

ブロックスと N 人囚人のジレンマとの関係の研究

岩崎隆一^{†1} Reijer Grimbergen^{†2}

概要

3 人で囲碁を行う多人数囲碁は簡単な協力戦略を採用するプレイアウトや協力をを用いた報酬を組み合わせることにより、対戦の組み合わせ次第では協力戦略を使用しないアルゴリズムの性能を大幅に上回ることが出来た。そこで、多人数ゲームの協力と裏切りの関係を調べるため、囚人のジレンマをモデルとし、ブロックス上で協力 AI と裏切り AI を実装し自動対戦を行った。対戦の結果、協力 AI と裏切り AI が混在している場合、裏切り AI が協力 AI の成績を上回り、多人数ゲームにおける N 人囚人のジレンマの可能性を示すことが出来た。

On the Relation between N-Person Prisoner's Dilemma and Blokus

RYUICHI IWASAKI^{†1} and Reijer Grimbergen^{†2}

Abstract

In three-player Go, it has been shown that a cooperative strategy based on the sharing of rewards outplays other strategies. However, it is unclear how general this result is. To analyze the difference between cooperative and non-cooperative strategies in multi-player games, we propose to use the Prisoner's Dilemma model and apply it to the multi-player game of Blokus. We have implemented a cooperative and a non-cooperative player for Blokus and played these programs against each other. It was found that when both cooperative and non-cooperative players were played against each other, the non-cooperative strategy was better than the cooperative strategy. This shows that the Prisoner's Dilemma can be used as a model for multi-player games.

1. はじめに

2 人ゲームにおける人工知能の研究は進んでおり、特にチェスやチェッカーなどのゲームは既に人間が勝つことの出来ないレベルに到達している。しかし、多人数ゲームの領域においても、2 人ゲームと同様の成功を収めるには多人数ゲーム特有の問題を解決する必要がある¹⁾。

どのような方法を用いて探索を行うかは多人数ゲームの領域の中で大きな問題の一つである。多人数ゲームの場合、どの手を選んでもある相手プレイヤーの評価値は同じだが、そのプレイヤー以外の評価値はそれぞれ異なるという状況が考えられる。この場合、どの手を選択するかは相手プレイヤー次第となる。この時、探索を正確に行うには相手モデルが必要となるが、完璧な相手モデルを定義する事は困難である。

しかし、コンピュータ囲碁における主流のアルゴリズムとなりつつあるモンテカルロ木探索は、多人数ゲームにおいても有効であることが判明し²⁾、多人数で行う囲碁の場合、協力戦略が有効であることが分かった³⁾。

また、N 人囚人のジレンマは囚人のジレンマを多人数に

拡張したものである。この場合、裏切る人数が増えるごとに一人が得ることが出来る利得は減少し、最終的に全員が裏切った場合、全員が協力を採った時の利得よりも少なくなる⁴⁾。

本稿では、多人数ゲームと N 人囚人のジレンマの関係性に着目し、ブロックスと呼ばれる 4 人で行う多人数ゲーム上で協力的な行動をとる協力 AI と自己中心的な行動をとる裏切り AI をそれぞれ実装した。また、実装した AI 同士で、自動対戦を行い、N 人囚人のジレンマが多人数ゲームにおいても起こりうるか、検証を行った。

2. 関連研究

2.1 多人数ゲームにおける探索

多人数ゲームにおける評価値を用いた探索手法として \max^n アルゴリズムと paranoid アルゴリズムが存在する¹⁾。 \max^n アルゴリズムは各手番のプレイヤーが自身の評価値が一番高い手を選択するという前提で探索を行うのに対し、paranoid アルゴリズムはルートノードのプレイヤー以外のプレイヤーがルートノードのプレイヤーの評価値が一番低い手を選択するという前提で探索を行う。上記の 2 つの探索手法を用いた AI をダイヤモンドゲームで対戦させた結果、paranoid アルゴリズムが \max^n アルゴリズムの性能を上回る結果となった。

