

GPCC 報告 (2008 年)

Games and Puzzles Competitions on Computers

<http://hp.vector.co.jp/authors/VA003988/gpcc/gpcc.htm>

藤波順久*

酒井香代子†

1 2008 年の課題

2008 年の GPCC では、以下の 3 個の課題を取り上げた。

Celtis 6 角形のボード (一辺 8) を使った 7 目並べである。駒は鈍角に曲がった 3 連で、表が青青緑、裏が緑緑青となっている。駒はすべて同じで、鏡像対称のものはない。

ブロックデュオ 二人で行うボードゲームである。ボードは 14×14 のマス目でできていて、駒はサイズが 1~5 のポリオミノ全種類 (一人あたり 21 ピース) である。以下のルールに従って交互にピースを置いていく。置けなくなった人はパスで、二人とも置けなくなったら終了である。ボードを覆うマス目の多い人が勝ちである。

- ボードの一つの対角線の、両端から 5 番目の位置 (2 箇所) に印がついている。それぞれ 1 手目のピースはこのマス覆うように置かなければならない。
- 2 手目以降のピースは、自分のピースの角同士が接し、辺同士は接しない位置に置かなければならない。相手のピースとはどのように接してもよい。

固定ピースヘキサミノ ヘキサミノ 35 種類を箱 (19×11 の長方形の長辺の中央に正方形が飛び出した形) に詰めるパズルは、プラパズル No.600 として知られているが、解の数が膨大である (推定 $10^{20} \sim 10^{25}$) [1][2]。そこで、ピースをいくつか先置きして固定し、解が一つになるようにするのがこの問題である。固定ピースの数が少なければ少ないほどよい解であるとする。

2 2008 年の進展

Celtis については、前回のプログラミング・シンポジウムの夜のセッションで、人間同士の対戦が行われた。元々のルール (長連は勝ちではない) では、引き分けになることが多かったため、ルールを変更しながら対戦して、以下のように、1 回の鈍角曲がりを含む 8 連も勝ちするとよさそうだということがわかってきた。

- 直線の 7 連は勝ち
- 直線の 8 連以上は勝ちではない

*株式会社ソニー・コンピュータエンタテインメント、GPCC chair

†中央大学大学院理工学研究科、GPCC co-chair

お名前	プログラム名	勝敗数	SB	順位
坂本 邦彦 さん	hmmmm	13勝 1敗	38.5	1
柴原 一友 さん	EN-FIS	12勝 2敗	31.0	2
築地 毅 さん	Mate_with_pawn_drop Club	9勝 5敗	20.5	3
但馬 康宏 さん	パクリ上等	9勝 5敗	17.5	4
山口 文彦 さん	GU	7勝 7敗	10.5	5
佐々木 健太 さん	ぶーん(^_^)	4勝 10敗	2.0	6
東 祐子 さん	(_ _).o0	1勝 13敗	0.5	7
佐藤 直人 さん	ぼんち	1勝 13敗	0.5	-
古山 大輔 さん	full	(4敗)		-

表 1: 応募者名、プログラム名、成績 (-は参考記録)

- 1回の鈍角曲がりを含む8連は勝ち
- 1回の鈍角曲がりを含む9連以上は勝ちではない
- 勝ちと勝ちでないパターンの両方ができたら、勝ちが優先

別の言い方にすれば、以下の4パターンの一つを作ると勝ちというルールである。ここで、1~8を自分の色、#を相手の色、空き、盤外のどれかとする。

- a #-1-2-3-4-5-6-7-#
- b #-1-2-3-4-5-6-7(鈍角曲)7-8-#
- c #-1-2-3-4-5-6(鈍角曲)6-7-8-#
- d #-1-2-3-4-5(鈍角曲)5-6-7-8-#

bのパターンは、直線方向が長連の場合に必要な。

ブロックデュオについては、昨年と同様に東京農工大学の小谷研究室の大崎さんを中心とする方々に大会の運営をしていただき、4月20日対戦日程公開、9月13日応募要領公開、10月3日募集締切、10月25日大会実施という日程でプログラム対戦を実施した。以下3節で大会の結果を示し、4節で交流会での優勝プログラムと人間の対戦を解説する。5節では2位のプログラムのアルゴリズムを、6節では優勝プログラムに唯一1勝した3位のプログラムのアルゴリズムを説明する。

