

# 不利な情報を与えないように行動する algo プレイヤの構築

佐々木 一郎<sup>1</sup> 藤田 茂<sup>2</sup>

**概要:** 不完全情報ゲームでは、プレイヤは共有されていない他プレイヤ情報を知ることによって有利にゲームを進めることができる。逆に、自身のみが保有している情報を他プレイヤに知られないことにより、同じように有利にゲームを進めることができる。本研究では、algo におけるプレイヤのみが持つ情報を他プレイヤに知られないようなプレイを行うプレイヤ構築を行なった。

## A Design of Algo Player to Act Not to Give Unfavorable Information

ITIRO SASAKI<sup>1</sup> SHIGERU FUJITA<sup>2</sup>

**Abstract:** In the incomplete information game, the player can advantageously advance the game by knowing other player information that is not shared. Conversely, it is possible to advance the game in the same way advantageously by not being able to know the information possessed solely by other players. In this research, we constructed a player that played so that the information held only by the player in algo is not known to other players.

### 1. はじめに

ゲームにおいて自身や他プレイヤの情報が全て公開されており、それらの情報を全てのプレイヤが確認することができるゲームは完全情報ゲームと呼ばれる。ゲーム木を考えたときに、ある局面に置いてプレイヤにとって識別可能なゲーム状態の集合をそのプレイヤの情報集合 (information set) というが、完全情報ゲームでは、全てのプレイヤの情報集合が1つに定まるので、ゲーム状態が確定している。このとき情報とは、プレイヤの持つカードの種類やこれまでとられた行動などといった、ゲームを再現する上で重要な情報を指す。チェス・オセロ・囲碁のような完全情報ゲームでは、人工知能 (以降、AI) は人間に勝利した [1][2]。しかし、他プレイヤの持つ情報を確認できない不完全情報ゲームでは、制限を設けない場合、人間に勝ち越すほどの勝率はまだ収めていない。[3] 不完全情報ゲームは、全プレイヤの全情報集合の要素が2以上であるため、プレイヤは自身がどの情報集合に属しているかを判

断することはできない。そのため、存在するかわからない情報に関しては推測するしかないため、考慮すべきゲームの状況が極めて多く、ゲーム理論やモデル化において多くの課題が残されている。

不完全情報ゲームでは、他プレイヤの行動から持ちうる情報を予測し、かつ自身の持つ情報を秘匿または誤認識させるような論理的かつ心理的な戦略を構築することで、ゲーム状況を自身にとって有利な局面へ動かし、勝利に近づかせることが必要である。これらのゲームには様々な種類があり、その中でも有名な麻雀やポーカーなどといったゲームでは、プレイヤの手札が他プレイヤに見えないようになっている。そのためプレイヤたちは他のプレイヤの手札を予測しながら、自身の手札はできる限り知られないように行動しつつ、ゲームに勝利しなければならない。本研究では不完全情報ゲームとして algo を題材とし、他プレイヤにできるだけ自身の手札を推測されないような行動選択を行うプレイヤの構築を行う。

### 2. 関連研究

将棋や麻雀を用いた研究では、未来局面を提示することにより初心者の学習支援を行ったり [4]、役作成のヒントを示すことにより麻雀初心者プレイヤに新たな役を発見する

<sup>1</sup> 千葉工業大学 情報科学研究科 情報科学専攻, Graduate School of Computer Science, Chiba Institute of Technology JAPAN

<sup>2</sup> 千葉工業大学 情報科学部 情報工学科, Department of Computer Science, Chiba Institute of Technology JAPAN

手助けを行っている。[5] これらの研究はプレイヤーを有利にするものと捉えられるが、プレイヤーに不利な手を選ばせないような研究は少ない。本研究では有利な局面を作るのではなく、不利な場面を作らないことに焦点を当てて研究を行った。

### 3. 研究対象

本研究では不完全情報ゲームとして、カードゲーム algo を対象とし、プレイヤーの構築を行なった。algo は算数オリンピック委員会、東京大学数学科の学生有志、数学者ピーター・フランクル氏らが共同で発明・開発したゲームで、小学校や学習塾でも教材として使われている。以下に algo のルールを示す。



図 1 製品版 algo

#### 3.1 algo のルール

algo ではアルゴカード 24 枚とチップ 40 枚を使用する。アルゴカードは白と黒の 2 種類があり、それぞれ 0 から 11 までの数字が書かれている。ゲーム開始時に人数に応じてアルゴカードを配る。余ったカードは山札として、プレイヤー間の中央に置いておく。

