

## コンピュータ将棋機械学習における高頻度な駒組み合わせ抽出法の研究 Study on extraction method of high frequent piece combination in computer shogi

後藤 嵩幸<sup>†</sup> 橋本剛<sup>‡</sup>  
Takayuki Goto<sup>†</sup> Tsuyoshi Hashimoto<sup>‡</sup>

### 1. はじめに

コンピュータ将棋では、保木により機械学習を用いた評価関数の自動学習法 (Bonanza Method) が開発され<sup>1)</sup>, 将棋をよく知らない人でも強いプログラムを作れるようになった。現在のコンピュータ将棋の強豪ソフトはほとんど自動学習を利用しており、大会の上位を占めている。コンピュータ将棋自体のレベルも向上しており、将棋電王戦では米長邦雄永世棋聖や現役プロに勝利するなど人間のレベルに迫っている。しかし、中盤から終盤にかけてはプロレベルではあるが、序盤はアマチュア程度の実力しかない。

人間は、上級者ほど多くの駒の配置を塊として認識している<sup>2)</sup>。しかし、コンピュータは 2 駒か 3 駒程度の配置を点数化しているので、特に序盤の細かな違いなどを認識することができない。

高次元な組み合わせ評価を実現するために矢野等は盤面をグラフとして扱い、結合度を用いて評価する手法を提案した。しかし、この手法は計算量が多く実用的とはいえない。そこで、本研究では対局中に現れる頻度を基に評価項目の抽出を目指す。

2 章ではコンピュータ将棋評価関数の現状を説明する。

3 章では矢野等による項目数の増加を抑えた高次元な組み合わせ評価を実現するための研究を紹介する。

4 章では、新しい評価項目抽出法についてのアイデアを提案する。重要な組み合わせほど対局中に現れる頻度は高いと考えられるため、対局中に現れる各組み合わせの頻度を調べる。しかし、全ての組み合わせをカウントすると 3 駒以上ではメモリが足りなくなるため、ランダムに選んだ組み合わせをカウントした。実際にその考えに基づいてカウントを行った結果を示す。

5 章では、現在の課題を述べると共に今後の展望を示す。

### 2. コンピュータ将棋の評価

以前のコンピュータ将棋では、評価値は開発者が決めていた。そのため、多くの項目を評価することはできず、主に駒の位置のみを評価していた。

<sup>†</sup> 松江工業高等専門学校 電子情報システム工学専攻  
Advanced Engineering Faculty of Electronic and Information  
Systems, Matsue College of Technology

<sup>‡</sup> 松江工業高等専門学校 情報工学科  
Department of Information Engineering, Matsue  
College of Technology

評価関数の調整に機械学習が導入されたことがきっかけとなり、多くの項目を調整することが可能になった。その結果、駒の組み合わせを評価できるようになり、コンピュータ将棋のレベルは格段に向上した。しかし、駒の組み合わせ数を増やすと特徴数は爆発的に増える。例えば、GPS 将棋は評価項目として 2 駒の組み合わせ評価に加えて、駒の自由度や利きによる評価を加えており、その総数は約 300 万もある<sup>3)</sup>。Bonanza は駒割りと王を含む 3 駒の配置、約 9300 万個項目を評価している。盤面を評価するたびにこれらの項目全てを足し合わせるため非常に時間がかかる。また、学習にも膨大な時間を費やす。

しかし、対局中に現れる盤面は限られており、すべての組み合わせを評価値として持つ必要はない。よって、重要な項目のみを評価値として保持するようにすれば項目数の削減及び計算時間の短縮につながるため高次元な組み合わせ評価も実現できると考えられる。

### 3. 関連研究

矢野等は盤面を駒や升によって構成されるグラフと仮定し、各エッジに結合度を定義することにより組み合わせの絞り込みを行った<sup>4)</sup>。2 駒間の結合度を学習により定義し、それらの積を複数駒の結合度として、重要な組み合わせを抽出した。この手法は、理論的には特徴数の増加を防ぎつつ、高次元な組み合わせ評価を行うことが可能となる。しかし、実際は特徴の計算や盤面のグラフへの変換に時間がかかるため、既存の手法に比べ学習時間が爆発的に増加する。そのため、実用的とはいえない。

### 4. 評価項目の自動抽出

#### 4.1 駒組み合わせのカウント

本研究では、重要な組み合わせを選別するために、各組み合わせの対局中に現れる頻度を調べた。その際、全組み合わせをカウントすることは不可能であるため、カウントする駒の組み合わせは対局中の局面からランダムに決定する。

一見、ランダムに組み合わせを選んだ場合、高頻度な組み合わせを正しく抽出できないように思える。しかし、ランダムに選択する手法は確率的勾配降下法 (SGD 法)<sup>5)</sup>と呼ばれる最適化アルゴリズムでも用いられており、その有用性が実証されている。SGD 法とは自然言語処理など大規模データを扱う分野で用いられている手法で、目的関数の最適化の際に教師データ全てを扱うのではなく 1 つのみを扱う。この際使うデータをランダムに選択しても試行回数を増やすことにより通常の勾配法と同様な収束性を期待できる。

