

HBM: 基于双向蒙特卡洛树和混合贪心策略的序列反转求解方法

刘威伟, 王颖, 李云菲, 王希胤*

华北理工大学理学院, 河北 唐山

收稿日期: 2026年2月2日; 录用日期: 2026年2月23日; 发布日期: 2026年3月2日

摘要

基因序列重排是计算生物学中的经典组合优化问题, 目标是以最少的子串反转操作将给定序列转化为目标序列, 其在解析基因组进化机制方面具有重要意义。然而, 该问题具有NP难特性, 求解极具挑战。为提高求解效率与质量, 本文提出一种基于双向蒙特卡洛树与混合贪心策略的序列反转求解模型(HBM)。该模型首先通过预处理去除已匹配的前后缀, 压缩问题规模; 进而采用分层递进求解框架, 依次利用双向广度优先搜索获取初始路径, 在未获解时启用混合贪心策略, 融合五种启发式指标动态选择局部操作, 最终通过双向蒙特卡洛树搜索平衡探索与利用, 以获取高质量解。在特定DNA序列数据集上的实验表明, HBM在反转路径长度、时间-质量性价比及解质量稳定性方面均优于现有方法。

关键词

序列反转, 蒙特卡洛树搜索, 双向广度优先搜索, 混合贪心策略, 多目标优化, 动态调整权重, 组合优化

HBM: A Sequence Reversal Solving Method Based on Bidirectional Monte Carlo Tree and Hybrid Greedy Strategy

Weiwei Liu, Ying Wang, Yunfei Li, Xiyin Wang*

School of Science, North China University of Science and Technology, Tangshan Hebei

Received: February 2, 2026; accepted: February 23, 2026; published: March 2, 2026

Abstract

Gene sequence rearrangement is a classic combinatorial optimization problem in computational

*通讯作者。

文章引用: 刘威伟, 王颖, 李云菲, 王希胤. HBM: 基于双向蒙特卡洛树和混合贪心策略的序列反转求解方法[J]. 计算生物学, 2026, 16(1): 1-14. DOI: 10.12677/hjcb.2026.161001

biology, which aims to transform a given sequence into a target sequence with the minimum number of substring reversal operations and plays a crucial role in deciphering the mechanisms of genome evolution. However, this problem is characterized by NP-hardness, making its solution extremely challenging. To improve the efficiency and quality of the solution, this paper proposes a sequence reversal solving model (HBM) based on a bidirectional Monte Carlo tree and a hybrid greedy strategy. The model first compresses the problem scale by preprocessing to remove matched prefixes and suffixes; then adopts a hierarchical progressive solving framework, sequentially using bidirectional breadth-first search to obtain an initial path, activating a hybrid greedy strategy when no solution is obtained, integrating five heuristic metrics to dynamically select local operations, and finally balancing exploration and exploitation through bidirectional Monte Carlo tree search to acquire high-quality solutions. Experiments on a specific DNA sequence dataset show that HBM outperforms existing methods in terms of reversal path length, time-quality cost-effectiveness, and solution quality stability.

Keywords

Sequence Reversal, Monte Carlo Tree Search (MCTS), Bidirectional Breadth-First Search (BBFS), Hybrid Greedy Strategy, Multi-Objective Optimization, Dynamic Weight Adjustment, Combinatorial Optimization

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

基因序列重排是计算生物学中的经典组合优化问题，其核心在于通过最少的子串反转操作，将给定基因序列转化为目标序列。该问题可有效模拟基因组进化中的大规模重排事件(如染色体倒位)，在基因组进化分析[1]、比较基因组学等领域具有重要理论与应用价值[2]。然而，该问题已被证明属于 NP 难问题[3]，随着序列长度增加，解空间呈指数级扩展，使得寻找最优反转路径极为困难[4]。

为应对上述挑战，学者已提出多种启发式与元启发式算法[5]。早期方法多依赖单一启发策略，如基于断点分析或排列分组的贪心算法，虽计算效率较高，但易陷入局部最优[6]。随后，遗传算法、蚁群优化等随机化策略被引入以增强全局搜索能力，却往往难以兼顾求解质量与时间成本[7]。近年来，MCTS [8] [9]因其在组合优化中的出色决策能力受到关注，并开始用于序列重排。然而，传统 MCTS 在处理基因序列中高频碱基、长程依赖等复杂结构时，仍存在搜索效率低、收敛速度慢的问题。

尽管已有诸多进展，现有方法仍存在以下局限：一是难以在搜索过程中平衡全局探索与局部利用[10]；二是面对不同结构特征(如高频元素占比差异)的序列时，求解稳定性不足[11]；三是在处理大规模或复杂序列时，时间成本与解质量之间的性价比仍有待提升[12]。

为此，本文提出一种分层递进式求解模型 HBM，融合 BBFS [13]、混合贪心策略与双向蒙特卡洛树搜索。该模型首先通过预处理缩减问题规模；进而采用分层策略：利用 BBFS 进行初始搜索，将单一连通通道转变为多连通通道，随后借助混合贪心策略综合多种启发式指标并动态调整权重以优化局部搜索获得初始解，最后以双向蒙特卡洛树搜索为核心，通过其双向协同与探索 - 利用平衡机制搜索高质量反转路径。在模拟 DNA 序列数据集上的实验结果表明，与现有方法相比，HBM 能够更有效地获得更短的反转路径，并在时间 - 质量性价比与解质量稳定性方面表现更优。

2. 模型及相关理论介绍

2.1. 整体模型框架

设起始序列 $P = (p_1, p_2, \dots, p_n)$ ，目标序列 $Q = (q_1, q_2, \dots, q_n)$ ，其中 $p_i, q_i \in \{A, T, C, G\}$ 。反转操作定义为将序列中 i 到 j ($1 \leq i < j \leq n$) 的子序列逆序。即 $rev(i, j, P) = (p_1, \dots, p_{i-1}, p_j, p_{j-1}, \dots, p_i, p_{j+1}, \dots, p_n)$ ，其中 $rev(i, j, \cdot)$ 表示逆序操作。问题的解为反转操作序列 O_1, O_2, \dots, O_k ，需满足 $rev(O_k, \dots, rev(O_2, rev(O_1, P))) = Q$ ，且 k 最小。

本文提出的 HBM 模型采用分层递进式架构，包含预处理层、双向广度优先搜索层(BBFS)、混合贪心层以及蒙特卡洛树搜索层(MCTS)四个核心模块，依托缩减问题规模、定位全局路径、优化局部操作及执行智能决策的分步协同，完成问题求解过程。模型流程图见图 1。