アルゴカードはカードを伏せて並べる必要があり、このときの並べる順番は左から数字について昇順である。同じ数がきた場合は、黒の方を小さい値とし並べる。プレイヤーは自身の手番がきた際に、山札があれば山札からカードを 1 枚手札に加える。加えたカードは先述した通りに並べなおす必要がある。手番プレイヤーは、他のプレイヤーの手札を推測する「アタック」という行動を行う。手番プレイヤーは自身の手札のカードのうち 1 枚を選択し、表にせずに出す。そこから他プレイヤーのカードを 1 枚指定し、カードの数字を宣言する。数字が当たって入れば指定されたプレイヤーはカードを開示する。数字が外れていた場合は、アタックの際に選択したカードを開示し、そのプレイヤーの手番を終える。アタックは成功する限り何度でも続けて良いが、最低でも 1 回は行わなければならない。これらの動作を繰り返し、1 人を除いた全てのプレイヤーのカードが開示された時点でゲームは終了する。このとき、開示されていないカードを持っているプレイヤーがそのラウンドの勝利者となる。40 枚のチップは複数回ラウンドを行う際に使用でき、プレイヤーのアタックの成功時やラウンド勝利時にプレイヤー間でやりとりをする。本研究ではこのゲームにおけるアタックを行う際に、アタックが失敗しても不利にならないようなカード選択を行うゲームプレイヤーの構築を行う。

### 4. 提案手法

本研究で提案する手法では、不利な手を選ばない algo プレイヤーの構築を行うために、はじめに相手プレイヤーが持っているカードを推測する。推測して得られた情報や、それぞれのプレイヤーの行動から得られた情報を基に、自身のカードそれぞれの重要度を計算していく。その中で最も重要度が低いカードを選択し、それを最も不利にならないカードとしてアタックに使用する。

カードの推測は以下の手順で 3 つに分けて行う。

- (1) 手札を用いた推論
- (2) 発言を用いた推論
- (3) 他プレイヤーから行う手札を用いた推論

それぞれの推論について説明を行う。(1) では algo のルールに則って論理的にカードの推論を行う。デモンストレーションとして、公式サイトに掲載されている詰め algo 5 問 (図 2 など) を解かせたところ、問題なく推論が行えた。(2) ではプレイヤーがアタックした際に得られた情報を基に、(1) の状態からさらにカードの情報を絞る。例として、アタックに失敗したとき、指定したカードは宣言された数字でないことが確定するので、その数字を推論から外すことができる。(3) で、(1)、(2)、で得られた情報をもとに、他プレイヤーの目線でカードの推論を行う。これにより、自分のカードがどれほど推論されているかを確認し、最も推論されづらくなるであろうカードを確定する。これらの情報を用いて、アタックの際に使用しても不利にならないカードを決定する。これを用いることで、アタックに失敗した際でもできるだけ不利益な情報を漏らさないようなカード選択ができる。また、研究を進めるにあたって以

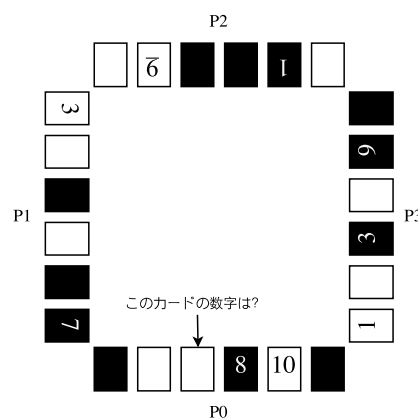


図 2 問題例 公式詰め algo 第 4 問

下の要素を推測項目から除外した。

- (4) 発言の信用判定
  - (5) 他プレイヤーの行動からなる手札予測
- (4) は algo のルール上、ブラフをかけて得られる利益より、必ず 1 枚カードを後悔してしまうことになるという不利益の方がはるかに大きくなるため、今回は考慮をしない。ま

た(5)に関しては、プレイヤーのモデル化を行う必要がでてくるものの、それを行ってまで得られる情報は少ないため、同じように今回の実験では考慮をしない。

## 5. 評価実験

それぞれ別々のカード選択を行う思考ルーチンをもった、対コンピュータ戦行う。対コンピュータでは、アタックカードの選択法の違うプレイヤーを使って対戦をさせる。選択法は以下の3つを用意する。

