



一种军棋机器博弈的多棋子协同博弈方法

张小川, 王宛宛, 彭丽蓉

引用本文:

张小川, 王宛宛, 彭丽蓉. 一种军棋机器博弈的多棋子协同博弈方法[J]. 智能系统学报, 2020, 15(2): 399–404.

ZHANG Xiaochuan, WANG Wanwan, PENG Lirong. A multi-chess collaborative game method for military chess game machine[J]. *CAAI Transactions on Intelligent Systems*, 2020, 15(2): 399–404.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.201812012>

您可能感兴趣的其他文章

大数据智能：从数据拟合最优解到博弈对抗均衡解

Big data intelligence: from the optimal solution of data fitting to the equilibrium solution of game theory
智能系统学报. 2020, 15(1): 175–182 <https://dx.doi.org/10.11992/tis.201911007>

动态规划求解中国象棋状态总数

A method for calculating the total number of states of Chinese chess on the basis of dynamic programming
智能系统学报. 2019, 14(1): 108–114 <https://dx.doi.org/10.11992/tis.201803008>

基于半张量积的企业创新网络演化博弈

Evolutionary enterprise innovation networked game based on the semi-tensor product of matrices
智能系统学报. 2018, 13(5): 776–782 <https://dx.doi.org/10.11992/tis.201706064>

中国象棋机器人棋子定位与识别方法

Chess-piece localization and recognition method for Chinese chess robot
智能系统学报. 2018, 13(4): 517–523 <https://dx.doi.org/10.11992/tis.201709020>

基于棋型的藏族“久”棋计算机博弈研究

Tibetan JIU computer game research based on chess form
智能系统学报. 2018, 13(4): 577–583 <https://dx.doi.org/10.11992/tis.201609023>

计算机博弈的研究与发展

Research and development of computer games
智能系统学报. 2016, 11(6): 788–798 <https://dx.doi.org/10.11992/tis.201609006>

微信公众平台



关注微信公众号，获取更多资讯信息

DOI: 10.11992/tis.201812012

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20200323.1314.004.html>

一种军棋机器博弈的多棋子协同博弈方法

张小川¹, 王宛宛², 彭丽蓉³

(1. 重庆理工大学两江人工智能学院, 重庆 400054; 2. 重庆理工大学 计算机科学与工程学院, 重庆 400054;
3. 重庆工业职业技术学院 信息工程学院, 重庆 401120)

摘要: 针对在军棋博弈不完全信息对弈中, 面对棋子不同价值、不同位置、不同搭配所产生的不同棋力, 传统的单子意图搜索算法, 不能满足棋子之间的协同性与沟通性, 同时也缺乏对敌方的引诱与欺骗等高级对抗能力。本文提出一种结合 UCT 搜索策略的高价值棋子博弈方法, 实现高价值棋子协同博弈的策略。实战经验表明: 高价值多棋子军棋协同博弈策略优于单棋子军棋博弈策略。

关键词: 机器博弈; 军棋; 协同博弈; Q 学习算法; 攻守平衡; 维度灾难; UCT; 高价值棋子

中图分类号: TP311.5 **文献标志码:** A **文章编号:** 1673-4785(2020)02-0399-06

中文引用格式: 张小川, 王宛宛, 彭丽蓉. 一种军棋机器博弈的多棋子协同博弈方法 [J]. 智能系统学报, 2020, 15(2): 399-404.

英文引用格式: ZHANG Xiaochuan, WANG Wanwan, PENG Lirong. A multi-chess collaborative game method for military chess game machine[J]. CAAI transactions on intelligent systems, 2020, 15(2): 399-404.

