Vol.2015-GI-33 No.7
2015/3/5

UPP による駒価値評価関数に基づいた NEGOgeister AI

三塩武徳[†1]　　藤田桂英[†1]

本論文では二人不完全情報ゲームである「ガイスター（geister）」に独自の交渉ルールを追加した「NEGOgeister」を対象とする．本ゲームでの交渉を有利に進めるためには相手と自分の駒の価値を相対的に正しく評価し自分の得になる選択をする必要がある．一方，教師データとなるガイスターの棋譜は極めて少ないため駒の価値を評価する関数を作成することが難しい．そこで本手法では UPP 的な概念を用いて，シミュレーション結果の差異によって相手の駒の価値を推定する．これは正体のわからない相手の駒がゲーム中どれだけの価値があるのかを評価する特徴の一つとなる．さらに，評価実験によりこの特徴が相手の駒の価値を正確に反映しているかを確認した．

NEGOgeister AI based on evaluation function of pieces using UPP

MISHIO TAKENORI[†1]　FUJITA KATSUHIDE[†1]

In this paper, we study the “NEGOgeister” including our own negotiation rules to the “geister” which is a bilateral incomplete information game. A player needs to evaluate the values of our and opponent’s pieces relatively, and select the effective strategies. However, it is difficult to decide functions of evaluating the values of the pieces because the records of “geister” which is used for learning data are little. Therefore, our proposed method estimates the values of the opponent’s pieces with differences of simulation results using the UPP concept. It is the one of important features to evaluate the values of unknown opponent’s pieces. We evaluate and discuss them with the objective of evaluating the values of opponent’s pieces, accurately.

1. はじめに

人工知能に関する研究において，強いゲームプログラムを作ることは重要な課題である．特に，「二人不完全情報ゲーム」は将棋やチェスのような完全情報ゲームとは異なり，相手の駒の情報が隠されているために正しく判断することが難しいことから，近年活発に研究が行われている．ポーカーに CFR を用いた Martin Zinkevich ら 1)、NA Risk ら 2)、MCCFR を用いた Neil Burch ら 3)の研究が一例である。実際に，局面の評価を行う評価関数の精度と局面の探索の効率が強さにかかわる要素であることは完全情報ゲームと同じであるが，相手の選択や戦略から適切に隠された駒を判定することが重要になっている．

本論文で対象にしている「NEGOgeister」は，「ガイスター（geister）」というゲームを基にしている．これは「ゴースト」「ファンタスミ」とも呼ばれる 1980 年にドイツのアレックス・ランドルフ（Alex Randolph）によって発表された二人零和確定不完全情報ゲームである 4)．ガイスターは 6×6 マスの盤面とプレイヤーごとに 8 個の駒を用いて行うゲームである．駒は良い駒と悪い駒の 2 種類，4 個ずつある．1993,1994 年の GPCC において南雲夏彦らが 5)6)，また 2014 年の第 31 回ゲーム情報学研究会においては筆者らが先行研究を発表している 7)．しかし，人間同士のガイスターでは交渉を通して互いにコミュニケーションを取り合い，駒の情報を取得するケースが多数存在する．しかし，既存研究ではこのような交渉を考慮したゲーム AI に関する研究はほとんどない．

本論文では，通常のガイスターに交渉ステップというプレイヤー同士がお互いにコミュニケーションを取り合うステップを導入した新しいゲーム「NEGOgeister」を提案する．今回提案する NEGOgeister は交渉ステップを単純化することでゲームプログラムを作成することが可能となっている．

また，教師データとなるガイスターの棋譜は極めて少ないため NEGOgeister での駒の価値を評価する関数を作成することが難しい．そこで本論文では UPP 的な概念を用いて，シミュレーション結果の差異によって相手の駒の価値を推定する手法を提案する．これは正体のわからない相手の駒がゲーム中どれだけの価値があるのかを評価する特徴の一つとなる．さらに，実験によりこの特徴が相手の駒の価値を正確に反映しているかに関して示す．