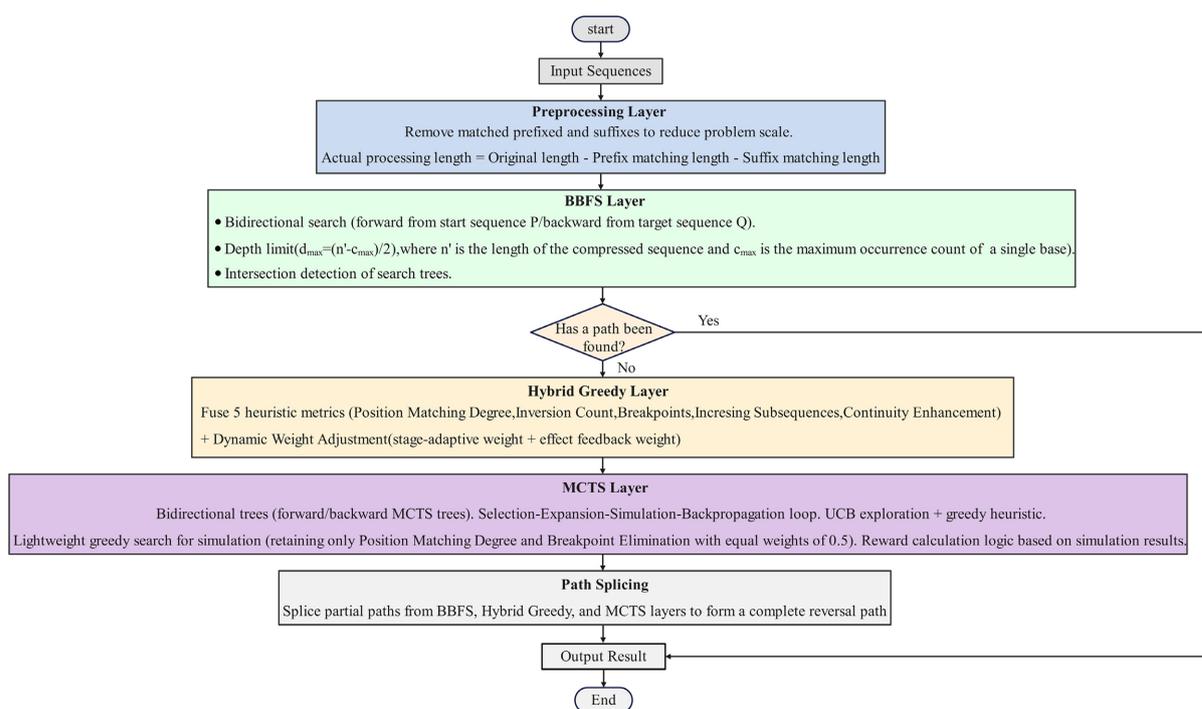


Figure 1. Model flow chart

图 1. 模型流程图

预处理层：基于局部匹配原理，分析输入序列并剔除前后缀已匹配部分以缩小问题规模。设序列长度为 n ，前缀匹配长度为 $l_p = \max\{k \mid p_1 = q_1, \dots, p_k = q_k\}$ ，后缀匹配长度为 $l_s = \max\{k \mid p_{n-k+1} = q_{n-k+1}, \dots, p_n = q_n\}$ ，则实际需要处理的子序列长度为 $n' = n - l_p - l_s$ 。

BBFS 层：从 P 和 Q 同时启动深度限制的广度优先搜索，构建起始搜索树 \mathcal{T}_S 和目标搜索树 \mathcal{T}_T ，搜索过程中设置扩展深度。含重复元素序列的全排列集合中，任意两个序列可达到最大的最少反转次数为 $n' - c_{\max}$ (c_{\max} 为序列中某一碱基的最大出现次数) [14]。由于 BBFS 采用双向协同搜索模式，起始端与终止端的搜索深度可分摊这一最大反转距离，故设置扩展深度的最大值为 $d_{\max} = \left\lfloor \frac{n' - c_{\max}}{2} \right\rfloor$ 。该设定通过量化“有效可操作序列长度”，既避免高频元素导致的“无效反转”，又平衡搜索空间膨胀与双向路径的交集检测效率。

混合贪心层：当 BBFS 层在扩展深度范围内无路径交集时启动，融合位置匹配度、逆序数、断点数、递增子序列及连续性增强评分，通过动态权重组合的综合评估函数，快速定位局部最优反转操作，为 MCTS 层提供初始解。

蒙特卡洛树搜索层：作为核心智能搜索模块，结合 UCB 探索策略[15] [16]和贪心启发式引导，通过“选择 - 扩展 - 模拟 - 回溯”循环平衡路径探索与利用，解决高频元素占比高、局部最优陷阱等复杂场景下的全局最优解搜索问题。

2.2. BBFS 算法

BBFS 算法从起始节点(P)和目标节点(Q)双向展开广度优先搜索，节点代表序列，两节点若可通过一次反转转换则连边，路径相遇时终止，以缩小搜索空间、提升效率。BBFS 层框架图见图 2。

设当前搜索树节点集合为 N_t ，下一层集合为 N_{t+1} 。搜索过程控制在扩展深度内。对于每个节点 $v \in N_t$ ，生成所有可能的反转操作 $rev(i, j)(1 \leq i < j \leq n')$ 。得到新状态 $s' = rev(i, j, s)$ ，若 s' 未在已访问集合 V 中，则加入 S_{t+1} 与 V 。

每层扩展后检测 T_s 的 N_t^s 与 T_T 的 N_t^T 交集：若 $N_t^s \cap N_t^T \neq \emptyset$ ，选择连接点 $s_{link} = \arg \min_{s \in N_t^s \cap N_t^T} (d_s(s) + d_T(s) - 1)$ ， $d_s(s)$ 、 $d_T(s)$ 分别为 P、Q 到 s 的反转次数，通过回溯父节点指针计算。最终重构路径 $Path = Path_s \oplus reverse(Path_T)$ ，其中 \oplus 为路径连接操作， $reverse(Path_T)$ 为目标端路径逆序处理。