A multi-chess collaborative game method for military chess game machine

ZHANG Xiaochuan¹, WANG Wanwan², PENG Lirong³

(1. Liangjiang Institute of Artificial Intelligence, Chongqing University of Technology, Chongqing 400054, China; 2. School of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China; 3. Faculty Information Engineering, Chongqing Industry Polytechnic College, Chongqing 401120, China)

Abstract: Owing to incomplete information on the military chess and the different strengths of chess pieces with different values, positions, and combinations, the traditional single-intention search algorithm cannot satisfy the coordination and communication requirements of chess pieces and lacks advanced confrontation capabilities, such as temptation and deception of the enemy. This study proposes the combination of the high-value chess piece game method and the UCT search strategy to achieve a high-value chess piece cooperative game strategy that can be used to solve the problems of the military chess game. Practical experience shows that the high-value multipiece military chess game strategy is better than the high-value single-piece military chess game strategy.

Keywords: computer game; military; collaborative game; Q learning algorithm; balance of attack and defend UCT; dimension disaster; high value chess piece;

机器博弈是人工智能领域重要的研究方向, 通过训练计算机下棋来衡量机器的智能程度, 具有人-机、机-机对弈 2 种形式, 2016 年 Google 的

Alpha Go 战胜韩国围棋大师李世石使得人机博弈变得家喻户晓^[1], 紧接着网络注册名 Master 线上挑战中日韩围棋高手, 以及升级版 Alpha Go Zero 以 4:0 战胜世界围棋第一人柯洁等标志性事件, 掀开了人机博弈中机器取胜的新篇章^[2]。但是机-机博弈依然是人类在人工智能领域不断探索的新

收稿日期: 2018-12-11. 网络出版日期: 2020-03-23.

基金项目: 国家自然科学基金项目 (61702063); 重庆理工大学研究生创新基金项目 (ycx2018244).

通信作者: 王宛宛. E-mail: 1033104010@qq.com.

坐标。

机器博弈根据博弈过程中信息的透明度可以分为完备信息和非完备信息博弈。完备信息博弈双方信息完全透明,对弈的双方完全了解彼此的博弈信息,主流的完备信息博弈有中国象棋、围棋、五子棋等。而非完备信息博弈双方信息不完全透明,其中存在一定的欺诈信息,主流的非完备信息博弈有军棋、德州扑克、桥牌、斗地主等^[3]。由于非完备信息博弈信息的不透明性导致其存在博弈诱导、欺诈等行为,使得非完备信息博弈难以攻克^[4]。

军棋博弈游戏中共有13类棋子,包括军、师、旅、团、营、连、排、司令、工兵等9类军人棋子以及地雷、炸弹、军旗等3类工具棋子。军棋博弈游戏过程中由于棋子棋力不同,位置也不相同,因此在不是一样的位置、不一样的棋子搭配的情况下,棋子所展现出来的战斗力也不相同。如何根据棋子特性搭配出最佳棋力成为研究的重点。

1 多棋子协同策略

多棋子协同博弈通过利用人类先验知识,制定出既定的目标和可实行的作战规划,发挥出相应的作战能力。其研究的目的是实现多棋子间协同,发挥群体作战功能,而不是单兵能力的体现与发挥。多棋子协同博弈的研究内容主要包括:多棋子优化协作、多棋子行为规划以及进行反馈学习^[5]。

1.1 多棋子联合定义

在军棋机器博弈游戏中,各个棋子间的协作关系是多种多样的,当博弈系统需要完成一个大目标时,就需要多个棋子进行协同作战。多棋子协同作战过程中会导致棋子间发生冲突,仅考虑棋子单独作战将会错过最优的作战计划^[6]。基于此,军棋机器博弈需要消除多方冲突的条件下产生最优的合作结果,使得机器博弈系统作战能力达到最优。军棋机器博弈系统中多棋子协同博弈策略取得好的作战力的时候,其组合策略就会得到一定的奖赏,遇到相同的情况下,相同的策略趋势便会增强,强化学习的方法为多棋子协同作战提供了高鲁棒性的学习方法,无监督的情况下,棋子可以在不同的棋局下不断优化更新出最优的协同策略^[7-8]。

