

Qhapaq_conflated の PR 文 for 第五回将棋電王トーナメント

SAWADA Ryoto, ITO Yuki

Mayu's room Shogi Club

September 24, 2017

ソフト名とその由来

「かぱっく こんふれいてっど」と読みます。Qhapaq はケチュア語で”偉大なもの”を指す言葉で、本ソフトが多くの巨人の肩に乗っていることを示しています。第四回将棋電王トーナメントではベスト 8、WCSC27 では 10 位でした。conflate は英語で”融合する”という意味を持ち、本ソフトが様々な人の知恵を融合させたものであることを示しています。

本大会の意気込み

数理と人間の閃きを武器に超貧弱な計算資源でも将棋ソフトを強くできることを示したいと思います。今回は特に、評価関数の学習/合成とその最適化、優勢/劣勢に応じた戦略の動的な調整に力を入れています。

恐らく、四駒には挑戦しないと思います。メモリ高いのよ……

開発者紹介

開発者 1 : SAWADA Ryoto

物理が大好きな某社 (and 某大学) 研究員。物理のプレゼン大会で入賞した経験がある。本大会では学習、定跡、探索の設計と実装を担当。WCSC27では独自の評価関数学習アルゴリズムと定跡生成ルーチンを開発しデスクトップ PC1 台で浮かむ瀬から 170 近くのレート上昇に成功しました。

開発者 2 : ITO Yuki

高速、高精度計算が得意なフレンズ。グラハム数とかにも手を出しているらしい。本大会では Gentoo の導入を始めとした超高速化を担当。前回電王トーナメントでは1日の準備期間で8%前後の高速化と省エネ(3日目参加ソフト内ではダントツの低消費電力)を、WCSC27では10年前レベルの化石OSでやねうら王を動かシクラスタ化するというデスマを達成しました。

電王トーナメントのメタゲーム：elmo 学習と評価関数の合成

elmo 式学習

elmo が WCSC27 で優勝する際に使った学習手法。自己対局の局面に対し深く読んだ際の評価値のみでなく、その対局の勝敗に依存したボーナスを加えることで評価関数の高精度化に成功しています。

評価関数の合成

やねうらお氏が公開している評価関数などと elmo の評価関数を合成（各パラメータを $X:1-X$ で足し合わせる）することで、elmo のレートが 100 前後上昇することが知られています。

やねうら王によるライブラリ化と有志開発の躍進

これまでも、各種自己対戦ツールによって、有志による評価関数や定跡のレート測定が行われてきましたが、やねうら王による elmo 学習のライブラリ化とマニュアルの充実に伴い、有志による評価関数開発という新機軸が生まれつつあります。野生の読み太、まふ関数などが特に知られており、中には elmo よりレート 100 以上強い関数もあります。

有志開発の工夫

基本的に elmo の手法をベースとしていますが、自己対局の初期局面の選び方（アマチュアの棋譜を使う、ランダム手でぐちゃぐちゃにする）、自己対局させる元の評価関数（elmo, qhapaq, 読み太 etc..）、学習させた評価関数に追加で行う合成処理を中心に有志開発者ならではの挑戦的な試みがされているようです。

全探索など許されない

elmo 式学習や合成を全て試すとなると...

手元のデスクトップ PC では 6 手読みのプレイアウトで 12 億局面作るのに一週間弱かかります。50 億なら一ヶ月弱。また、elmo 時代の評価関数は自己対戦に 12 スレ 1 手 1 秒以上の時間が必要です。レート 30 程度の上昇を調べるのには 1000 試合程度必要。評価関数一つを試すのに二日弱。更に、評価関数の他にも、探索部や定跡の生成も必要です。とてもじゃないけどリソースが足りません。

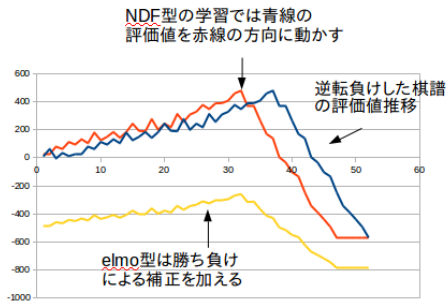
そこで理論予想の出番です！

過去の実験データやより簡単な予備実験から、各種学習法の効果を探り、更に、より少ないコストでより強いソフトを作るための指針を立てることでコンピュータ将棋を重課金ゲーから開放スルノデス.....！

他ゲームの強化学習理論を elmo 学習に導入

各局面が何点であるべきかの推測

左下の図はとあるプレイアウトの評価値と、対局結果を受けて学習部が与える、対局中の各局面は何点であるべきかの推移です。elmo 式は従来型に比べ、頓死した対局に対し、良い/悪い変化を評価値に早く取り入れる効果があります。



・頓死している以上、30-40手目付近の評価値は負であるべきな気がする

・NDF型だと40手目近くの評価値しか大きな変化はしない(疑わしきは罰せず)

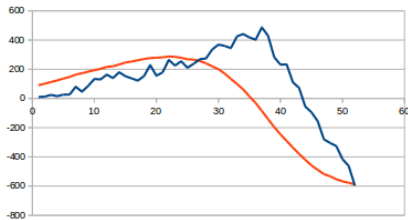
・elmo型では負けた局面は全て評価値を大きく下げる(疑わしきは罰せよ)