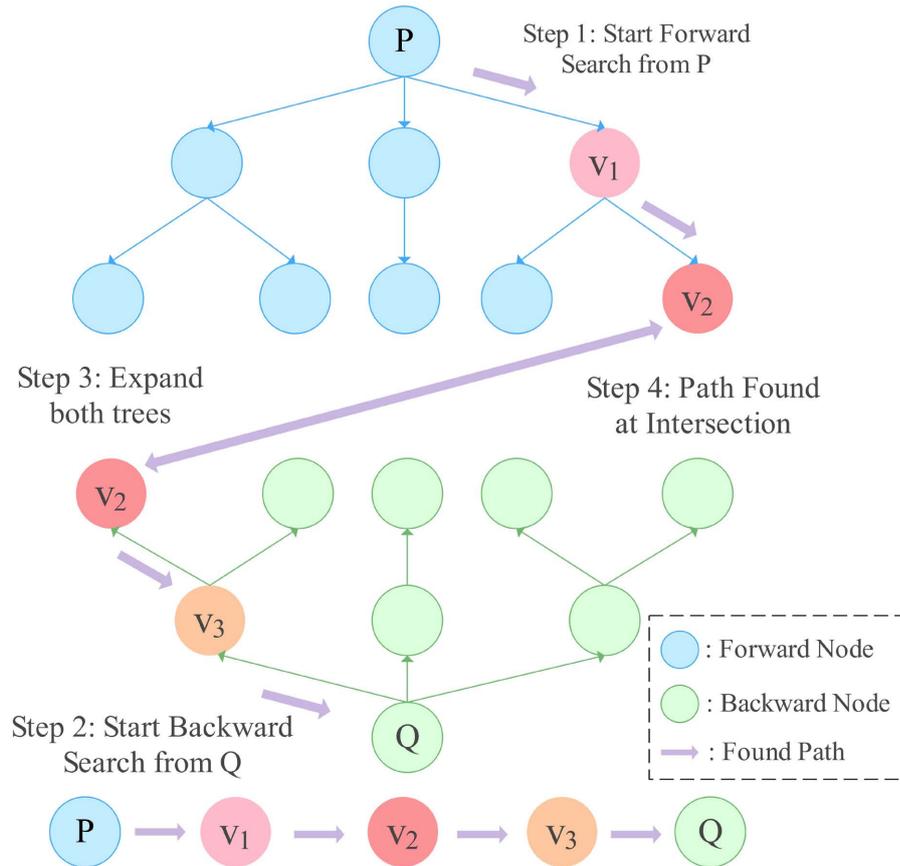


Figure 2. BBFS layer schematic diagram
图 2. BBFS 层架构图

2.3. 混合贪心策略

混合贪心搜索通过多目标优化选择局部最优反转操作，核心是融合多启发式指标的综合评分函数与动态权重机制。混合贪心策略层架构见图 3。

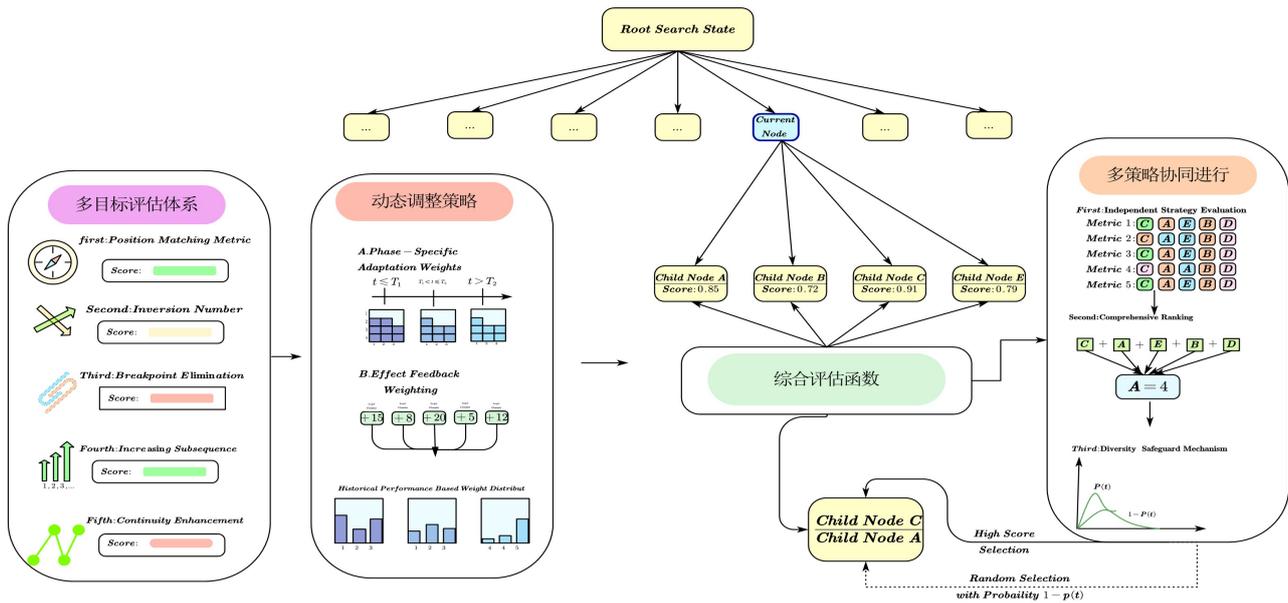


Figure 3. Hybrid greedy strategy layer schematic diagram
图 3. 混合贪心策略层架构图

2.3.1. 多目标评估体系

设当前状态为 s ，目标状态为 Q ，对应的排列为 π (序列元素映射为有序索引)，长度为 n' 。对于候选反转操作 $rev = (i, j)$ ，定义五个核心评估指标：

