

将棋AIにおける指し手の意図の 解釈向上手法の提案

金沢工業大学 情報工学科 中沢研究室 4EP4-13 大田 皐介

研究背景

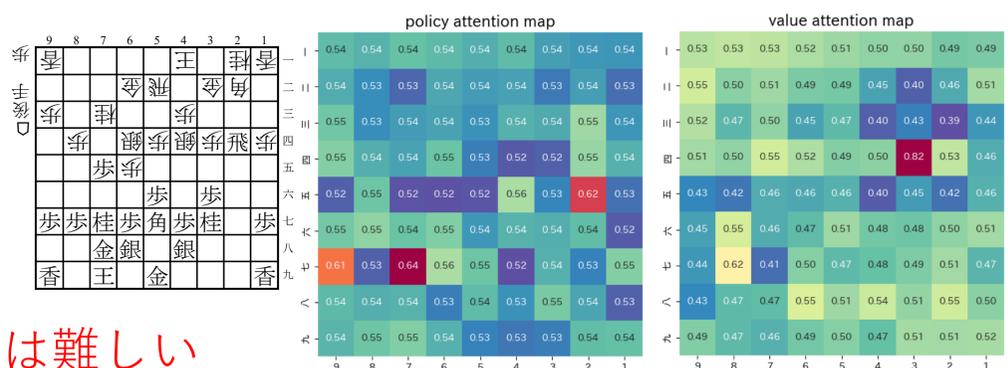
- 将棋AIは局面の評価値と指し手を出力するのみである。
- 将棋AIから出力される指し手を解釈することや指し手の意図を推測することは難しい

将棋AI(dlshogi)

- 深層学習を用いた将棋AI
- 局面の勝率と指し手を予測する ResNetを用いたニューラルネットワーク
- モンテカルロ木探索

既存手法：Attention Branch Network

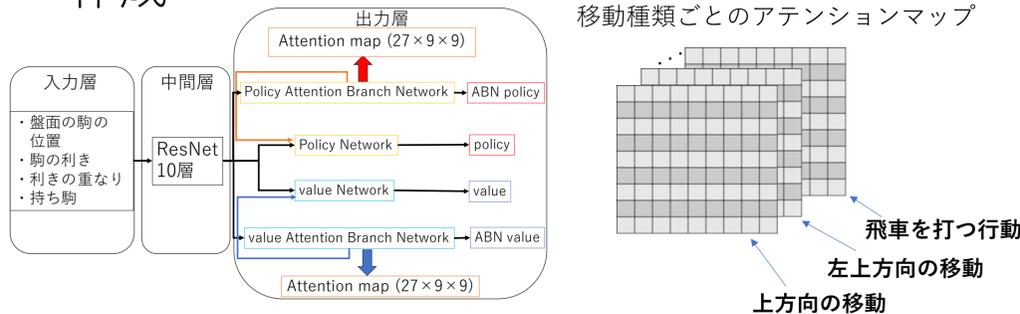
- Attention Branch Network
CNNが推論時に着目している領域を可視化することができるネットワーク構造
- policyとvalueを推論時のアテンションマップを出力し、局面のマスマスの重要度を可視化



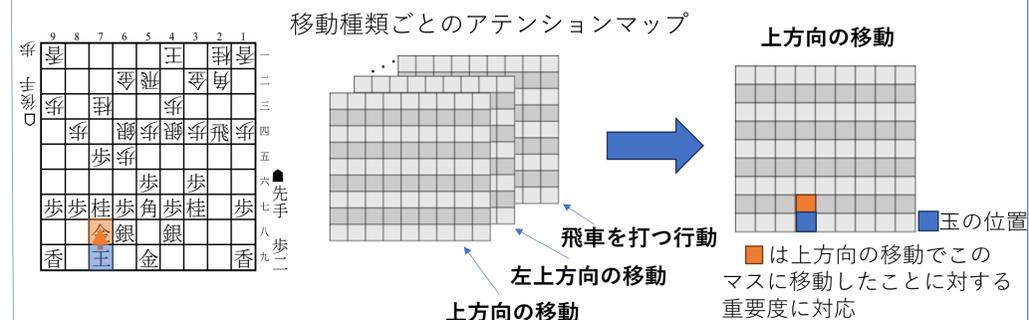
問題点: アテンションマップから指し手の解釈は難しい

提案手法: グラフ表現を用いた利きの重要度の可視化

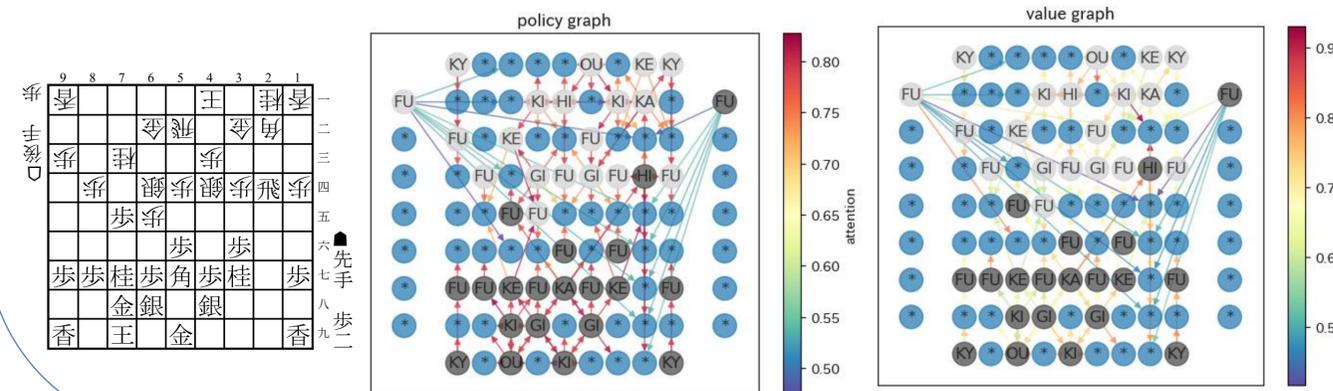
- Attention Branch Networkを用いて、駒の移動種類ごとのアテンションマップを作成



- 作成したアテンションマップから駒の利きに応じた重要度を抽出



- 駒の利きを辺とした局面を表す有向グラフの辺の重みに抽出した重要度を乗せることで可視化



駒の利き(駒の移動範囲)の重要度を可視化

将棋AIがどの駒をどう動かしていきたいのかを可視化し、指し手の意図の可視化に繋がる

評価

- 提案手法によって指摘された駒の利きが真に重要な駒の利きであるかを評価する。
- 局面sと提案手法によって指摘された駒の利きを除いたs'の将棋AIの出力の変動を計測する。

表. 出力の変動の計測方法

指し手(policy)	勝率(value)
KLダイバー ジェンス	value出力の差の絶対値