

『やねうら王 with お多福ラボ』

磯崎 元洋 (お多福ラボ 技術顧問)
浜道 崇 (お多福ラボ 代表取締役)
和泉 龍徳 (お多福ラボ エンジニア)
森田 康太郎(お多福ラボ アドバイザー)

人間の棋譜を用いない評価関数の学習

やねうら王では、プロの棋譜を用いずに評価関数パラメーターの学習に成功した。これは、科学ジャーナル『Nature』誌に掲載されてもおかしくないレベルの快挙だと言う人もいるようなのだが、当方、ねーちゃんに興味はあっても『Nature』には興味がないので、論文を書く気は毛頭ない。あしからず。

参考記事) 人間の棋譜を用いずに評価関数の学習に成功

<http://yaneuraou.yaneu.com/>

2017/06/12/%E4%BA%BA%E9%96%93%E3%81%AE%E6%A3%8B%E8%AD%9C%E3%82%92%E7%94%A8%E3%81%84%E3%81%9A%E3%81%AB%E8%A9%95%E4%BE%A1%E9%96%A2%E6%95%B0%E3%81%AE%E5%AD%A6%E7%BF%92%E3%81%AB%E6%88%90%E5%8A%9F/

リゼロ評価関数

人間の棋譜を用いない評価関数の学習、すなわち、ゼロから評価関数パラメーターを学習させる試みは、私はいままで何度かチャレンジしてきたが、最新のソフトと肩を並べるまでには至らなかった。

しかしWCSC27(第27回 世界コンピュータ将棋選手権, 2017年5月開催)で優勝したelmoのソースコードが公開されたのを機に、再度挑戦してみた。題して『Re:ゼロから始める評価関数生活』である。(以下、『リゼロ評価関数』と略す。) ちなみに、この「Re:」はリベンジのReである。

このリゼロ評価関数が、WCSC27で優勝したelmoと同等ぐらいの強さになったので、公開した。今回の将棋電王トーナメントのライブラリとしても使える。

リゼロ評価関数epoch8公開しました。

<http://yaneuraou.yaneu.com/>

2017/06/21/%E3%83%AA%E3%82%BC%E3%83%AD%E8%A9%95%E4%BE%A1%E9%96%A2%E6%95%B0epoch8%E5%85%AC%E9%96%8B%E3%81%97%E3%81%BE%E3%81%97%E3%81%9F%E3%80%82/

全自動雑巾絞り機『Shivoray』

評価関数をゼロから学習させられるようになったので、これを誰でもお手軽に体験してもらおうと、実行するだけで教師局面を生成し、そこから学習するのを繰り返す、全自動雑巾絞り機『Shivoray』(シボレー)を公開した。これにより、自分で評価関数を作って遊ぶという新しい将棋ソフトの遊び方が誕生した。

参考記事) 『Shivoray』(全自動雑巾絞り機)公開しました

<http://yaneuraou.yaneu.com/>

[2017/06/28/%E8%A9%95%E4%BE%A1%E9%96%A2%E6%95%B0%E3%81%AE%E3%82%AD%E3%83%A1%E3%83%A9%E5%8C%96%E3%82%B3%E3%83%9E%E3%83%B3%E3%83%89%E5%85%AC%E9%96%8B%E3%81%97%E3%81%BE%E3%81%97%E3%81%9F/](http://yaneuraou.yaneu.com/2017/06/28/%E8%A9%95%E4%BE%A1%E9%96%A2%E6%95%B0%E3%81%AE%E3%82%AD%E3%83%A1%E3%83%A9%E5%8C%96%E3%82%B3%E3%83%9E%E3%83%B3%E3%83%89%E5%85%AC%E9%96%8B%E3%81%97%E3%81%BE%E3%81%97%E3%81%9F/)

評価関数のキメラ化コマンド

二つの評価関数を1:1でブレンドすると、元となった評価関数の特徴を兼ね備えた評価関数が誕生することを発見した。そこで、評価関数をブレンドするためのコマンドを公開し、誰もが手軽に使える状態にした。

