通过与市场环境的交互学习优化的交易策略，以实现收益最大化或风险最小化。

E. 智能推荐系统:

强化学习可以用于个性化推荐系统，根据用户的反馈学习改进推荐策略，提供更加精准的推荐结果。

F. 工业控制与优化:

在工业领域，强化学习被应用于生产调度、设备维护、能源管理等方面。智能体学习优化的生产策略，以提高生产效率和降低成本。

G. 教育和训练:

强化学习可以用于个性化教育和培训。通过与学习者的交互，智能体能够根据学习者的表现调整教学策略，提供更有效的学习体验

4 总结

A. 发展背景:

强化学习的概念最早由心理学家提出，用于描述动物如何通过与环境交互学习行为。

20 世纪 80 年代，经典强化学习算法（如 Q-learning）的出现标志着强化学习研究的开始。

随着计算机技术和数据处理能力的提高，以及深度学习技术的兴起，深度强化学习技术在 2010 年代初开始蓬勃发展。

B. 技术现状:

强化学习是一种通过智能体与环境进行交互学习的方法，智能体根据环境提供的奖励信号来调整行为策略。

深度强化学习技术的兴起使得智能体能够从高维复杂的环境中学习和决策，取得了在多个领域的重大突破。

C. 技术发展趋势:

深度强化学习技术的不断发展和完善，如基于模型的强化学习、多智能体强化学习、迁移强化学习、可解释性强化学习等。

强化学习算法的稳定性、可解释性、泛化能力等方面的研究和改进。

强化学习与其他领域的融合，如深度学习、神经科学、心理学等，以及在多领域中的新应用。

D. 应用场景:

强化学习在智能游戏、自动驾驶、机器人控制、金融交易、智能推荐系统、工业控制与优化、教育和训练等多个领域有着广泛的应用。不同领域的应用场景和需求推动了强化学习技术的不断发展和创新，同时也为强化学习提供了丰富的实践和应用机会。

5 参考

5.1 参考文章

1. 一文概览人工智能 (AI) 发展历程: <https://zhuanlan.zhihu.com/p/375549477>.
2. 浅析 2022 年强化学习 20 个重要进展: <https://zhuanlan.zhihu.com/p/596276958>
3. 深度强化学习的现在, 将来与未来: <https://zhuanlan.zhihu.com/p/367326621>
4. Machine Learning 系列一文带你详解什么是强化学习 (Reinforcement Learning):
https://blog.csdn.net/m0_63947499/article/details/131529656?spm=1001.2101.3001.6650.15&utm_medium=distribute.pc_relevant.none-task-blog-2
5. Reinforcement Learning 一: 历史发展背景与介绍:
https://blog.csdn.net/qq_20499063/article/details/78762596#:~:text=Reinforcement
6. 强化学习 (RL)——Reinforcement learning:
https://blog.csdn.net/qq_47527477/article/details/120462181?utm_medium=distribute.pc_relevant.none-task-blog-2~default~baidujs_baidulandingword~default-0-120462181-blog-78762596.235^v43^pc_blog_bottom_relevance_base3&spm=1001.2101.3001.4242.1&utm_relevant_index=1