本論文の構成は以下の通りである．まず，通常のガイスターの基本ルールと新しく提案する NEGOgeister のルールを説明する．次に，UPP などの不完全情報ゲームで取り入れられているシミュレーションに基づく手法について説明する．その後，今回提案する NEGOgeister のゲーム AI に採用している相手の駒の評価値決定手法を提案する．最後に，評価実験により，提案した手法が正しく相手の駒の価値を評価できていることを示す．

†1 東京農工大学
Tokyo University of Agriculture and Technology

2. NEGOgeister

2.1 「ガイスター（geister）」の基本ルール

- このゲームは 6×6 マスの盤面とプレイヤーごとに 8 個の駒を用いて行う．
- 駒は良い駒と悪い駒の二種類，4 個ずつある．
- ゲーム開始時に自分の 8 つの駒を 2×4 のマスで囲まれた陣地内に好きな並びで配置する．図 1 は初期配置の例を示している．
- ゲーム中に自分の駒の正体は確認できるが相手の駒の正体は確認できない．
- 相手の駒を取った時に相手の駒の正体を確認できる．
- 取った駒の種類から盤面に残っている相手の 2 種類の駒の総数は知ることができる．
- 盤面の 4 隅のうち自分の陣地から遠い 2 つは自分の出口となる．

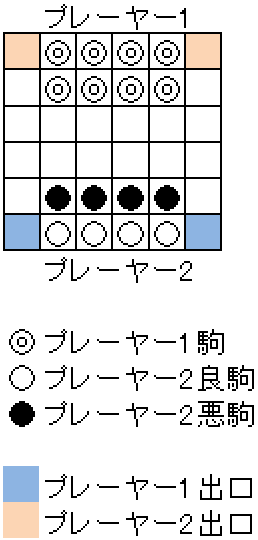

図 1　「ガイスター」初期配置例

2.2 手番の流れ

- 自分の手番では自分の駒の一つを上下左右の 4 方向いずれかに動かすことができる．
- 自分の駒の移動先に相手の駒があるときにはその駒を取る．
- 自分の駒を取るように移動することはできない．
- 自分の良い駒が自分の出口のマスの上にいる場合は移動の代わりに「脱出」を選択することができる．
- 移動か脱出を行うと自分の番は終わりとなる．

2.3 勝利条件

ゲーム中に以下のいずれかを満たした場合，そのプレーヤーの勝利となる．

- 「脱出」を行う（一つの駒のみでよい）．
- 相手の良い駒を全て取る．
- 自分の悪い駒を相手に全て取らせる．

2.4 NEGOgeister における交渉要素

NEGOgeister では各プレイヤーの着手の前に交渉ステップを行う．交渉ステップでは次のように行動する．

- 手番のプレイヤーが相手の駒を一つ選ぶ．
- 相手プレイヤーは手番のプレイヤーの公開する駒を一つ選択するか，交渉拒否を選択する．
- 手番のプレイヤーは提示された駒に対して同意か交渉決裂を選択する．
- 同意した場合はお互いの駒の正体を相手に公開する．

3. 不完全情報ゲームとシミュレーション

3.1 不完全情報ゲームに関する先行研究

思考ゲーム研究においてシミュレーションを利用することがある．例えば，Gelly らの作成したモンテカルロ法を囲碁に適用したモンテカルロ囲碁のプログラム Mogo などがある 8)．また，シミュレーションの中にある方策を取り入れることで性能を向上させようとする研究も存在し，Simulation Balancing はその一つである 9)．

不完全情報ゲームにおける問題を緩和しようとした先行研究が存在する．Whitehouse らは「Dou Di Zhu」というゲームで仮定を行う手法について述べており 10)，Zhang らは多人数ミリタリーチェスにおいて UCT アルゴリズムを適用する方法を考案した 11)．

また，筆者らは UPP という過去のプレイアウト結果と相手の着手から相手の駒の正体を推定し，現在の着手決定のためのプレイアウトの割り振りを偏らせるという手法を提案している 7)．「ガイスター」においてアルゴリズム UPP を取り入れたプレイヤーと各世界に平等にプレイアウトを割り振るプレイヤーを対局させたこの実験では，55%の勝率を上げることができ，このゲームにおいてアルゴリズム UPP が有効であることを明らかにした．

