elmoアピール文書 word版

瀧澤 誠 twitter: @mktakizawa

# 開発動機

元々はボナンザメソッドと呼ばれる機械学習の仕組み(識別でも回帰でも無い(pairwiseなrank学習))に興味を持ち、何かより良く出来るのではないかと考えたことがきっかけだったと思います

# 開発過程

## 調査

最初は既存技術の調査をしつつ今後の開発方針を考えます。半年位。今回はNNUEのネットワークやパラメータを変えてどう変わるかという調査と、高速なCNN実装について検討しました。

## 最先端へのキャッチアップ

具体的には公開されているillqha/NNUEkai/orqhaに追いつく作業ですが、教師データの生成に時間が掛かることと、思うような結果が出ず、上記検討していた新規実装は見送りになっています。この間あれこれ試した内容がアピールポイントとなっています(概ね不採用でしたが)。

## 新規仕組みの提案

上記状態のため断念しています。パワーポイント版のアピール文書に内容をぼかして(正直云うと少々外して)書いた内容です。評価関数部分で色々考えていますが探索部触りたい病とかが発病しているので今後はそっち方向に行ってしまうかもです。

# 開発内容

## elmo式の学習法

既に差別化要素ではないですが。教師データ生成時の勝敗情報と探索結果の評価値の両方を用いて学習する方式です。当時のアピール文書で簡単な説明と、やねうら王、Aperyで実装例があります。

<http://www2.computer-shogi.org/wcsc27/appeal/elmo/elmo_wcsc27_appeal_r2_0.txt>

<https://github.com/yaneurao/YaneuraOu/blob/master/source/learn/learner.cpp>

<https://github.com/HiraokaTakuya/apery/blob/master/src/usi.cpp>

## 教師データ生成法の変更

既存のDepth指定の学習方法では以下のような課題があります。

* 終盤の複雑な局面になるに連れて、探索ノード数が飛躍的に増加する(局面によって教師データの生成コストが大きく異なる)

特に、決着がつかないような長手数となった場合、散々計算した挙句教師として採用できない結果となることがあり効率が悪いです。現実的にはこれを回避するために評価値3000などで勝ったものと見做す設定をしますが、この場合宣言勝ちを学習することは困難となります。

elmoではtanuki-チームに倣い、Depth指定ではなくノード数指定で生成するようにして教師生成に掛かる時間を平準化しました(元々は時間指定としていましたがノード数の方が計算資源を有効に活用出来る、異なるCPUでも質を均一化出来る、CPU割り当ての不平等を解消出来る等のメリットがありそうです)。ただ、効率面での改善はあるものの相対的には序盤の探索数を増やして、終盤の探索数を減らす形になるので評価関数にもそのような傾向(序盤に強く、終盤に弱い)があるようです。

## 定跡生成法

昨年に引き続きですが以下、2つの手法で生成したものを合成して作成しています。

1. 対局結果からの採用
2. 深く読んだ際の指し手の採用

何れも大会版の評価関数を用いて定跡手を作成していますが1.の局面生成には他評価関数も利用しています。1.は1手1.2～2.4億ノードで対局した結果、評価値の逆転が無いまま勝利したものを採用しています。2.は定跡を意図的に外してくる対策として、depth33程度で広く探索した結果を採用しています。1.と2.で重複した場合は1が優先されます。1.は2年前に急ごしらえで作成した定跡生成法ですが、比較的少ない計算量で実践的な定跡が作成可能であり気に入っています。

# 実験結果

まず、定跡については短時間の思考時間では定跡を利用することで却って勝率が下がってしまうようでした(指しこなせない)。生成時の1手1億ノード以上で対局するのは十分な数が稼げないため、実験対象外としています。また、平手局面から開始したものとやねうら王互角局面集で比較していますが(前者の方が良いだろう)想定と異なり有意な差は確認出来ませんでした。

## vs illqha4 (elmoから見て 勝 – 引分 – 負 )

* 平手局面：　301 - 27 - 295
* 互角局面集：301 - 29 - 291

## vs水匠

* 平手局面：　812 - 59 - 718
* 互角局面集：260 - 15 - 239

## 対局条件

* 100万ノード/手
* 2thread(vs水匠), 4thread(vs illqha4)
* 探索部はdolphin1.01を利用
* 投了値 -999
* 250手で引き分け
* 定跡を利用する場合は12手まで
* 500局以上(水匠戦だけ多いですが特に意味無いです)

対局はAperyの対局ツール(初期版)を独自に書き換えたものを用いています。(以前からですが)一般に公開されているレーティングと乖離が出ていて悩ましいです。投了値のせいなのか、スレッド数のせいなのか。探索ノード数のせいでは無さそうなのですが。

# 追試について

やねうら王はkeyコマンドのレスポンスをhexで返しているのですが、上記対局ツールでは数値で返す必要がありました。