また、囲碁の分野で実装されているモンテカルロ木探索を多人数ゲームに応用した研究では探索に UCT の実装を

^{†1} 東京工科大学大学院 バイオ・情報メディア研究科
コンピュータサイエンス専攻

Graduate School of Bionics, Computer and Media Sciences, Tokyo University of Technology

^{†2} 東京工科大学 コンピュータサイエンス学部

School of Computer Science, Tokyo University of Technology

行った²⁾。研究で用いられたゲームで UCT は \max^n アルゴリズムと paranoid アルゴリズムの性能と同等もしくは上回る性能を発揮した。

2.2 多人数囲碁における UCT

囲碁におけるモンテカルロ木探索が有効であることを受け、囲碁を 3 人用に拡張した多人数囲碁に UCT を適用した研究が行われた³⁾。この研究では他プレイヤーと協力する手を選択するプレイアウトを実装した。また、協力しているプレイヤーを考慮した報酬を実装し、二つを組み合わせた AI は前述の二つを使用していない AI と対戦を行った。その結果、提案された AI は対戦の組合せ次第ではこれまでの AI を大幅に上回る性能を発揮した。

3. ブロックスについて

ブロックスは GPCC において 2007 年、2008 年、2009 年の課題となったブロックスデュオ⁵⁾を多人数で出来るようにしたゲームである。基本的なルールはブロックスデュオと同様であるが、4 人でやることを考慮し、以下の点で異なっている。

- ボードの大きさが 20x20 になっている
- ブロックを 4 色使用する
- 1 手目にピースで覆わなければならないマスが各 4 隅になっている

4. 囚人のジレンマ

囚人のジレンマはフロッドとドレジャーが考案し、タッカーによって定式化されたゲームである⁶⁾。行動とそれによって得られる利得の関係の例を表 1 に示す。

表 1 行動と利得の関係

| | | Player 2 | |
|----------|-----|----------|--------|
| | | 協力 | 裏切り |
| Player 1 | 協力 | 5, 5 | -4, 6 |
| | 裏切り | 6, -4 | -3, -3 |

この時、2 人のプレイヤーは協力を採れば、お互いに高い利得を得ることが出来るにも関わらず、相手が裏切りを採り、自分の利得が最小になることを恐れるあまり、お互いに裏切りを採ってしまうというのがこのゲームの特徴である。この時、2 人のプレイヤーを囚人とし、協力行動を黙秘、非協力行動を自白に例えたことから囚人のジレンマと呼ばれている。

また、3 人囚人のジレンマの場合、各プレイヤーの行動と行動によって 1 人のプレイヤーが得ることの出来る利得の関係は図 1 のように示される。

この時、1 人の利得は 1 人が非協力を採り、他の 2 人が協力を採った時に最も高くなる。逆に、1 人が協力を採り、2 人が非協力を採った時に最も低くなる。また、非協力を

採る人数が増えるごとに利得は減少し、全員が非協力を採ると、全員が協力を採った時よりも低い利得しか得られない結果となる⁴⁾。

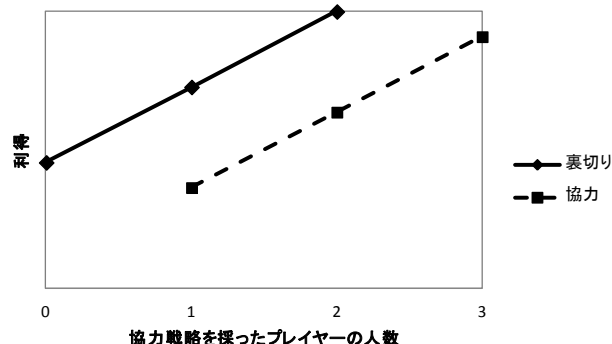


図 1 戦略とプレイヤーが得られる利得の関係⁴⁾

5. 提案手法

本研究では 4 節で述べた N 人囚人のジレンマがブロックスで起こりうるかを検証するため、協力をを行う AI と裏切りを行う AI を実装した。両方の AI は協力 AI と裏切り AI の特徴づけとして、モンテカルロ探索の報酬に対局終了時の残りブロック数を使用した。また、プレイアウトは全てランダムプレイアウトに設定した。

