

きのあ将棋の研究

2020 年コンピュータ将棋選手権提出用

作成 2020/03/31 山田元気

概要

現在(2020/03/31)のところ、まだ 2020 年コンピュータ将棋選手権参加のための将棋アルゴリズムの研究に着手できていません。これから急いで進めたいと考えてところです。

時間的余力と体力的かつ精神的余力があれば、実際に研究したことを元に本資料に更新したいです。

研究予定の具体的内容は、「きのあ将棋」と「きのあ囲碁」の共通で動作する独自の仕組みが不十分ながら機能しており、それら共通の仕組みを利用しながら研究を進める予定です。

なお「きのあ将棋」と「きのあ囲碁」における最近の研究は下記を実施。あるいは実施しました。

- ・進化モデル (きのあ囲碁において実施)
- ・いわゆるクラシック方式の抽象発展化 (きのあ将棋、きのあ囲碁において実施)
- ・テーブル参照の連鎖反応 (きのあ将棋でも試しているが特にきのあ囲碁において実施)
- ・ユーザさんの good/bad データ分析 (きのあ将棋、きのあ囲碁において実施)

1. 進化モデル (きのあ囲碁において実施)

自然界の生き残りモデルを参考に、下記の手順で進化モデルを模索した。

- ① 変更前の評価関数とパラメータ A と、パラメータにノイズを加えた B を用意。
- ② A と B を対局させ、勝敗は審判プログラムが判定。
- ③ 対局は複数回くりかえし、A と B から勝率の高い方を採用。
- ④ あらためて、①に戻り繰り返す。

結局のところこの方法は、下記の問題点から研究を断念した。

- ・パラメータ数が少ないスモールモデルであるにも関わらず、収束までに時間がかかった。
- ・収束の行き着くところとして審判プログラムをだます方向に進化することが多発した。

特に「審判プログラムをだます方向に進化」することがしばしばある問題は深刻だった。

たとえば連の数 128 をオーバーするように打ち、審判プログラムを誤作動させるように進化。

対策として連を 256 まで対応すると、それもオーバーするように進化。

あるいは、無駄に石を打ちパスパスを回避し 500 手ほどとしていた制限に逃げるように進化。

制限の手数を増やしても、それに対応するように無駄手をさらに打つように進化。

さらに制限の手数を増やすと、コウなどの形を複数作り対局が終わらないように進化。

問題の発生原因の要素として、系統が 1 つしかない中で進化しようとするので、審判をだます変なクセに一度進化するとそのスポットからなかなか脱出できなくなるのではないかと考えられます。

結局のところこのモデルは、放棄することにしました。

2. いわゆるクラシック方式の抽象発展化 (きのあ将棋、きのあ囲碁において実施)

今のところ 2020 年コンピュータ将棋選手権に、このモデルで参加予定する可能性が最も高いです。

いわゆるクラシックモデルを、評価項目を規格化 & 抽象化。各評価項目のパラメータを機械学習にかけるやりかたをしていて、これを発展させたいと考えています。

具体的には、

- ① 将棋の駒 (囲碁の石) を拡張し、盤面をマスごとに 64~512 パターンなど符号に変換。
- ② 隣接するマスごとに、64 パターンを元に演算を繰り返し、マスごとの評価を算出。
- ③ 集計した評価値を、自己対戦した結果データとの比較で機械学習。
- ④ 最初の①と②において、符号の組合せパターンは学習前にランダム設定し、優秀な結果を採用。

評価する仕組みは、ディープラーニングのように多層化構造を持つが、昔ながらの囲碁伝播ロジックのイメージで作成。

学習方式は順伝播を採用。逆伝播を採用していない。このため収束の効率の悪さを、パラメータの変動を 2 の倍数で行い、どのモデルがよいかの実験を効率化。実際に試す際には 2 の倍数を下げることで精度を担保する工夫を行っている。

どのような組合せが適切なのかは、④の仕組みで偶発的に発見できる仕組み。この評価関数の仕組みを疑似的に機械に構築させる仕組みが弱いながらも「きのあ囲碁」で機能する兆候が見られたため、この研究をおし進めていきたいと考えています。

とはいえ、このモデルは将棋と囲碁においてあわせて長年研究をしているが、最近ではディープラーニングの劣化版に感じられることが多い。

囲碁で例えるとオーラが出ている処理を繰り返すあたりが、ディープラーニングの多層構造と類似していて GPU を使っていないにも関わらず、ディープラーニングに対し優位な点が少ないことが課題と考えている。

