広さと深さのバランスに注目した こいこいのゲーム木展開制御手法の検討

栗田 凌太朗[1,a)]　横山 大作[1,b)]

概要：花札のこいこいは, 相手の手札を観測できない不完全情報ゲームであると同時に, 山札から札をめくる確率性を含む不確定ゲームである. こいこいに MCTS とチャンスノードを適用すると, 探索木の状態数が増加しやすく, 深い探索を行いにくい. そこで本研究では, チャンスノードの訪問数に応じてチャンス子数に制限を設けることで, 広さと深さのバランスを制御することを提案する. 多数回対戦と特定局面での手の選択割合によって評価した結果, チャンス子数の制御は深さをコントロールして与えることには成功したが, 全体の対戦性能を一貫して向上させるものではなかった. 一方で, 一部の局面における行動選択へ影響を与える可能性が示された.

A Study of Game-Tree Expansion Control for Koi-Koi with a Focus on the Balance Between Search Breadth and Depth

Kurita Ryotaro[1,a)]　Yokoyama Daisaku[1,b)]

Abstract: Koi-Koi, a Hanafuda card game, is both an imperfect-information game in which a player cannot see the opponent's hand and a stochastic game involving random card draws from the deck. Applying MCTS with chance nodes to Koi-Koi produces many possible states in the search tree, making deep search difficult. This study proposes limiting the number of chance-node children according to the visit count to control the balance between search breadth and depth. The method was evaluated through a large number of matches and move selection rates in selected positions. Controlling the number of chance-node children successfully increased the search depth but did not consistently improve overall match performance. However, the results suggest that it may affect move selection in some positions.

1. はじめに

本研究で対象とする花札のこいこいは, 相手の手札を観測できない不完全情報ゲームであると同時に, 山札から札をめくる確率性を含む不確定ゲームである. このようなゲームにモンテカルロ木探索 (Monte Carlo Tree Search, MCTS)[1] を適用する場合, プレイヤの意思決定に加えて, 非公開情報と確率的事象を扱う必要がある. その結果, 探索対象となる状態数が増加し, 限られたプレイアウト数で効率よく探索することが課題となる.

[1] 明治大学
Meiji University
[a)] ce265024@meiji.ac.jp
[b)] dyokoyama@meiji.ac.jp

こいこいでは，手札から出した札および山札からめくった札を，同じ月の場札と組み合わせて獲得し，獲得札によって役を成立させる．各手番では，手札から札を出す処理と山札から札をめくる処理が行われる．プレイヤは相手の手札および山札の順序を観測できないため，探索時には観測されていない札の所在を考慮する必要がある.

そこで本研究では, チャンスノード (Chance Node) を用いて確率的事象を探索木内で明示的に扱う. こいこいにおいてチャンスノードが扱う確率的事象は山札からのめくりであり, 未知札が相手の手札と山札の 2 か所に分かれることから, 非公開情報である相手の手札にも関係する.

こいこいでは, 次の自分の番までに相手の手番と 2 回の山札のめくりが行われる. チャンスノードを訪問するたびに新しいパターンを生成すると, 多くのプレイアウトがチャ

© 2026 Information Processing Society of Japan

ンス子の追加に使われ, 後続の決定ノードを十分に探索できない. この場合,1 パターンあたりの評価値はごく少数のランダムプレイアウトからなるため信頼性が低く, さらに木が 1 段程度で止まってしまうと MCTS である意味がなくなってしまう. 一方, 既存のチャンス子の再評価を優先しすぎると, 観測するパターンが偏り, 有利不利の誤認につながる可能性が高まる. したがって, 幅広いパターンを調べることと, 生成済みのパターンを追加で探索することの両方が必要となると考えられる.

そこで本研究では, チャンスノードの訪問回数に応じてチャンス子数の上限を増加させる手法を提案する. チャンス子数が上限未満の場合は新しいパターンを生成し, 上限に達している場合は既存のチャンス子を再評価する. これにより, 探索木の広さと深さのバランスを制御することを目的とする.

提案法の性能は, 複数のプレイアウト数による多数回対戦と, 対局中の各手の選択割合を用いて評価した. 比較対象は, チャンス子数の制限を設けない, 従来のチャンスノードを用いる手法を基準法として用意した. 実験の結果, 勝率や平均得点にはプレイアウト数を通じて一貫した方式間の差は見られなかった. 一方,1000 プレイアウトでは, 到達手番数 2 以上のフロンティアの割合が提案法では 71.8%, 基準法では 0.9%となり, 提案法が自分の次の決定ノードに対応する深さ以降まで探索する機会が増加した.

また, 複数の取得リスクが競合する一部の局面では, 方式間で手の選択傾向が大きく異なった. 菊に盃を別の種札へ置換した赤短取得順序局面では,2000 プレイアウトにおいて, 提案法は場の梅に赤短を取得する手を 70.3%, 基準法は 23.3%選択した. 一方, 手札の桜に赤短を取得する手は, 提案法が 26.6%, 基準法が 75.5%であった. この結果から, 提案法は全体の対戦性能を一貫して向上させるものではなかったが, 一部の局面では探索を深くすることで, 場の役札を相手に取られる危険と, 手札の役札の対応先を失う危険に対する優先順位を変化させる可能性が示された.

2. こいこい

花札は古くから遊ばれてきた日本の伝統的なかるたの一種である. 花札には多様な遊び方が存在するが, 本研究では「こいこい」が対象である. なお, こいこいについても役の種類や得点計算など多様にルールが分かれるため, 本節で実験の理解に必要な分の説明を行う.

