

# WCSC23 GA 将!!!!!!アピール文書

2013年03月31日

森岡 祐一

## 【GA 将!!!!!!とは】

- 私（森岡）がフルスクラッチで作成したコンピュータ将棋ソフトです。
- 読み方は"がしょう"です。
- Ver.1 では GA（遺伝的アルゴリズム）と言うものを評価関数のパラメータ調整に使っていたので、この様な名前にしました。
- 名前の最後の"!!"はメジャーバージョンを表します。現在は Ver.6 なので 6 個です。
- 元々は"!!"を増やすつもりはなかったのですが、Ver.2 で選手権に参加する際になんとなくノリで増やして以降、年々順調に増えていっています。
- プロ棋士やアマ高段者の棋譜 DB を用いずに、自己対戦の結果から評価関数のパラメータ学習を行うのが最大の特徴です（詳しくは後述）。

## 【コンセプト】

- 開発にあたっての最優先目標は「評価関数の精度向上」です。
- その為の手段として「強化学習による、自己対戦の結果からの評価関数のパラメータ学習」を用います。
- 評価関数と学習ルーチン以外の部分に関しては、不要なものは実装せず、必要なものも極力シンプルなロジックにしています。

## 【評価関数の実装】

- 以下の評価項目を用いた線形の評価関数を使用しています。
  1. 駒割
  2. 筋・段ごとの駒の位置評価
  3. 駒の位置評価
  4. 二つの駒の相対位置関係の評価
  5. 二つの駒の絶対位置関係の評価
  6. 三つの駒の相対位置関係の評価 (3×3のタイルに収まる関係のみ)
  7. 駒の自由度の評価
  8. 王将の移動可能範囲の評価
- パラメータ数は約 220 万個ですので、最近のプログラムにしては少ない方だと思います。
- 後述の通り、強化学習によるパラメータ学習を行なっています。
- 手作業でのパラメータ調整 (いわゆる「ドーピング」) はしていません。と言うか、作者にそこまでの棋力は無いです。

## 【探索ルーチン】

- 基本方針は「シンプル&手抜き」です。
- 基本構成は  $\alpha \beta$  全幅ベース + KFEEnd 流の 2 段階静止探索です。
- 探索深さは、静止探索は 1 段階目を 6 手・2 段階目は無制限で固定。全幅ベースの部分で反復深化を行なっています。
- Aspiration Search とか PVS とか MTD(f) とか、そんな賢いルーチンは使っていません。あくまで、ただの  $\alpha \beta$  です。
- ムーブオーダリングは置換表の手 (1 手のみ) を先頭にし、それ以降を History Heuristic の改良版でソートします。
- 枝刈りはせず、LMR で大雑把に短縮しています。他には、王手・Re-Capture のみ少し延長しています。
- 通常探索の内部ノードでは詰将棋ルーチンを呼ばず、ルートノードでのみ呼んでいます。
- 探索速度は Core i7 940 (2.93GHz) で 1 スレッドあたり 150kNPS 前後、25 分切れ負けだと深さ 4~5 手程度 (全幅部分) までしか読めません。

## 【定跡】

- 棋士の知識は極力使わない方針なので、定跡は学習時・対局時共に一切使用していません。

## 【学習アルゴリズム】

- 強化学習の一手法である **REINFORCE** がベースです。
- 将棋の学習の際には学習ルーチンに探索を組み合わせる必要があると考え、**REINFORCE** を拡張した **PGLeaf** というアルゴリズムを使用します。
- **GA 将!!!!!!**同士で数十万～数百万局ほど対局し、勝った場合の手は正例（正しい手）、負けた場合の手は負例（間違った手）と判断して、評価関数のパラメータを調整します。
- **Bonanza Method** 等とは違い、棋譜 **DB** を一切用いずに学習を行います。
- 本稿執筆時点では単独の評価関数の学習を行なっていますが、将来的には合議を行う前提で複数種類の評価関数を同時に学習する事も検討しています。

## 【学習アルゴリズムの具体例】

- 例えば初手 7 六歩を指して勝った場合は、7 六歩の評価値を上げ、他の手の評価値を下げます。（7 六歩は良い手だとの判断。）
- 具体的には、「7 六の歩の価値」「角の一筋横、二段上に歩がいる場合の価値」等を大きくします。
- 逆に、初手 9 八香で負けた場合は、9 八香の評価値を下げ、他の手の評価値を上げます。（9 八香は悪手だとの判断。）

## 【強化学習のメリット】

- **Bonanza Method** 等の教師あり学習と異なり、棋譜 **DB** を必要としない事。
- そのため、棋譜 **DB** の質・量に学習結果が左右される事が無くなる事。
- 学習に用いる棋譜を自己対戦で生成するので、必要であればいくらかでも数を増やせれる事（ただし、時間/PC の処理能力による制限はあります。）

### 【現状】

- えー、実は本将棋用のパラメータはまだ学習に成功していません。
- 5五将棋ではそこそこ強いのが出来たんですが、同じ設定で本将棋の学習をしようとしても何故か上手く行ってないです。
- という訳で、選手権まであと一ヶ月は必死こいて何とかします。
- …と言うか、何とかしないと本当にマズイですねえ。

### 【最後に】

- ブログ書いてますので、この文書を読んで GA 将!!!!!!に興味を持って頂ければ、ブログの方もよろしくお願いします。

<http://d.hatena.ne.jp/Gasyou/>