協力 AI は全てのプレイヤーの残りブロック数が少なくなる手を選択する。協力 AI はプレイアウト終了時、各プレイヤーの残りマス数の平均を計算する。そして 4 つの値を平均したものをプレイアウトの報酬とする。

裏切り AI は自分の残りブロック数を少なくする手を選択する。裏切り AI はプレイアウト終了時、自分の残りマス数の平均を計算し、プレイアウトの報酬とする。

6. 実験

6.1 実験概要

ブロックスと N 人囚人のジレンマの関係を調べるために、AI 同士の自己対戦を行った。協力 AI と裏切り AI の各人数、着手する順番の組み合わせで、16 通りの対戦を行い、対戦回数は各 10 回ずつに設定した。この時、4 つの AI が 1 位から 4 位の順位になった回数と終局時に残っていたブロックのマス数の平均を測定した。また、これらの結果が図 1 のグラフとどの程度一致するかによって、関係性を求めた。実験環境は以下の通りである。

- OS : Windows7
- CPU : Core i3-2130 3.4Ghz
- メモリ : 4GB

6.2 対戦結果

協力 AI の中に裏切り AI を 1 人ずつ追加して自動対戦を行った。一度、裏切り AI に設定したプレイヤーは変更せず、裏切り AI への変更はランダムに行った。全員が協力 AI の結果を表 2 に、裏切り AI が 1 人の結果を表 3 に、2

人の結果を表 4 に、3 人の結果を表 5 に、4 人全員が裏切り AI の結果を表 6 にそれぞれ示す。また、表 3 の裏切り AI は青、表 4 の裏切り AI は青と緑、表 5 の裏切り AI は青、赤、緑に設定されている。この時、裏切り AI がゲームの中に存在するとき、裏切り AI の成績は協力 AI の成績を上回っており、ゲームに参加する裏切り AI の人数が増えるとともに、裏切り AI の成績が徐々に低下した。

表 2 全員協力 AI の場合

| プレイヤー | 1 位 | 2 位 | 3 位 | 4 位 | 平均マス数 |
|-------|-----|-----|-----|-----|-------|
| 青 | 4 | 3 | 3 | 0 | 13.8 |
| 黄 | 2 | 5 | 2 | 1 | 13.3 |
| 赤 | 1 | 2 | 4 | 3 | 16.4 |
| 緑 | 3 | 0 | 1 | 6 | 16.1 |

表 3 裏切り AI が 1 人の場合

| プレイヤー | 1 位 | 2 位 | 3 位 | 4 位 | 平均マス数 |
|---------|-----|-----|-----|-----|-------|
| 青 (裏切り) | 9 | 1 | 0 | 0 | 4.9 |
| 黄 | 0 | 3 | 4 | 3 | 19.2 |
| 赤 | 0 | 2 | 4 | 4 | 19.7 |
| 緑 | 1 | 4 | 2 | 3 | 16.9 |

表 4 裏切り AI が 2 人の場合

| プレイヤー | 1 位 | 2 位 | 3 位 | 4 位 | 平均マス数 |
|---------|-----|-----|-----|-----|-------|
| 青 (裏切り) | 3 | 6 | 0 | 1 | 9.0 |
| 黄 | 0 | 1 | 6 | 3 | 22.3 |
| 赤 | 0 | 1 | 3 | 6 | 25.0 |
| 緑 (裏切り) | 7 | 2 | 1 | 0 | 6.3 |

表 5 裏切り AI が 3 人の場合

| プレイヤー | 1 位 | 2 位 | 3 位 | 4 位 | 平均マス数 |
|---------|-----|-----|-----|-----|-------|
| 青 (裏切り) | 4 | 1 | 3 | 2 | 14.4 |
| 黄 | 0 | 0 | 2 | 8 | 26.3 |
| 赤 (裏切り) | 4 | 4 | 2 | 0 | 8.9 |
| 緑 (裏切り) | 2 | 5 | 3 | 0 | 8.8 |