1) 位置匹配度指标[17]：基于汉明距离，衡量元素在正确位置的比例，序列的位置匹配度与位置匹配改进量分别定义为：

$$M(\pi) = \frac{1}{n'} \sum_{k=1}^{n'} \mathbb{I}(\pi(k) = k)$$

$$\Delta M(rev(i, j)) = M(\pi') - M(\pi)$$

其中， $\mathbb{I}[\cdot]$ 是指示函数，当条件成立时值为 1，否则为 0； π' 为反转后的排列。

2) 逆序数优化指标[18]：基于 Kendall Tau 距离，评估排列有序性。序列的逆序数、归一化逆序评分以及逆序数改进量分别定义为：

$$Inv(\pi) = \sum_{1 \leq k < l < n'} \mathbb{I}(\pi(k) > \pi(l))$$

$$S_{Inv}(\pi) = 1 - \frac{Inv(\pi)}{\binom{n'}{2}}$$

$$\Delta S_{Inv}(rev(i, j)) = S_{Inv}(\pi') - S_{Inv}(\pi)$$

3) 断点消除指标[19]：基于断点分析，用于识别序列中的不连续点。定义扩展排列：

$$\pi^+ = (0, \pi(1), \dots, \pi(n'), n'+1)$$

则断点集合、密度评分与断点消除改进量分别定义为：

$$B(\pi) = \{k \mid \pi^+(k+1) - \pi^+(k) \neq 1\}$$

$$S_B(\pi) = \frac{|B(\pi)|}{n'+1}$$

$$\Delta S_B(\text{rev}(i, j)) = S_B(\pi') - S_B(\pi)$$

4) 递增子序列指标[20]：基于动态规划，关注最长有序片段的扩展，递增子序列定义和递增子序列改进量分别定义为：

$$LIS(\pi) = \max_{1 \leq k \leq n'} dp[k]$$

$$\Delta S_{LIS}(\text{rev}(i, j)) = S_{LIS}(\pi') - S_{LIS}(\pi)$$

其中动态规划函数为：

$$dp[k] = 1 + \max_{1 \leq m < k, \pi(m) < \pi(k)} dp[m]$$

5) 连续性增强指标：基于局部连续性，强调相邻元素的递增关系，连续性增强、连续性评分以及连续性改进量分别定义为：

$$C(\pi) = \sum_{k=1}^{n'-1} \mathbb{I}(\pi(k+1) - \pi(k) = 1)$$

$$S_C(\pi) = \frac{C(\pi)}{n'-1}$$

$$\Delta S_C(\text{rev}(i, j)) = S_C(\pi') - S_C(\pi)$$

2.3.2. 动态权重调整机制

各种贪心策略通过时变权重系数进行组合，形成综合评估函数。这种方法基于自适应权重理论，综合评估函数为：

$$\text{Score}(\text{rev}(i, j), t) = \sum_{m=1}^5 \omega_m(t) \cdot \Delta S_m(\text{rev}(i, j))$$

其中 $\omega_m(t)$ 是第 m 个指标在时间步 t 的权重 ($\sum_{m=1}^5 \omega_m(t) = 1$)，由阶段适应权重 $\omega_m^s(t)$ 和效果反馈权重 $\omega_m^f(t)$ 组合而成：

$$\omega_m(t) = \alpha \cdot \omega_m^s(t) + (1 - \alpha) \omega_m^f(t)$$

其中 $\alpha \in [0, 1]$ ，表示控制阶段适应与效果反馈的相对重要性，此处 $\alpha = 0.7$ 。

阶段适应权重基于反转次数 d 调整，体现搜索阶段理论：

初始阶段 ($d \leq 0.3 \times 2d_{\max}$) 时，侧重全局优化，设置 $\omega_1^s = 0.1, \omega_2^s = 0.1, \omega_3^s = 0.4, \omega_4^s = 0.2, \omega_5^s = 0.2$ ；中间阶段 ($0.3 \times 2d_{\max} < d \leq 0.7 \times 2d_{\max}$)，侧重平衡探索与利用，设置 $\omega_1^s = 0.2, \omega_2^s = 0.2, \omega_3^s = 0.3, \omega_4^s = 0.2, \omega_5^s = 0.1$ ；收敛阶段 ($d > 0.7 \times 2d_{\max}$)，侧重局部优化，设置 $\omega_1^s = 0.3, \omega_2^s = 0.2, \omega_3^s = 0.1, \omega_4^s = 0.2, \omega_5^s = 0.2$ 。

效果反馈权重基于策略历史表现，体现强化学习原理：

$$\omega_m^f(t) = \frac{1 + \log(1 + c_m(t))}{\sum_{m'=1}^5 [1 + \log(1 + c_{m'}(t))]}$$

其中 $c_m(t)$ 是时间窗口 $[t-W, t]$ ($W = 20, t = 5$) 内指标 m 带来正向改进的次数。

2.3.3. 多策略协同执行

1) 独立策略评估: 各指标独立评估候选反转操作, 生成按改进量降序的排序列表, 定义为:

$$R_m = [rev_1^m, \dots, rev_M^m] \left(M = \left\lfloor \frac{n'}{2} \right\rfloor \right)$$

2) 综合排名聚合: 基于动态权重计算每个反转操作的综合得分, 定义为:

$$\text{Total Score}(rev) = \sum_{m=1}^5 \omega_m(t) \cdot r_m(rev)$$

其中, $r_m(rev) = 1 - \frac{\text{rank}_m(rev) - 1}{M - 1}$ ($\text{rank}_m(rev)$ 为 rev 在 R_m 中的排名)。

3) 多样性保障机制: 为了避免过早收敛, 引入多样性选择。以概率 $p(t)$ 选择综合排名最高的操作, 以概率 $1 - p(t)$ 从各策略的 $Top - k$ 操作中随机选择。

$$p(t) = p_0 \cdot e^{-\frac{t}{T}} + p_{\min} \quad (p_0 = 0.8, T = 100, p_{\min} = 0.3)$$