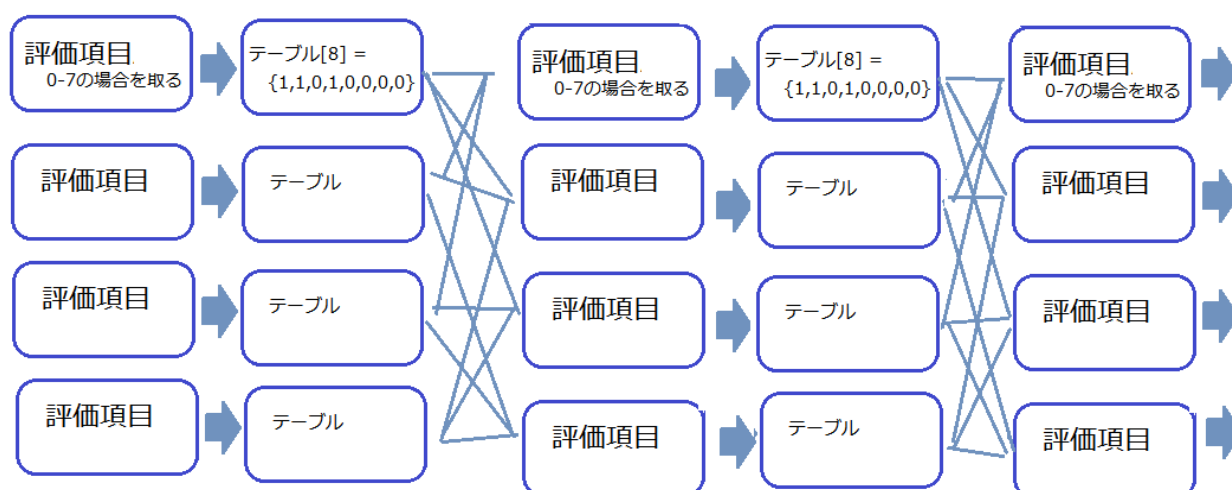
3. テーブル参照の連鎖反応(きのあ将棋でも試しているが特にきのあ囲碁において実施)

可能ならば 2020 年コンピュータ将棋選手権には、このモデルで参加したい。

このモデルも長年何度か挑戦しているものの、他の方法と比べて特に評価性能がでずに苦戦している。が、将来性があると考え断続的に研究を繰り返している。

最近ではディープラーニングに影響を受けて、各評価ユニットを 整数 0or1 のどちらかのみを出力するようにして模索している。以前は 0or1 ではなく、整数 0-7 までのような複数の値を出力し、 8×8 で 64 パターンの配列から、あらためて整数 0-7 を抽出し繰り返すというような方式だった。整数 0or1 にすることで、ある程度結果が安定するようになった。

モデルをシンプルにしたイメージは、下記の通り。



多層構造を持つのは、ディープラーニングと同じと言える。

大きな違いは、活性化関数を持たずに配列の値から 0or1 を各ユニットの値とする点。

また、各ユニットを連結するにはそれぞれのハッシュ値を求める点。

これにより、シナプスモデルとは大きく異なる性質をもつと考えられる。

メリットとして、

→1 層の処理で xor などの複雑な論理演算を出力可能。

→取りうる値が少なく、高い収束性能。

デメリットとして、

→ユニットが出力する値がデジタルなため、漸減的に機械学習が進まず不安定。

→学習進化が不安定なため、再現性が低くアルゴリズム改良確認が難しく、アプローチをしにくい。

→ディープラーニングよりも、内部構造が人間には意味不明であるという点。

4. ユーザさんの good/bad データ分析 (きのあ将棋、きのあ囲碁において実施)

「きのあ将棋」「きのあ囲碁」の共通で動作する、good / bad 投稿データの集計

ケース	囲碁		将棋	
	Good 率	投稿数	Good 率	投稿数
ユーザさんが、勝ちの時	59.6%	225	91.6%	11619
ユーザさんが、負けの時	86.5%	104	75.0%	2985
ユーザさんが、対局中の時	60.2%	226	66.3%	7603
待ったを、利用なし	67.0%	443	80.9%	19237
待ったを、利用あり	56.3%	112	79.7%	3048
手のアドバイス利用なし	未実装	未実装	80.5%	18164
手のアドバイス利用あり	未実装	未実装	84.1%	1476
9 路盤	82.6%	86	なし	なし
13 路盤	87.8%	41	なし	なし
19 路盤	59.1%	428	なし	なし
通常の局面から	63.7%	502	76.5%	14576
特殊の局面から	75.5%	53	87.2%	9424

・将棋、囲碁ともに、「待った」の利用ユーザの good 率が若干低い。

「待った」は、人を若干不幸にするのかもしれない。統計誤差の範囲かもしれない。
あるいは、待ったを使わざるを得ない環境で対局する方が、good 率が低いかもしれない。

・「手のアドバイス」利用で good 率が高い傾向がある。

アドバイス機能は、人に知見をもたらす幸福をもたらすかもしれない。

・初期局面を、特殊局面にすることで、good 率が上がる。

機械学習や自己対局による検証などでは、通常の局面でなく特殊な局面を試すことで機械学習効果の向上や自己対局の信頼性向上とすることがある。

人も同じように特殊局面から対局することで新鮮な気持ちで対局でき楽しい、または棋力の向上に役立つのかもしれない。

・good/bad 投稿データを元に機械学習した思考エンジンを研究開発。

昨年 2019 年の段階で学習を実験してみたが、機能しているのか機能していないのかよくわからなかった。単に研究時間がなかった、データが十分でなかったという問題もあるが、機械学習する前にデータの傾向をよく分析しておけば防げた問題かもしれない。

前ページの good/bad 集計データ期間について、

- 囲碁サイト 2019/12/01 の 21 時ごろ集計 (データ収集開始 2019/10/27 が終わる直前)
- 将棋サイト 2019/12/01 の 23 時ごろ集計 (データ収集開始 2019/02/13 の朝)
- いずれも、匿名として集計。1 人あたり対局開始局面ごとに、1 回まで投稿有効。

きのあ囲碁サイト <https://igo.qinoa.com/ja/>

きのあ将棋サイト <http://syougi.qinoa.com/ja/>

最後に、乱文となりましたことをお詫びいたします。