- (1) ランダムにカードを選ぶ
- (2) 端のカードから選ぶ
- (3) 中央のカードから選ぶ

それぞれを、Random, LR, Mid と名付け、不利な情報を与えないよう行動するプログラムを Agent とし、これらの4つのプログラムを同時に対戦させた。ゲームはサーバ上で行われ、他のプレイヤーのカードは常に秘匿されている。各プログラムは自身の手札と公開されているカードのみを用いて推論を行う。集計も同時にサーバで行われ、随時ログファイルにアタックの結果を書き込む。カード推論の思考ルーチンは全て同じであり、アタック時のカード選択の思考ルーチンのみ別にしてある。

また、プレイの順番による偏りがでないように、プログラムは毎回ランダムに配置換えを行う。あわせて、全てのプログラムがアタックをできるよう、アタックの回数は自分のターンに1回のみとした。(ルール上はアタックに成功した場合、連続でアタックすることができる)

このゲームを合計100試合行い、プレイ中に行われたアタックを集計し、その成功率と防御率を集計した。

## 6. 結果

200試合、3876回のアタックを集計し、成功率と防御率について表1、表2にまとめる。

表1 プログラム毎のアタック成功率

	成功数	攻撃数	成功率
Random	995	937	0.941
LR	963	921	0.956
Mid	908	923	0.943
Agent	913	956	0.955

表2 プログラム毎の防御率

	被攻撃数	防御数	防御率
Random	1071	52	0.048
LR	808	44	0.054
Mid	1026	41	0.039
Agent	971	60	0.061

## 7. 考察

プログラム毎のアタック成功率については、大きな差異はみられなかった。これはどのプログラムも同じ思考ルーチンで行動を行なっているためである。詰めアルゴからも

みられるように、algoの醍醐味の大部分は見えないカードの推論である。そのため、プログラムにプレイさせて、成功率が9割を超えるほど高くなってしまってもおかしいことではない。

防御率についても、さほど大きな差はなかったが、Agentの防御率が最も高くなっている。Agentに次いで高いのはLRである。ゲーム序盤における中央のカードは端のカードが開かれていても推測が難しく、逆に端のカードは推測より当てずっぽうで当てることの方が多い。これは0, 1や10, 11などといった、推測されても対して困らないカードが多いことにより、残りのカードの匿名性をあげている。万が一アタックに失敗してしまっても、リスクの少ない端のカードを使うことによって、中央のカードの推測をさせないようにしていると考えられる。

Midは4つのプログラムの中でも特に低い。中央寄りのカードはプレイヤーのカードたちの値の傾向を判断する材料になる上に、ゲーム終盤では勝敗のカギを担うことも多々ある。よって、中央のカードを使ったアタックは、失敗のリスクがかなり大きい。

Agentは、アタックに使うカードが公開されてしまったとき、連鎖的に推測できる他のカード数を計算することによって使うカードを選んでいる。情報が漏れるのは仕方ないが、その被害を最小限に留めるために、相手プレイヤーの視点から推測を行い手が決まる。場の状況から、臨機応変にカードの決定を行うことができた。

本研究ではカードゲーム algoを対象とし、相手プレイヤーに対して不利な情報を与えないよう行動するプレイヤーの構築を行なった。カード選択が勝率に及ぼす影響は小さなものではあるが、不完全情報ゲームではその小さな影響がゲームを左右する。今後の研究では様々な思考ルーチン、カード選択法などを追加し、アタックのカード選択のみならず、アタック先のカードの選択やアタックの回数などのパラメータも利用して、より不利になるような情報を与えずにすむプレイヤー構築の研究を行う。

## 参考文献

- [1] 松原仁, “Deep blue の勝利が人工知能にもたらすもの,” 電子通信情報学会, 7月, 1997年.
- [2] David Silver et al., “Mastering the game of go with deep neural networks and tree search,” *Nature*, vol. 529, pp. 484–503, 2016.
- [3] 古居敬大, 三輪誠, 近山隆, “不確定不完全情報展開型多人数ゲームにおける相手モデル化による搾取相手の選択,” ゲームプログラミングワークショップ2011論文集, vol. 2011, pp. 46–53, 10月2011年.
- [4] 伊藤毅志, “未来局面を用いた将棋の学習支援システム,” Tech. Rep. 7, 電気通信大学, Jul 2017.
- [5] 早川大貴, 上野末貴, 井佐原均, “麻雀初心者に向けた役作成のヒント例示システムの構築,” Tech. Rep. 8, 豊橋技術科学大学大学院, 情報メディア基盤センター, 情報メディア基盤センター, Jul 2017.