

---

# 人工智能再次战胜人类玩家

作者：writer 来源：科学网

本文原地址：<https://www.iikx.com/news/progress/21149.html>

**本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！**

人工智能再次战胜人类玩家。下棋、玩游戏，在与人类的博弈中，人工智能(AI)正在不断成长。近日，英国DeepMind公司的AI智能体DeepNash，就在西洋陆军棋(Stratego)中达到了专业级人类玩家的水平。该成果12月1日发表于《科学》。

事实上，在DeepNash攻克Stratego前，就有报道称AI能够玩强权外交这款类似的经典桌游。这是一款具有外交深度的游戏，玩家间存在合作与竞争关系，必要时须通过谈判结交盟友。

近年来，AI掌握本质不同的游戏的速度相当惊人。美国密歇根大学计算机科学家Michael Wellman表示，Stratego和强权外交这两款游戏大不相同，但它们都极具挑战性，与此前AI已经掌握的游戏也截然不同。

Stratego是一款需在信息缺失情况下进行战略思考的棋盘游戏，类似于中国的军棋。它要比此前AI掌握的国际象棋、围棋或扑克复杂得多。

在游戏中，双方各有40枚棋子，且彼此看不到棋子的真实身份。双方轮流移动棋子，来消灭对手的棋子，最终夺得对方军旗或消灭所有能动的棋子的一方获胜。因此，玩家需要进行战略部署、收集信息，并与对方博弈。

一局Stratego游戏中会有10535种可能的布局。相比之下，围棋的布局有10360种可能。此外，在Stratego中，AI需要推理对手超过1066种部署策略，这使德州扑克中初始的106种可能情况相形见绌。

Stratego中可能出现的结果数量之多及其复杂性，意味着在信息完全的游戏中表现出色的算法，甚至在扑克游戏中表现良好的算法，在这款游戏中都不起作用。DeepMind研究员Julien Perolat说。

所以Perolat和同事开发了DeepNash，该命名致敬了提出纳什均衡的美国数学家约翰·纳什。

纳什均衡是博弈论中一种解的概念，指满足以下条件的策略组合：任何一位玩家在此策略组合下单方面改变自己的策略(其他玩家策略不变)，都不会提高自身的收益。

DeepNash将强化学习算法与深度神经网络相结合，以找到纳什均衡。强化学习包括为游戏的每个状态找到最佳策略。为了学习最佳策略，DeepNash已经与自己进行了55亿次博弈。

---

今年4月，DeepNash在网络游戏平台Gravon上与人类Stratego玩家进行了两周的比赛。在50场比赛后，DeepNash目前在所有Gravon Stratego玩家中排名第三。

我们的研究表明，像Stratego这样复杂的涉及不完善信息的游戏，不需要通过搜索技术来解决。团队成员、DeepMind研究员Karl Tuyls表示，这是AI迈出的一大步。

而曾于2019年报告了玩扑克的AI——Pluribus的Meta AI研究员Noam Brown团队，则将目光投向了一个不同的挑战：建立一个可以玩强权外交的AI。

强权外交是一个最多可由7位玩家参与的游戏，每位玩家代表第一次世界大战前欧洲的主要力量，游戏目标是通过移动部队控制供应中心。重要的是，该游戏需要玩家间进行私人交流和合作，而不是像围棋或Stratego那样进行双人博弈。

当进行超过两人的零和博弈时，纳什均衡思想对游戏就不再有用。Brown说，他们成功训练出了AI——Cicero。在11月22日发表于《科学》的论文中，该团队报告称，在40场游戏中，Cicero的平均得分是人类玩家的两倍多，在玩过一场以上游戏的参与者中排名前10%。

Brown认为，可以与人类互动并解释次优甚至非理性人类行为的游戏AI，可以为其在现实世界的应用铺平道路。(来源：中国科学报 许悦)

相关论文信息：<https://doi.org/10.1126/science.add4679>

作者：Julien Perolat 来源：《科学》

更多科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发