花札の札は一組 48 枚で構成される.1 月から 12 月または対応した植物の名前で分類することができ, 各月に 4 枚ずつ存在する. また, 札は絵柄に応じて光札, 種札, 短冊札, カス札の札種に分類される.

こいこいとは, 親 (先攻) と子 (後攻) に分かれた 2 人のプレイヤが札を取り合い, 相手よりも早く役を作ることを目指すゲームである. はじめに各プレイヤに 8 枚ずつ, 場に 8 枚配り, 残りの札は山札として使用する. プレイヤは各手番に以下の 2 つの行動を順に行う.

① 手札を 1 枚出す
自分の手札から好きな 1 枚を出す. 出した札の月と同じ月の札が場札にある場合, 出した札と対応する場札を手元の取札に追加する. 同じ月の場札が 2 枚ある場合はいずれか 1 枚を選択し,3 枚ある場合は出した札と合わせて 4 枚すべてを取札に加える. 同じ月の札が場札にない場合は出した札をそのまま場札に追加する.

② 山札を 1 枚めくる
山札の上から 1 枚をめくり, 手札から札を出した場合と同じ規則で場札との対応を処理する. 同じ月の札が場にない場合, めくった札は場札として残る.

手札と山札の処理が終了した時点で, 取札に特定の組合せが揃うと役が成立し, あがりとなる. 本実装では, 五光を 10 点, 四光を 8 点, 雨四光を 7 点, 三光, 花見で一杯, 月見で一杯, 猪鹿蝶, 赤短, 青短を各 5 点とする. 猪鹿蝶, 赤短, 青短では, 成立に必要な枚数を超えて同じ札種を取得するごとに 1 点を加算する. 赤短と青短がともに成立した場合は, 個別に加算せず両者を合わせて 10 点を基礎点とする. タネとタンはそれぞれ 5 枚, カスは 10 枚で 1 点となり, 以後は同じ札種を 1 枚取得するごとに 1 点を加算する. 菊に盃は種札とカス札の両方として数える. 複数の役が成立した場合は得点を合計し, 合計が 7 点以上の場合は 2 倍とする.

プレイヤが観測できる情報は, 自分の手札, 場札, および両プレイヤの取札である. 一方, 相手の手札と山札の並びは観測できない.

本研究では手札からの行動選択に焦点を当てるため, 一部のルールを簡略化する. 初期配札で手四またはくっつきが成立した場合は, 札を配り直す. また, 役が成立直後に局を終了し, こいこい宣言による局の継続は行わない.

3. 関連研究

3.1 基礎理論

モンテカルロ木探索（Monte Carlo Tree Search,MCTS）は, ランダムなプレイアウト（シミュレーション, ロールアウトとも呼ばれる）の結果に基づいてゲーム木を成長させ, 次の一手を決定する探索アルゴリズムであり, 様々なゲームに適用されている [1]. MCTS は, 探索方策に従って葉ノードまで進む選択, 子ノードを追加する展開, ゲーム終端までプレイするシミュレーション, 得られた報酬を探索経路上へ反映する逆伝搬から構成される.

UCB（Upper Confidence Bound）は, 多腕バンディット問題を解く際の代表的手法であり, 期待報酬と探索回数に基づいて行動を選択する [2]. UCB を MCTS の探索方策に適用した手法が UCT であり, 期待報酬の高い行動の活用と, 訪問回数の少ない行動の探索を両立する [3]. また, 報

© 2026 Information Processing Society of Japan

酬の分散を考慮した実用的な改良手法として UCB1-Tuned が提案されており, 本研究ではこれを使用する [2].

カードの配布やサイコロの出目などの確率的な状態遷移は, プレイヤが選択できない結果を表すチャンスノードとしてゲーム木に導入できる [1]. プレイヤの行動を表すノードでは探索方策によって子ノードを選択するのに対し, チャンスノードでは結果の生起確率に従って子ノードを選択する. 一方, 生じ得る結果が多い場合には分岐数が増加し, 限られたプレイアウトが多数の子ノードへ分散することで, その先の探索が進みにくくなる.

このような大きな分岐を扱う方法として,Progressive Widening（PW）が提案されている. PW では, 訪問回数が少ない間は展開する子ノード数を抑え, 訪問回数の増加に伴って新しい子ノードを追加することで, 探索の広さと各子ノードより先の探索に割り当てる回数の均衡を図る. Couëtoux ら [4] は, 行動分岐に加えて, 確率的遷移によって生じる後続状態にも PW を適用する Double Progressive Widening（DPW）を提案した.

また,Sokota ら [5] は, 確率的遷移によって MCTS が浅い探索へ留まる問題に対し, 状態間の類似性を利用して抽象化を段階的に細分化する方法を提案している. この方法は, すべての遷移に同じ展開規則を適用する PW とは異なり, 状態空間の構造を利用して子ノードを統合する. これに対して本研究では状態の抽象化は行わず, 有限だが多数存在するチャンスノードの結果を明示的に保持したうえで, その追加数を制御する.

3.2 不完全情報ゲームの探索

不完全情報ゲームに対する方法の一つとして, 観測できない情報をサンプリングして完全情報の状態を生成し, その状態に対して探索を行う Determinization がある. しかし, サンプリングした状態ごとに独立した探索を行う場合, 同じ行動に関する探索が複数の木へ分散することや, 実際には識別できない状態ごとに異なる戦略を選択する問題が生じ得る.

