ゲーム木の分数手探索を用いた 局面を複雑化する勝負手の生成法

大槻 知史　　合原 一幸

東京大学大学院新領域創成科学研究科

チェスライクな2人ゲームにおいて人間のプレイヤーはしばしば「勝負手」といわれる指し手を選択することはよく知られている。「勝負手」という用語は、一般には「少ない可能性ではあっても逆転を狙う手」といった主観的かつ曖昧な意味で用いられることが多い。これに対し本稿では、どのような探索範囲における指し手を選択するかという観点から「勝負手」の定義を与える。さらに、ゲーム木の探索範囲を柔軟に変形できる分数手探索のモデル化を行い、これを用いた将棋における具体的な「勝負手」の生成法に言及する。本手法による勝負手探索実験により、不利な局面において、通常の深さ打ち切り探索に比べて局面を回復する指し手をより多く生成し、局面をより複雑化することを、統計的観点から確認した。

Game-Tree Search with Flexible Control and Its Application to Generate 'Syoubu-te'

Tomoshi Otsuki　　Kazuyuki Aihara

School of Frontier Sciences, The University of Tokyo

In 2-player chess-like games, human players sometimes select a kind of 'Syoubu-te'. This term generally means 'a move that might recover a drastically better position with a very little possibility', which seems subjective and vague. We first propose a new definition of 'Syoubu-te' from the viewpoint of search control. This paper next gives the model of searching game trees with flexible control, which realizes the way of generating 'Syoubu-te' and other kind of moves. Experimental results with shogi have stastically illustrated that our 'Syoubu-te' is more effective and makes positions more complicated than best moves by ordinary search.

1 はじめに

チェスライクな2人ゲームのゲーム木の探索を行う場合、ある程度以上複雑なゲームにおいては、終局が近付いたノード以外では全探索を行うことは現実的に不可能であるため、通常はある段階で探索を打ち切らざるを得ない。そのため、いかにして重要なノードをより深く探索するかということがゲーム木探索の一つの主要な課題である。これに対し最善手の0.5手延長アルゴリズム [1] や手筋の利用 [2] 等の改良はよく知られている。一方、実現確率打ち切り法 [3] では、指し手のカテゴリ分けを行い各々のカテゴリに遷移確率の値を付与することで、よりありそうな局面をより深く探索できることが知られている。

またより一般的な枠組としては、指し手のカテゴリ分けに基づくSearch Control[4] を用いて探索範囲を柔軟に変化させる手法が知られている。本稿ではまず、このSearch Controlを用いた分数手探索モデルを与える。

この分数手探索では探索範囲 (スコープ) を状況に応じて変化させることができるため、例えばチェス・将棋等において従来は困難であった「勝負手」、「駒落ち時の指し手」、あるいは「相手を指導するための指し手」といった通常の最善手以外の指し手の生成が可能となることが期待される。

この内「勝負手」に関する研究としては、「不利な局面における指し手のまずさ」をコンピュータによる探索の弱点であると指摘し、勝負手のパターンを分析したもの [5] が知られているが、具体的な生成法には

言及されていない。そこで、本稿では分数手探索を用いたコンピュータ将棋における「勝負手」の具体的な生成法を提案する。本手法による勝負手探索実験の結果、通常の深さ打ち切り探索に比べ、不利な局面においては局面を回復する指し手をより多く生成し、また統計的には局面をより複雑化することが確認された。

2 分数手探索

本章では、Bjornsson et.al. による深さパラメータを用いた Search Control の手法 [4] を用いて、分数手探索のモデル化を行う。

2.1 スコープ

限られた時間の中で、あるノード以下のゲーム木の探索を行い、より良い指し手を選択したい場合、主に以下の 3 要素が重要である。

1. 末端の静的評価関数
2. 各ノードにおける前向き枝刈り
3. 探索打ち切り条件 (深さ閾値) の調整

これらを探索の 3 要素と呼ぶことにする。

ここで、「1, 末端の静的評価関数」及び「2, 各ノードにおける前向き枝刈り」の 2 要素はゲームの性質に大きく依存する部分であり、また $\alpha\beta$ ベースの探索を考える場合には「3, 探索打ち切り条件」とは独立した要素であると捉えられのでここでは取り上げない。

