

不完全情報型ゲーム”ぷよぷよ”の勝率向上手法

上田 圭造[†] 梶並 知記[‡] 鈴木 孝幸[†] 清原 良三[†]

神奈川工科大学[†] 文教大学[‡]

1. はじめに

近年、ゲーム AI は急速に発達し、チェス、将棋などの完全情報ゲームでは、AI の強さは人間のトッププレイヤーを上回るようになった。これに対して、麻雀、ぷよぷよなどの不完全情報ゲームでは、未だ AI の強さはそれらのプレイヤーと比べると劣っている。その原因の一つは、ゲームの状態の特徴量の取得やルールの特示化が難しいことである。

“ぷよぷよ”では落下する塊は完全にランダムであり、数に制限がないため、どの塊が出るかは常に独立の事象となる。また連鎖によって発生する“おじゃまぷよ”は対戦相手に影響を与えることができ、最適な戦略を単純に計算できなくなるという特徴がある。

本論文では、落下型パズルゲーム“ぷよぷよ”においてコンピュータ側のプログラムを“COM”、“連鎖を開始する”ことを“連鎖を発火する”と呼ぶこととし、最適な戦術を実現することを目的とした“ぷよぷよ”の状態特徴量を用いた戦術的アルゴリズムを提案する。

2. 関連研究

杉本ら[1]は COM を対戦相手とした 2 人対戦型ゲームにおいてゲーム固有の知識を用いることなく一般にプレイヤーの強さを推定し、それに応じた強さでありながら不自然さを感じさせないような COM を用意する手法を作成し、プレイヤーが COM を相手とした対戦ゲームにおいて、より楽しさを感じるようにすることを目標としている。

また、最善の戦略が存在せず、戦略同士に相性があるような複雑な対戦型ゲームにおいて、COM の調整にはプレイヤーの戦略を含めたプレイ内容の推定が必要である。本論文では人間のプレイヤーを対象とした戦略ではなく、COM を対象にし、その性能向上を目的とした戦略を提案する。

3. ぷよぷよとは

“ぷよぷよ”のルールを以下に示す。

1. 図 1 に示すようにフィールドは縦 12、横 6 マスの格子で構成され、1 マスにつき 1 つの“ぷよ”と呼ばれる石を置くことができる。

2. “ぷよ”は 4 色あり、フィールド上部からランダムな色の“ぷよ”が 2 つ 1 組で落下する。プレイヤーはその”ぷよ”に対して、回転、横移動、高速落下のいずれかの操作を行う。
3. 縦横に同色の“ぷよ”が置かれると、それらは結合する。4 つ以上結合した場合に消滅し、得点となる。
4. “ぷよ”の消滅により、上にあった“ぷよ”は落下する。このとき再び“ぷよ”が 4 つ以上結合すると消滅し、連鎖が起きる。連鎖数が多いほど得点は高くなる。
5. 得点に応じて、相手のフィールドの上部に“予告ぷよ”が表示される。“予告ぷよ”は相手プレイヤーが“ぷよ”を落とした直後、“おじゃまぷよ”としてフィールドに落下する。この“おじゃまぷよ”は上下左右に隣接する“ぷよ”が消えると同時に消える。
6. 自分のフィールドに“予告ぷよ”が表示されている場合に連鎖を行うと、“予告ぷよ”の量を減らすができる。これを相殺という。また、“予告ぷよ”がある状態で相手の得点より大きな連鎖を行うと相手との得点の差分が相手のフィールドの上部に“予告ぷよ”として表示される。
7. 左から 3 列目の 12 段目が“ぷよ”で埋まるとゲームオーバーとなる。

次に基本的な戦術を説明する。ぷよぷよの対戦における勝利条件は相手のフィールドの左から、3 列 12 段目を埋めることである。相手のフィールドを埋めるには連鎖を発火し、“おじゃまぷよ”を相手のフィールドに送る必要がある。“ぷよぷよ”における相殺システムは、相手と自分の連鎖の大きさの差分を“おじゃまぷよ”

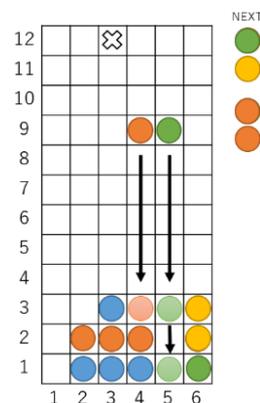


図 1 連鎖開始前状態

Improvement Strategy of “PuyoPuyo” which is a game with imperfect information

KEIZO UEDA[†], TOMOKI KAJINAMI[‡], TAKAYUKI SUZUKI[†], RYOZO KIYOHARA[†]