Cowling ら [6] は, 非公開情報をサンプリングしながら, 複数の状態における探索結果を情報集合に対応する一つの木へ集約する Information Set Monte Carlo Tree Search（ISMCTS）を提案した. これにより, サンプリングした完全情報状態ごとに独立した木を構築する場合と比べ, 行動に関する統計を共有できる.

また,Li ら [7] は, 観測情報と整合する複数の状態に共通の行動系列を適用し, 各状態の葉で得られる方策と評価値を集約する方法を提案している. これらに対し, 本研究では探索開始時に複数の状態を生成して評価を集約するのではなく, 探索中に非公開情報と確率的事象を局所的にサンプリングし, その結果をチャンスノードの子として明示的に保持する.

3.3 こいこいに関する研究

こいこいに関する研究は少ないが, 他の不完全情報ゲームと同様に, 主に学習的アプローチと探索的アプローチが存在する. 本研究は後者に属する研究である. また, こいこいには様々なルールが存在し, 採用されている役や得点計算などが研究によって異なるため, 勝率や得点を研究間で直接比較できない場合がある.

学習的アプローチとして, 佐藤ら [8] は方策勾配法と Neural Fitted Q Iteration（NFQ）を適用した. その結果, 方策勾配法ではルールベースプレイヤ以下の強さとなった一方で,NFQ ではルールベースプレイヤに勝ち越す強さを示した. Guan ら [9] は, 札をトークンとして表現し,Transformer を用いて札同士の関係を扱う強化学習プレイヤを提案した. 作成されたプレイヤは経験のある人間プレイヤに勝ち越しており, こいこいへの Transformer を用いた強化学習の有効性を示した.

探索的アプローチとして, 髙岡ら [10] は, 手札から出す札の選択に UCT を適用した. 木の成長を自分の 1 手分である 1 段で止め,2000 プレイアウトによる実験を行った結果, 役の完成を第一に考えるグリーディプレイヤに勝ち越している. 長峰ら [11] は, 高岡らの研究をもとに, 報酬設計によって点取り, 早あがり, 妨害といった異なる戦略を実現することを提案した. 早あがり戦略は他の戦略より早くあがる一方で, 平均得点については点取り戦略が高いという結果を示している.

Cruz ら [12] は, 不確定性を含むゲームの例として簡略化したこいこいを採用し, チャンスノードを含む Expectiminimax によって探索を行った. 一方で, 生じ得る分岐を展開することによるノード数の増加から, 探索を深さ 2 から 3 で打ち切っている. 作成されたプレイヤは, 評価関数と探索によってランダムプレイヤおよび 2 種類のグリーディプレイヤに勝ち越している.

従来のこいこいを対象とした探索研究では, プレイヤの強さや報酬設計が主な対象となっており, 幅広い確率結果を調べることと, その先の意思決定に対する探索資源の配分は十分に検討されていない. そこで本研究では, 山札から札をめくる処理を明示的なチャンスノードとして表現し, 訪問回数に応じて子ノード数を制御する.

4. 提案手法

4.1 全体像

本研究の探索プレイヤは MCTS をベースに構成される. 図 1 に示すように, 探索木は各プレイヤの行動を扱うノードと, 山札からのめくりおよび非公開情報の整合性を扱うチャンスノードによって構成される. こいこいでは手札から札を出した後に山札から 1 枚をめくるため, 各プレイヤのノードの直後にチャンスノードが挿入される.

未展開の決定ノードへ到達した場合は, その状態から現

© 2026 Information Processing Society of Japan

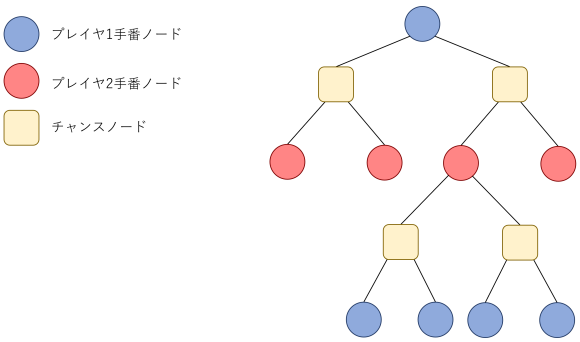


図 1　ツリーイメージ図

在の局が終了するまでランダムプレイアウトを 1 回行い, その後に子ノードを展開する. 各プレイアウトの報酬には, 髙岡ら [10] に倣って得点を用いる. ただし, 探索プレイヤが勝った場合は探索プレイヤの得点, 敗れた場合は相手プレイヤの得点の負値, 引き分けの場合は 0 とする. 得られた報酬は探索経路上の各ノードへ逆伝搬され, 訪問回数と報酬が更新される. 決められたプレイアウト数の探索後, ルート直下で訪問回数が最も多い行動を実際の手として選択する.

4.2　プレイヤの手番ノード

4.2.1　有効手の削減

本研究では, チャンスノードによる分岐に加えて決定ノードにおける分岐も抑えるため, 探索対象とする有効手を削減する. この処理は基準法と提案法の両方に共通して適用する.

有効手の削減は,「場札と合わせられる札がある場合は, その札を優先した方が基本的に強い」というドメイン知識に基づく. こいこいでは取札の組合せによって役を完成させるため, 札を取得できる行動は, 札を場へ捨てる行動よりも役の成立につながりやすいと考えられる. そのため, 場札と合わせられる手札が一つでも存在する場合は, 手札と対応する場札の組を有効手とする. 合わせられる手札が存在しない場合は, すべての手札を有効手とする. 同じ手札に対して同月の場札が 2 枚存在する場合は, それぞれの取得先を異なる行動として扱う.

