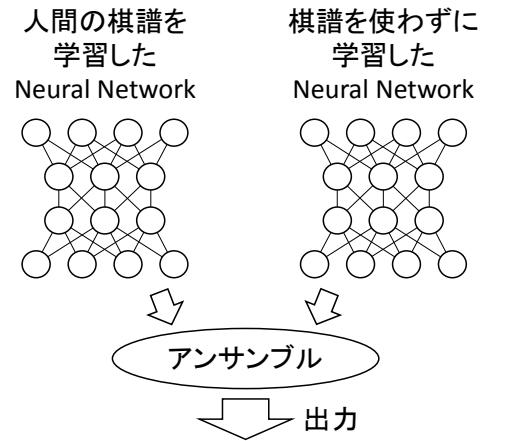


囲碁プログラム「Maru」

Maru は多層ニューラルネットワークを使ったコンピュータ囲碁プログラムです。基本的には、AlphaGo と同様に、Policy Network を用いて探索木を作成し、Value Network を用いて勝率を予測します。ただし、Maru では、AlphaGo とは異なり、最終的な領域（地）を予測する Territory Network も使います。Territory Network を用いることにより、勝率だけではなく、その局面で何目差となっているかを予測できます。そこで、Maru では、Territory Network の領域予測結果をもとに、終局や投了を判定しています。例えば、パスをしても領域予測が変化しない場合、その対局は終局したと判定します。また、Value Network によって予測された勝率が低く、Territory Network の予測領域に大差がある場合、その対局を投了します。そのため、勝率が極めて低くなったとしても、地が僅差となっている対局の場合は終局まで対局を続けます。

2018 年 6 月より、人間の棋譜データを使わない囲碁プログラム「MaruZero」の開発を行っています。MaruZero では、AlphaGoZero と同様に、棋譜生成と学習を繰り返す強化学習方式を採用しており、人間の棋譜データを使わないことにより、人間の考えとは異なる戦略・戦術を見つけようとする試みです。ただし、AlphaGoZero のプロジェクトとは異なり、個人所有が可能な計算機資源のみを用いて学習を実施するため、MCTS ではなく、Territory Network を使った強化学習となっています。現在まで、個人所有が可能な PC (CPU: Intel Core i7-7700K, GPU: Nvidia GTX 1080Ti × 2 個) で約 6 ヶ月間の学習計算を実行し、gnugo に対する勝率 90%以上を達成しています（ここまでに生成した棋譜の数は AlphaGoZero の約 2 時間分に相当し、使用した計算機資源は AlphaGoZero の約 0.7 時間分に相当します）。

今回の AI 竜星戦では、MaruZero のニューラルネットワークを使います。しかし、MaruZero のニューラルネットワークは学習途上であり、特に布石や序盤の定石を学習できていません。そこで、人間の棋譜を学習したニューラルネットワークも用意し、この 2 個のニューラルネットワークの計算結果をアンサンブルしたものを使って盤面の探索や勝率の予測などを行います。



- (1) Policy : 候補手予測 ⇒ 探索対象を決定
- (2) Value : 勝率予測 ⇒ 候補手を評価
- (3) Territory : 領域予測 ⇒ 終局・投了を判定