「3, 探索打ち切り条件の調整」に関しては、例えば将棋の場合、最善手の 0.5 手延長アルゴリズム [1] や手筋の利用 [2] 等によって、重要な手順をより深く読む手法はよく知られている。また、実現確率打ち切り法 [3] も探索打ち切り条件を調整していると考えられる。この探索打ち切り条件をより一般化したスコープの概念を以下のように与える。

def. スコープ: ルートノード以下のゲーム木全体の中で、実際に探索範囲に含まれるノードの集合

例えば、深さ 5 手打ちきりの場合、ルート局面から 5 手以内で到達できる全てのノードがスコープに含まれる。このスコープを調整して、より重要そうなノードをより深く、そうでないノードを比較的浅く探索することが、制限時間のある実際のゲーム木探索においては非常に重要である。従来の深さ打ち切り法の場合、ゲーム木の各枝に相当する深さは 1.0 手であり、重要と思われるカテゴリの指し手に関しては深さの和の打ち切り閾値を延長することによりスコープを調整していると考えられる。

2.2 深さパラメータを用いたスコープの調整

分数手探索では、各 (カテゴリ i に属する) 指し手に対して分数手の深さパラメータ w_i を与え、この深さパラメータの和を閾値とする打ちきり条件を考える。この w_i を 1.0 一定としたのが深さ打ちきり法であり、逆に分数手探索では打ちきり閾値を固定し、深さパラメータをカテゴリごとに増減することによりスコープを調整できる。

例えば図 1 のようにルート局面 A からの探索を考える場合、単純な深さ 4 手の打ち切り条件では、点線の範囲の 4 手以内で到達できるノードが探索される。一方、深さパラメータの和 4.0 を閾値とする分数手探索では、深さパラメータが大きい指し手を含む $A-B-C-D$ の経路では、ルート局面から 3 手後の時点でルート局面からの深さ閾値の和が 4.0(1.0+2.0+1.0) に到達するために、探索は打ち切られる。これに対し、深さパラメータが小さい指し手を含む $A-E-F-G-H-I-J$ の経路では、ルート局面からの 6 手後に初めて深さパラメータの和が 4.0(0.5+1.0+0.5+0.5+1.0+0.5) になり、探索が打ち切られる。すなわち、より深く探索したいカテゴリ i に関しては、深さパラメータ w_i を小さくすることで、カテゴリ i の指し手を含む手順の探索手数は相対的に長くなる。

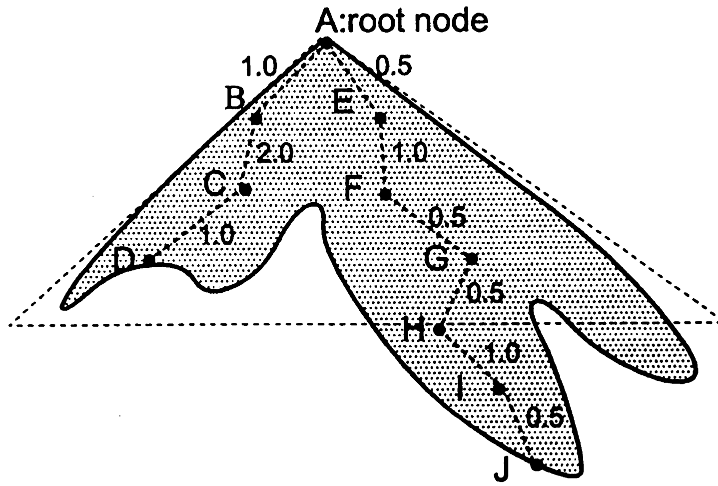

図 1: 深さ打ち切り探索のスコープ (図の点線内) と分数手探索のスコープ (図の網部) の比較。数字は各指し手の深さパラメータ。

2.3 分数手探索のモデル化

以上から、ノード r 以下の分数手探索モデル M は次の変数及び関数により、$M = (K_r, C, \mathbf{Scope}(W, d), \mathbf{Prune}, \mathbf{Score})$ により与えられる。