この知識の例外として, 山札からより良い札を得るために, あえて場札と合わせずに手札を捨て, あがりを遅らせる戦略などが考えられる. 本研究の探索ではこのような行動は候補から除外される. また, 予備実験として, このドメイン知識の有無のみを変更したランダムプレイヤ同士の多数回対戦を行い, 有効手を削減することによる成績の改善を確認している.

4.2.2　子ノードの選択

探索プレイヤのノードでは, 有効手の数だけ子ノードを作成する. 選択時には, まだ訪問されていない子ノードを優先する. すべての子ノードが訪問済みの場合は,UCB1-Tuned[2] を基にした探索値を計算し, その値が最も大きい子ノードを選択する. このとき, 少数の高得点が行動評価を過度に左右することを抑えるため, 平均報酬項には符号付き得点をシグモイド変換した値を用いる.

相手プレイヤのノードでは, 相手が即座に役を完成できる場合はその行動を優先し, それ以外の場合は有効手からランダムに行動を選択する. 一方, 未展開ノードから行うプレイアウトでは, 探索プレイヤと相手プレイヤの両方が有効手からランダムに行動する.

4.3　基準法におけるチャンスノード

相手の手札は探索プレイヤから観測できないため, 探索開始時に山札と相手手札からなる未知札全体から, 仮想的な相手手札を与える. さらに各チャンスノードでは, その時点で仮定されている相手手札を未知札の集合へ戻し, 仮想相手手札と山札からめくられる札を改めて決定する. このように, 本研究の探索では非公開情報を探索開始時に一度だけ決定して固定するのではなく, 各チャンスノードで局所的に再サンプリングする.

チャンスノードでは, 仮想相手手札, 山札からのめくり札, および場札との取得先の組合せを一つのパターンとして扱い, それぞれを異なるチャンス子として保持する. 場札との取得先は本来プレイヤが決定する行動であるが, 本研究では探索木を単純化するため, 山札からめくられた札に同月の場札が 2 枚存在する場合の取得先をチャンスノードでサンプリングする.

基準法では, チャンスノードを訪問するたびにパターンをサンプリングする. サンプリングした仮想相手手札, めくり札, 取得先が既存のチャンス子とすべて一致する場合は, その子を再利用する. 一致する子が存在しない場合は, 新しいチャンス子として追加する.

4.4　提案法におけるチャンス子数の制御

提案法は, 概ね基準法と同様にチャンスノードを扱うが, チャンス子数を, チャンスノードの訪問回数に応じて段階的に増加させる点が基準法と異なる. Couëtoux ら [4] の progressive widening を参考に, チャンスノードの訪問回数を n としたとき, チャンス子数の上限を式 1 によって定める.

$$K(n) = \max\left(1, \lfloor cn^{\alpha} \rfloor\right), \qquad c > 0, \quad 0 < \alpha < 1 \quad (1)$$

現在のチャンス子数が $K(n)$ 未満である場合は, 上限に達するまで新しいパターンをサンプリングし, チャンス子として追加する. 一方, チャンス子数が上限に達している場合は, 新しい子を追加せず, 既存のチャンス子を選択してその先を再評価する.

© 2026 Information Processing Society of Japan

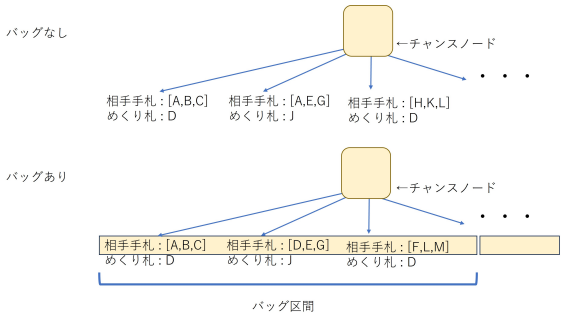


図 2 サンプリングイメージ図

新しいチャンス子を生成する際は, 未知札をバッグとして保持する. 仮想相手手札を作るときは, このバッグから相手の手札枚数分の札を非復元抽出する. 抽出した札は, バッグが空になるまで次の仮想相手手札の生成には使用せず, バッグが空になった場合は未知札全体からバッグを再構成する. これは, 独立に相手手札をサンプリングする場合と比べて, 同じ札が短期間に繰り返し仮想相手手札へ含まれることを抑え, 生成するチャンス子の短期的な多様性を確保することを目的としている. 仮想相手手札を決定した後, 残った未知札から山札のめくり札と取得先をサンプリングする.

さらに既存のチャンス子を再評価する際, 追加された時期の違いによって特定の子へ訪問が偏ることを抑えるため, 訪問回数の少ない子を優先する. チャンス子 i の訪問回数を n_i としたとき, その子を選択する確率を式 2 によって定める.

$$P(i) = \frac{(n_i + 1)^{-\beta}}{\sum_j (n_j + 1)^{-\beta}} \tag{2}$$

ここで,β は既存のチャンス子に対する訪問回数の偏りを調整するパラメータである. c,α,β はいずれも経験的に設定したため, 理論的な検証は行っていない. 本研究の実験では, 基本設定として $c = 1.0$,$\alpha = 0.75$,$\beta = 1.0$ を使用する.

4.5 チャンス子の制御の意味

基準法はチャンスノードを訪問するたびにパターンをサンプリングするため, 限られたプレイアウト数の中で幅広い確率結果を観測しやすい. 一方で, チャンス子が増加し続けると, 各チャンス子より先に存在する決定ノードの訪問回数が少なくなり, 複数手先の行動選択を十分に探索できない可能性がある.

