

Research on Einstein Chess Algorithm Based on Heuristic Pruning Optimization Monte Carlo Tree Search

Haodi Yang

College of Computer and Software Engineering, Liaoning University of Science and Technology, Anshan, Liaoning, 114051, China

Abstract

To address the inefficiency and excessive simulation overhead of standard Monte Carlo Tree Search (MCTS) in Einstein Chess, this paper proposes a heuristic-pruned MCTS algorithm (HP-MCTS). The proposed method preserves the classical four-stage MCTS loop (selection, expansion, simulation, and backpropagation), while introducing heuristic guidance in selection and dynamic pruning in expansion. Specifically, an improved Heuristic UCT (HUCT) is designed by adding a heuristic correction term to the conventional exploitation-exploration balance. In expansion, low-value actions are filtered out using a dynamic threshold with a safety-retention mechanism. Considering the stochastic and constrained characteristics of Einstein Chess, a composite heuristic function is constructed with three factors: mobility probability weight, target-approach gain, and defensive-penalty term. Parameter tuning and complexity analysis are also provided. Under a fixed decision time budget (5 seconds per move), HP-MCTS is compared with standard MCTS through repeated matches. Experimental results show that HP-MCTS reduces unnecessary node expansion, increases effective search depth, and improves winning rate. Under heavy pruning, the winning rate reaches 82.0%, while expanded nodes are reduced by 58.2% compared with standard MCTS. The study confirms the practical value of heuristic pruning for complex stochastic game search and provides useful insights for Einstein Chess AI design.

Keywords

Einstein Chess; Monte Carlo Tree Search; Heuristic Pruning; Game AI; algorithm optimization

基于启发式剪枝优化蒙特卡洛树搜索的爱因斯坦棋算法研究

杨浩笛

计算机与软件工程学院辽宁科技大学, 中国·辽宁 鞍山 114051

摘要

针对标准蒙特卡洛树搜索 (Monte Carlo Tree Search, MCTS) 在爱因斯坦棋中存在的搜索冗余高、有效深度不足和时间预算利用率低等问题, 本文提出一种基于启发式剪枝的改进算法 (Heuristic-Pruned MCTS, HP-MCTS)。该方法在保留 MCTS “选择—扩展—模拟—回传” 主流程的基础上, 将局面启发式评估融入选择与扩展环节; 在选择阶段构建改良 HUCT (Heuristic UCT) 公式, 在探索项与经验价值项之外增加启发式修正项; 在扩展阶段采用动态阈值与安全保留机制进行分支剪枝, 从而减少低价值分支的后续模拟开销。围绕爱因斯坦棋 “随机骰子驱动+单步方向约束+近似替代选子” 的博弈特征, 本文构建了由 “可行动概率权重” “目标逼近收益” “关键防线惩罚” 组成的综合评估函数, 并给出参数标定和复杂度分析。实验在统一思考时限 (5s/步) 条件下, 将 HP-MCTS 与标准 MCTS 进行多轮对弈比较。结果表明: HP-MCTS 可显著降低扩展节点总数, 提升平均搜索深度与对战胜率; 在重度剪枝配置下, 算法胜率达到 82.0%, 扩展节点较标准 MCTS 下降 58.2%, 验证了启发式剪枝在复杂随机博弈中的有效性。本文研究可为爱因斯坦棋 AI 与同类随机对抗问题的搜索优化提供参考。

关键词

爱因斯坦棋; 蒙特卡洛树搜索; 启发式剪枝; 博弈 AI; 算法优化

1 引言

爱因斯坦棋 (Einstein Chess) 是一类兼具随机性与策略性的双人零和棋类博弈。与国际象棋、围棋等完全信息确定性博弈不同, 爱因斯坦棋的行动由骰子结果触发, 玩家每回合可行动作受 “点数匹配或就近替代” 规则约束, 这导致

局面评估不仅要考虑空间控制与攻防态势, 还要考虑未来可行动概率分布。也正因为如此, 爱因斯坦棋在学术上被视为研究 “随机机制下策略搜索” 的典型平台。

蒙特卡洛树搜索 (MCTS) 在大规模博弈决策中具有重要地位。其不依赖复杂手工评估函数, 通过大量随机模拟估计动作价值, 适合高维状态空间问题。但在爱因斯坦棋场景下, 标准 MCTS 存在三个突出瓶颈:

(1) 分支利用率低: 扩展时对合法动作近似等价处理,

【作者简介】杨浩笛 (2004-), 男, 中国辽宁沈阳人, 在读本科生, 从事机器博弈与智能搜索算法优化研究。

导致大量弱势分支被重复探索；

(2) 有效深度不足：固定时间预算下，冗余模拟压缩了关键路径的搜索深度；

(3) 随机性放大误差：骰子导致短期回报波动较大，若无先验引导，早期采样噪声易误导搜索方向。

为解决上述问题，本文提出基于启发式剪枝的 HP-MCTS 算法。核心思想是：在不改变 MCTS 统计学习本质的前提下，把“领域可解释先验”注入搜索过程，以更少计算预算获得更高质量决策。