- K_r: ノード r 以下のゲーム木全体。
- C: 指し手のカテゴリの集合。$C = \{1, 2, \cdots\}$
- W: 深さパラメータの集合。$W = \{w_i\}$: w_i はカテゴリ i の深さパラメータ。
- $\mathbf{Scope}(W, d) \in K_r$: 深さパラメータ W から得られるスコープ (閾値: d)。
- $\mathbf{Prune}(n)$: ノード n の全ての枝の中から、次に展開する枝を選択する関数。
- $\mathbf{Score}(n)$: ノード n の静的評価値を与える関数。

分数手探索では、深さパラメータ w_i が得られれば自動的にスコープが得られ、打ち切り条件が定まるので、従来通り $\alpha\beta$ 法による探索アルゴリズムを適用できる。また深さパラメータの和の閾値を漸時増加することにより、通常と同様の反復深化を行うことができるなど、通常の $\alpha\beta$ 法の改良手法をほぼそのままの形で用いることができる。

2.4 深さパラメータの値

適当なカテゴリ分けを行った上で、適当な深さパラメータ w_i を与える手法としては、山登り法を用いたものが知られている [4] が、カテゴリの数が増えると困難である。また鶴岡らが指摘しているように、実現確率打ち切りの手法は深さパラメータを $w_i = -\log p_i$(p_i はカテゴリ i の実現確率) とする近似であると解釈できる [3]。

ここでは深さ重みを状況に応じて調節できる、より一般的なモデルを与えているが、どのようにして適当なカテゴリ分けを行い、さらにどのようにして各カテゴリの深さパラメータを与えるかは今後の重要な課題の一つである。

3　勝負手の定義

一般に、不利な局面において逆転を狙う勝負手の生成を考える場合、相手の探索モデル M_o を考慮することは、相手モデルの既知・未知にかかわらず有効であると考えられる。しかし本稿では、前に挙げた探索の3要素のうち、相手モデルにおける「1, 末端の静的評価関数」及び「2, 各ノードにおける前向き枝刈り」の違いを考慮する勝負手は、相手のミスを期待する消極的な勝負手として、ここでは考えないことにする。

def. 消極的な勝負手: 味方の探索モデルより劣る相手の探索モデルの Score,Prune を考慮に入れた (期待する) 指し手

「勝負手」の定義としては、松原らによる「不利な時あるいは評価値が減少傾向にある時に作為的に最善手を避けて指す代替手」としたものが知られている [5] が、ここでは、ゲーム理論上の「真の最善手」と「あるスコープにおける最善手」を明確に区別した上で定義を与える。
まず真の最善手は以下のように定義される。

def. 真の最善手: Scope をルートノード以下のゲーム木全体とした最善手

真の最善手は多くの場合、当該ノード以下のゲーム木全体を探索しない限り発見できないため、実際上は限られたスコープにおける探索によるみかけの最善手を最善手とみなす。

def. みかけの最善手: 通常の深さパラメータ W から得られる Scope(W) における最善手

これに対し、勝負手の定義を以下のように与える。

def. 勝負手: 通常とは異なる深さパラメータ W' から得られる Scope(W') における最善手

つまり我々の定義する勝負手は、単なる相手のミスを期待するはめ手という意味ではなく、通常探索による「みかけの最善手」を上回る「真の最善手」となる可能性も含んだ意味で用いる。どのような W' を採用すべきかは状況により異なると考えられる。具体例の一つを次章以下で述べる。