図1にランダムに選ぶイメージ図を示す。

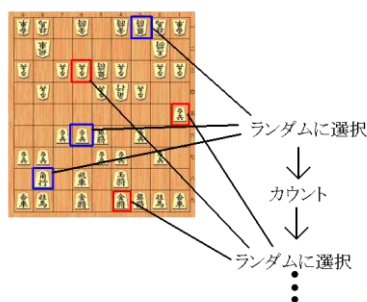


図1. ランダム選択のイメージ図

## 4.2 検証

4.1であげたアイディアに基づいて3駒,4駒,5駒の組み合わせのカウントを行った。カウントに用いた棋譜はBonanzaの学習で利用されているプロ棋士の対局約50000局分で、組み合わせの記憶用に用意したハッシュのサイズは8192である。カウントは各局面20回行った。

予備実験として1駒及び2駒組み合わせのカウントを行った。全ての組み合わせとランダムに選んだ組み合わせの両方をカウントした結果、両者とも同様な結果が得られた。

図2はランダムに3駒選んだ組み合わせのカウント数分布を表している。縦軸がカウント数で横軸は選ばれた組み合わせのインデックスである。

表1, 表2, 表3は結果から3箇所を抽出したものであり、それぞれ図2のグラフの①②③に対応する。表中に「v駒名」とあるものは後手の駒を表す。また、今回は王はカウントの対象としていない。

表1はカウント数が一番多い組み合わせである。香車や桂馬など対局中ほとんど初期位置から動かない駒が現れている。このことから単純に高頻度な組み合わせを評価項目としても、重要な特徴をうまく抽出できない事が分かる。

表2はカウント数が平均値周辺の組み合わせである。この辺りの組み合わせは囲いの組み合わせや守りに重要な駒配置などが見られる。評価項目としてはこの辺りが重要な組み合わせだと思われる。

表3は非常にカウント数が少ない組み合わせである。これらの組み合わせは、対局中にたまたま現れたもので、ゲームの優劣にほとんど影響を与えないものと考えられる。

4駒,5駒の場合も3駒の時と同様な傾向となった。今回の結果から評価項目として重要な組み合わせは現れる頻度の高い①のグループではなく中間辺りの②のグループであるといえる。

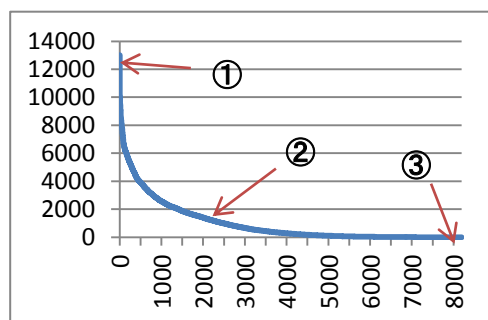


図2. カウント数の分布

表1

| 1駒目    | 2駒目    | 3駒目    | カウント数 |
|--------|--------|--------|-------|
| 91 v 香 | 99 香   | 19 香   | 13003 |
| 11 v 香 | 99 香   | 91 v 香 | 12668 |
| 21 v 桂 | 19 香   | 11 v 香 | 10666 |
| 11 v 香 | 21 v 桂 | 91 v 香 | 10246 |

表2

| 1駒目    | 2駒目    | 3駒目    | カウント数 |
|--------|--------|--------|-------|
| 99 香   | 88 角   | 39 銀   | 1093  |
| 77 銀   | 36 歩   | 78 金   | 1093  |
| 87 歩   | 53 v 歩 | 44 v 歩 | 1093  |
| 21 v 桂 | 85 v 歩 | 57 歩   | 1092  |

表3

| 1駒目    | 2駒目    | 3駒目    | カウント数 |
|--------|--------|--------|-------|
| 37 v 馬 | 46 金   | 24 v 歩 | 2     |
| 82 v 馬 | 35 飛   | 77 銀   | 2     |
| 47 歩   | 45 v 角 | 88 銀   | 1     |
| 88 銀   | 39 銀   | 65 歩   | 1     |

## 5. 今後の課題

現在は、組み合わせを局面から選ぶ際に全ての駒の中から完全ランダムで組み合わせを選んでいく。そのため、動かない駒が上位を占める結果となったと考えられる。今後は、より関連の高い組み合わせが選ばれやすくなるような工夫も必要と考えられる。

また、現在は駒のカウントまでは終えているが、その結果から評価項目の抽出及び学習までは行っていない。今後はこれにより抽出された組み合わせを基にBonanzaを用いて学習を行い、抽出された評価項目がどの程度有効かを調べて行きたい。

### 参考文献

- [1] 保木邦仁, “局面評価の学習を目指した探索結果の最適制御”, 第11回ゲームプログラミングワークショップ, (2006)
- [2] 伊藤毅志, 松原仁, ライエル・グリーンベルゲン, “将棋の認知科学研究(1) 記憶実験からの考察”, 情報処理学会論文誌, Vol.43, No.10, (2002)
- [3] 金子知適, 山口和紀, “将棋の棋譜を利用した大規模な評価関数の学習”, 情報処理学会論文誌, Vol.51, No.12, (2010)
- [4] 矢野友貴, 三輪誠, 横山大作, 近山隆, “ゲーム構成要素を組み合わせた特徴の最適化”, 第15回ゲームプログラミングワークショップ, Vol.2010, No.15, (2010)
- [5] Tong Zhang, “Solving Large Scale Linear Prediction Problems Using Stochastic Gradient Descent Algorithms.”, In ICML '04: Proceedings of the twenty-first international conference on Machine learning, (2004)