固定ピースヘキサミノについては、進展はなかった。

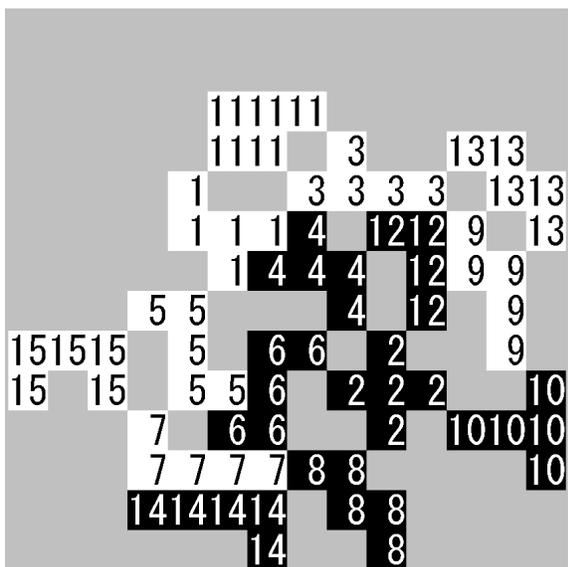
3 前回優勝者の連覇、しかし...—ブロックデュオ プログラム対戦

今回は応募数が9件と少なめであったため、原則として先手後手総当たりで対戦を行った。勝ち星が同じ場合には、SB(勝利した相手あげている勝ち星の総数)を使用して順位を決めた。応募された方のお名前、プログラム名、成績を結果を表1に示す。すべての対戦の勝敗および棋譜については、<http://hp.vector.co.jp/authors/VA003988/gpcc/08g2b.htm>を参照してほしい。

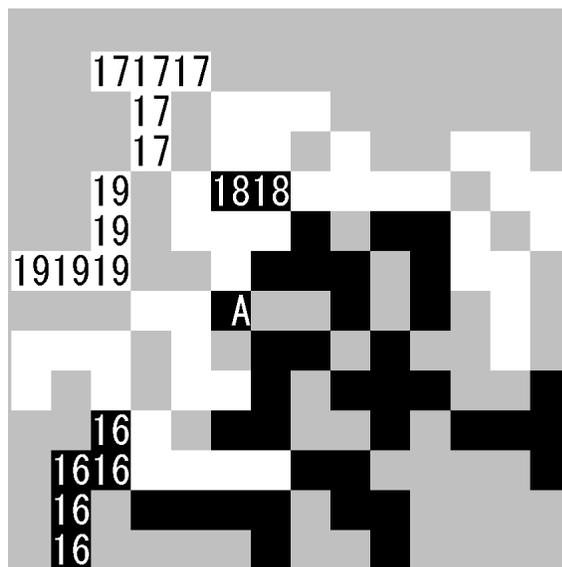
対戦時の手の入力と表示に「4文字コード」¹を用いることを、今回から義務化したため、棋譜を迅速に公開することができた。応募された方、運営された方のご協力に感謝したい。

優勝したのは、前回と同じ坂本邦彦さんのプログラムであった。「思考ルーチンは去年とほぼ同じもので、前向き枝刈りにMulti-ProbCutを使うようにしたのが主な変更点です。」とのことで

¹ピースの位置を英数字2文字、種類を英字1文字、向きを数字1文字で表す。棋譜には図と4文字コードの両方を掲載した。詳細は<http://hp.vector.co.jp/authors/VA003988/gpcc/08g2.htm>を参照。



