

ポケモン対戦に対するISMCTSアルゴリズム活用の可能性

1.はじめに

- 囲碁は盤面の評価が難しく、将棋などの他のゲームで使われていた手法が通じなかった
→2006年に画期的なアルゴリズム、モンテカルロ木探索(MCTS)が登場

- 2016年に囲碁のチャンピオンに Deep MindのAlphaGOが勝利
→もはや完全情報ゲームではAI>人間



- 不完全情報ゲームでは決定化(determinization)を使ったMCTSがいくつかのゲームで結果を残す
→しかしながら決定化は計算効率の悪さや戦略集合(strategy fusion)の影響を受けるなどの弱点があるためいくつかの不完全情報ゲームではうまく働かなかった

- 2012年に登場した情報集合を探索木のノードにするISMCTS(Information Set MCTS)は決定化でうまくいかないいくつかの不完全情報ゲームにも働いた



より戦略融合の影響が強く複雑な不完全情報ゲームでもあるポケモン対戦にもISMCTSを活用したい

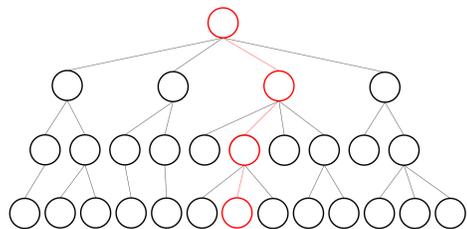
2.MCTSアルゴリズム

古典的モンテカルロ(MC)との相違点

- 有望そうな手を優先して探索する
→MCでは全ての手に均等に探索回数を振り分ける
有望そうな手を選択するときUCB1を使うのがUCTアルゴリズム
- ある手の探索回数が一定値を超えたら手を展開し、次の手も木に追加して探索するようにする
→MCはゲーム木が描けないほど大きな探索空間のゲームにも対応するために木探索の良さを失っていたが、MCTSは時間と共に木探索をするようになる

MCTSの手順

- 根ノードから出発し、評価値が高い手を辿って木を下っていく
- 末端ノードまで来たときに訪問回数が閾値を超えていればノードを展開する
- 展開できない末端ノードに到達した後は乱数により手を決定してゲームセットまで進める(通称プレイアウト)
- 辿ったノードに勝敗をフィードバックする

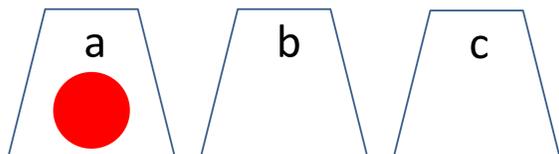


3.情報集合(information set)

情報集合とはあるプレイヤーから区別できない状態の集合のこと

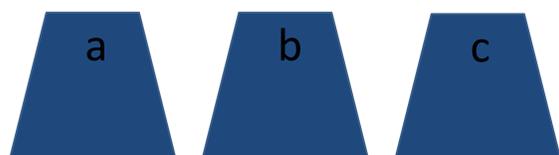
例えば、プレイヤー1は三つのカップの内、どれか一つにボールを隠しプレイヤー2はボールが入っているカップを当てるゲームを考える

- カップが透明な場合



プレイヤー2はプレイヤー1の取った行動の区別がつくため簡単にカップを当てることができる

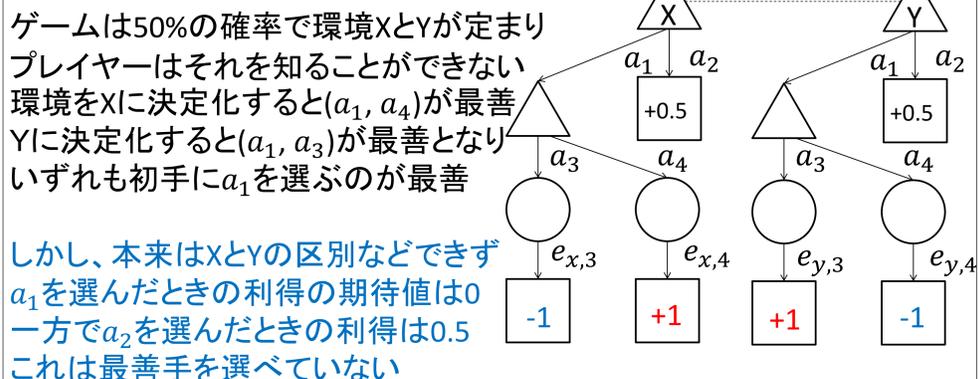
- カップが不透明な場合



プレイヤー2はプレイヤー1の取った行動の区別がつかないつまり、プレイヤー2が到達した情報集合に三つの状態が含まれている

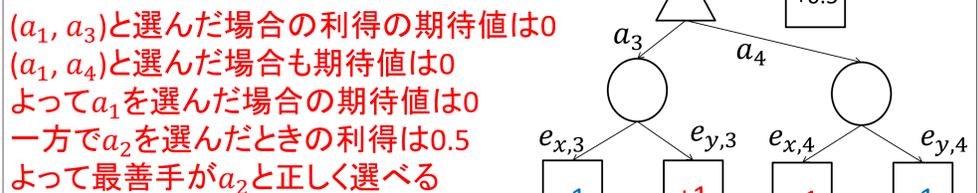
4.戦略融合(strategy fusion)

戦略融合を理解するために以下のゲームを考える
決定化を使用すると a_2 が最適解であるのに a_1 を選んでしまう



根ノードを情報集合に置き換える

XとYの区別ができた場合どうするか
考えるのをやめ、分からないものはそのまま考える



情報集合を使うと戦略融合の影響を抑えられることもある

5.ISMCTSアルゴリズム

ISMCTSはMCTSの探索木のノードを情報集合に置き換えたもの
提案者は以下の三つの変形を示した

- Single-Observer Information Set MCTS (SO-ISMCTS)
MCTSのノードを情報集合に置き換えたISMCTSの基本形
決定化を使ったMCTSよりも戦略融合の影響を受けにくい
部分的に観測可能な手(Partially Observable Moves)を含んだゲームだと戦略融合の影響が残る
- Single-Observer Information Set MCTS with Partially Observable Moves (SO-ISMCTS + POM)
SO-ISMCTSにPOMを組み込んだ形
決定化を用いずに相手のPOMをランダムな手と仮定するため
SO-ISMCTSよりも戦略融合の影響を受けにくい
相手のPOMをランダムだとするため相手のモデリングが不十分
- Multiple-Observer Information Set MCTS (MO-ISMCTS)
異なる情報を持つプレイヤーの数だけ探索木を作る変形
探索木に現れたPOMはPOMを打つプレイヤーの探索木で考えるため
相手のPOMに対するモデリングに優れる
木を複数作ることによって計算資源が分散される

これら三つの変形はいずれもメリットデメリットがあるため最優はない
ゲームの特性や先手後手でもアルゴリズムの実力が変わる

6.実験設定と結果

ISMCTSの実装は間に合わず、今回実験できなかったが
ポケモン対戦に戦略融合の影響があることを確かめる実験を行った

| | ランダム | MC | MCTS |
|-----------|-------|-------|-------|
| ランダムプレイヤー | - | 0.215 | 0.140 |
| 古典的モンテカルロ | 0.785 | - | 0.510 |
| 決定化+MCTS | 0.860 | 0.490 | - |

探索木を持たない古典的なモンテカルロ法と決定化を使ったMCTSの実力差がほぼなかったことから戦略融合の影響が木探索のメリットを打ち消していると考えられる
今後とも研究を重ねてより強いポケモン対戦のAIを考えていきたい