4　将棋における勝負手の生成法

本章では、分数手探索を用いたコンピュータ将棋における具体的な勝負手の生成法を提案する。
コンピュータ将棋プログラム同士の対戦はジリ貧負けになりやすいことが指摘されている [5] が、その原因の一つとして、有利側・不利側双方が似たようなスコープの探索を行なっていることが挙げられる。探索時間が限られている場合に、不利側がより不利になるような手順を数多く探索することは、しばしば無駄であると思われる。例えばある状況においては、味方が攻められる展開には時間を割かず、相手玉を攻める手をより深く読むことが逆転につながる場合もあり得る。
以上のような状況に応じた選択的な探索を実現することは従来の探索では困難であった。それは従来よく用いられる探索法の多くがみかけの最善手のみを得ることを目的としているからである。分数手探索においては、深さパラメータを増減することでスコープを調節し、異なるスコープにおける指し手を得ることができる。
例えば、通常のスコープの深さパラメータが $w_i \in W$ であるときに、勝負手を探索したい場合は、深さパラメータ $w_i' \in W'$ を以下のように定めることを考える。

$$w_i' = \begin{cases} aw_i \quad (0 < a < 1) & \text{if 指し手が勝負手条件を満たす} \\ w_i & \text{if 指し手が勝負手条件を満たさない} \end{cases}$$

勝負手条件: 味方番の敵玉から3マス以内に進める手及び相手番の自玉から3マス以内に進める手

この W' から得られるスコープによる探索では、W の場合に比べ、相手玉周辺の指し手に関してはより長い手数探索され、それ以外の指し手に関しては相対的に探索手数が短くなる。このような探索により、従来の探索ではスコープに入らない手順を発見し、上で述べたような逆転の「勝負手」が生成されることが期待される。以下この W' のスコープによる探索のことを「勝負手探索」と呼ぶ。

5 勝負手生成実験

本章では、前章の勝負手条件を仮定した下での勝負手探索と、通常の深さ打ち切り探索のそれぞれによる先手の指し手の生成実験を行い、得られた指し手の評価を2手後のスコアに関する統計的な観点から行う。なお、後手番は常に通常探索を行うとする。

5.1 実験の手順

- テスト局面は人間同士の対局の75手目の先手番局面とする
- 深さパラメータ以外は同一のプログラムを用いて先手・後手共に一手づつ進める
- 後手番の探索では深さパラメータを全て1とした深さ打ち切り探索を行う
- 先手番の探索では勝負手条件を満たす場合に深さパラメータを $a(\leq 1)$、満たさない場合は1とする
- 先手・後手の思考時間はそれぞれ3[sec] 打ち切りとし、$\alpha\beta$ 法ベースの探索を行う
- 最初のテスト局面と2手進めた後の先手番局面のスコアの差分を得る

5.2 実験条件

- 前節の深さパラメータ a 値に関して、1,1/2,1/3,1/4の4種類を比較する
- テスト局面が先手不利な場合 (スコアが-1000未満) と不利でない場合 (-1000以上) を比較する
- 以上の各条件につき、それぞれ1000[個] のテスト局面を用い、それらの統計値を得る

5.3 スコアに関して

各局面のスコアは当該局面において15[sec] の探索を行った場合の最善応手の動的評価値を採用した。このスコアは局面の優劣を完全に正確には反映していないと考えられるが、この不確実さを軽減するために、スコアの算出の際には、指し手生成に与えた考慮時間の5倍である15[sec] の探索による評価値を採用し、また統計値としては1000局のテスト局面の平均値を採用した。

ここで探索の末端における静的評価関数 Score(n) は、局面 n が先手有利ならば正、後手有利なら負の整数値を与え、各駒の価値・各駒の玉との相対位置・玉の安全度、駒のはたらき等を考慮している。例えば、先手の歩の標準価値は100である。

6 結果及び考察

表1は、勝負手探索と通常の深さ打ち切り探索のそれぞれに関して、2手進めた後と進める前のスコアの差分の統計値をまとめたものである。$a=1$ の場合は、全ての指し手が同じ深さパラメータになるので深さ打ち切り探索となり、以下 a の値が減少するにつれて、より極端に相手玉周辺の指し手を重点的に展開するようなスコープによる探索が行われる。いずれの場合も探索時間を3[sec] 一定としているので、制限時間内に探索される合計ノード数はほぼ一定であると考えられる。スコアの密度分布は正規分布に近い分布になるという実験結果が得られたため、表1には密度分布の平均及び標準偏差の値を示した。また、1000局面のうちスコアの差分が+500以上及び+1000以上になる頻度も示した。

