3 種類の理論値推定法の提案

北 隼人 [1],Alessandro Cincotti[1] ,飯田 弘之 [1,2]

[1] 北陸先端科学技術大学院大学

[2] 科学技術振興事業団さきがけ研究 21「機能と構成」領域

E-mail: h-kita@jaist.ac.jp, cincotti@jasit.ac.jp, iida@jaist.ac.jp

概要

理論値はゲームの最も重要な性質のひとつであり,N 人零和完全情報ゲームには必ず存在する.しかし,複雑なゲームにおいてはそれを求めることはほとんど不可能といってよい.そこで我々は自己対戦実験の結果を対象とする 3 種類の理論値推定法を提案する.その推定法をいくつかのゲームに適用し,それぞれ結果について議論する.

Proposal of three game-theoretical value prediction methods

Hayato Kita[1], Alessandro Cincotti[1], and Hiroyuki Iida[1,2]

[1]Research Unit for Computers and Games, JAIST, Japan

[2]PRESTO, Japan Science and Technology Agency, Japan

Abstract

One of the most important characteristic of a game is represented by its theoretical value. And Every N-player zero sum game with perfect information has a unique theoretical value. However, to find out the theoretical value of a game is time expensive and it is almost impossible to establish it in complex games. Therefore, we propose three different methods to predict the outcome of the game under ideal play. We apply these methods to some games and we discuss the validity of every single method.

1. はじめに

N 人零和完全情報ゲームには必ず固有の理論値が存在するが,その理論値を求めることは非常に困難である.理論値を求める場合,基本的にしらみつぶしの探索によらなければならないが,三目並べのような単純なゲームならともかく,将棋や囲碁のような複雑なゲームをしらみつぶしで解いてしまうのは現状では不可能なことである.例えば,6×6 マスのリバーシを解くのに要した時間は 2 週間であり,その際に探索したノード数は 4×10^{10} ノードであった[2].これが,8×8 マスのリバーシになれば 10^{58} ノードの探索が必要になると言われている.既に現実的な数字ではない.さらにこれよりも探索空間が大きい将棋や囲碁では,理論値を求めることは絶対に不可能である.

理論値は非常に重要な性質である. しかし,複雑なゲームの理論値を求めることはほぼ不可能である.理論値を求めたければ,何か別のアプローチが必要になる.そこで,我々は理論値を推定するための 3 種類の方法を提案する.3 種類の方法とは,Diminishing return を利用する方法,勝率の変化を回帰して理論値を推定する方法,得点差の変化を回帰して理論値を推定する方法である.いずれの方法も自己対戦実験によって得たデータを利用する.

これらの 3 種類の方法の利点や欠点,また,その正当性を議論するために既に解かれたゲームである 6×6 マスのリバーシや tic-tac-toe を題材とする.

2. 自己対戦実験について

本稿で題材として扱うゲームは,チェス,tic-tac-toe,6×6 リバーシ,8×8 リバーシの 4 種類である.これらのゲームのデータは全て自己対戦実験で得られたものを使用する.

ゲーム名	実験者	使用プログラム
チェス	E Heinz	FRITZ 6
Tic-tac-toe	梶原など	自作プログラム
6×6 リバーシ	北	自作プログラム
8×8 リバーシ	北	WZebra

表 1:各ゲームの実験者など

チェスはE Heinzが行った実験のデータを使った[1].この実験はX手読み対X+1手読みのハンデキャップマッチで行われたものである.

三目並べは梶原らが行った実験のデータを使用した.これはセミランダムプレイで行われたものである.それぞれの深さで10000試合ずつ行っている.

6×6および8×8リバーシは筆者が実験を行った.6×6では自作のプログラムを使用した.評価関数は位置評価,着手可能数,および角と辺の確定石のみを考慮しており,その中でも確定石を最も重視したものになっている.また,ある程度の結果の幅を出すため,一手ごとに最大で確定石1.5個分の幅の乱数を入れてある.8×8リバーシはWZebraを使用し,定石のみ変化させて実験を行った.6×6リバーシはそれぞれの深さで1000試合ずつ,8×8リバーシは1500試合ずつ行った.

3. 3種類の方法について

我々が提案する3種類の方法について簡単に説明する.

まず,最初の方法はDiminishing returnを利用する方法である.コンピュータとゲームの間にもDiminishing returnが見られることに注目し,その逓減の様子から理論値の収束を予想する.