3.2 不完全情報ゲームとシミュレーション

シミュレーションを利用した考え方は不完全情報ゲームにも取り入れることができるが，シミュレーションを行う前に不完全情報を仮定する必要がある．不完全情報の仮定が手法にどのように影響を及ぼすのか説明する．

図 2 と図 3 は，完全情報ゲームと不完全情報ゲームにおいてモンテカルロ法を用いる場合の簡単なゲーム木の図である．

完全情報ゲームでは不完全情報の仮定を行う必要がないために，図 2 のように単純なモンテカルロ法を用いて探索することができる．一方，不完全情報ゲームでは図 3 のように世界の仮定（不完全情報の仮定）を想定し，より複雑な状況をモンテカルロ法などで探索する必要がある．

本論文では不完全情報を仮定した局面のことを世界と呼

ぶ．世界はその仮定のされ方によって局面ごとに複数存在する．世界の仮定を行うとプレイアウトを（1／世界の数）しか割り振ることができなくなるために報酬に信頼がおけなくなる．

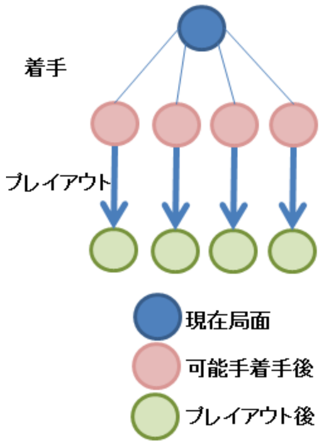

図 2　完全情報ゲームにおけるモンテカルロ法

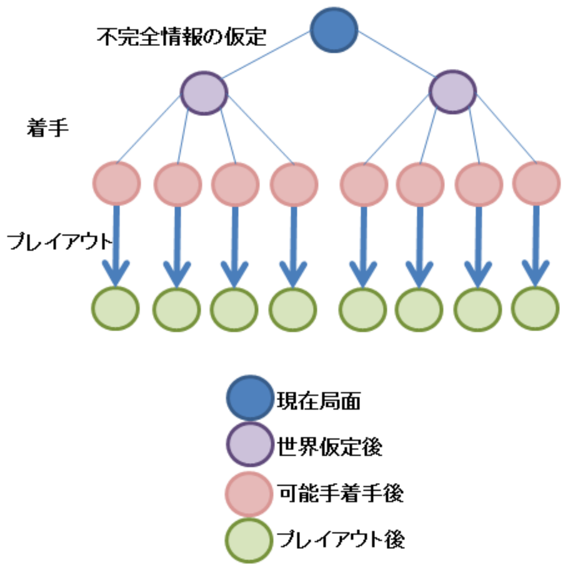

図 3　不完全情報ゲームにおけるモンテカルロ法

世界ごとにプレイアウトの結果は異なる場合がほとんどであり，ある世界において勝率の高い手が別の世界では悪手とされることもある．各世界における同じ着手の異なる報酬をどのようにまとめればよいかという課題がある．

4. NEGOgeister に対応したゲーム AI

「NEGOgeister」で交渉を行うにあたり，適切な戦略の決定が必要となる．戦略の決定には駒価値の評価関数を必要とし，評価関数の作成にはいくつかの特徴を定義する必要がある．

4.1 特徴

特徴は駒の価値を評価するために必要な要素である．以下に本論文で使用する特徴である，出口からの距離と他の駒との相対距離の特徴に関して説明する．

1	2	3	3	2	1
2	3	4	4	3	2
3	4	5	5	4	3
4	5	6	6	5	4
5	6	7	7	6	5
6	7	8	8	7	6

図 4　出口からの距離による特徴

図 4 で示されるのは手前プレイヤーの駒の価値を評価するための特徴である．出口からの距離が近いほど価値が高く設計されている．

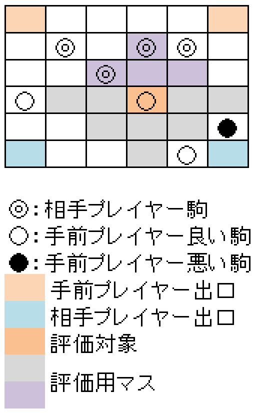

図 5　他の駒との相対距離による特徴

図 5 では上から 4 マス，左から 4 マスの位置にある手前プレイヤーの良い駒の価値を評価している．評価する駒からちょうど 2 マスの距離にある駒の種類によって評価値が変化する．他の駒との相対距離による特徴は上下左右の対称な 4 マス（図 5 の濃い紫の部分）の組み合わせで考えることができる．

