

同時着手ゲームにおける Alpha-Beta 探索

小森成貴, Alessandro Cincotti, 橋本剛, 飯田弘之
北陸先端科学技術大学院大学

概要

組み合わせゲームは二人のプレイヤーが交互に着手を行い、通常のルールでは最初に合法手なくなったプレイヤーが敗者となる。同時着手ゲームにおいてプレイヤーは同時に盤面に着手を行う。結果として、両プレイヤーの合法手なくなったとき引き分けに終わる可能性が存在する。Synchronizied Domineering は、古典的な二人組み合わせゲーム Domineering の同時着手版である。本稿では、Alpha-Beta 探索による Synchronized Domineering のコンピュータプログラムを紹介する。現在得られている実験結果では、Alpha-Beta 探索によるプレイヤーはモンテカルロ法によるプレイヤーよりもやや強い。

Alpha-Beta Search on Synchronized Game

Shigetaka Komori, Alessandro Cincotti, Tsuyoshi Hashimoto, Hiroyuki Iida
Japan Advanced Institute of Science and Technology

Abstract

In combinatorial games two players take turns in alternation and in normal play convention the first player unable to move is the loser. In synchronized games the players make their moves simultaneously, and, as a consequence, there exists the possibility to get a draw when both players are unable to move. Synchronized Domineering is the synchronized version of Domineering, a classical two-player combinatorial game. In this paper, we present a computer program for Synchronized Domineering using an Alpha-Beta search approach. Experimental results so far obtained show that our Alpha-Beta approach is a bit more stronger than a Monte-Carlo approach.

1 はじめに

同時着手ゲームは中村らによって提案され [9], その後, Cincotti らによって, 組み合わせゲームである Cutcake, MaundyCake, Domineering に適用された [4][5][6]. 組み合わせゲームは, 最初に合法手なくなったプレイヤーが敗者となるゲームであるが [2], 同時着手の概念を適用することにより, 通常では起こりえなかったゲームが引き分けで終わる可能性が生まれた. また, 同時に着手することにより, 先手勝ちや後手勝ちと言った手番に起因する有利性がなくなる.

これまでの同時着手ゲームに関する研究において, 理論的な解析はなされてきたが, コンピュータプレイヤーについては議論されてこなかった. そこで筆者は組み合わせゲームである Domineering に同時着手の概念を適用した Synchronized Domineering [6] についてモンテカルロ法を用い

てコンピュータプレイヤーを実現した [10].

本研究では, 探索と評価関数を用いて Synchronized Domineering のコンピュータプレイヤーを実現し, モンテカルロ法によるプレイヤーとの比較を行う.

2 Domineering と Synchronism

本研究では組み合わせゲームである Domineering に同時着手の概念を適用した Synchronized Domineering を対象としてコンピュータプレイヤーの実現を行った. 本節において Domineering および Synchronized Domineering の紹介を行う.

2.1 Domineering

Domineering は Crosscram や Dominoes という別名でも知られている二人零和完全情報ゲーム

であり, Göran Andersson によって 1973 年ごろに提案された [7]. 二人のプレイヤーは通常 Vertical と Horizontal と呼ばれ, ボード上に 2×1 のドミノを交互に置いていく. Vertical はドミノを垂直に置くことのみが許可され, Horizontal はドミノを水平に置くことのみが許可される. ドミノが重なることは許可されず, 最初にドミノを置けなくなったプレイヤーが敗者となる. 既存研究により Domineering は様々なボードサイズについて解かれている [1][3][8].

2.2 Synchronized Domineering

Domineering に同時着手の概念を適用したゲームが Synchronized Domineering である. Synchronized Domineering のルールは Domineering とほぼ同じであり, 唯一の違いは, 二人のプレイヤーが交互ではなく同時に着手を行うという点にある. 従って, 着手時にドミノの一部が重なることが考えられ, これを許可する. 詳細は [6] を参照されたい.

3 コンピュータプレイヤーの設計

コンピュータプレイヤーを実現するために主に二種類の方法が使われている. ひとつは将棋やチェスなどのゲームで使われている, 探索アルゴリズムと評価関数を用いる方法であり, もうひとつは囲碁などのゲームで使われている, モンテカルロ法を利用した方法である. 本研究では Synchronized Domineering のプレイヤーとして探索と評価関数を用いる方法を考える.

探索と評価関数によるプレイヤーには次節で示すような特徴があり, 必ずしも常に良いプレイヤーであるとは言えるわけではないが, 同時着手ゲームにおける最初の探索プレイヤーを考えるうえで最も単純であるため探索を考えた.

Synchronized Domineering におけるモンテカルロ法に関しては先行研究 [10] を参照されたい.

3.1 プレイヤーの特徴

Synchronized Domineering は二人のプレイヤーが同時に着手を行うため, 探索の結果より着手を

行ったとしても, 探索の結果の局面の到達するとは限らない. つまり, 利得としては得られた評価値を最低限保証し, さらに相手の着手によってはそれ以上の利得を得られるという特徴がある.