2 相关理论基础

2.1 爱恩斯坦棋规则与博弈特性

本文采用 5×5 棋盘、双方各 6 枚编号棋子的标准规则。任意时刻，玩家由骰子点数 r 驱动行动：优先移动编号为 r 的棋子；若该编号棋子已被吃，则选择与 r 最近且仍存活的编号棋子行动。每枚棋子仅可向对方大本营方向移动一格，可选方向为横、竖、斜（受边界与占位限制）。玩家满足以下任一条件即获胜：

任一己方棋子到达对方起始角；

吃光对方全部棋子。

该规则导致博弈具有两类耦合复杂性：

状态-动作耦合：局面变化会改变未来可行动编号集合，进而改变概率结构；

攻防-概率耦合：高机动棋子并不总是高价值棋子，需结合可掷中概率与位置威胁综合判断。

2.2 蒙特卡洛树搜索基本原理

标准 MCTS 由四步构成循环迭代：

选择：从根节点沿 UCT 公式向下选择，直至叶节点；

扩展：对未终止的叶节点生成合法子节点；

模拟：从新节点随机走子至对局结束，得到胜负结果；

反向传播：将结果回溯更新路径上节点的访问次数与胜率。

MCTS 无需复杂评估函数，但在高分支、高深度场景下模拟开销巨大。

UCT 常用形式为：

$$UCT(v_i, v) = \bar{Q}(v_i) + C \sqrt{\frac{\ln(N(v_p) + 1)}{N(v_i) + 1}}$$

其中，

$$\bar{Q}(v_i) = \frac{W(v_i)}{N(v_i)}$$

v_p 为 v_i 的父节点。UCT 在“利用—探索”之间平衡良好，但未利用领域先验，扩展阶段也缺乏筛选机制。

标准 MCTS 优点是通用性强、实现简洁，但在高分支随机博弈中容易出现无效模拟比例高的问题。

2.3 启发式剪枝基本思想

启发式剪枝并非替代 MCTS，而是在关键环节对动作进行价值预判：

对明显劣势动作降低被选择概率或直接剪去；

保留高潜力动作获得更多仿真样本；

在有限预算下提升“有效模拟比例”。

其关键在于：评估函数要可计算、可解释、与胜负相关，且剪枝策略需具备安全兜底，避免过度剪枝造成策略退化。

3 基于启发式剪枝的 MCTS 算法设计

3.1 算法整体框架

本文提出 HP-MCTS 算法，在 MCTS 的选择与扩展阶段加入启发式评估：

- 选择阶段：使用改进 HUUCT 公式，引入启发式先验信息；

- 扩展阶段：根据启发式分值进行动态剪枝，剔除低分节点；

- 模拟与回溯流程保持 MCTS 原有机制。

3.2 启发式评估函数设计

对局面 s 定义综合评分：

$$H(s) = \alpha \sum_{i \in Alive} w_i + \beta \sum_{i \in Alive} (D_{max} - d_i) - \gamma \cdot Penalty(s)$$

其中：

Alive：当前存活棋子集合；

w_i ：棋子 i 的行动概率权重（基于骰子与“最接近替代”规则估计）；

d_i ：棋子 i 到对方目标点的切比雪夫距离；

D_{max} ：棋盘最大距离常数（ 5×5 中取 4 或经归一化处理）；

Penalty(s)：关键防线暴露、可被一步吃子、核心路径被封锁等惩罚项；

α, β, γ ：权重系数。

为避免量纲差异，本文在实现中对三项进行区间归一化后再线性融合。该设计兼顾“走得动”“走得快”“不冒进”三类目标，较单一距离启发式更稳健。

3.3 改进选择策略

在中加入启发式修正项，得到公式：

$$HUUCT(v_i, v) = \bar{Q}(v_i) + C \sqrt{\frac{\ln(N(v_p) + 1)}{N(v_i) + 1} + K} \cdot \frac{\tilde{H}(v_i)}{(N(v_i) + 1)^{\lambda}}$$