途中図: 15手目 (29n6) まで



参考図: 16手目 (2C14) 以降 43p4,75b2,37q1 と指したところ Aは68a0

図 1: hmddd 対人間の途中図と参考図

ある [3]。前回は全勝優勝であったが、今回は3位のプログラムに後手番で1敗を許してしまった。大会後の感想戦で注目の一戦となった。

4 人間側敗北—交流会での優勝プログラムとの対戦

大会後の交流会では、恒例により優勝した hmddd と人間 (相談あり) の対戦が行われた²。今回は人間側が後手を選んだ。

5手目 (59s3) で早くも hmddd が長考。161秒かけて5手先を読んだ。その後 hmddd が15手目 (29n6) を指したところまでは、人間側に余裕があった (図1左)。hmddd の包囲網は75b2または68a0 (図1右の18またはA) で突破でき、その際29n6の上の面が平らなのがhmddd に不利に働くと見られたからである。しかしここで hmddd が43p4を指すとまずいことがわかった。人間側が75b2を置いても hmddd が37q1とすれば、指せる手は非常に限られてしまう (図1右)。実際に43p4は、hmddd の17手目の候補に挙がった。しかし hmddd は約2分かけて12手先を読み、8Bd0を指した³。

人間側は長考をしたが、5手先をどう読んでもうまいかないことが判明した。思い返してみると、11手目の73m7が効いているのであった⁴。人間側は負けを覚悟して指し続ける。21手目 (42u0) で突然、hmddd 思考時間が短くなり、甘い手を指した。勝ちを読み切ったのではないかと思われる。以後 hmddd は短時間で指し続け、最終的に7ポイント差となった (図2)。

対戦者によると、「人間側の敗着を挙げるとするならば、10手目のDBp3でしょう。自分の地に

²対戦には比較的高性能なノートパソコンが使われた (大会で使用したPCとは仕様が異なる)。

³対戦者より: 15手目の29n6は最初意味がわかりませんでした。しかし、手を進める内に人間側が苦しいことがわかりました。おそらく6手先 (21手目の42u0) のことを考えていたのでしょう。29n6の段階で人間側は負けると思いました。あとは、hmddd のミスを誘うような手しか打つことができませんでした。

⁴対戦者より: 11手目に hmddd がこの対局で初めて受けに回ります。それまでは人間の地を奪うような手の連続でした。ここで人間側は攻めのピース、12手目のB7f5を置きます。当然、hmddd はD5r0で受けました。

2433	21	292929	23	3131
24	212121		232323	3232
24242421	111111	233030	32	
	22	1111	330301313	
	1922	1	3 3 3 328	1313
	1922	1 1 1 4	1212	9 13
19191922		1 4 4 4	12	9 9
202020	5 5		4	12 9
15151518	5	6 6	2	9
15	1518	5 5 617	2 2 2	10
	16 7	6 61717	2	101010
1616	7 7 7 7 8	82525		10
162714141414		8 8252526		
16272727	14	8	262626	

- 1:66t0
- 2:AAu0
- 3:95o6
- 4:87t4
- 5:59s3
- 6:7As2
- 7:5Ck2
- 8:9Dr0
- 9:D715
- 10:DBp3
- 11:73m7
- 12:B7f5
- 13:D5r0
- 14:6Dk6
- 15:29n6
- 16:2C14
- 17:8Bd0
- 18:49b0
- 19:37q1
- 20:28c2
- 21:42u0
- 22:45e0
- 23:B2p2
- 24:13q0
- 25:BCi0
- 26:DEg6
- 27:4Ef2
- 28:C5a0
- 29:71c2
- 30:B3i1
- 31:E1b2
- 32:E2d3
- 33:21a0

hmmmm: 72
人間: 65

図 2: 最終図

こだわり過ぎた一手でした。今思えば相手の地に踏み込む74o2など⁵が良かったでしょう。昨年より格段と強くなったと実感しました。」とのことである。

5 EN-FISのアルゴリズム

(本節は、2位となったEN-FISのアルゴリズムについて、作者の柴原一友さんに書いていただいたものである。)