多棋子协同博弈策略在棋子强化学习过程中的关键问题是确定棋子是否处于联合状态,并具有与其他棋子间的协同动作。多棋子进行协同博

弈同时选择棋子组合,棋子不能及时沟通将会导致系统无法确定多棋子协同博弈的动作。针对此问题,机器博弈系统采取模拟走步解决此问题,即根据MCTS模拟的结果结合强化学习结果进而制定棋子协同策略。因此,军棋机器博弈协同强化学习系统由模拟预测单元和强化学习单元构建而成,其结构流程图如图1所示。在多棋子协同博弈强化学习系统中,模拟预测单元利用博弈树搜索进行MCTS模拟,返回高价值走步,并反馈出其他棋子的走步及预测结果,并基于其反馈的结果完成多棋子强化学习算法。强化学习单元将反馈回来的协同策略返还给动作预测单元进行协同策略预测模型更新^[9]。

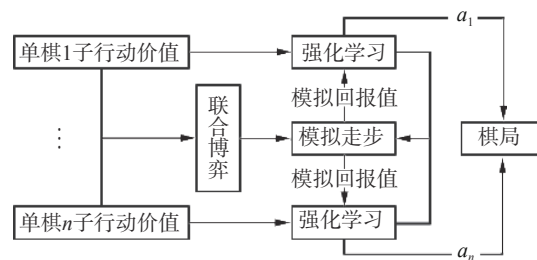


图1 强化学习模型

Fig. 1 Reinforced learning model

1.2 Q学习算法应用

机器博弈双方和多方下棋的过程可以看作一个随机过程,服从马尔可夫序列过程,用元组 $\langle S, a^1, a^2, \dots, a^n, f^1, f^2, \dots, f^n, g^1, g^2, \dots, g^n \rangle$ 来表示一个多棋子协同博弈的随机策略,元组中 n 为军棋博弈系统中棋子的个数,即25个; S 为博弈当前的局面,是一个 12×5 的矩阵; a^i ,其中 a^i 表示棋子 i 可以选择的作战行为序列的集合 $i=1, 2, \dots, n$ 。军棋机器博弈系统协同作战的行动序列为 $a^1 \times a^2, \dots, \times a^n$ 。 f^i ,其中 $i=1, 2, \dots, n$,通过利用局面和行动的积得到棋子的行动概率,即 $S \times a \times S \rightarrow [0, 1]$, g^i ,其中 $i=1, 2, \dots, n$,通过局面及行棋动作返回一个强化收益函数,即 $S \times a \times S \rightarrow R$;在军棋机器多棋子协同博弈系统中,每个棋子的行动概率都是根据所有棋子协同作战的结果得出的,强化收益函数也是根据棋子协同作战后返回最高收益,根据棋局局面映射到棋子作战行为,采取协同策略 π 。多棋子协同博弈强化学习方法转化为在棋子局面状态空间到棋子协同作战行为的映射学习。

Q强化学习算法是一种无监督的通过系统不断探索经验进而提高自身作战能力的算法,不需要人为干预和状态转移函数。当系统进行行为决策时只需要选出返还结果 $Q(s, a_i)$ 的最大值,进而

简化了多棋子协同博弈系统的决策过程, Q 强化学习方法扩展为多棋子强化学习算法。

$$Q_t^i(S_t^i, a_t) = (1 + \alpha)Q_{t-1}^i(S_t^i, a_t) + \alpha[r_t^i(S_t^i, a_t) + \gamma\pi^1(S_{t+1}), \dots, \pi^n(S_{t+1})Q_{t-1}^i(S_{t+1})]$$

$$\pi^1(S_{t+1}), \pi^2(S_{t+1}), \dots, \pi^n(S_{t+1})Q_t^i(S_{t+1}) = \sum_{a^i \in a} \dots$$

$$\sum_{a^i \in a} P_t^1(S_{t+1}, a^1), P_t^2(S_{t+1}, a^2), \dots, P_t^n(S_{t+1}, a^n) \times Q_{t-1}^i(S_{t+1}^i, a^1, a^2, \dots, a^n)$$