参考記事) 評価関数のキメラ化コマンド公開しました

<http://yaneuraou.yaneu.com/>

[2017/06/28/%E8%A9%95%E4%BE%A1%E9%96%A2%E6%95%B0%E3%81%AE%E3%82%AD%E3%83%A1%E3%83%A9%E5%8C%96%E3%82%B3%E3%83%9E%E3%83%B3%E3%83%89%E5%85%AC%E9%96%8B%E3%81%97%E3%81%BE%E3%81%97%E3%81%9F/](http://yaneuraou.yaneu.com/2017/06/28/%E8%A9%95%E4%BE%A1%E9%96%A2%E6%95%B0%E3%81%AE%E3%82%AD%E3%83%A1%E3%83%A9%E5%8C%96%E3%82%B3%E3%83%9E%E3%83%B3%E3%83%89%E5%85%AC%E9%96%8B%E3%81%97%E3%81%BE%E3%81%97%E3%81%9F/)

relmoの世界

評価関数のブレンドコマンドを公開してしばらくすると、評価関数をブレンドすることで、元の評価関数より強くなることが発見された。例えば、elmoとリゼロ評価関数(やねうら王のGitHubで公開していたもの)とをブレンドするとelmoよりR100程度強くなるというのだ。原理はいまだ解明されていない。

つまり、私は意図せずして、elmoより+R100強い評価関数を公開してしまったことになる。今回の電王トーナメントで、これらはライブラリとして使用可能であるから、予選通過レベルが、WCSC27の優勝レベル+R100+ α であっても全然おかしくはない。とんでもないことをしてしまったものである。

参考記事) relmoの世界

<http://yaneuraou.yaneu.com/2017/07/05/relmo%E3%81%AE%E4%B8%96%E7%95%8C/>

KPP_KKPT型評価関数

従来のKPPT(手番つき3駒関係)型の評価関数は、KK(玉と玉の位置関係)、KKP(玉×玉×それ以外の駒の位置関係)、KPP(玉×それ以外の駒×それ以外の駒の位置関係)の3つから構成されている。このうちKPPには手番評価はなくても良いのではないかと思ひ、この手番評価を無くした評価関数を用意した。それがKPP_KKPT型評価関数である。

パラメーターの数が半分程度になったので必要な教師局面の数も半分程度で学習出来るようになり、また、探索時の計算量も減ったため、KPPT型と比べてさほど棋力のダウンはないようである。

また、評価関数のパラメーターが半分程度になったため、以前は動作しなかったメモリの少ないスマホや、RaspberryPiでも動作するようになった。

参考記事) やねうら王、KPP_KKPT型評価関数に対応しました

<http://yaneuraou.yaneu.com/>

2017/09/02/%E3%82%84%E3%81%AD%E3%81%86%E3%82%89%E7%8E%8B%E3%80%81kpp_kkpt%E5%9E%8B%E8%A9%95%E4%BE%A1%E9%96%A2%E6%95%B0%E3%81%AB%E5%AF%BE%E5%BF%9C%E3%81%97%E3%81%BE%E3%81%97%E3%81%9F/

機械学習用のPCにメモリを増設したが…

「現在のKPPT型評価関数が表現の限界を迎えているぞ」という電波を宇宙からキャッチしたため、新しい評価関数を設計した。

その評価関数の名前を仮に『アカシックレコード評価関数』と呼ぶ。この評価関数は、莫大な情報量を持ち、2TB程度の評価関数テーブルが必要で、その学習のために12TB程度のメモリが必要となる。ちなみに鉄腕アトムのアトムの電子頭脳の記憶容量は15兆8000億ビット(おおよそ2TB)らしい。

この時代にはまだ発見されていないはずの未来の12個のテクノロジーを以てすればこれが32GB程度に圧縮が可能である。ところが、その学習のためにメモリが200GBほど必要になる。そこで、昨年から倍ぐらいに値上がりしているというこの状況においてメモリの増設を決めた。

Xeon dual(40c80t)×6台のうち、1台をメモリ32GB×8枚 = 256GBに増設した。残り5台が96GB。後者は自己対局と教師局面の生成のために使う。