EN-FISでは通常のゲーム木探索による着手決定ではなく、モンテカルロアルゴリズムによる着手決定を行っています。モンテカルロアルゴリズムとは、近年囲碁で多く適用されるようになったアルゴリズムであり、評価関数や探索アルゴリズムを必要としない方法です。簡単に説明すると、モンテカルロアルゴリズムでは、ある局面の価値を、「その局面からランダムに着手をしつづけた際、たどり着く終了局面の勝敗を数多く集計し、勝つ確率が高い局面ほど価値が高い局面」とみなす方法です。詳しくは[4]を参照していただきたいと思います。

モンテカルロアルゴリズムでは、ランダムで着手しつづけたときに辿り着く終了局面を得る一回の試行をPlayoutと呼んでいます。このPlayoutを、より優勢な局面に集中させて試行することが重要であり、そのための様々な方法が提案されています。今日で最も有名なアルゴリズムとしてUCTがあり、これまでに得られた各局面の勝敗の統計値を元に、勝率が高くなりそうな局面に絞ってPlayoutを実行することを可能にしています。

EN-FISではPlayoutをより多く実行する方法として、まずPlayout自体の高速化を行っています。ブロックデュオは着手の生成に時間がかかりやすく、Playoutの回数が少なくなりがちなのですが、現在のEN-FISでは手元のPCで、秒間約5,000回のPlayoutを実行出来るようにしています。

⁵これで11手目の73m7は置けなくなる。

また、Playout を集中させるアルゴリズムとして、今日有名な UCT を使用せず、UCT 以前に提唱されていたアルゴリズムを使用しています [5]。このアルゴリズムは、各候補手の勝敗分布を正規分布と仮定して、勝率が最も高い候補手に Playout を実行することに加えて、その候補手より大きな価値を期待できる確率を各候補手に対して算出し、その確率に応じて Playout を実行するというものです。UCT を使用したアルゴリズムも作成してみましたが、こちらの方がよさそうに感じたので、こちらを使用しています。

また、通常勝ったか負けたか、つまり勝敗を統計値とするのが一般的なのですが、EN-FIS では最終局面でのスコアを代わりに使用しています。つまり、最終スコアが最も高い候補手より大きな価値を期待できる候補手に対して Playout を多く実行します。最終スコアを使用することで、勝敗を使用する場合より多くの情報を得ることが出来ますが、統計的に不安定になりがちなので、手の Playout 試行回数が一定数以上にならないと、最終的な指し手として決定しないようにしています。

基本的なアルゴリズムは以上であり、極めて単純なアルゴリズムで実行されています。また、モンテカルロアルゴリズムでは盤面を広く網羅できる手を序盤に指しやすく、特に I 型 (1 × 5 のピース、GPCC の 4 文字コードでは j) を置きやすいようです。これは人間からみると明らかな悪手であることが多いため、序盤ではこの I 型のピースを置く手を生成しないなどの処理も行っています。

6 Mate_with_pawn_drop Club のアルゴリズム

本節は、優勝プログラムに唯一 1 勝した「Mate_with_pawn_drop Club ver.22 僕と私と打ち歩詰めの会」のアルゴリズムについて、作者の築地毅さんに書いていただいたものである。）

本プログラムは以下のアルゴリズムで構成されている。

- 評価関数による 探索による着手選択
- 反復深化は行わず、序盤 3 手・中盤 5 手・終盤 7 手読み
- hmmm を参考にした地の価値と、座標ごとの駒価値・有効マス価値・死角マス価値を手作業により与えそれに基づく評価関数の設計
- 序盤と終盤で用いる評価関数を変化

