

ポケモン対戦に対するUCTアルゴリズムの有効性の評価

目的

ポケモン対戦におけるUCTアルゴリズムの評価と有効な運用法の調査

結論

類似した特徴を持つ情報ゲームと同じ運用法が最も効果的だとわかったがUCTアルゴリズム自体がそれほど有効でない可能性があることが分かった

1.はじめに

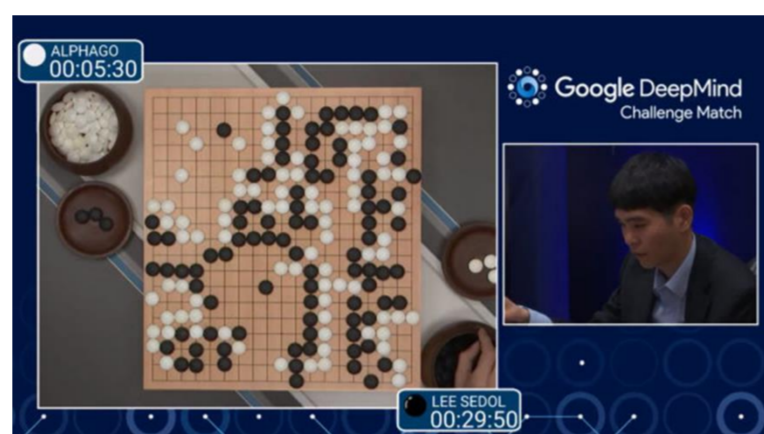
• 囲碁は盤面の評価が難しく、将棋などの他のゲームで使われていた手法が通じなかった

→2006年に画期的なアルゴリズム、モンテカルロ木探索(MCT)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



AlphaGOと対戦する囲碁チャンピオン

• 相手の状態が完全に分かる完全情報ゲームではAIの方が人間よりも強くなってしまったが、相手の状態が完全には分からないポーカー、麻雀、人狼などではまだまだ勝てない

→不完全情報ゲームや不確定情報ゲームには研究の余地あり

• ポケモン対戦は世界一有名なturn-based RPG

そのゲーム特性は不完全不確定情報ゲーム

であり、学術的に興味深いポケモン対戦の

ゲームAIはまともに研究されていない

→ポケモン対戦は大いに研究の余地あり



ポケモンの代表的なキャラクター
©株式会社ポケモン

ポケモン対戦のAI研究の第一歩として囲碁で成功を収めたUCT(MCTの中の代表的な手法)の有効性を調査しよう!

3.比較手法

不完全情報ゲームにUCTアルゴリズムを使うためには

1. 局面を仮定して不完全情報を完全情報にする

→相手の技やステータスなどの分からない情報を仮置きしないとポケモン対戦ではゲームが成り立たない

2. 探索中に遷移不能なノードが出た場合は一時的に削除

→局面の仮定をすることで複数のゲーム木が一本に集約され局面の仮定次第では指せない手もゲーム木に含まれてしまう

3. 探索中に新たな手が打てるようになった場合はゲーム木に追加

→局面ごとに打てる手が異なるため、新しい手は常に木に追加する

他の不完全情報ゲームではこれら三つの工夫が取り入れられることで完全情報ゲーム向けのUCTを不完全情報ゲームでも扱えるようになったしかし、ポケモン対戦ではこれらの工夫を取り入れた場合には**メリットもデメリットも存在する**ため、これらの工夫はポケモン対戦にも有効なのかどうか、複数のパターンのUCTを戦わせて実験した

	局面の仮定	ノードの追加	ノードの除外
MC	○	×	×
UCT①	×	×	×
UCT②	△	×	×
UCT③	○	×	×
UCT④	○	×	○
UCT⑤	○	○	○