ところがこの新しく設計した評価関数、すこぶる弱かった。

評価関数の表現力が劇的に上がったのは素晴らしいことなのだが、学習のための教師局面が従来の300倍ぐらひ必要になる。現時点までですでに電王トーナメントの賞金総額以上にお金を使っている私であるが、さすがに300倍も用意できない。全くもって人類には早すぎた評価関数であった。

探索部の改良でR100伸ばす

WCSC27の前日にやねうら王のGitHubにcommitした探索部の改良でR100程度棋力が向上していたことがその後の調べで判明した。これがWCSC27でのelmo優勝の原動力になったと言われている。

このように探索部には無限の可能性があるように思うし、R100ぐらいまだまだ余裕で伸びるんじゃないかと個人的には思うのだが、二匹目のドジョウはいまだ発見できていない。

独自性をアピールしないと！

やねうら王はオープンソースである。今年は多くの新しい技術を発明し、探索部、評価関数、学習部に改良を施し、将棋ソフト界を牽引してきた。それは、WCSC27でのやねうら王ライブラリ勢の多さ(決勝に残った12チームのうち、8チームがやねうら王ライブラリを使用)が物語っている。そして、WCSC27以降も、上に書いたような多くの技術(将棋ファンのためのおもちゃ?)を提供してきた。

ところが残念なことに、将棋電王トーナメントのルール上、ライブラリは思考部には含まないのだそうだ。上に書いてきたような改良、発明、ツール類はすべてやねうら王プロジェクトのものであり、これは将棋電王トーナメントの使用可能ライブラリに属する。

つまり、これだけたくさんやってきたが、ルール上、現時点ではやねうら王には独自性は一切ないのだ。

ΩΩΩ<な、、なんだってー!

つまりは、このままではルール上、参加すらままならない。さすがにこれは、ルールがおかしいのではないだろうか。この場をお借りして、ルールを作った奴の首根っこをとっ捕まえて一言申したいのだが、そろそろやねうら王は、顔パスでもいいぐらいではなかるうか。というか、やねうら王には、もう参加しなくとも賞金をくれてもいい頃合いなんじゃないだろうか。それか、優勝ソフトが決まったあとに親善試合として3局ほど優勝ソフトとやねうら王が対局して、その対局料として1,620万円ほどくれても誰もおかしいとは言わないんじゃないだろうか。(ちなみに1,620万円とは、竜王戦の敗者側に支払われる対局料)

evaluation function ∇ (nabla)の発見

『アカシックレコード評価関数』は大失敗であった。中二病っぽい名前が良くなかったのであるうか…。256GBのメモリはまったくの無駄であった。

いずれにせよ、KPPT型評価関数はもはや限界近くであり、ちょっとやそつでは強くないという状況であるようだ。KPPT型の評価関数より、表現力を持った評価関数が必要なことは確かであるが、計算量が闇雲に増えるのは良くないようであるし、学習させるための計算資源も足りない。

そこで、KPPT型より軽量でかつ、KPPT型より表現力が少しだけ高いような評価関数なら良いのではなからうか？果たして、そんなことが可能なのであろうか？

可能である。現代ではまだ発見されていない(中略)の7個のテクノロジーを投入すれば容易い。

それが ∇ (ナブラ)評価関数である。この ∇ は、ベクトル解析などで出てくる、あの ∇ であるが、参加チームの皆様は、せいぜい、やねうら王に ∇ (ナブラ)れていただきたい。そういう邪(よこしま)な気持ちを込めて命名した。

ちなみに2017年9月20日現在、まだ1行もコード書けていない。

まとめ

今回の電王トーナメントは、relmoがライブラリとして使える以上、大波乱は必至である。relmo + いま公開されているやねうら王を探索エンジンに用いた場合、(レーティング計算上)WCSC27時点でのponanzaと同等ぐらいであるらしい。

relmo + やねうら王 \approx ponanza(WCSC27) > ponanza(SDT4)

※ SDT4とは昨年の将棋電王トーナメントのこと

であるから、昨年の電王トーナメントの優勝ソフトのponanzaがそのままの強さで今回の大会に出てきた場合、結構な確率で予選落ちするという熾烈な状況である。無論、今回、やねうら王が予選落ちしても全く不思議ではない。