式中: S_t^i 为多棋子协同博弈系统中棋子 i 的状态变量; $a_t = \{a^1, a^2, \dots, a^n\}$ 为多棋子协同博弈系统中多棋子的协同行为序列; S_{t+1} 为行动后的下一步棋盘的局面, 局面状态通过函数 $S_{t+1} = f_t^i(S_t, a_t)$ 进行转移; 棋子 i 的作战行为通过模拟回报策略 π_i 得到; $P_t^i(S_{t+1}, a^i)$ 是棋子 i 在局面 S_{t+1} 的情况下模拟选择出 a^i 走步的概率。

2 UCT 指导激活棋子

棋子间联合作战, 通过反馈学习进行小规模棋子合作规划, 实现对敌方欺诈是可以实现的。但是, 军棋机器博弈系统中, 每方都有 25 个棋子, 棋子种类又分为 12 种, 假设在军棋机器博弈过程中, 每个棋子有 A 种走步, 而在 Q 学习的过程中, 25 个棋子都会返回一个 $Q_t^i(s_t, a_t)$ 表, 那么其全部状态空间大小为 A^{25} 。由以上分析可以看出, 棋盘上面棋子数越多, 空间复杂度就越高, 并且成指数倍增长, 进而导致军棋机器博弈系统出现“维度灾难”问题^[10]。

军棋计算机博弈系统根据军棋棋子数量多造成的“维度灾难”问题, 系统参考神经元特征, 例如举一下胳膊时, 并不会激活身体所有神经元, 提出一种优先利用 UCT 搜索算法剪掉大部分棋子的行动序列, 而把搜索学习算法重点放在几个高价值棋子之间的协作上面的方法。

2.1 UCT 算法

2006 年匈牙利学者 Levente Kocsis 和加拿大大学者 Csaba Szepesvari 共同合作提出一种 MCTS 算法与 UCB 公式相结合的 UCT 搜索算法。UCB(upper confidence bound) 搜索算法最初设计的目的是用来解决多臂老虎机之类的问题。多臂老虎机问题是指在多臂老虎机上, 通过摇动每个老虎机手臂都可以得到一定的收益, 如果玩家想要摇到最大的收益就要根据以前摇臂的经验去判断摇哪个手臂得到的收益值最大。UCB 搜索算法是一种离线的强化学习策略, 即根据以往的经验知识去帮助系统做出未来的决策, UCB 算法中的

UCB 值又被称为上限置信区间值, UCB 公式为

$$UCB_i = \bar{V}_i + k \sqrt{2 \ln(n/n_i)}$$

式中: \bar{V}_i 作为老虎机第 i 支手臂获得的平均收益; $k \sqrt{2 \ln(n/n_i)}$ 是老虎机在搜索其他手臂时探索的未知收益, 存在一定未知风险; k 值作为一个风险参数, 当 k 值设大时, 注重对其他最优结果的探索, 当 k 值设小时, 老虎机博弈系统注重以前高收益手臂, 系统偏向于保守。 N 是所有手臂的访问次数, n_i 则表示任意 i 支手臂的访问次数。

UCT 搜索算法是在蒙特卡洛搜索的基础上, 根据模拟出来的结果优先选择出胜率高的根节点, 最后根据选择出来的胜率较高的根节点进行多次访问, 增加高价值节点的访问次数, 以最高访问数的节点作为最优值^[11]。

2.2 发现高价值棋子

军棋机器博弈系统采用多棋子协同博弈, 开局时, 局面中有 50 个棋子, 60 个棋位。并非所有棋子都进行系统协同作战, 系统根据不同的布局采用局部令、强、中、弱棋子的搭配进行配合出战, 而 UCT 搜索算法则在系统中起到筛选作用, 为棋子协同作战提供指导方向, 军棋机器博弈多棋子协同作战系统将搜索重点布局在高价值棋子上面, 进而减少搜索的棋子数量, 避免了搜索系统出现“维数灾难”问题。通过多棋子 Q 强化学习方法加强棋子之间策略学习, 以使得多棋子可以配合出战, 例如采用强子和弱子配合引诱对方的中子来吃弱子, 而灭掉对方的中子^[12]。

UCT 搜索算法指导激活多棋子强化 Q 学习算法的流程如下所示:

- 1) 初始化系统的值函数、学习因子以及马尔可夫过程的状态集合 $Q_t^i(s_t, a_t)$;
- 2) 将当前节点作为博弈树系统的根节点, 从根节点进行扩展;
- 3) 通过 UCT 公式计算各个节点的 UCB 值, 并为节点赋值, 选择出 UCB 较高的节点作为扩展结点进行搜索;
- 4) 依次搜索节点, 并检查是否为叶子节点, 若不是叶子节点则重复第 3) 步搜索;
- 5) 核查该节点被搜索次数是否达到系统规定访问次数, 如果达到规定的访问次数, 则将该节点进行扩展, 并在访问次数上加 1;
- 6) 筛选出当前高价值节点状态下激活 1 的棋子, 按照系统策略 $\pi^*(s_t) = \arg \max_{a_t} Q_t^i(s_t, a_t)$ 计算出当前状态下的动作 a_t , 并对下一局面状态 S_{t+1} 进行观察;
- 7) 根据多棋子强化学习 Q 值迭代公式迭代

出新的局面-行为对的值函数 $Q_t^i(s_i, a_i)$;

8) 判断是否满足学习的终止条件, 若满足则结束学习, 不满足则返回第2)步继续迭代。判定当前状态是否满足多棋子Q学习算法的终止条件, 如果满足, 则停止学习, 否则返回第2)步。

2.3 构建军棋博弈系统

由于军棋在开局的时候对敌方信息处于未知状态。因此在军棋博弈系统中先跟据人类先验知识以及蒙特卡罗模拟对棋盘进行完备化^[13]。搜索模块采用UCT搜索算法模拟走步, 返回价值高并且模拟次数多的走步, 并激活行动价值高的走步。通过Q学习算法对棋子协同矩阵优化, 选择出最优走步。军棋博弈系统流程图如图2所示。

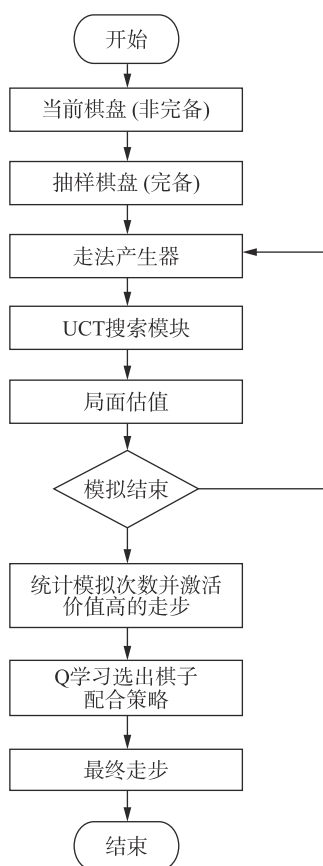


图2 军棋博弈搜索系统流程

Fig. 2 Flow chart of the military chess game search system

3 实验结果分析

为了测试多棋子Q强化学习算法在实战中是否占据优势, 本文采用与其他算法系统进行对打模式测试本算法的实用性, 通过分先后手对战极大极小值搜索算法、alpha-beta搜索算法、MCTS搜索算法以及UCT搜索算法各1000局并不断调整多棋子Q强化学习算法的参数, 以得到最优值, 实验结果如图3~7所示。