次の方法は,勝率の変化を回帰する方法である.自己対戦実験を行うと,プレイヤの強さが変化することによって,勝率や引き分け率が変化する.その変化の度合いに着目し,それを直線に回帰することによって結果を予測する方法である.

3番目の方法は,得点差の変化を回帰する方法である.プレイヤの強さの変化に伴って変化する得点差に注目する.2番目の方法と基本的にやることは同じであるが,さらに細かく予想をすることができる. ただし,得点のないゲームには使えない.

4. Diminishing returnを使った方法

この方法は,Diminishing returnの法則を利用して理論値の収束を予想する方法である.Diminishing returnの法則とは,経済学の用語であり,資源や金銭を投入すればするほど,それに見合う収穫は得がたくなるというものである.この章で述べる方法の考え方は,Diminishing returnの法則をコンピュータプレイヤの強さに適用し,強さが収束しきったところで出てくる結果が理論値に等しいというものである.この場合で言う資源は,コンピュータのスペックや,先読みの深さなどである.この考え方が基本的に正しいということを示しているのが,Heinzによって行われたチェスでの自己対戦実験である.Heinzの実験では,X手読み対X+1手読みで自己対戦実験を行っているが,Xが大きくなるにしたがって,両者のレーティング差は縮まっている.つまり,Xが小さいときに得られていた一手分深く先読みできるプラスの効果が徐々に得られなくなっているということが言える.こうした現象を利用して理論値の推測ができないだろうかというのが基本的なアイディアである.

深さ	W:D:L/Total	ELO
6⇔5	1,686 : 915 : 399 / 3,000	+159
7⇔6	1,643 : 1,066 : 291 / 3,000	+169
8⇔7	1,457 : 1,212 : 331 / 3,000	+137
9⇔8	1,093 : 1,133 : 274 / 2,500	+118
10⇔9	434 : 509 : 107 / 1,050	+112
11⇔10	404 : 539 : 107 / 1,050	+101
12⇔11	375 : 550 : 125 / 1,050	+84

表2:チェスの自己対戦実験の結果

tic-tac-toeの結果を見ると,3手読み対4手読み,4手読み対5手読み,5手読み対6手読みと先読みが深くなるに従い,増減がなくなっていき,引き分け率が50パーセント付近に収束したあと,6手読み対7手読みのときに引き分け率が急激に増えている.これは後手がいきなり読みきってしまったため,後手側に負けがなくなっているからである.引き分け率のみ3~5手読みの際に上昇しており,先手勝率後手勝率はともに減少している.この結果と実際のtic-tac-toeの理論値とを照らし合わせると,このグラフがどのようなカーブを描くかを考えることができる.

深さ	先手勝ち	draw	後手勝ち
0	39.98%	8.15%	51.97%
1	13.84%	20.37%	65.79%
2	31.13%	44.82%	24.05%
3	30.8%	49.19%	20.01%
4	29.68%	50.8%	19.52%
5	0%	88.67%	11.33%
6	0%	100%	0%
7	0%	100%	0%
8	0%	100%	0%
9	0%	100%	0%

表3:tic-tac-toe(X手読み対X+1手読み)の結果

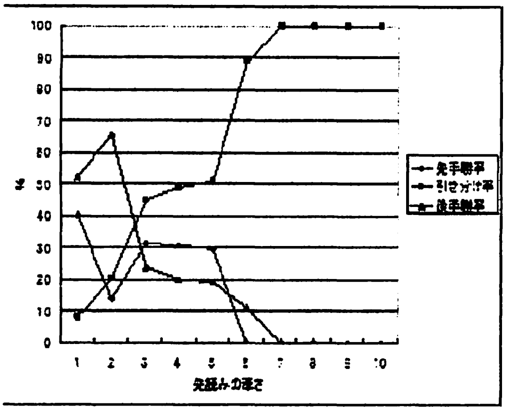

図 1: tic-tac-toe の引き分け率の推移

それぞれの率の変化に着目して考えると,理論値になりうる結果の推移は以下のような特徴を持った曲線になるのではないかという仮説が立てられる:

- 安定した先読みができる深さまではばらばらに散らばる(tic-tac-toe でいう 3 手読み)
- 安定した先読みができる深さ以降に Diminishing return が見られる
- その後増減は限りなく 0 に近づいていく
- Diminishing return が見られるものの,理論値とそれ以外では徐々に差が付く
- 増減が 0 になる寸前で読みきり,そこまでで最も確率の高かったものが理論値になる

この仮説に沿えば,チェスの場合は引き分けと予測することができる.多少のばらつきは見られるものの,引き分け率は徐々にその増分を減らしながら上昇しており,先手勝率,後手勝率はその逆になっているからである.