4.2 駒価値の評価関数

3.1 で挙げた特徴を利用すると，駒の価値を評価する関数を作成することができる．

$$\mathrm{value}(\mathrm{f1}, \mathrm{f2}, \dots) = \sum_{i=1} w_i \, \mathrm{f}_i \quad -(1)$$

(*f1,f2,...*：特徴　*w1,w2,...*：重み)

数式(1)は，いくつかの特徴に対する重みづけ線形和を示している．より高い勝率を上げることのできる重みを定義することで，適切な相手の駒の価値を決定することができる．

4.3 UPP による特徴

UPP による特徴は，まず，現在の局面から多数シミュレーションを行い，その勝敗を保存する．次に，評価したい相手の駒を良い駒と仮定した世界でのシミュレーション結果と悪い駒と仮定した世界でのシミュレーション結果に判別する．その後，それぞれのシミュレーション結果の勝率を算出し，良い駒と仮定した世界でのシミュレーション結果と悪い駒と仮定した世界でのシミュレーション結果の差を特徴量とする．以下に UPP による特徴の式を示す．（数式(2)）

$$f_{op} = |winp_{good} - winp_{bad}| \quad \text{-}(2)$$

$winp_{good}$：相手の駒を良い駒と仮定した世界でのシミュレーション結果

$winp_{bad}$：相手の駒を悪い駒と仮定した世界でのシミュレーション結果

一般的に，本特徴量の大きい駒はその正体が判明することでゲームの結果が大きく異なる重要な駒である．一方，本特徴量が小さい駒は正体が判明することでゲームの結果にあまり影響を与えない駒である．したがって，交渉時に本特徴量の大きい駒を選択することでより駒の正体が判明することで勝率を大きく上げることができる駒を判別することが可能となる．

4.4 戦略の決定

今後，本論文で提案した特徴に基づく評価関数による交渉の戦略を提案する必要がある．戦略には駒の選択基準と駒の公開基準の二つが必要となる．

駒の選択基準には最も価値の高い駒を選択するだけでなく，以下のような戦略が考えられる．

- 自らが手番プレイヤーの場合には自分の持つ駒の価値よりも著しく価値の高い相手の駒は選ばない．
- 手番プレイヤーでない場合には相手の選択した駒の価値よりも著しく価値の高い相手の駒は選ばない．これは，交渉成立の可能性を上げるために，初めから価値の釣り合わない提案をしない戦略である．

また，駒の公開基準も以下のような戦略が考えられる．

- 自分の駒の価値が高い場合には公開しない．
- 自分の駒の価値が高くてもその差が閾値以内ならば公開する．
- 相手の駒の価値が高くてもその差が閾値以内ならば公開しない．

これらの戦略の優劣や評価関数の組み合わせに対する評価は今後の重要な課題である．

5. 実験

新しく提案した UPP による特徴(4.3)が正しく相手の駒の価値を評価できているかどうか評価を行う．

5.1 実験概要

テストデータとして 200 回の局面を用意しそれぞれの局面で最も価値の高い相手の駒を 1 つずつ決定した．

局面は「NEGOgeister」のゲーム中で起こりうるすべての局面の中からランダムに作成したが，常に手番プレイヤーとして相手の駒の価値を評価する．

正解となる最も価値の高い相手の駒は，事前に人力で判断した．この時，人力では価値の高い駒を決めかねるような難しい局面は排除した．

正解となる駒を決めることのできた 200 局面に対し UPP による特徴量を計算し，最も特徴量の高かった駒が正解と一致するかどうかを確認した．

駒を評価するためのプレイアウトは，一つの局面に対し 11,000 回ずつ行い，プレイアウト内での着手決定は完全ランダムとした．また，局面に残っている駒の総合計から大きく局面を前半と後半の 2 種類に分け，それぞれの場合において正解率が変化しているかどうかも確認した．総合計が 13 個以上の場合は前半，それ未満の場合は後半とした．