実験は原始的なMCを比較のために加え、UCTの5パターンと合わせた6パターンのAIで実験を行った

ポケモン対戦でも不完全情報ゲームにUCTを用いるときの三つの工夫が有効であるならば表の下の方にあるUCTほど強いはずである

2.UCTアルゴリズム

MCTはモンテカルロ法(MC)に木探索を盛り込んだ探索方法

モンテカルロ法との相違点

1. 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にUCB1を使うのがUCTアルゴリズム

2. 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、MCTは時間と共に木探索をするようになる

UCTの手順

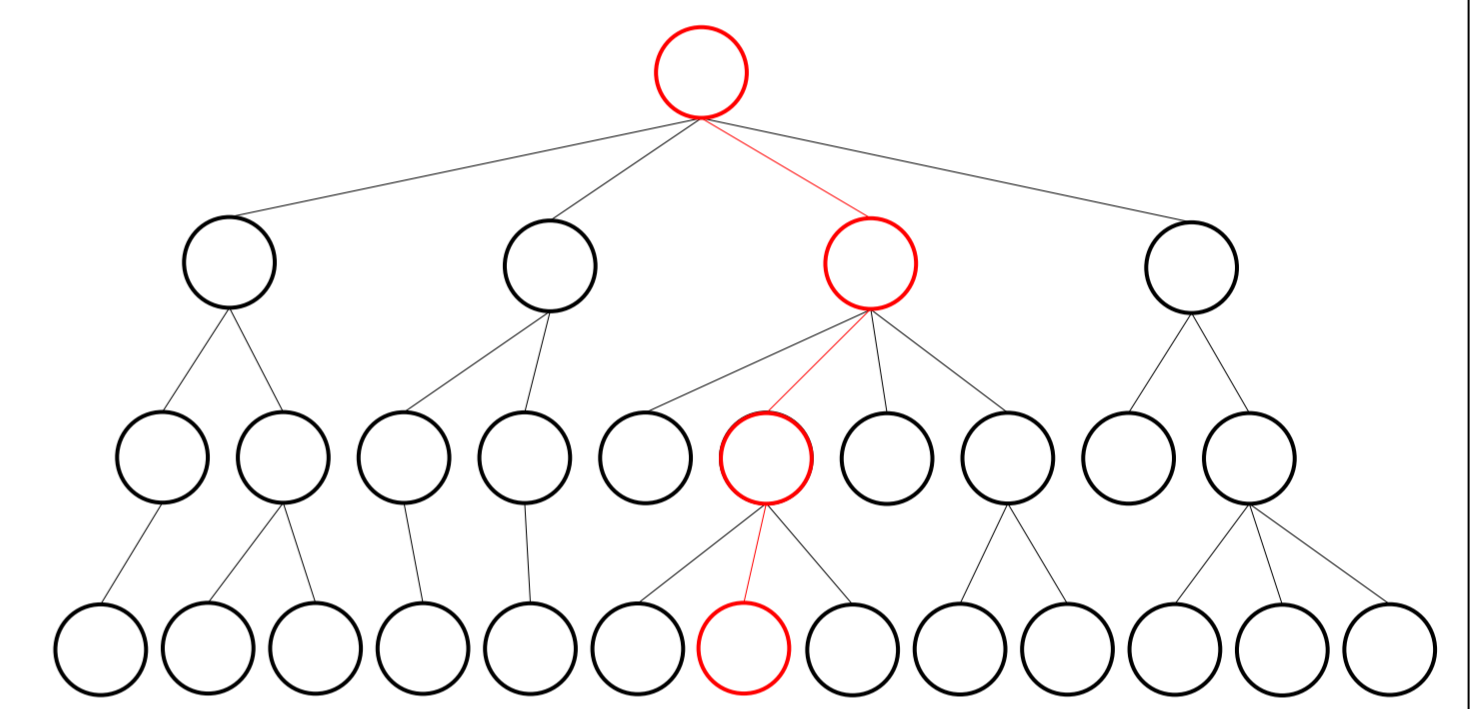
1. 根ノードから出発し、UCB1値が高い手を辿って木を下っていく

2. 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

3. 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)

4. 辿ったノードに勝敗をフィードバックする



4.実験設定と結果

• 実験はゲーム機で回数をこなすことが困難なことから筆者が作成したシミュレーション上で行った

• 各手法同士を200回ずつ対戦させて勝率により強さを比べた

	vsMC	vsUCT①	vsUCT②	vsUCT③	vsUCT④	vsUCT⑤	順位
MC		0.665	0.665	0.605	0.565	0.510	1位
UCT①	0.335		0.455	0.460	0.375	0.340	6位
UCT②	0.335	0.545		0.485	0.490	0.330	5位
UCT③	0.395	0.540	0.515		0.465	0.445	4位
UCT④	0.435	0.625	0.510	0.535		0.455	3位
UCT⑤	0.490	0.660	0.670	0.555	0.545		2位

実験から判明したこと

1. UCT①から③までの三つの結果から**局面の仮定を行うことはポケモン対戦において重要**

→局面の仮定を多く行った手法の方が強いことから言える

2. UCT①から⑤までの五つの結果から**三つの工夫は全てポケモン対戦に対して有効**

→工夫を多く取り入れた手法の方が強いことから言える

3. 原始的なMCがUCTよりも強かったことから**UCT自体がポケモン対戦に不向きである可能性がある**ことが言える

→ポケモン対戦は局面ごとに有効な手が全く異なるため、前回までの局面で有効だった手を今回の局面でも有効だと仮定して探索を行うUCTとは相性が悪かった可能性がある

原始的なMCがUCTよりも強かったことは大変興味深いので今後はこのような結果になった原因をしっかりと調査したい