5. 勝率の変化を使った方法

この方法は,強さを変化させながら行う自己対戦実験で得た勝率データから予測するものである.それぞれの強さごとにある程度自己対戦実験を行い,それぞれの平均を出す.そして,その平均のデータの変化を最小二乗法で直線に回帰してやることで,その先の理論値を予測しようというものである.例えば,8×8 マスのリバーシであれば,60 手読みができれば理論値を出すことができるので,自己対戦実験で得たデータから作った回帰直線を利用し,60 手読みのときの結果を予測することになる.そのようにして作成したグラフが図 2 と図 3,図 4 である.図 2 は 6×6 リバーシのものである.

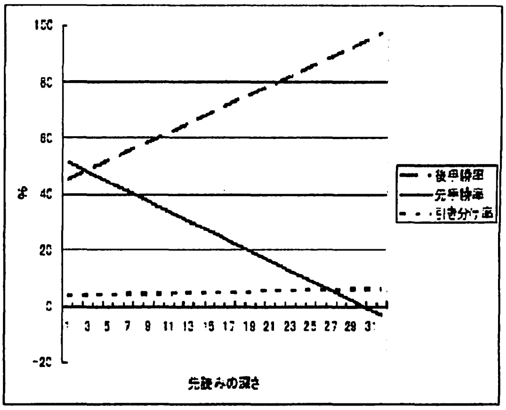

図 2: 6×6 リバーシに適用した結果

先読み	先手勝ち	引き分け	後手勝ち
1	48.3%	3%	48.7%
2	45.7%	3.6%	50.7%
3	45.8%	3.2%	51%
4	50.4%	3.7%	45.9%
5	48.7%	2.9%	48.4%
6	44.9%	4.7%	50.4%
7	55%	3%	42%
8	32.2%	7.1%	60.7%
9	27.2%	12.6%	60.2%
10	42.1%	2.5%	55.4%
11	19.7%	2%	78.3%
12	32.9%	1%	66.1%
13	29.4%	2.4%	68.2%
14	51.2%	0.5%	48.3%
15	12.1%	10.6%	77.3%

表 4: 6×6 リバーシの自己対戦の結果

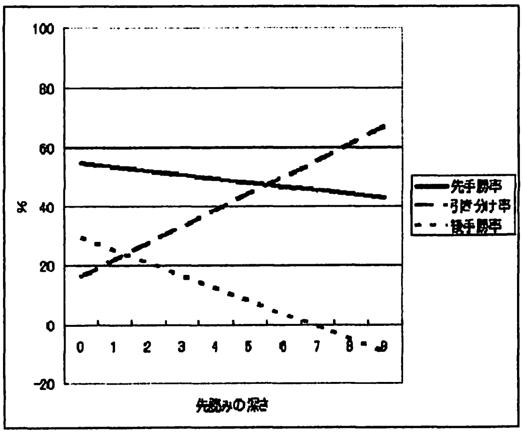

図 3: tic-tac-toe に適用した結果

図2や図3のように結果が出てくれば,簡単に後手が勝つだろうと予測することができる.しかし,図4の8×8リバーシのようになると,二通りの解釈ができるため,簡単に言い切ることができない.
その解釈とは

- 60手目で一番確率の高い白の勝ちである
- 一番上昇度が高い引き分けである

という二つである.このどちらがより正しいかという選択になる.