5.2 実験結果

表 1　正解率の比較

	問題数	正解数	正解率
全体	200	79	0.395
前半	113	40	0.353982
後半	87	39	0.448276

正解率の比較の結果を表 1 に示す．問題全体では 4 割程度の正解率で，駒数の少ない後半の盤面のほうが高い正解率を出していることが分かる．

以下に，それぞれ図に局面，表に相手の駒の特徴量一覧を載せる．

正解できた局面

正解できた局面の例を示す．

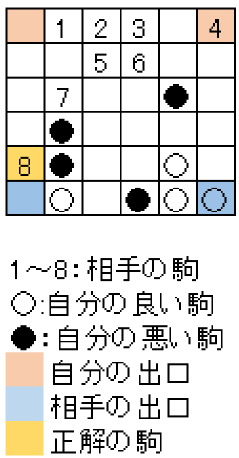

図 6　問題例 1

表 2　問題例 1 の特徴量

	駒の特徴量
1	0.10
2	0.05
3	0.50
4	0.81
5	1.61
6	2.12
7	2.68
8	7.78

図 6 と表 2 の例では正解とした駒 8 の特徴量が最も大きくなっているだけでなく，次に取ることのできる駒 6，7 の特徴量も大きくなっている．次に取ることのできる駒の正体を知ることができれば着手を決定する助けになるため，正しく評価できている．

また，自分の駒から遠い駒 1 から 4 の中では相手の出口をふさいでいる駒 4 の特徴量が最も高くなっている．脱出の最後の障害になりうる駒の正体がゲームに影響を与えうることを反映している．

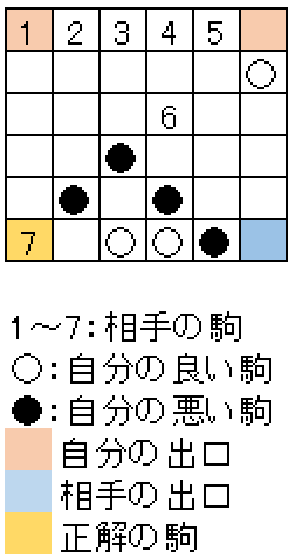

図 7　問題例 2

表 3　問題例 2 の特徴量

	駒の特徴量
1	4.02
2	3.87
3	4.15
4	4.59
5	4.06
6	5.16
7	32.86

図 7 と表 3 は非常に極端な例を示している．現在のこちらの手番ではどのように動いたとしても相手の駒 7 が良い駒であった場合には負けてしまう．このような例では，駒 7 が良い駒であるか悪い駒であるかはゲームに多大な影響を及ぼす．特徴量も駒 7 はほかの駒に比べて非常に大きいため，正解している．

ただし，実際のゲームにおいては駒 7 を悪い駒であると断定しなければどのような着手も意味がなくなってしまうため，駒 7 の正体を知ることが必ずしもこちらを有利にするとは言えない．今回の例では駒 6，もしくは駒 5 を選択するというのも戦略として考えられる．

正解できなかった局面

以下に，正解できなかった局面について述べる．

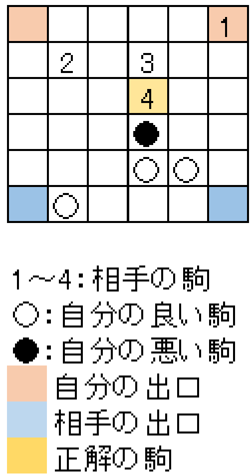

図 8　問題例 3

表 4　問題例 3 の特徴量

	駒の特徴量
1	6.96
2	5.95
3	4.66
4	0.09

図 8 と表 4 では次の一手で取ることのできる駒 4 の正体を知ることでゲームを有利に進めることが出来る．しかし駒 4 の特徴量は最も低く，駒 1 が最も高い特徴量を持っている．自分の出口をふさいでいる駒とはいえ，この駒の正体を知ることが駒 4 の正体を知ること以上に有利に働くことは考えづらい．