具体的には 8 割の確率で負けるが 2 割の確率で勝てる手 [a] と, 勝つことはできないが負けることもなくゲームが引き分けで終わる手 [b] があった場合, 探索を行うプレイヤーは, 最低の利得を保証するので, たとえ引き分けに終わっても負けることのない手 [b] を選ぶ.

しかし, 逆に 8 割の確率で勝つが 2 割の確率で負ける, というような勝ちやすい手がある場合にも負ける可能性があればその手は選ばない.

そのため, このような探索によるプレイヤーが最良であるとはいえず, どのような探索プレイヤーが最良なのかは未知数である.

本研究では同時着手ゲームにおける探索プレイヤーを考えるスタートとして以下のアルゴリズムにより探索を実現した.

3.2 探索アルゴリズム

Synchronized Domineering では, 二人のプレイヤーが同時に着手を行うが, 本研究では, 擬似的に自分が先手で対戦相手が後手と仮定することによって探索を行った. しかし, この探索において先手が着手した後の局面で後手の着手を考えてしまうと, それは手番を交代する Domineering となってしまう. そのため, 後手の着手を考えるときは, 先手が着手した後の局面ではなく, その前の局面において考えるようにしなければならない.

図 1, 2 に例を示す. 現在局面において Vertical には [V1, V2, V3] という 3 つの合法手が, Horizontal には [H1, H2] という 2 つの合法手が存在する例である. このとき, 次に現れる局面は [V1H1, V1H2, V2H1, V2H2, V3H1, V3H2] の 6 種類のいずれかになる. つまり, 図 1 の局面遷移となる. ここで, 擬似的に Vertical が着手 [V1 or V2 or V3] を行ってから Horizontal が着手 [H1 or H2] を行うと考えると, 図 2 の木構造とみなせる. この木構造を考えた場合, 従来と同様の探索を行うことが可能である. なお,

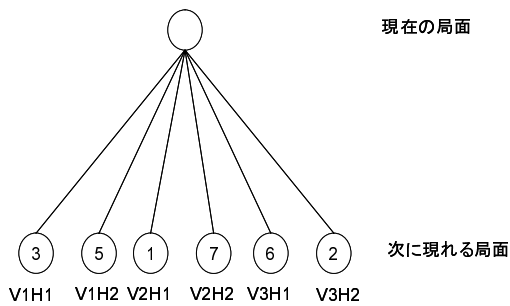


図 1: 実際の局面遷移

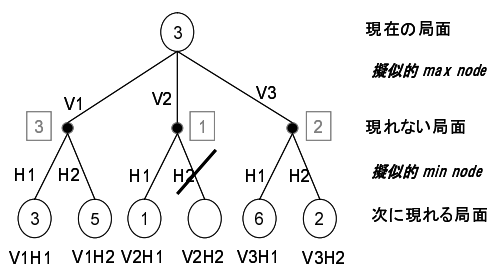


図 2: 擬似的に手番を考えたときの木構造

Horizontal の合法手 [H1, H2] は実際には現れない の局面での合法手ではなく、もとの局面における合法手である。

この例で Vertical が 探索を行うと、局面 [V1H1] の評価値は 3, [V1H2] の評価値は 5 であり、相手は自分にとって都合の悪い着手をしていくという仮定のため、Vertical の合法手 [V1] の評価値は 3 となる。次に局面 [V2H1] の評価値は 1 であるため、局面 [V2H2] の探索はカットすることができ、合法手 [V2] の評価値は 1 となる。さらに局面 [V3H1] の評価値が 6, [V3H2] の評価値が 2 となるため、合法手 [V3] の評価値は 2 となり、これらのなかから評価値が最大の合法手を選ぶと [V1] の 3 となるため、Vertical は合法手 [V1] を選ぶ。

この図 2 の例においては Vertical が 探索を行っているが、もしも Horizontal が 探索を行う場合、Horizontal が先手で Vertical が後手となる擬似的な探索木を考える。

3.3 評価関数

Synchronized Domineering は、手番を交代する Domineering と同様に、着手できなくなったほうが負けというゲームである。そのため、[3] を参考に残り着手可能数 rm と安全着手数 sm を利用して評価関数を作成した。ここで、残り着手可能数 rm とは、現在の局面において対戦相手が着手をしてこないという仮定のもと、自分があと何回着手できるかを示した数であり、安全着手数とは、対戦相手の着手にかかわらず自分が着手できる回数 (対戦相手に邪魔されずにドミノを置ける場所の数) を示した数である。

評価関数は、自分の残り着手可能数 rm_{player} および安全着手数 sm_{player} 、対戦相手の残り着手可能数 $rm_{opponent}$ および安全着手数 $sm_{opponent}$ を用いて以下とした。

$$value = (rm_{player} - rm_{opponent}) + (sm_{player} - sm_{opponent})$$

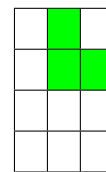


図 3: 3 × 4 ボードの例

図 3 は 3 × 4 というボードサイズにおいて 1 回同時着手した後の局面の例である。この局面において、Vertical の残り着手可能数 rm は 4 であり、安全着手数 sm は 1 である。同様に、Horizontal の残り着手可能数 rm は 2 であり、安全着手数 sm は 0 である。

