

sakura アピール文書

2021/3/27 Yuhei Ohmori

大事なこと

- DeepLearningです
- dlshogiクローンです
- 独自の工夫ポイントのほぼすべて学習部分です
- 強化学習ではなくほぼ教師あり学習です

ポイント

- NNUEとKPPTのdepth8/12/20で作成した局面を元に教師あり学習しています
- 初期局面として、floodgateやAobaZeroやNNUEの自己対局の局面を使用しています
- GPUリソースを持っていないのでギリギリまで教師ありで強くし、どうしても強くならなくなったときに強化学習をやる予定です
- seresnetの20ブロックを予定しています
- ただ学習が全然進んでいないのでresnet + swishの10ブロックかもしれません
 - こちらはいろいろな実験用にGCTと同じモデルで学習させているものです

学習

- dlshogiで使用しているActor-Criticは使用しておらず、Policyは単純に教師の指し手を学習しています
 - 単純にこちらのほうがPolicy損失の結果が良かったからです
- 単純に教師ラベルを1ほかを0とするのではなく、label smoothingしています
 - NNUeでMultiPVである程度指し手の分布を出すのかも考えたのですが、うまくいかなそうかなと思いやっていません
- 教師局面が終盤に偏っているため、Value損失に進行度によりweightをかけました
 - 学習序盤のdepth8の教師データのみ
- Policy損失とValue損失のバランスを取るため一定の係数をかけています
- Data Augmentationとしてランダムで局面を左右反転しています
- なんかいろいろやったのですがほとんど意味はなく、いかに良い教師を作るかを注力したほうが良さそうだと今の所思っています

使用している外部リソース

- dlshogiを使用しています
 - 学習部分はそこそこ書き換えました
 - 探索部分はまだ手を出していません
- NNUEの教師局面は電竜戦のnozomiを使用しています
 - 水匠2ベースの追加学習です
- dlshogiチームが公開している教師局面を一部使用します
- floodgateならびにAobaZero等の棋譜を使用しています
- リソースではないですが、dlshogiチームの山岡さん、加納さん、2番絞リチームの芝さんのブログ等での情報を大いに参考にさせていただきました