提案法はチャンス子数の増加を制御し, 生成済みのチャンス子を再評価することで, パターンの新規生成に用いるプレイアウトの一部を後続の決定ノードの探索へ配分する方法である. 行動価値を正確に推定するためには, 多様なパターンを扱う幅広さと, 各パターンの先にある行動を十分に探索する深さの両方が必要である. 提案法はチャンス子数を制御することで両者への探索資源の配分を変更するため, その効果は局面に応じて異なると考えられる.

5. 実験

5.1 共通条件

本研究では,4 節に示した基準法と提案法を比較する.

探索でのプレイアウト数は 500,1000,2000 の 3 条件を用いた. また, 方式間で乱数の影響をできるだけ揃えるため, 同じ実験の比較条件では同一のシードを用いた.

5.2 全体性能評価

5.2.1 実験方法

提案法が実対局の成績に与える影響を確認するため, 各方式の探索プレイヤと即時あがりを優先するランダムプレイヤを対戦させた. 即時あがりを優先するランダムプレイヤは, 手札から札を出すことで直ちにあがれる場合は, 得点が最大となる行動を選択する. 直ちにあがれる行動がない場合は,4.2.1 目で示したドメイン知識に基づいて有効手を限定し, その中からランダムに行動を選択する.

先後の影響を分けて確認するため, 各方式・各プレイアウト数について, 探索プレイヤを先手と後手に配置した対戦をそれぞれ 1 万局行った. 各条件の 1 万局では, 各対局の初期配札にシード 1 から 10000 を割り当てて実行した.

まず対戦性能の評価指標として, 両者の勝率や平均得点, および探索プレイヤ視点の 1 局当たり得失点差を調べる. 各指標は, 探索プレイヤが先手の場合と後手の場合に分けて集計するとともに, 先後を統合した値も求めて比較する.

さらに探索性能の評価指標として, 探索プレイヤが先手番であった 1 万局のログにおける探索木のフロンティアを調べる. 本研究では, 探索方策による選択が終了し, ランダムプレイアウトまたは終端状態の評価へ移った位置をフロンティアとする. ログ上の深さからルート直下の 1 層を除いた値を到達手番数とし, 探索開始直後を 0 手先と数える. この到達手番数は, 探索プレイヤだけでなく相手プレイヤの手番進行も含むが, チャンスノードは含まない.

有効手が 2 手以上あり, 実際に MCTS が行われた探索だけを集計対象とする. 評価指標は, 各探索における深さごとのフロンティア割合に加え, 平均到達手番数および各探索の最大到達手番数を全対象探索について平均した値とする. なお, フロンティアにはランダムプレイアウトの開始位置とゲームの終端位置が混在するため, 到達手番数は探索が到達した深さの傾向を表す指標として用いる.

5.2.2 実験結果

表 1 に対戦性能評価の結果を示す. 先後統合勝率の方式間差は,500,1000,2000 プレイアウトでそれぞれ -0.44,$+0.30$,-0.22 ポイントとなり, プレイアウト数によって大小関係が変化した. 平均得点と得失点差についても同様に一貫した差は見られず, 先手と後手に分けた場合にも

© 2026 Information Processing Society of Japan

表 1　対戦性能評価の結果

プレイアウト数	方式	勝率 (%) 探索/相手	平均得点 探索/相手	得失点差
探索プレイヤが先手				
500	提案法	60.63/37.88	2.893/1.648	1.245
500	基準法	60.29/38.12	2.872/1.651	1.221
1000	提案法	61.42/36.91	2.844/1.606	1.238
1000	基準法	61.03/37.34	2.927/1.606	1.322
2000	提案法	61.17/37.19	2.906/1.582	1.324
2000	基準法	61.45/36.96	2.950/1.585	1.365
探索プレイヤが後手				
500	提案法	53.06/45.35	2.550/1.976	0.574
500	基準法	54.27/44.14	2.619/1.918	0.701
1000	提案法	54.44/43.65	2.595/1.891	0.704
1000	基準法	54.23/43.97	2.557/1.893	0.664
2000	提案法	55.46/42.55	2.598/1.846	0.752
2000	基準法	55.63/42.57	2.700/1.877	0.823
先後統合				
500	提案法	56.84/41.62	2.722/1.812	0.910
500	基準法	57.28/41.13	2.745/1.784	0.961
1000	提案法	57.93/40.28	2.719/1.748	0.971
1000	基準法	57.63/40.66	2.742/1.749	0.993
2000	提案法	58.32/39.87	2.752/1.714	1.038
2000	基準法	58.54/39.76	2.825/1.731	1.094

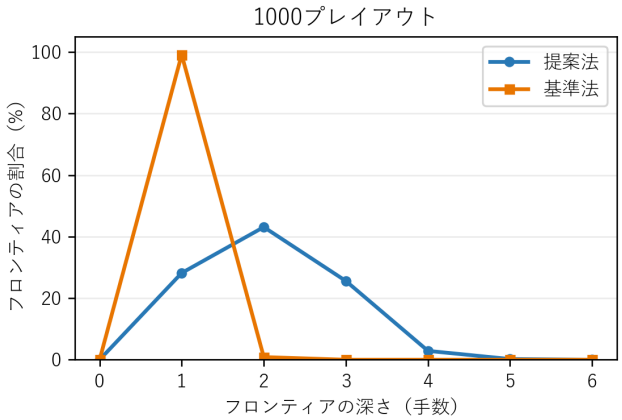


図 3　1000 プレイアウトでのフロンティア深さ比較

特定の方式が常に上回る結果とはならなかった.

