



中国共产党新闻网 >> 理论 >> 理论期刊 >> 《中国发展观察》 >> 每期精选

人工智能与博弈论——从阿尔法围棋谈起

安波

2016年03月22日10:00 来源：人民网-理论频道

原标题：人工智能与博弈论——从阿尔法围棋谈起

谷歌AlphaGo（阿尔法围棋）在与棋手李世的人机大战中，最终以4：1赢得胜利。这一人类智慧和人工智能的对决在世界各地掀起了对人工智能空前的关注热潮。

AlphaGo是一款围棋人工智能程序，由谷歌Deep Mind团队开发。AlphaGo将几项技术很好地集成在了一起：通过深度学习技术学习了大量的已有围棋对局，接着应用强化学习通过与自己对弈获得了更多的棋局，然后用深度学习技术评估每一个格局的输赢率（即价值网络），最后通过蒙特卡洛树搜索决定最优落子。同时谷歌用超过1000个CPU和GPU进行并行学习和搜索。

在过去20多年中，人工智能在大众棋类领域与人类的较量一直存在。1997年，IBM公司研制的深蓝系统首次在正式比赛中战胜人类国际象棋世界冠军卡斯帕罗夫，成为人工智能发展史上的一个里程碑。然而，一直以来，围棋却是个例外，在这次AlphaGo取得突破性胜利之前，计算机围棋程序虽屡次向人类高手发出挑战，但其博弈水平远远低于人类，之前最好的围棋程序（同样基于蒙特卡洛树搜索）被认为达到了业余围棋五、六段的水平。

这其中的一个原因就是围棋的棋局难于估计，对局面的判断非常复杂。另外一个更主要的原因是围棋的棋盘上有361个点，其搜索的宽度和深度远远大于国际象棋，因此，求出围棋的均衡策略基本是不可能的。AlphaGo集成了深度学习、强化学习、蒙特卡洛树搜索，并取得了成功。

我们这里顺便说一说人工智能和人类在另一项棋类项目——德州扑克的较量。德州扑克于20世纪初开始于德克萨斯洛布斯镇，后来在全美大面积流行起来。德州扑克以其易学难精的特点，受到各国棋牌爱好者的青睐。世界德州扑克系列大赛（WSOP）是一个以无上限投注德州扑克为主要赛事的扑克大赛，自上世纪70年代登陆美国以来，比赛在赌城拉斯维加斯的各大赌场举行。其中，以冠军大赛的奖金额最高，参赛人数最多，比赛最为隆重，北美各地的体育电视频道都有实况转播。有史以来第一次人类和计算机无限注德州扑克比赛于2015年4月24日到5月8日在美国宾夕法尼亚匹兹堡的河边赌场举行，组织者为卡内基梅隆大学的Tuomas Sandholm教授，包括微软研究院等多家机构提供了奖金支持。该比赛共有两组玩家，一组是电脑程序“Clau-do”，另一组是该类扑克游戏的顶级专家Dong Kim、Jason Les、Bjorn Li和Doug Polk。Clau-do是之前Tartanian（2014美国人工智能大会电脑扑克大赛冠军所用的程序）的改进版本。该比赛一共进行了8万回合，最后扑克专家以微弱的优势获得了胜利，学术界认为Clau-do取得了很大的成功。

和AlphaGo不同的是，Clau-do的策略基于扑克博弈的近似均衡。围棋比赛本身是一种完全信息博弈，而扑克是不完全信息博弈（玩家不能观测到对手手中的牌），因此比完全信息博弈更难解决。Clau-do通过下面这三个步骤决定其策略。第一步：原始博弈被近

热点关键词

建党95周年	学习路上	改革创新案例征集
中国政要资料库	“两学一做”	党建云平台

似为更小的抽象博弈，保留了最初博弈的战略结构。第二步：计算出小的抽象博弈中的近似均衡。第三步：用逆映射程序的方法从抽象博弈的近似均衡建立一个原始博弈的策略。Clau-do的成功必须归功于算法博弈论最近几年的进展。在2015年年初《科学》杂志发布的一篇论文中，加拿大阿尔伯塔大学计算机科学教授Michael Bowling带领的研究小组介绍了求解有上限投注德州扑克博弈均衡的算法，基于该均衡策略的程序 Cepheus是接近完美的有上限投注德州扑克计算机玩家，以致于人类玩家终其一生也无法战胜它。这并不是说 Cepheus一局也不会输，但是从长期来看，结果只能是平手，或者计算机获胜。需要注意的是，有上限投注德州扑克博弈比无上限投注德州扑克博弈要容易求解。由于围棋和扑克在本质上都是博弈问题，我们这里谈谈博弈论以及作为求解扑克博弈的算法博弈论。1944年，John von Neumann与Oskar Morgenstern合著《博弈论与经济行为》，标志着现代系统博弈理论的初步形成，因此他被称为“博弈论之父”。尽管历年来，博弈论与计算学科学不时有显著的重叠，但在早期，博弈论主要为经济学家所研究应用。事实上，博弈论现在也是微观经济学理论的主要分析框架。 博弈论在经济教科书中的应用非常广泛。在经济科学领域，很多杰出的博弈理论家曾荣获诺贝尔奖，如2012年诺贝尔经济学奖得主罗斯和沙普利。

就在博弈论理论出现不久后，人工智能领域紧随其后得到开发。事实上，人工智能的开拓者如von Neumann 和Simon 在两个领域早期都有杰出贡献。博弈论和人工智能实际上都基于决策理论。例如，有一个著名观点把人工智能定义为“智能体的研究和构建”。从20世纪90年代中期到后期，博弈论成为计算机科学家的主要研究课题，所产生的研究领域融合计算和博弈理论模型，被称为算法博弈论。近几年来，算法博弈论发展尤为迅速，得到了包括哈佛大学、剑桥大学、耶鲁大学、卡内基梅隆大学、加州伯克利大学、斯坦福大学等世界各大著名研究机构的重点研究，该领域的会议如雨后春笋般出现，并与多智能系统研究融合，其普及程度已经在缓慢地追赶人工智能。算法博弈论的主要研究领域包括各种均衡的计算及复杂性问题、机制设计（包括在线拍卖、在线广告）、计算社会选择等，并在包括扑克等的很多领域得到应用。过去几年，算法博弈论在安全领域的资源分配及调度方面的理论——安全博弈论逐渐建立并且在若干领域得到成功应用。

与算法博弈论求解均衡策略或者近似均衡策略不同，基于学习以及蒙特卡洛树搜索的AlphaGo无法在理论上给出赢棋的概率。考虑到将博弈抽象的思想应用到扑克博弈上的成功，是否可能将围棋博弈抽象成小规模博弈，求解（近似）均衡策略，并产生原始博弈问题的策略？即使这种策略不能有赢棋概率的保证，这些基于均衡产生的策略有可能对提高AlphaGo的性能提供帮助。从另外一个角度，深度学习技术是否会为求解大规模博弈问题提供帮助也值得探索。也许我们无法证明基于深度学习的策略能够形成某种均衡，但是可能会从实验模拟结果来说接近均衡策略。因此，AlphaGo的成功不仅会引爆人工智能研究的热潮，也会促进人工智能与算法博弈论的进一步交融与发展。

（作者单位：新加坡南洋理工大学计算机工程学院）



使用微信“扫一扫”功能添加“学习微平台”

热门标签

热门搜索：焦扬 王侠 赵克志 陈敏尔



点击进入

中纪委领导	中组部领导	全总领导	公安部领导
国家发改委领导	省市区党委书记	省级党委秘书长	
省级团委书记	省级政法委书记		