2.4. 蒙特卡洛树搜索

蒙特卡洛树基于多臂赌博机理论, 平衡探索与利用, 适用于 BBFS 无交集、混合贪心搜索陷入局部最优的场景。双向蒙特卡洛树架构图见图 4。

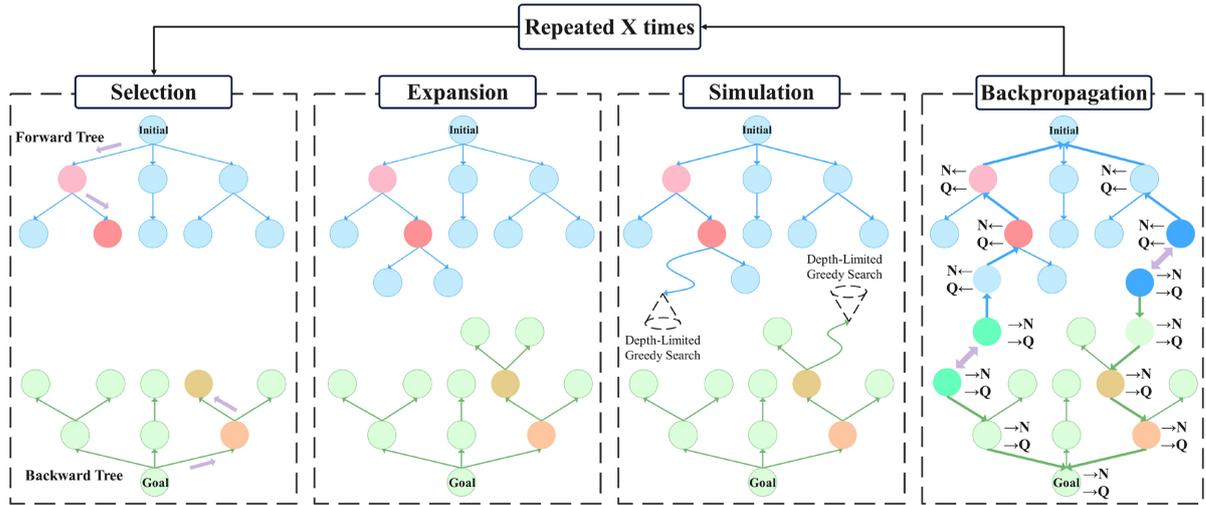


Figure 4. Bidirectional Monte Carlo tree

图 4. 双向蒙特卡洛树

2.4.1. 树节点结构

每个节点 u 包含: 核心状态(序列 s_u 、排列 π_u)、路径信息(父节点指针 $parent(u)$ 、反转操作 $op(u)$)、统计信息(访问次数 $N(u)$ 、累计价值 $V(u)$)、启发式信息(贪心评分 $H(u) = \text{Total Score}(op(u))$)、辅助信息(深度 $depth(u)$ 、唯一标识 $id(u)$)。

2.4.2. 混合选择策略

节点选择基于混合价值函数, 平衡 UCB 探索项和贪心启发式:

$$Value(u) = (1 - \beta) \cdot UCB(u) + \beta \cdot H(u)$$

其中 β 为贪心权重，用于引导贪心的强度 ($\beta = 0.5$)，UCB 探索项为：

$$UCB(u) = Q(u) + c \cdot \sqrt{\frac{\ln N(\text{parent}(u))}{N(u) + \varepsilon}}$$

其中： $Q(u) = \frac{V(u)}{N(u)}$ 是节点的平均价值； $c = \sqrt{2}$ 是探索常数，控制探索强度； $\varepsilon = 1.0 \times 10^{-5}$ ，防止除零错误。

2.4.3. 模拟与回溯机制

模拟阶段采用轻量级贪心策略快速评估节点潜力，避免全指标计算导致的效率损耗。从节点 u 出发，以 s_u 为起点进行深度限制搜索；搜索过程中，使用简化的综合评估函数(仅保留了位置匹配度和断点消除两个核心指标，权重均为 0.5)，用于快速选择反转操作；到达终止状态或最大深度时，计算奖励 R ，定义如下：

$$R = \gamma^{d_{sim}} \cdot (\omega_1 M(\pi_{sim}) + \omega_2 (1 - S_B(\pi_{sim})))$$

其中 d_{sim} 为限制搜索深度， γ 为折扣因子 ($\gamma = 0.8$)，且对于较深的模拟路径进行奖励衰减， ω_1, ω_2 为奖励权重。

回溯阶段更新路径上所有节点的统计量，对于从根节点到模拟终止节点路径上的每个节点 n ，更新每个节点的访问次数与累计价值：

$$V(n) \leftarrow V(n) + R$$

$$N(n) \leftarrow N(n) + 1$$

2.4.4. 双向协调机制

起始端(从 P 出发)与目标端(从 Q 出发)MCTS 独立运行且共享信息：每隔 50 次迭代同步边界节点状态，检测到边界节点在对方已访问集合时验证连接可行性，按搜索效率动态分配资源，确认连接后逆序拼接路径得到完整解。

3. 实验设计

3.1. 实验数据

本研究所用的所有 DNA 序列均为计算机模拟生成，并非来源于生物实验或测序数据。具体而言，我们生成了长度为 16 个碱基(A、T、C、G 的数量均为 4)且随机排列的 DNA 片段，并剔除了包含最少反转次数小于 5 的样本，共计 9327 条样本。

3.2. 实验结果

为验证 HBM 模型的综合性能，本实验从最少反转路径长度、时间 - 质量性价比及解质量稳定性三个核心维度，将 HBM 与位置匹配度、断点消除、连续性增强、逆序数四种传统启发式算法展开对比，以下为具体实验结果分析。

1) 不同模型下最少反转结果的对比

依据理论最少反转长度对实验样本进行分组，划分出反转长度分别为 5、6、7、8 的多个子集(覆盖不同难度的序列反转场景)；随后针对每个子集，分别得出 HBM 模型与位置匹配度贪心、断点消除贪心、

连续性增强贪心、逆序数贪心四种传统策略的输出反转次数，将各组中各个样本的结果以可视化形式呈现。从整体趋势来看，HBM 模型在所有反转长度区间内，均能得到更短的反转路径，且优势随着反转长度增加而愈发显著。不同策略下的反转结果图见图 5。