図 3 に,1000 プレイアウトにおけるフロンティアの到達手番数ごとの割合を示す. 基準法ではフロンティアの 99.1%が到達手番数 1 にあり, 相手の次の決定ノードに対応する. 自分の次の決定ノードまたはそれより深い位置に当たる, 到達手番数 2 以上の割合は 0.9%であった. これに対し提案法では, 相手の次の決定ノードに対応する到達手番数 1 が 28.2%, 自分の次の決定ノードに対応する 2 が 43.1%, その次の相手の決定ノードに対応する 3 が 25.6%であった. 到達手番数 2 以上は 71.8%であり, 提案法では多くの探索が自分の次の決定ノード以降まで到達していた.

500,1000,2000 プレイアウトにおける平均到達手番数は, 提案法がそれぞれ 1.96,2.04,2.11, 基準法が 1.00,1.01,1.02 であった. 各探索の最大到達手番数を平均した値は, 提案法が 4.28,4.53,4.71, 基準法が 1.17,1.25,1.36 であった. したがって, 提案法はすべてのプレイアウト数で基準法より深い位置へ探索を進めていた. なお, フロンティアには終端状態の評価も含まれるため, 各到達手番数のすべてが対応するプレイヤの決定ノードでランダムプレイアウトを開始したことを意味するものではない. 一方, 表 1 では方式間に一貫した対戦性能の差が見られなかったため, 探索の深さの増加が全体の対戦成績の向上へ直接つながったとはいえない.

表 2　赤短取得順序局面で観測できる札

月	自分の手札	場札
松	松のカス，松に鶴	松に赤短
梅	梅にうぐいす	梅のカス，梅に赤短
桜	桜に赤短	桜のカス
萩	萩のカス	萩のカス
菊	菊のカス	菊に盃，菊のカス
紅葉	紅葉のカス	なし
柳	柳に燕	柳のカス

5.3　個別局面における評価

5.3.1　前提

全体性能評価だけでは個々の局面での選択方針がわからないため, 対局ログから序盤・中盤・終盤の複数局面を選び, 基準法と提案法を 500,1000,2000 プレイアウトでそれぞれ 1,000 シードずつ実行し, 各手の選択割合を比較した. 多くの局面ではプレイアウト数を増やしても方式間の違いが小さかった一方, 赤短の取得順序に関わる局面では比較的大きな差が見られたため, 以下ではこの局面を詳しく調べる.

5.3.2　赤短取得順序局面

この局面は局の第 1 手であり, 自分の手札に桜に赤短, 場札に梅に赤短と松に赤短が存在する. 手札の桜に赤短は手札に保持している間は相手に取られない一方, 場札の赤短は取得しなければ相手に取られる可能性がある. また, この局面では場札の菊に盃を取得でき, 赤短とは異なる有力な行動候補も存在し, 全体としては 9 通りの有効手が存在する. 観測できる札のリストを表 2 に示す. 基準法と提案法を 500,1000,2000 プレイアウトでそれぞれ 1,000 シードで実行し, 全有効手の選択回数と選択割合を求めた.

5.3.3　赤短取得順序局面の結果

図 4 に,2000 プレイアウトにおける各手の選択割合を示す. 菊に盃の選択割合は提案法 57.7%, 基準法 55.0%であり, 両方式で最も多く選択された. 一方, 提案法では梅に赤短を取得する行動が 30.4%, 桜に赤短を取得する行動が 9.7%であったのに対し, 基準法ではそれぞれ 10.4%,33.1%であり, 赤短を取得する 2 手の優先順序が逆になった. こ