表 6 全員裏切り AI の場合

| プレイヤー | 1 位 | 2 位 | 3 位 | 4 位 | 平均マス数 |
|---------|-----|-----|-----|-----|-------|
| 青 (裏切り) | 2 | 3 | 2 | 3 | 17.3 |
| 黄 (裏切り) | 4 | 3 | 2 | 1 | 11.2 |
| 赤 (裏切り) | 3 | 1 | 4 | 2 | 15.5 |
| 緑 (裏切り) | 1 | 3 | 2 | 4 | 16.6 |

6.3 N 人囚人のジレンマとの比較

ブロックスが N 人囚人のジレンマであるか確かめるため、ゲームに参加している各 AI の人数とゲーム終了時に残った平均マス数の推移を図 2 に示す。この時、同じ AI が複

数人ゲームに参加している場合は同 AI の平均とした。また、平均マス数は少ないほど良い成績であるため、図 2 では Y 軸の向きを反転させてグラフを描画している。

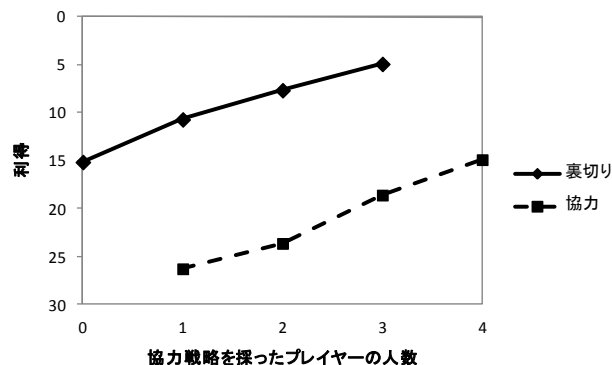


図 2 各 AI の人数と成績の推移

7. おわりに

本稿では多人数ゲームが N 人囚人のジレンマと同様の状態になるか、ブロックスを用い、協力 AI と裏切り AI の人数を増減させて自動対戦を行った。自動対戦の結果、協力 AI と裏切り AI が共にゲームに参加している場合、裏切り AI が協力 AI の成績を上回った。また、ゲームに参加する裏切り AI の人数が増加するごとに裏切り AI の成績が低下した。だが、協力 AI の成績は総じて低く、4 人全員を協力 AI にしても、4 人全員が裏切り AI の時の成績と同程度という結果となってしまった。しかし、裏切り AI が 1 人の時、裏切り AI が一番良い成績となり、グラフの傾きも図 1 と類似性が見られることから、多人数ゲームにおける N 人囚人のジレンマの可能性を示すことが出来た。

今後の課題としては、協力 AI の改良を行って、ブロックスにおける協力戦略の優位性を見出したい。また、協力 AI は全てのプレイヤーと協力していたが、今後、協力 AI は 1 人もしくは 2 人と協力を行う、また、その中に裏切り AI を入れるなどして、より柔軟なチーム組んで実験を行いたい。

参考文献

- 1) Sturtevant, N: "A Comparison of Algorithms for Multi-player Games," pp. 108-122, Springer-Verlag Berlin Heidelberg, 2003.
- 2) Sturtevant, N: "An Analysis of UCT in Multi-Player Games," ICGA Journal, vol. 31, no. 4, pp. 195-208, 2008.
- 3) Cazenave, T: "Multi-player Go," Proceedings of the 6th International Conference on Computers and Games(eds. H.J. van den Herik, X.Xu, Z.Ma, and M.H.M Winands), vol. 1531 of Lecture Notes in Computer Science(LNCS), pp. 50-59, Springer-Verlag, Heidelberg, Germany, 2008.
- 4) 西山賢一, "三人の囚人のジレンマ," 勝つためのゲームの理論 適応戦略とは何か, 木村次郎 (編), pp.70-74, 講談社, 1986.
- 5) 藤波順久, 酒井香代子, "GPCC 報告(2007 年)," プログラミング・シンポジウム報告集.
- 6) 岡田 章, "囚人のジレンマ," ゲーム理論・入門 人間社会の理解のために, pp.88-89, 有斐閣, 2008.