4 性能評価

探索と評価関数によるプレイヤーを評価するために、先行研究 [10] で作成したモンテカルロ法によるプレイヤーとの対戦実験を行った。

4.1 実験のデザイン

ボードサイズは、Domineering における標準的なサイズである 8 × 8 と、それより一回り小

表 1: Experimental results

		7 × 7		8 × 8	
V \ H		4	6	4	6
10000	V win	87	40	3	28
	H win	139	90	77	81
	Draw	774	870	920	891

さい 7 × 7 の 2 種類とした。これらは理論値が引き分けになると予測されているボードである [6][10]。

それぞれのプレイヤーの設定として、モンテカルロ法のプレイアウトを 10000 回と、探索の深さを 4 および 6 とした。ここで探索は擬似的に先手と後手を設定しているため、探索の深さ 4 とは 2 回同時着手した後の局面、探索の深さ 6 とは 3 回同時着手した後の局面を指す。

この条件において、Vertical をモンテカルロ法によるプレイヤー、Horizontal を探索と評価関数によるプレイヤーに設定し、それぞれ 1000 ゲームの対戦実験を行った。

4.2 結果と考察

実験結果を表 1 にまとめる。なお、1 ゲームの平均思考時間は、探索と評価関数によるプレイヤーが 7 × 7 の探索深さ 4 で 0.185sec、探索深さ 6 で 7.3141sec、8 × 8 の探索深さ 4 で 0.4641sec、探索深さ 6 で 23.5061sec であり、モンテカルロ法によるプレイヤーが 7 × 7 で 4.5904sec、8 × 8 で 8.0502sec であった。

対戦結果より、どのケースにおいても 75% 以上ゲームは引き分けに終わっている。また、モンテカルロ法によるプレイヤー (Vertical) と探索と評価関数によるプレイヤー (Horizontal) を比較すると、Vertical の勝ち数よりも Horizontal の勝ち数のほうが多い。よって、理論値が引き分けとなると予測されているというボードで、先行研究 [10] において有力なプレイヤーであることが示されたモンテカルロ法と同等以上の強さがあるといえる。

探索の深さを 4 から 6 にすることによって思

考時間は圧倒的に増えるが、その結果、自分の勝ち数が大きく増えるということではなかった。ボードサイズ 7 × 7 においては、両プレイヤーの勝ち数が減り、引き分けで終わる回数が増えた。これは、探索深さを一段階増やしても、もともとの評価関数の精度が高く、深くもぐってもアドバンテージを得られなかったため、勝ち数が大きく増えなかったのではないかと考えられる。

以上より、今回作成した探索と評価関数によるプレイヤーは、評価関数の精度が高かったことにより強いプレイヤーになったと考えられる。

5 まとめと今後の課題

本研究では同時着手ゲーム Synchronized Domineering について、探索と評価関数を用いてコンピュータプレイヤーを実現した。そしてモンテカルロ法によるプレイヤーとの対戦実験を行い、互角以上に勝負できることが確認できた。

しかし、この探索と評価関数によるプレイヤーには 3.1 節で述べたような特徴があり、常に最良のプレイヤーであるとは限らない。そのため、最高利得を期待するようなプレイヤーや、平均利得を考慮するプレイヤーなどを調べていきたい。

参考文献

- [1] E. R. Berlekamp, "Blockbusting and Domineering", *Journal of Combinatorial Theory Ser. A*, vol.49, pp. 67-116, 1988.
- [2] E. R. Berlekamp, J. H. Conway, R. K. Guy, "Winning Ways For Your Mathematical Plays", Vol.1, Academic Press 2nd Edition, 2001.
- [3] D. M. Breuker, J. W. H. M. Uiterwijk, and H. J. van den Herik, "Solving 8 × 8 Domineering", *Theoretical Computer Science*, vol. 230, pp. 195-206, 2000.
- [4] A. Cincotti and H. Iida, "The Game of Synchronized Cutcake", in *Proceedings of the IEEE Symposium on Computational Intelligence and Games (CIG 2007)*, pp. 374-379, 2007.
- [5] A. Cincotti and H. Iida, "The Game of Synchronized Maundy Cake", *Hawaii International Conference on Statistics, Mathematics and Related Fields*, pp. 422-429, 2008.
- [6] A. Cincotti and H. Iida, "The Game of Synchronized Domineering", *Computers and Games 2008*, pp.241-251, 2008.
- [7] M. Gardner, "Mathematical games", *Scientific American*, vol. 230, pp. 106-108, 1974.
- [8] M. Lachmann, C. Moore, and I. Rapaport, "Who Wins Domineering on Rectangular Boards", in *More Games of No Chance*, edited by R. Nowakowski, Cambridge University Press, 2002.
- [9] T. Nakamura, A. Cincotti, and H. Iida, "The Rebirth of Solved Games "8th International Conference on Computer Science and Informatics", 265-269, 2005.
- [10] 小森成貴, Alessandro Cincotti, 飯田弘之, "同時着手 Domineering におけるモンテカルロ法の適用", *IPSJ-SIG-GI-20*, 51-58, 2008.