© 2026 Information Processing Society of Japan

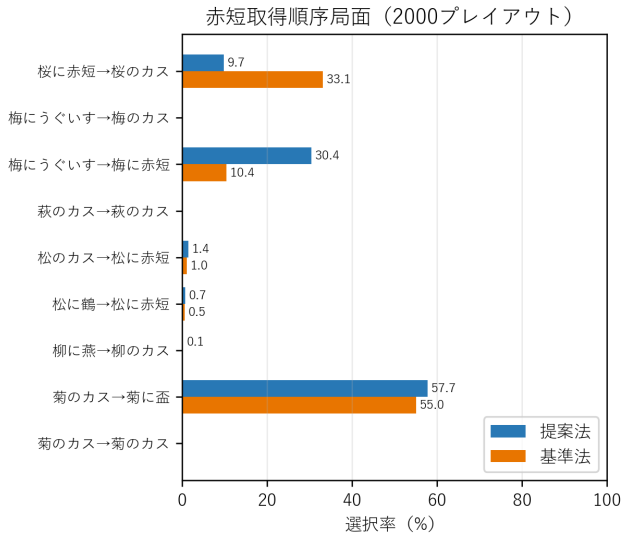


図 4　2000 プレイアウトでの赤短取得順序局面

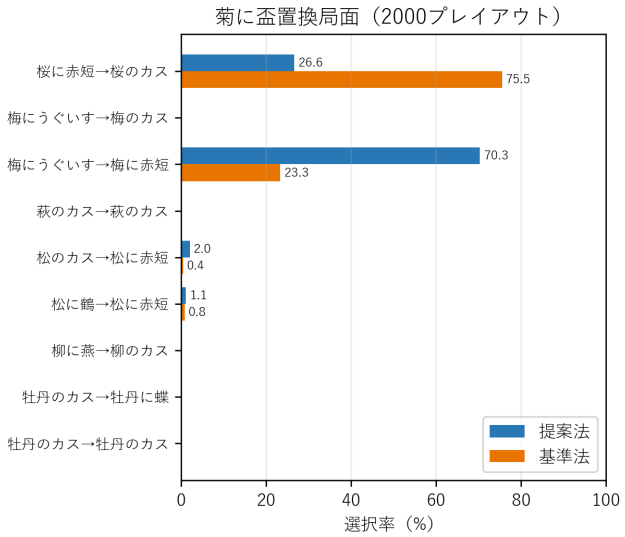


図 5　2000 プレイアウトでの菊置換局面

の順序は 500 プレイアウトと 1000 プレイアウトでも同じであった．500 プレイアウトでは梅に赤短が提案法 26.6%, 基準法 16.2%, 桜に赤短が提案法 17.3%, 基準法 31.4%であり,1000 プレイアウトではそれぞれ 27.7%と 14.8%,15.9%と 31.7%であった．

桜の行動では手札の桜に赤短と場のカス札を, 梅の行動では場の梅に赤短と手札の種札を取得する．桜に赤短は手札にある間は相手に取られない一方, 場の梅に赤短は相手に取られる可能性がある．したがって, 両行動には同時に取得する札種と, 赤短が手札と場のどちらにあるかという違いがある．さらに，桜は場に 1 枚, 梅は場に 2 枚提示されているという条件の違いもある．

5.3.4　赤短取得順序に関する要因評価

赤短取得順序局面では菊に盃が高い割合で選択されるため, 赤短の選択を詳しく比較するために, 菊に盃を別の種札へ置換した．具体的には, 局面内の 6 月と 9 月の札を, 札種と配置の対応を保ったまますべて交換した．この置換によって, 自分の手札にある菊のカスと, 場札にある菊に盃・菊のカスは, それぞれ牡丹のカスと, 牡丹に蝶・牡丹のカスに変更される．一方, 赤短を構成する札と, 有効手の手札・場札の位置対応は維持される．

この置換局面を基準として, 方式間の選択割合の差が生じる要因を確認するため, 梅の手札の札種と, 見えている桜の枚数を変更した 4 条件を用意した．梅の手札については, 梅にうぐいすをそのまま使用する条件と, 未公開の梅のカスへ置換する条件を用いた．桜については, 手札の桜に赤短と場札の桜のカスの 2 枚が見えている条件に加え, 場札の萩のカスを未公開の桜のカスへ置換し, 桜が 3 枚見えている条件を用いた．

いずれの条件でも, 梅は手札 1 枚と場札 2 枚の計 3 枚が見えており, 梅に赤短は場札, 桜に赤短は手札に配置されている．桜を 3 枚見えるようにした条件では, 桜に赤短から取得できる場札が 2 枚となる一方, 萩の組合せがなくなるため, 有効手の総数は 9 通りのまま維持される．同じ札の組合せとなる桜の 2 手については, 結果の集計時に選択回数を合算した．

表 3　菊に盃置換後の要因変更時の選択割合 (2000 プレイアウト, 単位:%)

桜枚数	梅手札	方式	桜赤短	梅赤短	その他
2 枚	種札	提案法	26.6	70.3	3.1
2 枚	種札	基準法	75.5	23.3	1.2
2 枚	カス札	提案法	46.5	48.6	4.9
2 枚	カス札	基準法	86.9	11.6	1.5
3 枚	種札	提案法	0.0	99.3	0.7
3 枚	種札	基準法	0.0	99.7	0.3
3 枚	カス札	提案法	0.0	98.3	1.7
3 枚	カス札	基準法	0.0	99.6	0.4

4 条件について, 基準法と提案法を 500,1000,2000 プレイアウトでそれぞれ 1,000 シードずつ実行し, 全有効手の選択回数と選択割合を求めた．

5.3.5　赤短取得順序に関する要因評価の結果

表 3 に,2000 プレイアウトにおける 4 条件の選択割合を示す．ここでは, 桜に赤短と梅に赤短以外の行動をその他としてまとめている．

4 条件のうち, 桜が 2 枚見えており, 梅の手札が種札である条件が図 5 と同じ菊置換局面に当たる．置換前は両方式とも菊に盃を最も高い割合で選択していたが, 菊に盃を置換すると, 提案法は梅に赤短を 70.3%, 基準法は桜に赤短を 75.5%選択した．この結果から, 両方式が共通して高く評価していた菊に盃を取り除くことで, 赤短の取得順序に関する差がより明確になった．

桜が 2 枚見えている条件で梅にうぐいすをカス札へ置換すると, 提案法の梅に赤短の選択割合は,500,1000,2000 プレイアウトで 43.2%,58.9%,70.3%から 37.0%,42.7%,48.6%へ低下した．　基準法でも 32.1%,29.5%,23.3%から

© 2026 Information Processing Society of Japan

23.1%,19.8%,11.6%へ低下した．したがって, 梅にうぐいすの種札としての価値は両方式の選択に影響していたと考えられる．

一方, カス札への置換後も, 提案法と基準法の梅に赤短の選択割合には,500,1000,2000 プレイアウトでそれぞれ 13.9,22.9,37.0 ポイントの差が見られた．したがって, 梅にうぐいすの種札としての価値は梅に赤短を優先する一因であるが, 方式間の差をそれだけで説明することはできない. また, 桜が 2 枚見えている 2 条件では, プレイアウト数の増加に伴って提案法では梅に赤短の選択割合が増加し, 基準法では低下した.

