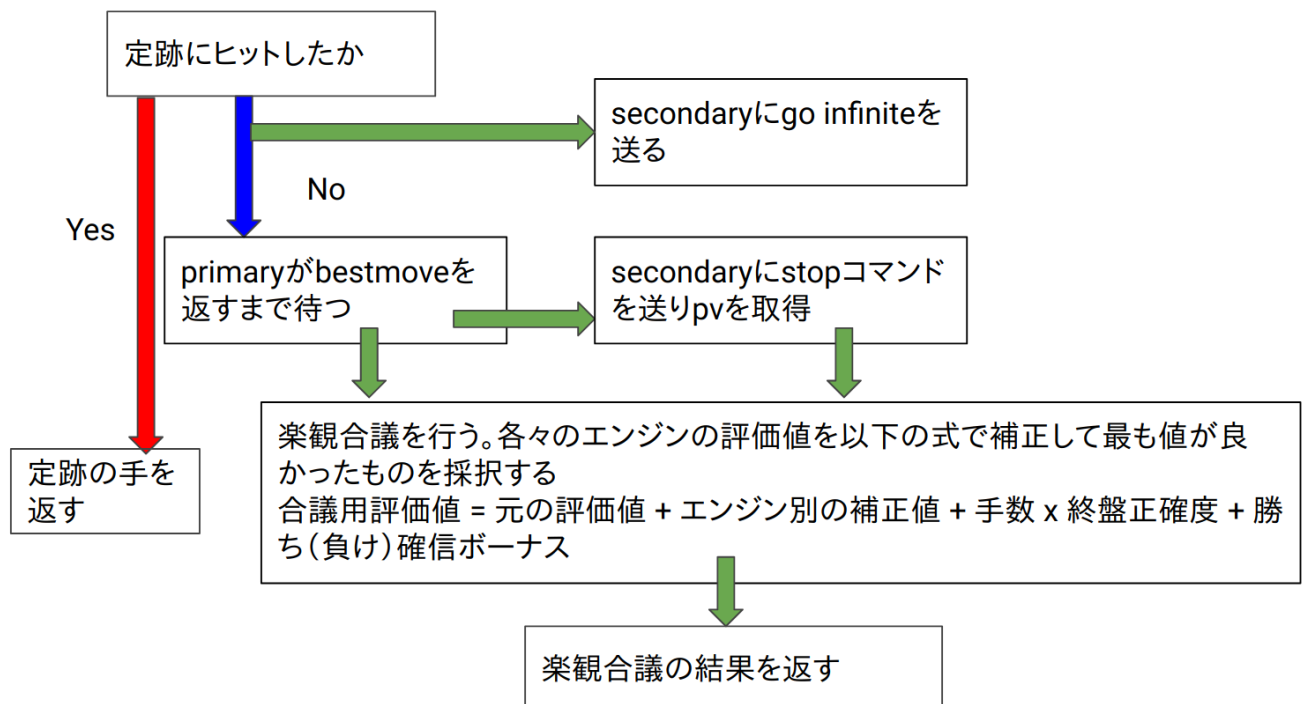


WCSC32 将 (Qha) 実装内容詳細

Quorax 将棋部 (Ryoto Sawada, Yuki Ito, Toshihiro Shirakawa and Keigo Nitadori)

【大会で実際に用いたモデル/クラスタについて】

本大会では、やねうら王とdlshogiによる楽観合議を用いた。以下のようなワークフローを用いることで、やねうら王-dlshogi間の評価値のスケールの違い、および、計算資源やモデルの精度の違いに由来した実力差を補正した。



・大会で使ったマシンのスペック

Yane: EPYC 7452 デュアルソケット,メモリ 512 GB

DL: Ryzen 5 3600X, メモリ 64GB, GeForce RTX 3080

マシンスペックの差とdlshogi側のモデルの精度の問題から、この条件で対局をするとeloレーティング300以上の差を付けてやねうら王側が勝つと思われる。

・合議用評価値の各パラメタの意味

エンジン別の補正 エンジンの強さに応じて一定の評価値を足す。過度に楽観的なエンジンと合議を組む際にそのエンジンの評価値にペナルティをつけるのに使える。本大会では、やねうら王、dlshogi双方に対して0とした

終盤正確度 手数に応じて各々のエンジンの評価値にボーナスをつける。例えば dlshogiの終盤力に不安があるなら、この値を負の値にすることで終盤のdlshogiの意見を無視するようなことができる。本大会ではdlshogi側に対して手数x8点のペナルティをつけるようにした

勝ち/負け確信ボーナス 各々のエンジンの評価値が過度に高い場合、何らかの勝ち筋を見つけた可能性があるとして、そのエンジンの評価値に大幅な(今回は+Mateと同じだけの)ボーナスを加える。本大会ではやねうら王側は評価値334以上で勝ちを確信しているとした。

・大会で使った定跡

HoneyWaffleの開発者が後悔している振り飛車用の定跡と、その定跡をベースにして生成した定跡を対戦相手によって使い分けていた。eloレーティングなどに対する影響はおそらく殆ど無い。長らく同じ定跡を使っているが対策された様子はない。居飛車側の振り飛車定跡に対する最良の対策は、相手の定跡に外すことで相手が飛車を振り損ねるようなことがないようにすることだと思ってる

【運用実績と個人的な感想】

今回調達したマシンではdlshogi側がやねうら王に比べて著しく弱いため、dlshogiの意見を取り入れることで得られるeloレーティングの向上は0~20程度と思われる。だが、試合中に双方のエンジンの意見を見られるのは面白い上に、やねうら王が諦めていても、dlshogiが諦めていないみたいな展開があると最後までやる気が出るので開発者の精神衛生的には非常に良かった。

【実験結果と再現性について】

合議の成績はやねうら王、dlshogi側のマシンのスペックに依存するため再現は容易ではない。合議そのものの追実験は、例えばfloodgateの手を正解とした教師データに対して、少ないノード数で探索した、やねうら王、dlshogiの探索結果を合議させることで可能である。例えば、2022年のfloodgateのrating 3500以上のソフトの指し手を当てる問題に対し、やねうら王10万ノードとdlshogi 1000ノードを合議させる場合、dlshogi側に300点程度のペナルティをつけることで手の一致率が0.5%程度上昇する。

【おまけ:PR文章で紹介したMuZeroの再現実装について】

PR文章で紹介したMuZeroの再現実装はdlshogiに比べて著しく弱かった、かつ、飛車を振らせるのが難しかったため採用を見送った。アーキテクチャとしてはpolicy, dynamics双方とも192ch、10層のresnetとした(dlshogiの10 layersのモデルと同じ)