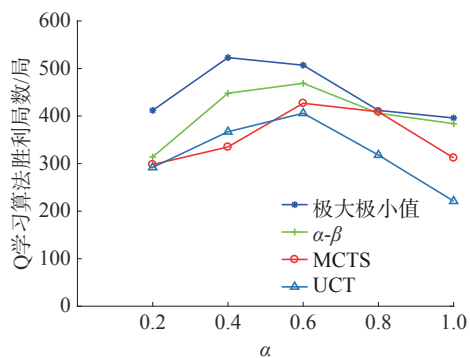


图3 γ 为固定值 0.2, α 取值参数对比

Fig. 3 The comparison chart of γ is the value of 0.2 with α is a parameter

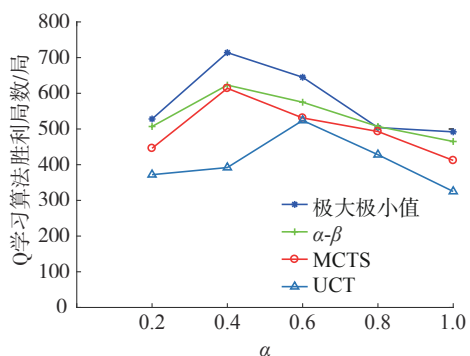


图4 γ 为固定值 0.4, α 取值参数对比

Fig. 4 The comparison chart of γ is the value of 0.4 with α is a parameter

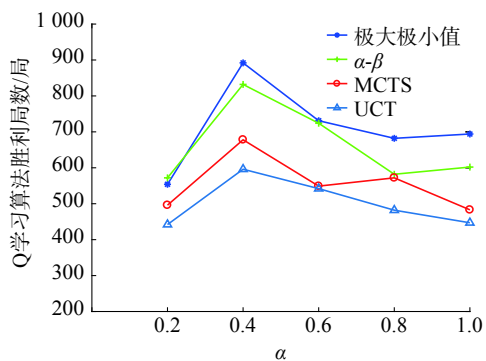


图5 γ 为固定值 0.6, α 取值参数对比

Fig. 5 The comparison chart of γ is the value of 0.6 with α is a parameter

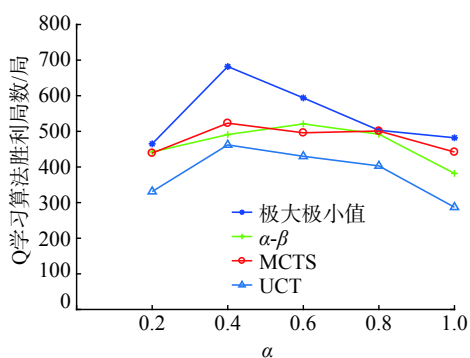
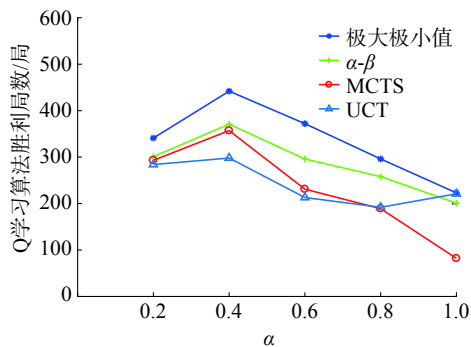


图6 γ 为固定值 0.8, α 取值参数对比

Fig. 6 The comparison chart of γ is the value of 0.8 with α is a parameter

图 7 γ 为固定值 1, α 取值参数对比Fig. 7 The comparison chart of γ is the value of 1 with α is a parameter

由以上实验结果可知, 参数 α 和 γ 对系统的作战能力影响极大, α 参数值影响着系统学习的

效率, 值越大, 学习效率就越高。 γ 参数影响着系统学习过程中经验的影响力, γ 值越大, 以往的经验就显得越重要。当 $\alpha=0.4$, $\gamma=0.6$ 时多棋子 Q 强化学习系统对战其它算法系统的结果相对最好。

单子作战系统重防守、轻攻击, 很难做到主动攻击。而加入多棋子协同作战策略后的军棋机器博弈系统能够自我组织协同策略, 多棋子联合作战, 增加了系统的主动进攻意识。本文为了测试军棋多棋子协同博弈系统主动作战意识是否提高, 采用了与其它搜索算法系统对战的模式, 并统计出棋子作战过程中主动进攻的步数进行对比研究, 研究发现该系统能够达到攻守平衡的状态, 统计结果如表 1 所示。

表 1 算法系统走步对比

Table 1 The comparison table of algorithm system walking

步

方式	极大极小搜索	α - β 搜索	MCTS搜索	UCT搜索	Q强化学习
进攻	7	9	13	12	21
防守	24	25	29	27	18
磨棋	29	23	24	26	19