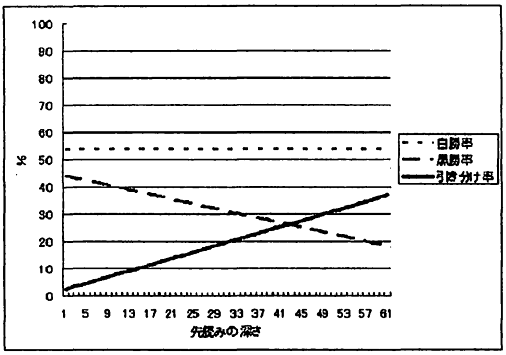

図4: 8×8リバーシに適用した結果

先読み	先手勝率	引き分け率	後手勝率
5	0.3767	0.05	0.5733
6	0.4713	0.0627	0.466
7	0.3807	0.0487	0.5707
8	0.48	0.0693	0.4507
9	0.3207	0.064	0.6153
10	0.4733	0.0873	0.4393
11	0.3687	0.086	0.5453
12	0.4393	0.1	0.4607
13	0.3593	0.0887	0.552
14	0.4233	0.104	0.4727
15	0.3413	0.1227	0.536
16	0.4107	0.1153	0.474
17	0.3287	0.11	0.5613
18	0.386	0.128	0.486
19	0.2973	0.1387	0.564
20	0.3787	0.1473	0.474
21	0.34	0.1507	0.5093
22	0.3733	0.142	0.4847
23	0.3133	0.15	0.5367
24	0.3693	0.1427	0.488
25	0.2947	0.1847	0.5207
26	0.3473	0.17	0.4827

表5: 8×8リバーシの自己対戦の結果

6. 得点差の変化を使った方法

この方法は 5 章で扱った勝率の変化を利用する方法と基本的な考え方は同じである.勝率の変化の代わりに得点差を使用するだけの違いである.しかし,当然ながらこの方法は適用できるゲームが限られる.チェスや将棋のようなゲームには適用できない.ただ,どちらが勝つかということだけでなく,そのときにどの程度の差で勝つかということまで予想できるのがこの方法の利点である.

図6は8×8リバーシの自己対戦結果にこの方法を適用したものである.60手目がちょうど0となっているので,素直に予測すれば引き分けである.

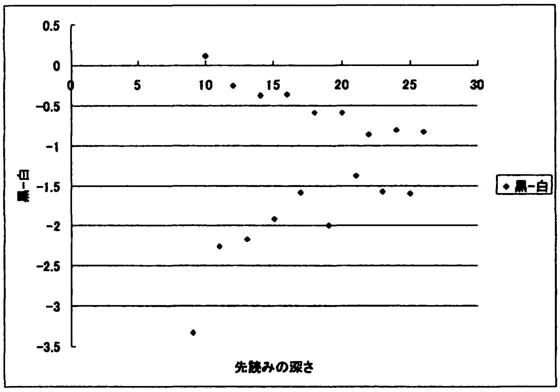

図5 : 8×8リバーシの石の差の平均の分布

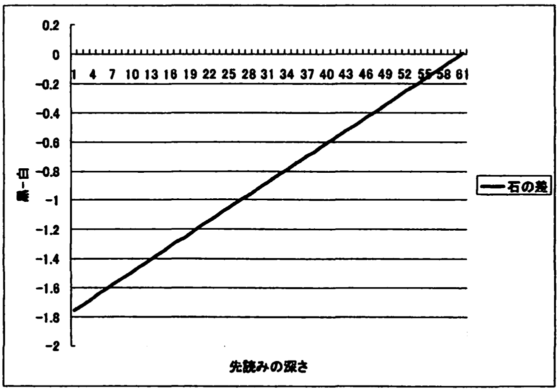

図6 : 図5の結果に適用した結果

7. 考察

ここまでに3つの方法をそれぞれ提案し,例を示したが,それぞれの利点と欠点,及びその方法が妥当かどうかを検討する.

7.1 誤差の発生要因

3つの方法は全てが自己対戦実験のデータを基にした予測である.つまり,自己対戦実験の精度や条件次

第で予測がずれる可能性が大きい.

最も条件の中で大きな影響を与えると思われるのが評価関数の精度である.特に,得点差の変化を利用する方法がその影響を大きく受けると思われる.

ここでその誤差を示すために6×6リバーシで,終盤までランダムで進め,残り手数 X だけを完全読みした結果に勝率の変化,石の差の変化を利用する方法を適用したものと,評価関数を使った自己対戦実験の結果に石の差の変化を回帰する方法を適用したものを示す.