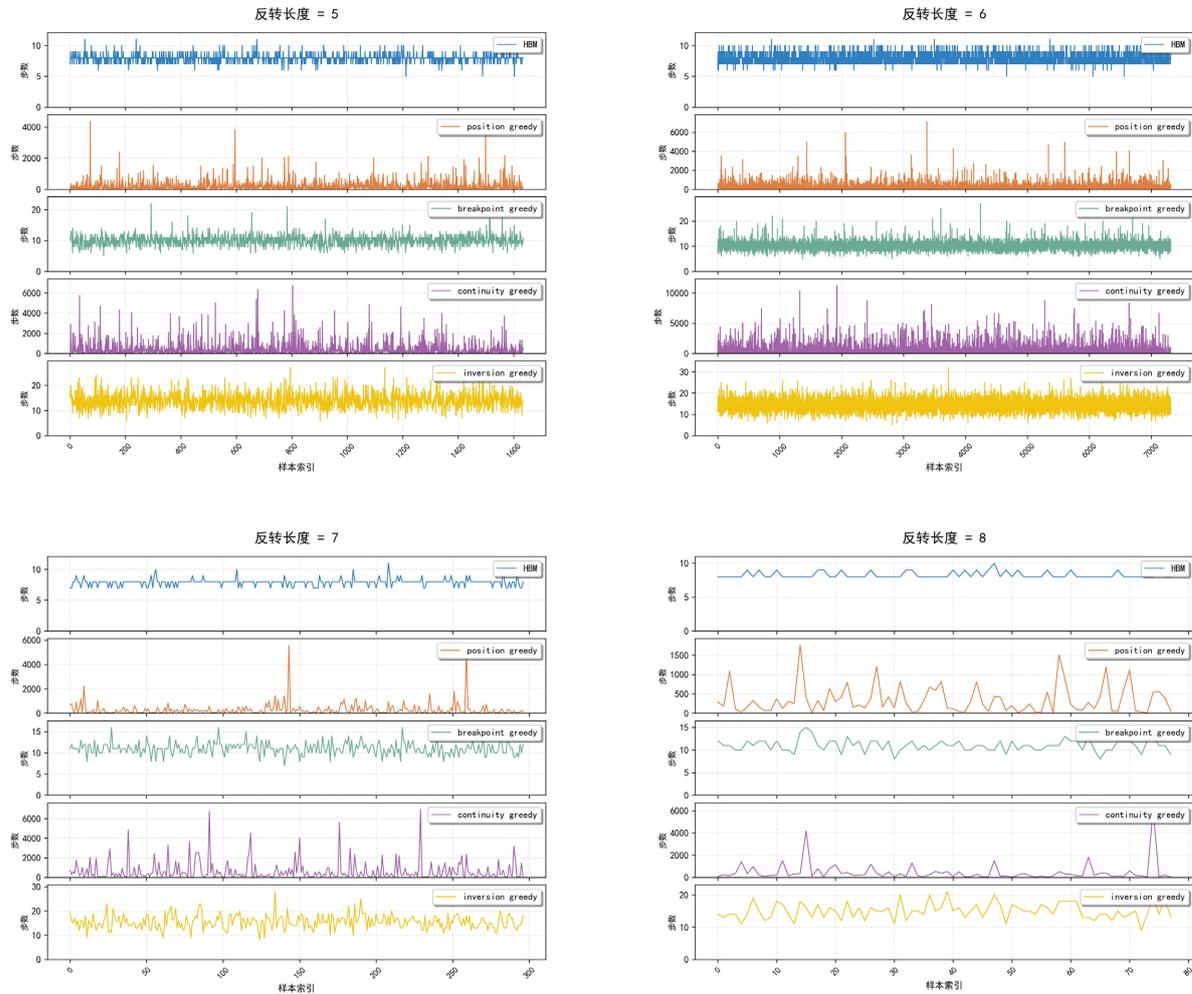


Figure 5. Reversal outcome diagram under different strategies
图 5. 不同策略下的反转结果图

Table 1. Comparison of average number of reversals under different strategies
表 1. 不同策略下的平均反转次数结果对比

Methods	5	6	7	8
位置匹配度	226.90	237.64	291.95	342.04
断点消除	10.02	10.25	11.01	11.01
连续性增强	454.49	494.95	635.74	463.21
逆序数	13.96	14.31	15.51	15.01
HBM	7.96	7.95	7.91	8.27

统计在理论最少反转次数下不同策略的平均反转次数。当目标长度为 5 时，HBM 的平均反转次数比

理论值高 2.96 次，而断点消除策略和逆序数策略所得到的平均反转长度则比理论值分别高出 5.02 次与 8.95 次。当反转长度提升为 8 时，HBM 的平均反转次数较理论值高出 0.27 次，比传统贪心策略的偏差更小，不同策略下的平均反转次数结果对比见表 1。

2) 时间 - 质量性价比(QTR)对比

时间 - 质量性价比(Quality-Time Ratio, QTR): 衡量算法“单位计算时间内提升解质量的效率”，核心反应算法在消耗单位时间时，其输出结果向“理论最少反转次数”逼近的程度，是横屏“解质量”与“时间成本”的关键指标，其计算公式为：

$$QTR = \frac{1 - \frac{S_{algo}}{S_{opt}}}{T_{algo}} \times 100\%$$

其中， S_{algo} 是算法输出的反转次数， S_{opt} 是基因序列重排的理论最少反转次数， T_{algo} 是算法的求解时间。同一算法在不同样本上的 QTR 波动越小，说明算法在“时间 - 质量平衡”上的表现越稳定。各算法的 QTR 分布图如图 6 所示：

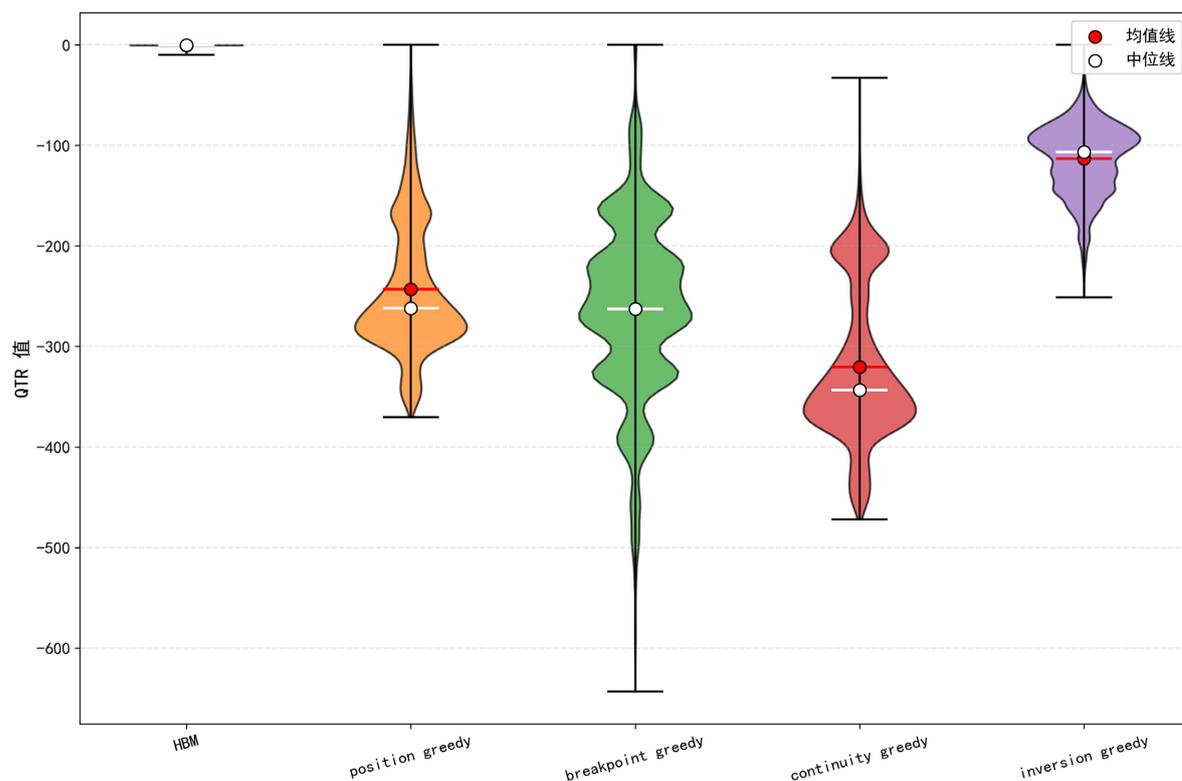


Figure 6. QTR results under different strategies

图 6. 不同策略下的 QTR 结果

从数值大小来看，HBM 的 QTR 均值均比传统贪心策略的值大，这意味着 HBM 每单位时间内实现的解质量提升普遍优于传统贪心策略；从稳定性看，HBM 的波动区间更小，传统贪心算法的波动普遍较大，这是因为 HBM 的预处理压缩与分层搜索机制可有效控制时间成本，保障单位时间的求解收益。