军棋多棋子协同博弈系统由于棋子种类多的问题导致的“维度灾难”问题成为系统对战成败的关键性因素, 针对此问题系统采用多种搜索算

法进行高价值棋子激活, 分别对战 500 局后, 在军棋博弈的 69 步范围之内计算出每一步的平均搜索时间如图 8 所示^[14]。

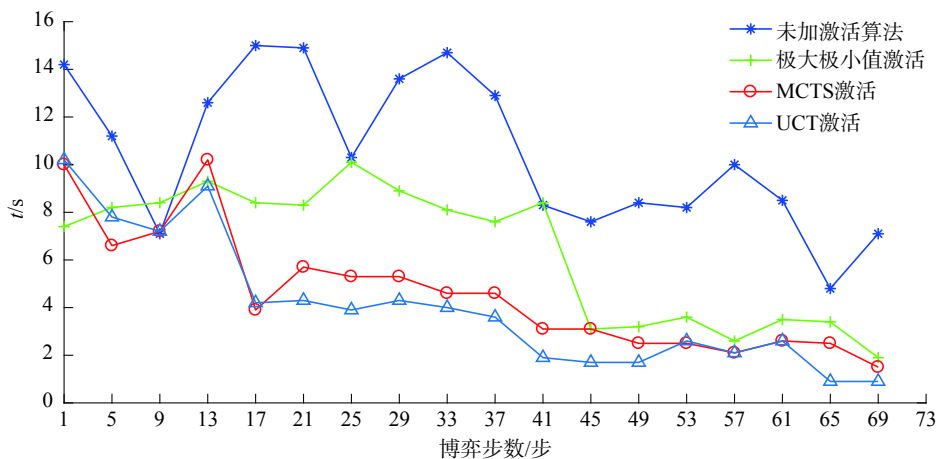


图 8 激活算法系统走步搜索时间对比

Fig. 8 The comparison chart of Activation algorithm system walking search time

由图 8 可知, 在未加激活算法时搜索的时间较长, 而军棋机器博弈比赛是在限定的时间内搜索到走步, 否则认为系统搜索超时判负, 加入各种激活算法后, 系统搜索时间明显降低, 尤其在加入 UCT 搜索算法时, 系统搜索时间降低最明显。

4 结束语

本文利用 Q 学算法优化迭代更新军棋计算机

博弈系统中多棋子协同矩阵, 增加了多棋子协同作战的意识, 增强了博弈系统主动出击诱导敌方的进攻趋势。利用部分重要棋子激活的方法解决了棋子数量巨大造成的“维度灾难”问题, 使得搜索时间得到提升, 避免了军棋计算机博弈系统在正式比赛过程中由于超时被判负的情况。未来将考虑利用计算智能算法加入此系统, 进一步优化多棋子协同矩阵^[15]。

参考文献:

