

1. 序論

京都将棋は縦横 5 マスの将棋盤で(香・と),(銀・角),(金・桂),(飛・歩), 王の 5 種類の駒を利用して対戦するミニ将棋である. 図 1 に京都将棋の初期配置を示す. 通常の将棋のルールと異なり, 王以外の 4 種類の駒は一手ごとに必ず裏返さなければならない, 持ち駒は裏表どちらでも打つ事が出来る. また通常の将棋で禁止されている二歩や, 行く場所のなくなる駒を打つことも出来る. 京都将棋はこのような独特なルールも相まって AI に関する研究が少なく, 学習データも少ない. そこで本研究では機械学習を用いて京都将棋の AI を開発する. AI 開発には Alpha Zero を参考に開発を行う. Alpha Zero とは DeepMind 社によって 2017 年に開発され, プロ棋士の棋譜データを用いることなく, 自己対戦により囲碁・チェス・将棋の学習が出来る AI である.

5	4	3	2	1	
香	香	王	香	と	一
					二
					三
					四
と	銀	玉	金	歩	五

図 1 京都将棋の初期配置

2. 研究内容

本研究では,python 言語を用いて Alpha Zero を参考に京都将棋 AI を作成する.

本研究では, 学習データの作成方法として, まずランダムに次の一手を選択する AI 同士を対戦させ,500 回分の対戦における学習データを作成する. この学習データを価値と方策を出力するデュアルネットワークにおいて学習を行う. 次に学習によって作成されたモデルと最新のモデルの対戦を 20 回行い, 最新モデルの勝率が 5 割を超えた場合, モデルの更新を行う. 最終的に学習データの作成, デュアルネットワークの学習, モデル同士の対戦を繰り返すことで強いモデルが作成される. モデル同士の対戦を行う際, 探索方法としてモンテカルロ木探索を利用する. 探索を行う際, 各ノードは情報として累計価値と試行回数を持つ. これらを評価関数に代入し評価の高いノードを選択し, AI 同士の対戦を繰り返すことにより, ノードの価値と試行回数を更新する.

3. 結果・考察

前章で述べた学習サイクルを 30 回行い, 学習させた bestAI とランダムに次の手を選択する RandomAI の対

戦を 100 回行った. 結果を以下の表 1 に示す. 表 1 が示すとおりランダムに手を選択する AI 相手には必ず勝てる強さの AI を作成することが出来たと考えられる.

次に人間と bestAI の対戦における局面の一部を図 2 と図 3 に示す. 図 2 の局面で人間側が持ち駒の角で, ▲ 1 三角と王手をかけたが, bestAI は図 3 のように△ 2 四と成と王手を回避しない手を選択した. このような手が選択される理由として学習回数が不足し特定の局面における学習が不十分であると考えられる.

表 1 RandomAI との対戦結果 (試行回数 100 回)

	bestAI 先手	bestAI 後手
勝	100	100
負	0	0



図 2 人間と AI の盤面 1

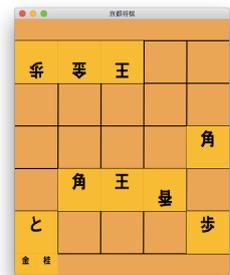


図 3 人間と AI の盤面 2

4. 結論

本研究では Alpha Zero を参考に京都将棋の AI 作成を行ったが, 人間との対戦では王手を回避しないなど脆弱な部分が確認された. 今回の学習は Google Colab を利用したため, 12 時間制限や GPU の利用上限などにより十分な学習時間が得られず, 次の手を探索する際に適切な評価が出来なかったことが原因であると考えられる.

今後の課題として王手を回避しない脆弱性の対策に, 合法手から王手放置の手を含めないプログラムを作成する必要があると考えられる. また学習回数を重ね, RandomAI や人間ではなくアルファベータ法や他の AI と対戦させ, 強さと脆弱性を検証する必要があると考えられる.

参考文献

- 1) 布留川 英一: Alpha Zero 深層学習・強化学習・探索 人工知能プログラミング実践入門, 株式会社ボーンデジタル (2019)