其中：

$\tilde{H}(v_i)$ 为对子节点对应局面的归一化启发式分值；

K 为启发式权重因子；

$\lambda \in [0, 1]$ 为衰减系数（访问增多后先验影响逐步减弱）。

改良理由：

若直接把启发式项 $\tilde{H}(v_i)$ 线性叠加到 UCT，早期可能过强、后期仍持续干扰；

引入 $(N+1)^{-\lambda}$ 后，先验只在“数据不足”阶段主导，样本充足时由经验价值 \bar{Q} 接管；

保持与 UCT 一致的可解释性和工程可控性，通过启发

式分值引导搜索方向，减少无效探索。

3.4 扩展阶段动态剪枝

对当前节点合法动作集合 $A(s)$ ，先计算每个动作后继局面的启发式得分 $h(a)$ ，再执行动态阈值剪枝：

$$\theta = \mu - \eta\sigma$$

其中 μ 为平均得分， σ 为标准差， η 为剪枝强度系数 $h(a) < \theta$ 若则标记为剪枝候选。

为避免极端误剪，加入安全机制：

最低保留数：至少保留个 $K_{min}=2$ 动作；

胜势保留：若动作可直接吃子或形成到达目标点的一步威胁，则强制保留；

差距保护：当动作分数与最优分差小于 δ 时不剪枝。

该机制在压缩分支的同时保证基本战术完整性。

3.5 算法执行流程

初始化根节点，设置时间预算 T ；

在限定时间内循环执行：

(1) 使用 HUCT 策略选择节点；

(2) 对候选节点计算启发式分值并剪枝；

(3) 对保留节点执行随机模拟至终局；

(4) 回传结果更新 N, W, \bar{Q} ；

3. 选择根节点下访问次数最多的动作作为最终决策。

3.6 复杂度分析

设单次决策总迭代次数为 M ，平均分支因子为 b ，剪枝后分支因子为 b' ($b' < b$)。标准 MCTS 扩展与模拟开销近似与 $M \cdot b$ 相关而 HP-MCTS 可降为 $M \cdot b'$ 。虽然每次扩展新增启发式评估（代价 C_{heu} ），但其计算远低于完整模拟开销 C_{sim} ，因此总体收益表现为：

$$\Delta Cost \approx M \cdot [(b-b') \cdot C_{sim} - b' \cdot C_{heu}]$$

当 $C_{sim} \gg C_{heu}$ 时，剪枝收益显著。

4 实验与结果分析

4.1 实验环境与实现设置

开发平台：Eclipse；

硬件环境：AMD R9 处理器，32GB 内存；

对弈约束：双方单步最大思考时间统一为 5s；

基础参数：标准 MCTS 探索常数 $C=2$ ；

HP-MCTS 参数（调优后）：

$$\alpha=1.5, \beta=2.0, \gamma=1.2, K=0.5, \lambda=0.5, \eta=1.0$$

剪枝安全机制：

$$K_{min}=2, \delta=0.15$$

为提升公平性，实验采用相同随机种子并交换先后手，结果按多局平均统计。

4.2 对比实验设计

1. 标准 MCTS：UCT 选择，无启发式项，无剪枝；

2. HP-MCTS（轻度剪枝）： $\eta=0.7$ ，保留动作相对较多；

3. HP-MCTS（重度剪枝）： $\eta=1.2$ ，更激进地压缩低分支。

评价指标包括：平均搜索深度、扩展节点总数、对战胜率、剪枝率。

4.3 实验结果与分析

表 1 HP-MCTS 与标准 MCTS 性能对比

算法类型	思考时间 / 步	平均搜索深度	扩展节点总数	对战胜率	剪枝率
标准 MCTS	5s	12.4	45210	50.0%	0%
HP-MCTS (轻度剪枝)	5s	16.8	31540	68.5%	30.2%
HP-MCTS (重度剪枝)	5s	21.3	18900	82.0%	58.2%

从结果看，HP-MCTS 在相同预算下表现出明显优势。与标准 MCTS 相比：

轻度剪枝已可将扩展节点减少约 30.2%，并把深度提升至 16.8；

重度剪枝把扩展节点减少到 18900，降幅约 58.2%，深度提升到 21.3；

胜率从 50.0% 提高到 82.0%，说明计算资源更集中于高价值路径后，决策质量显著提升。

4.4 结果讨论

效率提升原因：剪枝直接减少了低价值分支的扩展和后续模拟，释放出的预算用于关键节点反复采样，使统计估计更稳定。

胜率提升原因：改良 HUCT 在搜索早期即可引导到更有潜力的战术方向，降低了随机模拟噪声对策略的误导。

重要性与有效性：在安全保留机制约束下，激进剪枝并未导致战术性漏搜，反而提高了有限时间内的深度利用率。

潜在风险：若启发式函数偏差较大，或 η 设置过高，仍可能误剪“短期劣势、长期优势”动作，因此参数调优与保底机制至关重要。

5 结论

本文面向爱恩斯坦棋这一高随机、强约束、策略空间复杂的博弈场景，提出了基于启发式剪枝的 HP-MCTS 算法。与标准 MCTS 相比，本文方法通过改良 HUCT、动态剪枝双机制，在不改变 MCTS 核心统计本质的前提下，显著提升了搜索效率与对局胜率。实验结果表明，在统一 5s / 步预算下，HP-MCTS 可明显降低扩展节点规模、提升有效搜索深度，并在对弈中取得更高胜率，验证了启发式剪枝在随机博弈中的工程价值与可行性。

参考文献

- [01] 徐心和，邓志立，王骄，等。机器博弈研究面临的各种挑战[J]. 智能系统学报，2008, 3 (4): 288-293.
- [02] 王亚杰，邱虹坤，吴燕燕，等。结合神经网络的改进 UCT 在国际跳棋中的应用[J]. 重庆理工大学学报(自然科学), 2021, 35 (7): 112-118.
- [03] Silver D, Huang A, Maddison C J, et al. Mastering the Game of Go with Deep Neural Networks and Tree Search [J]. Nature, 2016, 529 (7587): 484-489.