[†] Kanagawa Institute of Technology

[‡] Bunkyo University

として送る。よって、一般に相手より大きな連鎖を作る必要がある。連鎖は発火してから次の“ぷよ”が操作可能になるまで、インターバルが生じ、連鎖数が多いほど長い。この間に降ってきた新しい“ぷよ”で相手より1連鎖大きくすれば有利状況を作ることができる。

既存 COM[2]のアルゴリズムでは、相手の連鎖の発火を確認し、対応する方がより堅実なプレイができるため、自ら連鎖を発火し、仕掛けることは少ない。しかし、自ら仕掛けた方が有利な状況を作れる状況で、連鎖を仕掛けないと勝率は向上しない。本提案では、連鎖実行時間の短い連鎖を保持し、相手に隙がある状態で、自分から保持している連鎖を仕掛ける。

4. 提案手法

アルゴリズムを実装した環境を表1に示す。提案アルゴリズムでは、お互いが一定の連鎖数以上を保持している場合に、連鎖実行時間の短い連鎖を作成・保持し、相手の連鎖作成時に一定の隙があるタイミングで、保持していた連鎖を発火する。これにより、相手フィールドに“おじゃまぷよ”を送り、有利状況を作る。

一定の連鎖数のある連鎖を保持しているか否かを判断は、対戦における中盤かどうかの判定に必要である。今回の提案手法では5連鎖以上を保持している場合を中盤と定義する。中盤以降は如何に相手に先に大きな連鎖を発火させるかが課題となる。また、一定の隙とは、「任意の連鎖を開始するために必要な個数の“ぷよ”」を引くまでの時間である。

提案するアルゴリズムの適用条件は、アルゴリズムによって作成した短い連鎖を発火してから連鎖終了までにかかる時間が、「相手が任意の連鎖を開始するために必要な個数の“ぷよ”」を引いてくるまでの時間より短いことである。既存 COM における、1組の“ぷよ”を引いて設置するまでにかかる時間は約1.9秒であり、短い連鎖の連鎖実行時間が長いと、相手が任意の連鎖を開始するために必要な個数の“ぷよ”を引く確率が高まり、作成した連鎖を発火する条件を満たさない場合がある。そのため、短い連鎖の連鎖実行時間の上限を、相手が“ぷよ”を3組までしか引くことができず、2連鎖を発火した際に生じる連鎖実行時間と等しい5.7秒と定義する。

表1 実装環境

CPU	Intel Core i7-7500U 2.70GHz
OS	Windows 10
メモリ	8GB

5. 評価

提案手法を適用した COM と従来手法の COM を対戦させ、アルゴリズム適用後の各対戦場面の有利・不利状況の検証を行った。

有利・不利状況の判断は、相手に送った“おじゃまぷよ”によって相手の保持している連鎖の選択肢の数の変化をもとに判断をする。連鎖の選択肢の数が減った場合、行動が制限されるため“おじゃまぷよ”がない場合に比べ、不利になったと言える。アルゴリズム適用後の各対戦場面の有利・不利状況を“おじゃまぷよ”が降る前と後で分析した結果を表2に示す。

この結果からわかるように相手に“おじゃまぷよ”を降らせ、相手の保持している連鎖の選択肢を減らすことに成功し、有利状況を作ることができた。しかし、アルゴリズムが適用される状況は限定的であり、得られたデータ数は少ない。原因としてアルゴリズムの適用条件を達成する頻度が少ないことが挙げられる。また、相手の連鎖の形によっては連鎖の選択肢を減らすことができず、ただ時間と連鎖を無駄にってしまう場合があることもわかった。

6. おわりに

本論文では、“ぷよぷよ”の対戦における戦略的アルゴリズムを提案した。既存 COM では対戦の中盤において自分から連鎖を発火することは少なく、勝率向上の妨げになっていた。このため、著者らは中盤以降に実行時間の短い連鎖を保持し、隙を窺って発火して有利状況を作り出すことで課題点の解決にあたった。しかし、アルゴリズムが適用される状況は限定的であり、得られたデータ数は少ない。また、実行時間の短い連鎖を作成・保持した後に発火する機会がなく対戦が終わってしまう場面も多々あった。アルゴリズムの適用条件を調整し、再実験を行い確かめる必要がある。

7. 参考文献

- [1] 杉本直樹, 鶴岡慶雅: 戦略の動的推定による2人対戦ゲーム接待AIの提案, The 23rd Game Programming Workshop 2018, pp.114-119
- [2] “GitHub - puyoai/puyoai: AI for puyo”. <https://github.com/puyoai/puyoai>, (参照 2020-05-04).

表2 おじゃまぷよの有無による連鎖の選択肢の変化

	おじゃまぷよの数	降る前	降った後
1	8個	2通り	2通り
2	24個	3通り	0通り
3	16個	3通り	0通り
4	10個	2通り	1通り
5	2個	3通り	2通り