

基于 AlphaGo Zero 对人工智能领域发展 趋势的思考

学号：1170300213

姓名：席奴瓦

目录

1. AlphaGo Zero 概述
2. AlphaGo Zero 对人工智能领域的影响
3. 与 AlphaGo Zero 相关的研究
4. 深度学习（Deep Learning）近三年研究热点及方向
5. 今后需要强化的重要专业课程及其对从事职业的作用
6. 参考文献

1. AlphaGo Zero 概述

1.1. AlphaGo 简介

AlphaGo 是位于伦敦的 Google DeepMind 团队开发的一款围棋人工智能程序，程序利用“价值网络”去计算局面，用“策略网络”去选择下子。AlphaGo 的主要工作原理是“强化学习”，其算法结合了神经网络、机器学习和蒙特卡罗树搜索技术。AlphaGo Zero 是最新版本。

1.2. AlphaGo Zero 战绩

AlphaGo Zero 完全从零开始，在不参考人类任何先验知识，仅输入围棋基本规则的情况下，在第三天实现对 AlphaGo Lee（打败李世石的 AlphaGo 版本）的碾压，对弈成绩 100:0；在第 21 天达到 AlphaGo Master（打败柯洁的 AlphaGo 版本）的水平，在线击败 60 位顶尖专业选手，并在 2017 年世界锦标赛上以 3:0 击败冠军柯洁；在第 40 天超过其他所有 AlphaGo 版本，成为世界上最好的棋手，达到这一水平没有人类干涉，没有使用历史数据，完全依赖于自学。

1.3. AlphaGo Zero 与前代对比：

- ◆ 算法：自对弈强化学习，完全从随机落子开始，不用人类棋谱。
- ◆ 数据结构：直接使用棋子位置做神经网络输入(下面会进一步叙述)，不再使用人工特征 (AlphaGo Lee/Master/Fan 用过的人工特征包括: 当前位置是否是征子/引征, 当前位置吃子/被吃子数目、本块棋的气数等)。
- ◆ 策略：基于训练好的神经网络进行简单的树形搜索。之前使用蒙特卡罗算法实时演算并加权得出落子位置。
- ◆ 模型：使用一个神经网络，初始训练时不再使用人类棋谱做有监督学习，而是直接从基于围棋规则的随机下法开始强化学习。之前使用策略网络学习人类棋风，局面网络计算当前局面下每一个不同落子的胜率。

1.4. AlphaGo Zero 创新技术

- ◆ 使用 ResNet 代替原来的 ConvNet；
- ◆ 合并 Policy Network 和 Value network;
- ◆ Reinforcement Learning 训练算法的改进;
- ◆ 取消人类棋谱知识和人工特征。

2. AlphaGo Zero 对人工智能领域的影响

- 2.1 证明了在某些领域训练 AI AGENT 的过程中，模型的价值超过了训练数据（先验知识）。如果这种模型可以运用到其他领域，可以期望 AI AGENT 会产生超过人类现有知识积累的新的创造性的知识。
- 2.2 证明了如果采样算法合理，即使只采样了模型空间的一个天文数字分之一的子集，也能很好描述问题领域。从特征工程(Feature Engineering) 到数据工程/数据增强 (Data Engineering/Data Augmentation) 再到采样工程 (Sampling Engineering)。传统人工智能系统的性能并非取决于分类器设计，而主要取决于数据预处理，特征提取和特征设计。到了深度学习时代，大数据采集，数据增强(根据某些先验知识对数据进行变化，以增加数据的数量和多样性)成了算法设计师和系统工程师的日常。能否借鉴 AlphaGo Zero 的思想，不直接做 Data Augmentation，而是对目标系统做部分或全部建模，然后在庞大的模型空间中通过特定采样算法进行采样？
- 2.3 当面对一个搜索空间巨大的决策问题时，神经网络能比一个真·蒙特卡洛搜索树（CrazyStone 的算法）更好的表示一个蒙特卡洛搜索树（可能是 node 之间 dependent，用一个真树有太多无用的 parameter）。

