

# オセロAI

## モンテカルロ木探索、ニューラルネットワークを試す

### 概要

ボードゲーム:オセロにモンテカルロ木探索やニューラルネットワークを利用してアルファゼロを実装しました。アルファゼロは事前の学習はせずに自己対戦することでどんどん強くなるもので、こちらを試してオセロに組み込んでみました。前提知識がない状態で試行錯誤を繰り返しているためまだまだ改善点がたくさんあります。

【開発環境】 ・ Python

### 実装機能

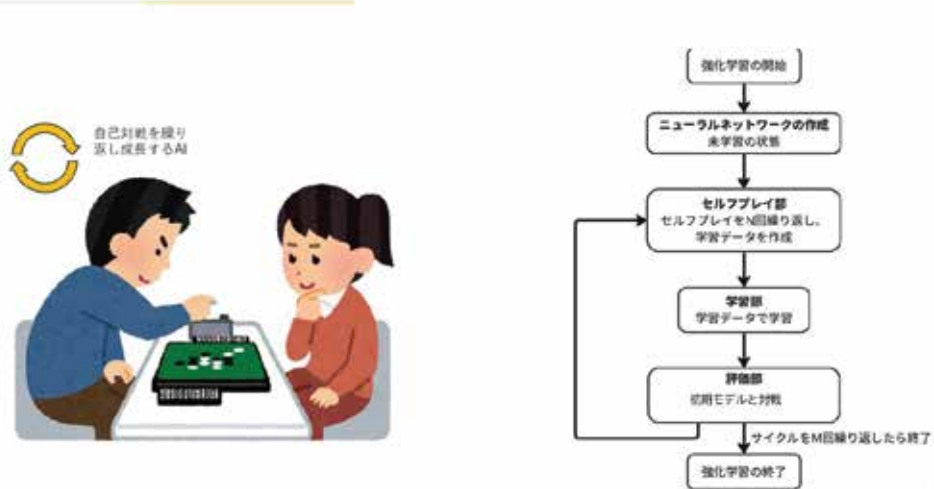
#### AlphaZeroの誕生

2006	モンテカルロ木探索	2016	AlphaGo誕生	2017	AlphaZero誕生
読み込みニューラルネットワークと強化学習の発展					
オセロ	1997年	ロジステロが世界チャンピオン村上健に勝利			
チェス	1997年	IBMのディープ・ブルーが世界チャンピオン ガルリ・カスパロフに勝利			
将棋	2013年	GPS得勝がA級棋士に勝利			
	2017年	Furukawaが最優秀名人に勝利			
囲碁	2016年	AlphaGoがイ・セドル九段に勝利			

AlphaZeroは三つの要素が組み合わさったモデル  
(探索) 先読みする力 ← 確率的探索アルゴリズム  
(直感) 最善手を予測する ← 読み込みニューラルネットワーク  
(経験) 自己対戦による経験 ← 学習サイクル  
AlphaZeroは、一定の深さまで確実に先読みを行い、その先をニューラルネットワークで予測外部データを一切使用せず、自身で生成したデータを学習に用いる点が特徴的