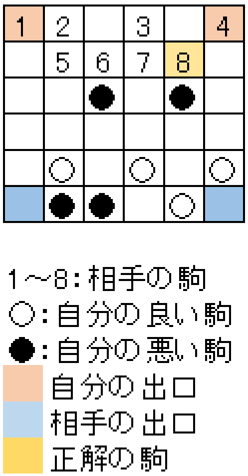

図 9　問題例 4

表 5　問題例 4 の特徴量

	駒の特徴量
1	1.34
2	1.02
3	0.92
4	0.64
5	0.14
6	1.36
7	1.69
8	0.73

図 9 と表 5 の局面では, 次に取ることのできる駒である駒 6 と駒 8 のうち, より双方の出口に近い駒 8 を正解とした. 駒 8 の正体によって真下の自分の駒で取ってしまうかもしくはかわすように出口に進むかを決定できると考えたためである．しかし，駒の特徴量は 7 が最も大きくなった．駒 6 や駒 8 に比べ駒 7 の正体を知ることがゲームを有利に運ぶとは考えにくい.

出口をふさいでいる駒 1 の特徴量が高めである．問題例 3 の結果と比べてみると，どうやら出口をふさぐ駒の特徴量が高くなる傾向にあると予想できる．これは，自分の駒が脱出を行う際に取らなくてはならない駒であるために，ほかの駒に比べ特徴量が高く出ているのではないかと考えた．この真偽の確認は今後の一つの課題である.

6. おわりに

本論文では，シミュレーション結果の世界ごとでの差異を利用して相手の駒の価値を評価する方法について提案した．この手法では人間が最も価値があると評価する駒を約 4 割で発見することが出来ることを示した．実験から，局面に駒が 13 個以上ある場合とそれ未満である場合では正解率に差異があることが分かった. 以上から UPP による特徴は局面に数が少ないほうが駒の価値に対して正確に現れることが分かった.

今回提案した特徴を用いた交渉フェーズや駒の動かし方に関する適切な戦略の提案が重要な課題である．提案した戦略の優劣や評価関数の組み合わせに関する評価もまた重要な課題である.

参考文献

1) Martin Zinkevich, Michael Bowling, Michael Johanson, Carmelo Piccione: Regret Minimization in Games with Incomplete Information, In Advances in Neural Information Processing Systems 20 (NIPS), pp. 905–912(2008)
2) Nick Abou Risk, Duane Szafron: Using counterfactual regret minimization to create competitive multiplayer poker agents, Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems(AAAMAS),pp.159-166(2010)
3) Neil Burch, Marc Lanctot, Duane Szafron, Richard G. Gibson: Efficient Monte Carlo Counterfactual Regret Minimization in Games with Many Player Actions, Advances in Neural Information Processing Systems 25(NIPS),(2012)
4) ガイスター
http://boardgamegeek.com/boardgame/2290/ghosts
5) 南雲夏彦：GPCC 報告「ゴースト」と「ループトラックス」と「ドット&ボックス」,34 回プログラミング・シンポジウム報告集,pp195-199(1993).
6) 南雲夏彦：GPCC 報告「ゴースト」と「ポジット」,35 回プログラミング・シンポジウム報告集,pp173-176(1994)
7) 三塩武徳，小谷善行：ゲームの不完全情報推定アルゴリズム UPP とそのガイスターへの応用,第 31 回 ゲーム情報学研究会（2014）
8) Sylvain Gelly, Levente Kocsis, Marc Schoenauer, Michèle Sebag, David Silver, Csaba Szepesvári, Olivier Teytaud: The grand challenge of computer Go: Monte Carlo tree search and extensions, Communications of the ACM volume55 Issue3 March 2012,pp106-113(2012)
9) David Silver, Gerald Tesauro: Monte-Carlo simulation balancing, ICML '09 Proceedings of the 26th Annual International Conference on Machine Learning, pp945-952(2009)
10) Daniel Whitehouse, Edward J.Powley, Peter I.Cowling: Determinization and information set Monte Carlo Tree Search for the card game Dou Di Zhu, 2011 IEEE Conference on Computational Intelligence and Games, pp87-94(2011)
11) Jiajia Zhang, Xuan Wang, Jing Lin, Zhaoyang Xu:UCT Algorithm in Imperfect Information. Multi-Player Military Chess Game, 11th Joint Conference on Information Sciences(2008)2