3.与 Alphago Zero 相关的研究成果

- ◆ 利用人工智能提高早期乳腺癌诊断率——MIT
- ◆ Fooling neural networks w/3D-printed objects.——MIT
- ◆ Understanding Black-box Predictions via Influence Functions.——Stanford
- ◆ Machines are getting schooled on fairness.——CMU
- ◆ Importance Sampling for Fair Policy Selection.——Stanford

4.深度学习（Deep Learning）近三年研究热点及方向

- ◆ **无监督学习**。无监督学习在深度学习刚火的那几年有很重要的作用，比如用无监督方式训练深度信念网络还有稀疏自编码器等，使用无监督学习主要是为了预训练，以得到一个较好的初始值，随后再使用有监督训练进行微调。但是随着计算能力的发展，人们发现只要在数据集足够大的情况下使用纯有监督学习也能得到较好性能，所以近几年无监督学习发展不是很大，Hinton 他们希望在未来无监督学习能有更大发展，因为人类和动物的学习在很大程度上都是无监督的：我们通过观察这个世界来学习，而不是有个老师在教我们这个世界的原理。
- ◆ **深度强化学习**。深度强化学习的主要思想简而言之就是将深度学习与强化学习相结合，是一种从感知到动作的端到端学习。简单的说，就是和人类一样，输入感知信息比如视觉，然后通过深度神经网络，直接输出动作，中间没有人工特征的工作。深度增强学习具备使机器人实现真正完全自主的学习一种甚至多种技能的潜力。深

度强化学习最突出的代表就是 DeepMind 公司了，该公司在 NIPS 2013 上发表的 [Playing Atari with Deep Reinforcement Learning](#) 一文，在该文中第一次提出深度强化学习这个名称，之后在 Nature 上发表了改进版的文章 [Human-level control through deep reinforcement learning](#)，引起了广泛的关注，深度强化学习从此成为深度学习领域的前沿研究方向。最近的李世石大战阿法狗中，阿法狗背后的技术也是深度强化学习，DeepMind 将阿法狗背后用到的技术发表在了 2016 年的 Nature 上 [Mastering the game of Go with deep neural networks and tree search](#)。

- ◆ 自然语言理解。自然语言理解也是深度学习在未来几年能大有作为的领域。使用深度学习技术的各种应用比如神经机器翻译，问答系统，文摘生成等都取得了不错的效果，效果的提升主要归功于注意力机制和循环神经网络相结合的强大能力。相信未来几年内还会有大量相关工作出现。

5.今后需要强化的重要专业课程

- ◆ Artificial Intelligence : Principles and Techniques
- ◆ Machine Learning
- ◆ Probabilistic Graphical Models:Principles and Techniques
- ◆ Natural Language Processing with Deep Learning

参考文献

- [1] AlphaGo Zero : Learning from scratch (<https://deepmind.com/blog/alphago-zero-learning-scratch/>)
- [2] 如何评价 AlphaGo Zero? (<https://www.zhihu.com/question/66861516>)
- [3] Online Meta-learning by Parallel Algorithm Competition, by [Stefan Elfwing](#), [Eiji Uchibe](#), [Kenji Doya](#)
- [4] Learning Deep Architectures for AI, by Yoshua Bengio
- [5] Using artificial intelligence to improve early breast cancer detection (http://www.csail.mit.edu/ai_cancer_detection)
- [6] Machines are getting schooled on fairness (<https://www.sciencenews.org/article/machines-are-getting-schooled-fairness>)
- [7] Fooling neural networks w/3d-printed objects(http://www.csail.mit.edu/fooling_neural_networks_with_3Dprinted_objects)
- [8] <http://www.cs.cmu.edu/~shayand/papers/UAI2017.pdf>
- [9] <https://www.zhihu.com/question/23140232>