アルファゼロは、前提知識がない状態で自己対戦を繰り返しながら対戦データを蓄積して強くなるAIです。

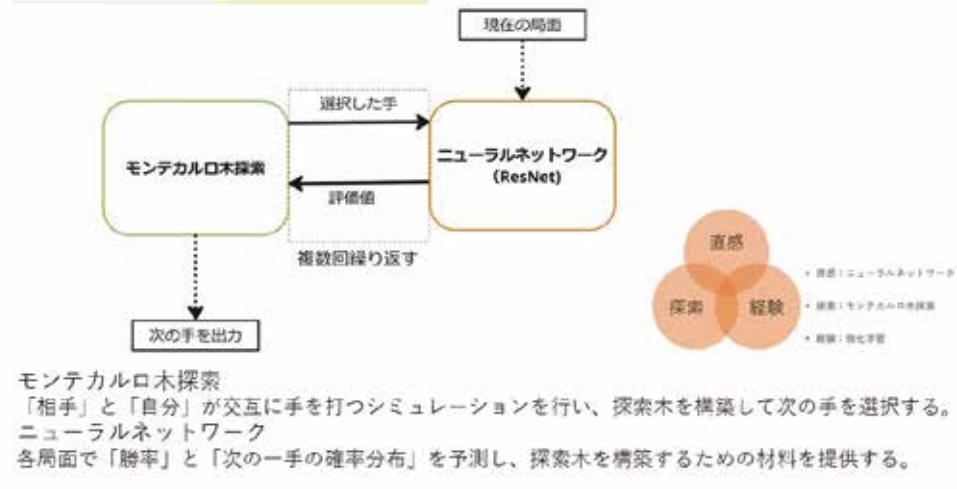
#### アルファゼロの強化学習サイクル



Aiが人間のように試行錯誤して学習します。

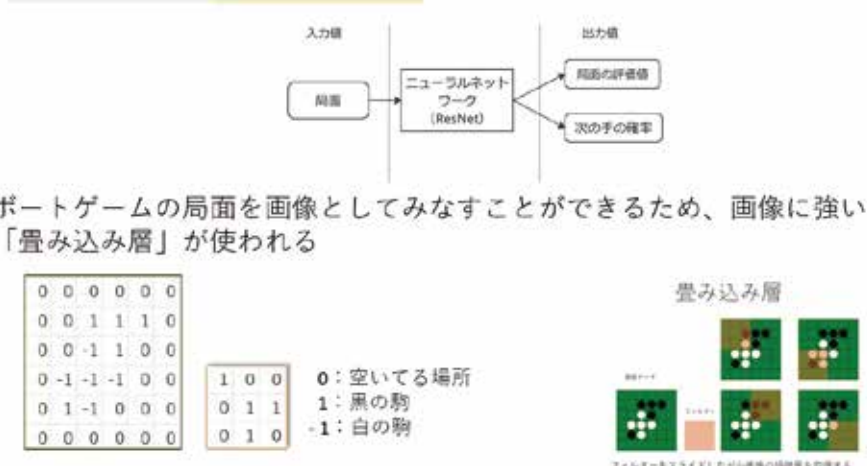
- 自分自身と対戦することを繰り返し勝敗のデータを蓄積する
- 蓄積されたデータを元にニューラルネットワークを更新する
- 未学習の状態と対戦することでモデルの強さを確認する
- セルフプレイ、学習、評価を繰り返しモデルが強くなる

#### AlphaZeroモデルの全体像



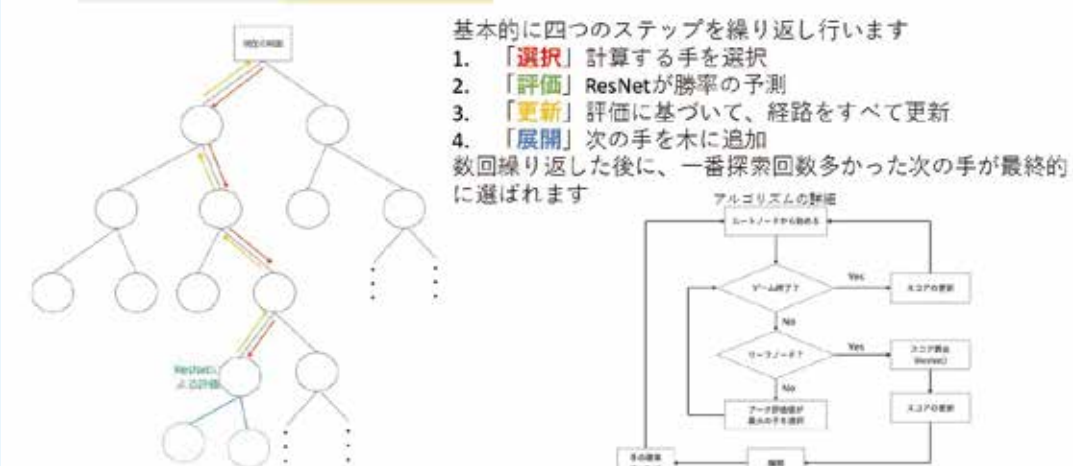
現在の盤面から有利になりそうと直観で思ったらこれを元に予想をする。  
次の一手を予想して相手はどう打ってくるかを比べながら一番いい手を選ぶ行為。

#### AlphaZero - ニューラルネットワークの部分



ニューラルネットワークは、人間の神経細胞を模倣して作られたアルゴリズムです。データの特徴量をニューロンとして表現しています。ボードゲームではResnetを利用して複雑な局面の画像の特徴量を抽出して学習します。カーソル(フィルター)を使ってスライドしながら特徴量を抽出して次の手に活かせる。

#### モンテカルロ木探索



モンテカルロ木探索は、次の4つのステップを繰り返して最善手を選びます。

- 現在の局面から始まり、訪問回数や勝率に基づいて次の手を選ぶ。
- 試していない展開に到達するとResNetを使って勝率を予測。
- 評価した結果を木構造全体に反映させる。
- 展開 評価後次の手を木に追加する。探索の幅が広がっていきます。最も探索された手が選択されます。

#### 今後の展開

ボードゲームは明確な「ルール」とゲームの終了時の「フィードバック」があるため、アルファゼロのような強化学習に適しています。

今後はボードゲームだけでなく、実用的な分野へ応用を目指したいです。現実でまだ解明されていない探索分野に活かせる可能性があります。

アルファゼロの応用範囲は、ゲームを超えて産業、医療、金融、科学など多岐にわたります。その汎用的な学習能力と効率的な意思決定は、複雑な問題解決において革新的な成果をもたらす可能性があります。