3) 相对解质量稳定性(rSQS)对比

相对解质量稳定性(Relative Solution Quality Stability, rSQS): 衡量算法在处理不同样本时, 解质量(即反转次数)的相对波动程度, 用于评估算法对样本差异的适配能力。相对稳定性越高, 算法在不同特征基因序列上的解质量波动越可控。

$$rSQS = \frac{1}{\bar{S}_{algo}} \sqrt{\frac{1}{N_{valid}} \sum_{i=1}^{N_{valid}} (S_{algo,i} - \bar{S}_{algo})^2}$$

其中, N_{valid} 是算法的有效样本数, $S_{algo,i}$ 是算法在第 i 个有效样本上输出的反转次数, \bar{S}_{algo} 是算法在所有有效样本上的平均反转次数。数值越小, 那么该算法的稳定性越高。各算法的 rSQS 数值对比结果如表 2 所示:

Table 2. Comparison of rSQS values under different strategies

表 2. 不同策略下的 rSQS 值对比

Methods	5	6	7	8
位置匹配度	1.500	1.487	1.806	1.065
断点消除	0.169	0.162	0.133	0.119
连续性增强	1.583	1.673	1.554	1.871
逆序数	0.222	0.200	0.133	0.119
HBM	0.085	0.081	0.072	0.066

HBM 的 rSQS 值在理论反转次数为 5、6、7、8 次的样本上均为最小, 远低于传统贪心策略值。这表明 HBM 的双向蒙特卡洛树双向协调机制与动态权重贪心策略可适配不同序列结构特征, 能够更好地控制解的波动。

3.3. 超参数敏感分析

为验证 HBM 模型核心超参数对求解性能的影响规律, 确保模型参数设置的合理性, 本研究选取动态调整权重机制中的平衡系数 α 、混合选择策略中平衡 UCB 和贪心指标权重 β 、模拟与回溯机制中的折扣因子 γ 作为关键超参数, 通过控制变量的形式展开敏感分析。实验中, 固定其他参数为基准值, 分别在各超参数的典型取值区间内设置梯度变量, 探究超参数变化对模型性能的影响趋势。

3.3.1. 超参数 α 的敏感分析

Table 3. Performance comparison under different α values

表 3. 不同 α 下的性能对比

α 取值	平均反转路径长度	QTR	rSQS
0.1	8.72	-2.76	0.102
0.3	8.01	-1.36	0.093
0.5	7.94	-0.65	0.084
0.7	8.02	-0.39	0.076
0.9	8.67	-0.69	0.101

超参数 α 为动态权重调整机制中阶段适应权重与效果反馈权重的平衡系数, 其取值范围为[0, 1], 核

心作用是调节搜索过程中阶段适配性与历史效果反馈的相对重要性($\alpha = 0$ 时仅依赖效果反馈权重, $\alpha = 1$ 时仅依赖阶段适应权重)。该超参数下的性能对比如表 3 所示。

当 $\alpha = 0.5$ 时, 所得到的平均反转路径长度最短; 当 $\alpha = 0.7$ 时, 所得到的 QTR 值与 rSQS 值均为最优。当 $\alpha < 0.3$ 或 $\alpha > 0.7$ 时, 平均反转路径长度、QTR 值与 rSQS 值均显著恶化, 表明权重失衡会破坏策略适配性与效果反馈的平衡。

3.3.2. 超参数 β 的敏感分析

超参数 β 为混合选择策略中平衡 UCB 探索项和贪心启发式的贪心权重, 取值范围为[0, 1], 核心作用是调节蒙特卡洛树搜索中探索未知路径与利用已知路径的强度($\beta = 0$ 时仅依赖 UCB 探索, $\beta = 1$ 时仅依赖贪心启发式)。该超参数下的性能对比如表 4 所示。

Table 4. Performance comparison under different β values

表 4. 不同 β 下的性能对比

β 取值	平均反转路径长度	QTR	rSQS
0.1	9.85	-3.25	0.135
0.3	8.57	-2.92	0.098
0.5	7.97	-1.27	0.085
0.7	8.05	-1.53	0.074
0.9	8.96	-1.97	0.121

当 $\beta = 0.5$ 时, 所得到的平均反转路径长度最短且 QTR 值最优; 当 $\beta = 0.7$ 时, 所得到的 rSQS 值均为最优。在其他取值中, 明显观察出该参数设置过小会导致分支探索增多而难以找到最优路径, 而参数设置过大又会导致过于偏向贪心搜索使得路径长度增加。

3.3.3. 超参数 γ 的敏感分析

超参数 γ 为模拟与回溯机制中的折扣因子, 取值范围为[0.5, 0.95], 核心作用是调节模拟路径中深度节点奖励的衰减程度(γ 越接近 1, 较深的奖励权重越大; γ 越小, 越侧重近期奖励)。该超参数下的性能对比如表 5 所示。

Table 5. Performance comparison under different γ values

表 5. 不同 γ 下的性能对比

γ 取值	平均反转路径长度	QTR	rSQS
0.5	8.53	-2.55	0.097
0.6	8.21	-0.73	0.090
0.7	8.02	-0.96	0.086
0.8	7.94	-0.23	0.073
0.9	7.95	-1.74	0.089

当 $\gamma = 0.8$ 时, 所得到的平均反转路径长度最短且 QTR 值与 rSQS 均最优。表明取该值下深度节点的奖励与浅层节点的奖励权重配比最合理, 既通过适度保留深度节点的奖励鼓励模拟探索具有长期优化潜力的路径, 又避免过度放大深度节点的奖励而导致评估出现偏差。