- [1] 陶九阳, 吴琳, 胡晓峰. AlphaGo 技术原理分析及人工智能军事应用展望 [J]. 指挥与控制学报, 2016, 2(2): 114-120.
TAO Jiuyang, WU Lin, HU Xiaofeng. Principle analysis on AlphaGo and perspective in military application of artificial intelligence[J]. Journal of command and control, 2016, 2(2): 114-120.
- [2] CHEN J X. The evolution of computing: AlphaGo[J]. Computing in science & engineering, 2016, 18(4): 4-7.
- [3] 李翔, 姜晓红, 陈英芝, 等. 基于手牌预测的多人无限注德州扑克博弈方法 [J]. 计算机学报, 2018, 41(1): 47-64.
LI Xiang, JIANG Xiaohong, CHEN Yingzhi, et al. Game in multiplayer no-limit texas hold'em based on hands prediction[J]. Chinese journal of computers, 2018, 41(1): 47-64.
- [4] 王亚杰, 邱虹坤, 吴燕燕, 等. 计算机博弈的研究与发展 [J]. 智能系统学报, 2016, 11(6): 788-798.
WANG Yajie, QIU Hongkun, WU Yanyan, et al. Research and development of computer games[J]. CAAI transactions on intelligent systems, 2016, 11(6): 788-798.
- [5] HONG Yiguang, CHEN Guanrong, BUSHNELL L. Technical communique: distributed observers design for leader-following control of multi-agent networks[J]. Automatica, 2017, 44(3): 846-850.
- [6] 滕雯娟. 基于虚拟遗憾最小化算法的德州扑克机器博弈研究 [D]. 哈尔滨: 哈尔滨工业大学, 2015.
TENG Wenjuan. Research on texas poker game based on counterfactual regret minimization algorithm[D]. Harbin: Harbin Institute of Technology, 2015.
- [7] VAN HASSELT H, GUEZ A, SILVER D. Deep reinforcement learning with double q-learning[C]//Proceedings of the 30th AAAI Conference on Artificial Intelligence. Arizona, USA, 2015.
- [8] 史晓茹, 侯媛彬, 张涛. 不完全信息博弈的机器人对抗决策 [J]. 智能系统学报, 2011, 6(2): 147-151.
SHI Xiaoru, HOU Yuanbin, ZHANG Tao. The decision-making system of robots based on an incomplete information game[J]. CAAI transactions on intelligent systems, 2011, 6(2): 147-151.
- [9] MICHAEL L. INCZE, SCOTT R. SIDELEAU, CHRIS GAGNER, and CHARLES A. PIPPIN. (2015). Communication and collaboration of heterogeneous unmanned systems using the joint architecture for Unmanned Systems (JAUS) standards[C]//OCEANS 2015- Genova. IEEE.
- [10] 乔林. 多智能体系统中的 Q 学习算法研究 [D]. 南京: 南京邮电大学, 2012.
QIAO Lin. Study of Q-learning algorithm in multi-agent system[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2012.
- [11] 梁国军, 谢垂益, 胡伶俐, 等. UCT 算法在不围棋博弈中的实现 [J]. 韶关学院学报, 2015, 36(8): 17-21.
Liang Guojun, Xie Chuiyi, Hu Lingli, et al. An implementation of UCT algorithm for nogo game[J]. Journal of Shaoguan university, 2015, 36(8): 17-21.
- [12] 孟坤, 王俊, 闫桐. 一种基于经验知识的军棋博弈算法设计与实现 [J]. 智能计算机与应用, 2017, 7(2): 66-69.
MENG Kun, WANG Jun, YAN Tong. Design and implementation of a military chess game algorithm based on knowledge experience[J]. Intelligent computer and application, 2017, 7(2): 66-69.
- [13] 孙英龙. 非完美信息博弈算法研究与军棋博弈系统设计与实现 [D]. 沈阳: 东北大学, 2013.
SUN Yinglong. The study on imperfect information game and design and implementation of military chess system[D]. Shenyang: Northeastern University, 2013.
- [14] 王学厚. 群体智能优化的计算模式和方法研究与应用 [D]. 北京: 华北电力大学, 2011.
WANG Xuehou. Research on computational mode of swarm intelligent optimization and applications[D]. Beijing: North China Electric Power University, 2011.
- [15] 张小川, 桑瑞婷, 周泽红, 等. 一种基于双通道卷积神经网络的短文本分类方法 [J]. 重庆理工大学学报(自然科学), 2019, 33(1): 45-52.
ZHANG Xiaochuan, SANG Ruiting, ZHOU Zehong, et al. A Short Text Classification Method Based on Two Channel Convolutional Neural Network[J]. Journal of Chongqing University of Technology(Natural Science), 2019, 33(1): 45-52.

作者简介:



张小川, 教授, 中国人工智能学会机器博弈专委会主任、《CAAI Transactions on Intelligence Technology》副主编, 人工智能系统研究所所长、两江人工智能学院副院长, 主要研究方向为机器博弈、智能机器人、软件工程。主持纵向项目 35 项、横向项目 45 项, 获省部级奖项 4 项。发表学术论文 90 余篇。



王宛宛, 硕士研究生, 主要研究方向为人工智能、机器博弈。