これに対し, 桜が 3 枚見えている 2 条件では, 梅の手札の札種にかかわらず, 両方式とも梅に赤短を高い割合で選択した．500 プレイアウトの時点で梅に赤短の選択割合は 85.2–91.8%であり,2000 プレイアウトでは 98.3–99.7%となった．2000 プレイアウトでは選択割合が上限付近に達しているため, この条件から梅の種札としての価値が選択へ与える影響の大きさを判断することは難しい.

桜が 2 枚見えている条件では, 手札の桜に赤短を後回しにすると, 場に 1 枚だけ存在する桜札が先に場から取り除かれ, 対応先を失う可能性がある．その一方で, 梅に赤短は場札であるため, 先に取得しなければ相手に取られる可能性がある．この 2 つの取得順序上の危険に対し, 提案法は場の梅に赤短を, 基準法は手札の桜に赤短を相対的に高く評価したと考えられる.

桜が 3 枚見えている条件では, 場に桜のカスが 2 枚存在するため, 一方が場から取り除かれても桜に赤短の対応先が残る．そのため, 手札の桜に赤短を直ちに出す必要性が低下し, 両方式が場の梅に赤短を優先したと考えられる.

以上から, 方式間の差は手札と場札の違いと場札の公開枚数によるリスクが競合する条件で現れる可能性がある. ただし, 本実験は各手の選択割合を比較したものであり, 梅に赤短と桜に赤短のどちらがより良い手であるかを直接評価したものではない.

6. 終わりに

本研究では, 不完全情報と不確定性を持つこいこいを対象に, チャンスノードの訪問回数に応じてチャンス子数を制御する手法を提案した. 新しい子の生成を制限し, 既存のチャンス子を再評価することで, 広さを残しながら, 深く探索する機会を増やした.

全体評価では,500,1000,2000 プレイアウトのすべてで, 提案法が基準法より深い位置まで探索した.1000 プレイアウトでは, 到達手番数 2 以上のフロンティアの割合が提案法で 71.8%, 基準法で 0.9%であった. 一方, 勝率や平均得点にはプレイアウト数を通じて一貫した方式間の差は見られなかった. 個別局面では, 菊に盃を別の種札へ置換した赤短取得順序局面において, 提案法は場の梅に赤短を 70.3%, 基準法は手札の桜に赤短を 75.5%選択した. 以上から, 提案法は全体の対戦性能を一貫して向上させるものではなかったが, 探索を深くすることで, 一部の局面における複数のリスクに対する優先順位を変化させる可能性が示された.

ただし, 個別局面の実験は各手の選択割合を比較したものであり, 提案法が選択した手の優位性を直接示したものではない. 今後は, 各行動を選択した後の勝率や得点を調べるとともに, 青短や猪鹿蝶など他の複数札役でも同様の傾向が見られるかを確認する必要がある.

また, 本研究では手札から出す札の選択に焦点を当てるため, こいこい内の一部ルールを簡略化している. こいこい宣言による局の継続や, 山札からめくった札による取得札の選択,12 局 1 セットでの親子交代や点数戦略など, より一般的なこいこいのルールでの評価が必要である.

参考文献

[1] A Survey of Monte Carlo Tree Search Methods, Cameron B. Browne, Edward Powley, Daniel Whitehouse, Simon M. Lucas, Peter I. Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, Simon Colton, 2012.

[2] Finite-time Analysis of the Multiarmed Bandit Problem, Peter Auer, Nicolò Cesa-Bianchi, Paul Fischer, 2002.

[3] Bandit Based Monte-Carlo Planning, Levente Kocsis, Csaba Szepesvári, 2006.

[4] Continuous Upper Confidence Trees, Adrien Couëtoux, Jean-Baptiste Hoock, Nataliya Sokolovska, Olivier Teytaud, Nicolas Bonnard, 2011.

[5] Monte Carlo Tree Search With Iteratively Refining State Abstractions, Samuel Sokota, Caleb Y. Ho, Zaheen Farraz Ahmad, J. Zico Kolter, 2021.

[6] Information Set Monte Carlo Tree Search, Peter I. Cowling, Edward J. Powley, Daniel Whitehouse, 2012.

[7] MAPLE: Multi-State Aggregated Policy Evaluation for AlphaZero in Imperfect-Information Games, Qian-Rong Li, Hung Guei, I-Chen Wu, Ti-Rong Wu, 2026.

[8] 花札のこいこいにおける方策勾配法と Neural Fitted Q Iteration の適用, 佐藤直之, 池田心, 2017.

[9] Learning to Play Koi-Koi Hanafuda Card Games With Transformers, Sanghai Guan, Jingjing Wang, Ruijie Zhu, Junhui Qian, Zhongxiang Wei, 2023.

[10] UCT 探索を用いた花札の戦略決定の実験的考察, 髙岡勇樹, 川上敬, 大江亮介, 三田村保, 木下正博, 2015.

[11] モンテカルロ法の報酬の変更による花札「こいこい」の戦略構築, 長峰拓己, 鈴木徹也, 2021.

[12] Amê: An Environment to Learn and Analyze Adversarial Search Algorithms Using Stochastic Card Games, Ana Beatriz Cruz, Sabrina Serique, Leonardo Preuss, Angélica Ogasawara, João Quadros, Eduardo Bezerra, Uéverton Souza, Eduardo Ogasawara, 2015.

© 2026 Information Processing Society of Japan