まず、いずれの探索の場合にも、不利な局面の場合にはスコアが減少し、標準偏差の値は不利でない局面よりも小さくなる傾向がみられた。これは一度不利になるとジリ貧負けになりやすいコンピュータプログラムの傾向を裏付けていると考えられる。また平均値のみをみると、微差ながら $a=1$ の通常探索の場合が最良の結果となった。

表 1: 深さパラメータ a に対するスコアの差分の統計値。A: +500 以上の頻度, B: +1000 以上の頻度。

テスト局面の形勢	a の値	平均	標準偏差	A [度数]	B[度数]
不利でない場合 (スコア-1000 以上) 計 1000[局面]	1	39	617	143	44
	1/2	35	623	162	56
	1/3	-30	644	140	39
	1/4	-18	642	147	48
不利な場合 (スコア-1000 未満) 計 1000[局面]	1	-50	558	118	37
	1/2	-55	614	131	45
	1/3	-40	629	136	39
	1/4	-81	617	123	37

次に $a \neq 1$ の勝負手探索の場合に着目すると、スコアが 2 手後に+500 以上、+1000 以上になる頻度では、$a = 1$ の場合に比べ多くの場合優っているが、これは標準偏差が増大した効果であり、不利になる頻度もより多くなっている。また不利な場合には、$a = 1$ の場合に比べスコアの平均減少量はそれほど変わらないにもかかわらず、標準偏差がより増大する傾向がみられた。これは、本手法により生成した指し手が局面を複雑化した結果であると解釈できる。このような傾向は、不利でない場合には顕著ではない。

以上のような統計的な性質は「不利な局面において、逆転の可能性は増やすもののリスクも大きい」といった一般に用いられる勝負手の性質と一致する点が多く、本稿で生成した指し手が「局面を複雑化する勝負手」であることを支持していると考えられる。

今回の実験の範囲では、不利でない場合は $a = 1/2$ の場合の結果が最も優れており、不利な場合は $a = 1/3$ の場合の結果が最も優れていると考えられる。テスト局面の形勢により、適当なスコープが異なるという点は非常に興味深いが、常に最適なスコープを与えるような、状況に応じた深さパラメータを与えることは今後の重要な課題であると思われる。

7 おわりに

本稿では、ゲーム木探索の概念を一般化した分数手探索のモデル化を行った。またスコープの観点から勝負手の定義を再考し、分数手探索を用いた勝負手の生成法を提案した。本手法により生成した指し手は、勝負手の目的の一つである「局面の複雑化」を少なくとも統計的には実現することを確認した。

しかし、今回仮定した勝負手条件は非常に単純なものであり改善の余地があると思われる。また本稿では通常探索を深さ打ち切りとして実験を行ったが、通常探索をナイーブな深さ打ち切り以外に置き換えて検証する必要がある。

一方、本勝負手探索手法では味方が攻められるような指し手をあまり深く探索しないために、かえって負けを早める場合も存在すると思われる。どのような局面で勝負手探索を行うかを判断することは今後の課題である。

参考文献

[1] 山下宏：YSS -そのデータ構造, およびアルゴリズムについて, コンピュータ将棋の進歩 2 pp.112-142, 1998.

[2] 棚瀬寧：IS 将棋のアルゴリズム，コンピュータ将棋の進歩 3 pp.1-14, 2000.

[3] 鶴岡慶雅, 横山大作, 丸山孝志, 近山隆: 局面の実現確率に基づくゲーム木探索アルゴリズム, ゲームプログラミングワークショップ pp.17-24, 2001.

[4] Y Bjornsson and TA Marsland. Learning search control in adversary games. In Advances in Computer Games 9, Editors van den Herik and Monien. University of Maastricht: pp.147-164, 2000.

[5] 松原仁, 飯田弘之: ゲーム・プレイイングにおける勝負手, 情報処理学会人工知能研究会 99-15, pp.111-118, 1995.