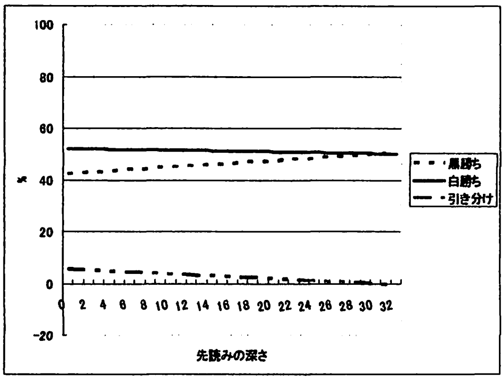

図7 :ランダムプレイ+X 手完全読みの 6×6 リバーシのデータに勝率の変化の方法を適用

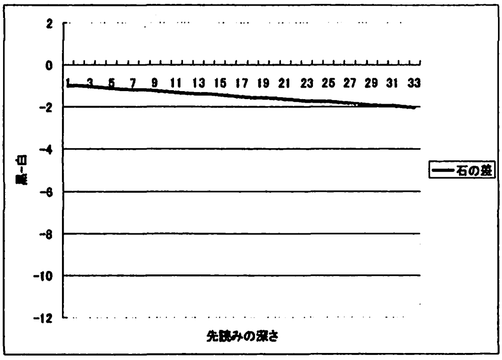

図8 :ランダムプレイ+X 手完全読みの 6×6 リバーシのデータに石の差の変化の方法を適用

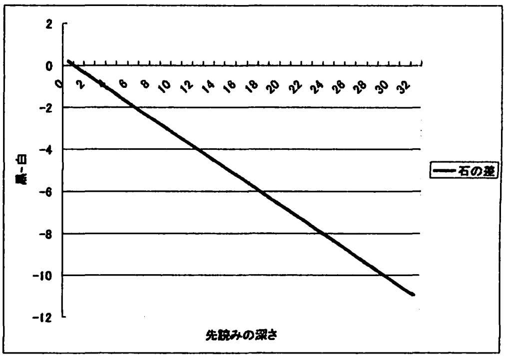

図9 : 6×6リバーシの石の差を回帰したもの

図7と図2を比べるとかなりの差が見て取れる.当然,理論値は白が勝ちであるので図2の方が正確である.つまり,この結果の精度は評価関数の精度にかなり依存することが分かる.一方,図8と図9を比べると,図8のほうが正解に近い(6×6 リバーシは後手 4 目勝ち).このことから,得点差を回帰してやる方法の正当性には疑問符が付く.

7.2 これらの方法は本質的には同じ

結局のところ,これらの方法は自己対戦実験の結果の変化を予測することで理論値を求めようとしているわけで,本質的には全て同じ方法であると考えることができる.Diminishing return を考慮せず,単純な直線回帰を行ったものが勝率の変化を利用する方法であるし,得点差の変化を利用する方法は,勝率の変化を利用する方法に重みをかけているものに過ぎない.

7.3 その他の考えられる方法

出現する棋譜を全て数え上げ,その増減から完全プレイの手筋を探し出すという方法が考えられるが,あまり現実的ではない.6×6 リバーシの場合,1 手読みから15手読みまで実験したが,評価関数を使用した場合も,乱数を使用して実験した場合もいずれの棋譜にも完全プレイの手筋は現れていない.もちろん,評価関数の精度が悪いこともあるが,困難なことであるのは間違いない.

8. まとめ

我々は,自己対戦実験の結果から理論値を推測する方法を 3 種類提案した.Diminishing return を利用する方法,勝率の変化を利用する方法,得点差の変化を利用する方法の 3 つである.これらの方法に対して,それぞれ例を挙げ,妥当かどうかの検討を行った.また,これら 3 つの方法は本質的には同じものではないかという考察を行った.

今後はこれら 3 つの方法を同一のものとみなし,融合させることによって正確に予測できる方法を検討する.

9. 参考文献

[1]Heinz, E. A. A New Self-Play Experiment in Computer Chess. Technical Report MIT-LCS-TM608 (2000)

[2] Kajihara, Y., Sakuta, M., and Iida, H. Semi-Random Play in Game Playing. Proc. of JICAST'99, pp.205.208. Zhejiang University, China (1999)

[3]Feinstein, J. Perfect play in 6x6 Othello from two alternative starting positions, http://www.maths.nott.ac.uk/personal/jff/othello/6x6sol.html ,(1993)

[4]コンピュータリバーシ研究会 http://dais.main.jp/reversi/index.html

[5]Gunnar Andersson, WZebra, http://radagast.se/othello/ ,(2005)