・0-30手目ぐらいは本当は悪い手では無い気がするが、その補正は対局回数を増やすことで行う。

他ゲームの強化学習理論を elmo 学習に導入 2

discount を使ったより高度な評価値推測

elmo の学習方法は勝敗に響いていない局面の点数も大幅に下げ（上げ）てしまいます。そこで、<https://arxiv.org/abs/1701.01724>（ポーカーの AI に関する論文）などで使われる discount を使うと、勝敗に響きそうな局面を選択的に変動させることができます。将棋の学習に於いて discount は $V(x) = \sum_{i=0,1,\dots} \frac{\gamma^i v(x)}{1+\gamma+\gamma^2+\dots}$ で表現できます。（ $V(x)$ は x 手目の評価値。 γ は割引率。 γ は 0-1 の値（よく見るのは 0.9 前後）を取る）



割引率を導入することで頓死した対局に対し、「元の評価値（青線）の30-40手目ぐらいの評価値は怪しいけど、0-30手目ぐらいはそこまで悪くないのでは」という人間的な局面の解釈が実現できている

その他、実戦的な微調整

評価関数の評価関数を使った評価関数の合成最適化

$$V(t, n, p) = V_{best}(p) + E(t, p) + S(n, p)$$

$$R(V_1, V_2) = -(E_1 + S_1) + (E_2 + S_2)$$

我々が作る評価関数は、最適値 $V_{best}(p)$ (p は局面、 $V(p)$ はその評価値) と比べ、教師 t のズレ由来の誤差 $E(p)$ と局面数 n の不足由来の誤差 ($S(p)$) を含んでいる。評価関数のレート差は E, S の差によって表されるとした場合、自己対局の結果から、最適な合成比率を予想することが出来ます

この式に基づいて合成するとそれなりの打率で強くなる。... のですが、リソースがあるなら全探索したほうが絶対いい気がします

一応企業秘密ということでは...

- adagrad の時代は終わりました (第二次 elmo-qhapaq 絞り。略して二次工口河童絞り)
- 矢倉は終わってないようです (河童パーク定跡)
- 評価値に応じた探索パラメタの変調を駆使し、序盤中盤終盤で隙がない差し回しを目指します (河童マジック)

スポンサー様のハードの力を最大化します！（解説：Ito）

Gentoo ちゃん、はじめました

今回の大会ではシステム全体のパフォーマンスを稼ぐために OS に Gentoo を採用し、最適化のかかったバイナリを生成しつつ、要らない機能を削ぎ落とします。今回の実行環境は `systemd/genkernel/X11/awesome/` にする予定です。グラフィックがまともに動くことを祈ります。大会前日に行われるであろう `GentooInstallBattle` にご期待ください。

今年は消費電力にもちゃんと触れてくれるんですよね！？

いつものことながらコンパイラには、“対局実行時に”最もパフォーマンスが出そうなものを採用する方針です。一部分野では Clang 4 系のパフォーマンスが異様に高い（icc の 2/3 の実行時間とか）ことが発覚しているためいろいろと楽しみです。

体力が許す限り、対局中であれ最新のコードを外部のサーバーで検証し、最もパフォーマンスが出るコードとコンパイラの組み合わせを探し続け、対局開始時ごとに最もパフォーマンスが出てるバイナリを使用します。

数理の力を思い知らせに行きます

- elmo+150 超えくらいの評価関数は作れそうです
- 定跡、探索でも強くすると同時に、何らかのメッセージ性を残せたら良いなと思います
- Gentoo 楽しいです (きっと)

アピール終了。お付き合いいただきありがとうございました！

web 上で定跡ファイルの一部や将棋ソフトに関する解説記事を書いています

- Qhapaq のブログ <http://qhapaq.hatenablog.com/> 大会で使える定跡ファイルや探索、学習部に関する数学的な解説をしています
- Qhapaq 開発者の twitter https://twitter.com/qhapaq_49 時々将棋に関することをつぶやいています
- Qhapaq の github <https://github.com/qhapaq-49> 今まで作った作品があります

本研究で特にお世話になったサイトです

- やねうら王 <http://yaneuraou.yaneu.com/> やねうらお氏による解説記事です。学習や探索のアルゴリズムに関する解説が豊富で、コンピュータ将棋の教科書代わりに使うことができます
- 将棋フリーソフト rating on single machines
<http://www.uuunuuun.com/> uuunuuun 氏によるコンピュータ将棋のレーティングサイトです。最新ソフト、定跡に関する統計的なデータとして活用できます
- コンピュータ将棋研究ブログ <http://www.fgf7.com/> fgfan 氏による解説記事です。戦型ごとのコンピュータの指し手の特徴に触れています。本研究では量子ガシャの動作確認に用いました

本研究で特にお世話になったサイトです

- Chess Programming Wiki
<https://chessprogramming.wikispaces.com/> チェスのルーチンについて簡単な解説がされています
- LaTeX + Beamer でスライドを作る
<http://xaro.hatenablog.jp/entry/2013/09/18/020615> スライドを作るのに逐一マウスを触りたくない人に役立つ情報が掲載されています
- ソフト開発者の twitter
<https://twitter.com/yaneuraoh/lists/shogidev/members> 誰かが時々有益なことをつぶやいています

and 将棋を楽しんでくれている全ての人に感謝！