探索は 探索を採用し、可能手が 200 手になるまでは深さ 3、200 手から 100 手までは深さ 5、それ以下は深さ 7 の深度で行った。可能手数による深さ設定により、時間切れの問題はなかったがもう少し時間をかけられたと思われる局面も多く見受けられた。

評価関数は座標ごとに価値を与えるという手法を取ったため、パラメータ数が非常に多いものとなった。しかしプログラマの棋力が値を調節するに足らなかったためほとんど調節はされていない。ただし、[6] による学習結果を参考に、以下に示す値の方針を取っている。

- 序盤は敵の裏に回り込む座標の価値が高い
- 終盤は盤の辺の価値が高い

また、終盤はほとんど駒価値のみの評価となるように調節した。さらに地の計算であるが、hmmm による地の計算を参考にした [7]。ブロックスは陣地を取り合うゲームであるため、地の計算は非常に重要であり、本プログラムにおける評価関数の中でも非常に重要なもののひとつである。hmmm による地の価値は、筆者の知るところ「パクリ上等」にも用いられており、概ね有効に働

いていると聞いている [8]。このような調節により、可能手が多いため深く探索ができない序盤においても、「敵の陣地を攻める」「自分の陣地を守る」着手を指せるようになったようだ。

序盤と終盤における異なった評価関数の扱いだが、以下の式によって計算されている。

$$E = (1 - t)E_{\text{序盤}} + tE_{\text{終盤}}$$

$$t = 0.028 \text{ 手数}$$

ブロックスは終了する最大手数が決まっているため、手数による進行度 t を用いて計算した。

本プログラムが hmmm に勝つことができたのは偶然の産物であると思う。ここまで述べたように、プログラムとしては非常に未熟なものであるため、改良を加えればさらに強化できる可能性は高いと考えている。考えられる改良点として以下のものが挙げられる。

- 反復深化・Scout などによる探索の強化
- 学習による評価関数の強化

特に評価関数の強化は必須であると考えられる。大会後の追加実験により、現状の評価関数に反復深化を組み込み探索深さが深くなったプログラムが、大会時のプログラムに 8 勝 12 敗と負け越したためである。これは、序中盤において適切な評価ができなかったために有利な局面に持ち込めないまま、ほとんど探索による利点の差がない終盤になってしまったためだと考えられる。このような評価関数で 3 位の結果を残せたのは、駒価値の割合が高くなる裏を返せば細かいパラメータ調節の必要がない中終盤の深さ 5-7 手探索が有効に働いたためだと考えている。更なる強化のためには、探索と評価関数の強化をバランスよく行っていかなければならないようだ。

参考文献

- [1] 寺田実: 複合モンテカルロ法による探索木の解総数の見積り, 情報処理学会論文誌, Vol.38, No.8, 1997.
- [2] プラズル No.600 (ヘキサミノ) の完全解を求めて, <http://hexomino.hp.infoseek.co.jp/>
- [3] irori の日記, <http://d.hatena.ne.jp/Irori/20081103/1225730691>
- [4] 美添一樹: 「コンピュータ囲碁におけるモンテカルロ法」(理論編), <http://minerva.cs.uec.ac.jp/~ito/entcog/contents/lecture/date/20080614.html>
- [5] Rémi Coulom, Efficient Selectivity and Backup Operators in Monte-Carlo Tree Search, 5th International Conference on Computer and Games (CG2006), 2006.
- [6] 大崎泰寛, 柴原一友, 但馬康宏, 小谷善行: モンテカルロシミュレーションを用いた強化学習法の提案, ゲーム情報学研究会報告, Vol.2008, No.28, 2008-GI-19, pp.37-44, 2008.
- [7] irori の日記, <http://d.hatena.ne.jp/Irori/20071104/1194151812>
- [8] 但馬康宏, 小谷善行: モンテカルロ法における勝率近似関数の組み込み方法, 第 13 回ゲームプログラミングワークショップ, pp.100-103, 2008.

各 URL は 11 月 30 日にアクセスできることを確認した。