4. 结论

本文提出的 HBM 模型在基因序列反转求解这一 NP 难问题上实现了系统性创新,其核心在于构建了一个分层递进式协同求解框架,该框架通过模块化协作显著提升了解题效率与质量:首先,模型利用预处理机制压缩问题规模,并基于理论最大反转次数设定搜索深度,采用 BBFS 进行全局底层搜索;若未获解,则启动融合动态权重调整的混合贪心策略,将五种启发式指标(位置匹配度、逆序数、断点数、递增子序列、连续性增强)按搜索阶段(全局优化、平衡探索、局部收敛)自适应组合,有效规避局部最优陷阱;最终,通过双向蒙特卡洛树搜索实现智能决策,其双向异步搜索与定期同步机制扩展了搜索范围,并通过路径验证与拼接优化解质量。

在模拟 DNA 序列数据集上的实验结果表明, HBM 模型相较于传统启发式方法展现出全面优势。具体而言,其在反转路径长度上更接近理论最优值,且优势随问题难度增加而愈发显著;在衡量效率的关键指标时间-质量性价比 QTR 上表现更优且更稳定;其相对解质量稳定性 rSQS 也远高于对比算法,证明该模型能有效适配不同序列结构特征,输出波动更小。

尽管 HBM 模型已取得良好成效,未来仍有可拓展方向。后续研究可致力于权重机制的进一步自适应优化,探索模型在更复杂序列结构(如非均匀分布、长序列)上的泛化能力,并推动其在与真实生物基因组重排数据结合中的应用验证,以深化其在计算生物学领域的实用价值。

参考文献

- [1] Liu, Y., Yi, C., Fan, C., Liu, Q., Liu, S., Shen, L., *et al.* (2023) Pan-Centromere Reveals Widespread Centromere Repositioning of Soybean Genomes. *Proceedings of the National Academy of Sciences*, **120**, e2310177120. <https://doi.org/10.1073/pnas.2310177120>
- [2] Jiang, Z., Zhu, X., Wang, X., Chen, Q., Zhang, W., Li, Z., *et al.* (2025) Chromosomal-Level Genome Assembly of *Trichogramma Japonicum* (Hymenoptera: Trichogrammatidae). *Scientific Data*, **12**, 1537-1537. <https://doi.org/10.1038/s41597-025-05910-z>
- [3] 张志良. 基于反转和删除操作的基因组重组全解空间研究[D]: [硕士学位论文]. 济南: 山东大学, 2012.
- [4] Sipel, A.C. (2003) An Algorithm to Enumerate Sorting Reversals for Signed Permutations. *Journal of Computational Biology*, **10**, 575-597. <https://doi.org/10.1089/10665270360688200>
- [5] 黄鑫, 张志佳, 孙平, 等. 基于深度强化学习的路径规划算法综述[J]. 机器人, 2026, 48(1): 196-216.
- [6] García, A. (2025) Greedy Algorithms: A Review and Open Problems. *Journal of Inequalities and Applications*, **2025**, Article No. 11. <https://doi.org/10.1186/s13660-025-03254-1>
- [7] 党予卿, 陈丁鹏, 李春宝, 等. 蚁群系统算法求解比例多处理器开放车间调度问题[J/OL]. 计算机工程与应用, 1-14. <https://link.cnki.net/urlid/11.2127.TP.20251128.0816.002>, 2025-12-06.
- [8] Petersen, J.E., Kapur, S., Gkantonas, S., Mastorakos, E. and Giusti, A. (2023) Modelling and Optimisation of Extinction Actions for Wildfire Suppression. *Combustion Science and Technology*, **195**, 3584-3595. <https://doi.org/10.1080/00102202.2023.2246195>
- [9] Wei, L., Yi, L., Yan, M., *et al.* (2023) A Self-Learning Monte Carlo Tree Search Algorithm for Robot Path Planning. *Frontiers in Neurorobotics*, **17**, Article ID: 1039644.
- [10] Dai, W., Wang, J., Chen, Y., Wang, X., Yang, X. and Zhang, H. (2026) An Improved PSO for Operational Optimization of Steam Power Systems in Petrochemical Enterprises. *Expert Systems with Applications*, **297**, Article ID: 129404. <https://doi.org/10.1016/j.eswa.2025.129404>
- [11] 张泽宁. 基于 Alphago-Zero 及图扩散模型的图构造方法研究[D]: [硕士学位论文]. 北京: 北京交通大学, 2024.
- [12] 李国良, 王磊, 张金玉, 等. GaussDB: 智能云原生分布式数据库[J]. 电子学报, 2025, 53(4): 1103-1122.
- [13] Baudet, C. and Dias, Z. (2010) An Improved Algorithm to Enumerate All Traces that Sort a Signed Permutation by Reversals. *Proceedings of the 2010 ACM Symposium on Applied Computing*, Sierre, 22-26 March 2010, 1521-1525. <https://doi.org/10.1145/1774088.1774416>
- [14] Christie, D.A. and Irving, R.W. (2001) Sorting Strings by Reversals and by Transpositions. *SIAM Journal on Discrete Mathematics*, **14**, 193-206. <https://doi.org/10.1137/s0895480197331995>

- [15] Bian, J., Li, J. and Wang, G. (2024) Automatic Move Program for Two-Player Military Chess Based on UCB&MCTS Algorithm and Move Evaluation. *International Core Journal of Engineering*, **10**, 388-396.
- [16] Tian, C. (2023) Monte-Carlo Tree Search with Epsilon-Greedy for Game of Amazons. In: ITM Department, Illinois Institute of Technology, Eds., *Proceedings of the 3rd International Conference on Signal Processing and Machine Learning (Part 3)*, Faculty of Engineering, Architecture & Information Technology, University of Queensland, 420-425.
- [17] 王孟珂. 2,4-表油菜素内酯对银杏类黄酮代谢的影响及相关基因的功能研究[D]: [硕士学位论文]. 南京: 南京林业大学, 2022.
- [18] Bader, D.A., Moret, B.M.E. and Yan, M. (2001) A Linear-Time Algorithm for Computing Inversion Distance between Signed Permutations with an Experimental Study. *Journal of Computational Biology*, **8**, 483-491. <https://doi.org/10.1089/106652701753216503>
- [19] Fang, B., Lai, J., Liu, Y., Liu, L., Yu, X., Li, X., *et al.* (2024) Hybrid Sequencing for Detailed Genetic Characterization of Human Adenoviruses. *Scientific Reports*, **14**, Article No. 29490. <https://doi.org/10.1038/s41598-024-78960-9>
- [20] 王学. 基因组中最大唯一匹配的查找算法研究[D]: [硕士学位论文]. 西安: 西安电子科技大学, 2009.