

知識指向型オセロゲームプログラム OTL, PLG†

田 島 守 彦††

知識指向型のオセロゲームプログラム OTL, PLG を作成した。通常の α - β 法を代表とする探索型の接近法に代え、多様な知識を主力とし必要に応じて探索を行う接近法による。人の柔軟な思考過程に類似した動作を行わせることで、より本質的な知能の研究を行うことを目的とする。基本的な概念として「変化に関する知識」を導入し、これを用いて候補手による局面の変化を見積もることで候補手の評価を行う手法を開発した。また探索の制御を行うために必要な知識を用意した。これらの知識は Prolog によって具体的に記述され利用される。結果的に無駄の少ない大変効率的な手の評価が可能になり、非常に少ない探索ですむにもかかわらず、上級者に匹敵する力をもつことが確認された。また計算内容の解析を行い、プレイの特徴を観察し、ミニマクス法との比較を行った。これらの結果から、本手法に基づく知識指向型の接近法の利点を明らかにした。

1. ま え が き

人工知能研究のテーマとして、ゲームは最も早くから採りあげられてきたものの一つである。チェスゲームを中心に多くの研究が、とくに探索法や評価関数の分野で行われてきた^{7),8)}。バックギャモンでは世界チャンピオンに匹敵する力をもつものもある⁹⁾。

本論文で取り扱うオセロゲームは19世紀の盤ゲーム、リバーシを現代的に再生したものである。その歴史の浅さにもかかわらず、規則が単純で取扱いやすく見えることから多くの計算機プログラムが作られている。筆者らも α - β 探索に基づくプログラムを作成している⁹⁾。最近アメリカでは計算機によるオセロゲームのトーナメントが行われている。1981年に行われた Santa Cruz Open Machine Othello Tournament²⁾ において優勝した IAGO⁶⁾ は現在までに開発されてきた種々の探索および評価関数に関する技法を集大成したプログラムで、チャンピオン水準の力をもっている。

これらの成功しているプログラムの強さはおもに次の2点によっている。すなわち、いくつかのあまり複雑でないサブ評価関数とそれらを統合する評価関数、および工夫をこらした α - β 探索⁵⁾ である。そこでは知識の役割は大きくない代りに、膨大な探索を行うのが特徴である。人の場合とは強さの原因が根本的に異なっており、知識の獲得、推論を含む非常に柔軟な知識の研究には直接参考にならない。

筆者は知識指向型のオセロゲームプログラム OTL, PLG を開発しており、このほどその第1版を完成し

た。これは中盤のプレイに特徴がある。手の評価は通常の α - β 探索によらず、知識を主力とし必要などきのみ探索を駆動する方式による。こうすることで人の思考に近い柔軟な動作を可能にしている。

知識の記述には論理型言語 Prolog を用いた。最近注目を集めている論理型プログラム⁴⁾ のスタイルは、ゲームプログラムの記述にも非常に有用であった。

2. 中・終盤の知識

OTL, PLG は中盤から終盤初期の知識指向型プログラムに特徴がある。OTL, PLG の使用する、プレイに必要な各種知識について順に説明する。以下、オセロゲームに関する用語は文献³⁾ に従う。盤上のマスの呼称を図1に示す。その他、次の用語を用いる。

可能手数 打着可能な手の数

序盤 おおむね、辺上の戦いが起きる以前。

中盤 辺での戦いを含む手数をめぐる戦いの段階。

終盤 隅や星をめぐる攻防がなされる時点以降。

権利手 打着者側には打着可だが敵側には打着不可のマスがある場合、そこへの手。

2.1 中盤の目標

オセロゲームの中盤では、隅を取ることに高い価値がある。隅を取るためには敵に悪手を打たせる必要がある。よって中盤の目標は敵を手詰りに追い込むことである。ゆえに敵の可能手数は少なく、自分の可能手数は多くなるようにするのが中盤の戦略になる。

2.2 知識設計の方針

オセロの中盤は打着可能手数が最大になり⁹⁾、また各可能手について調べるべき内容もその場面に依じてさまざまなので、その都度必要な知識を用いて必要な判断を行うようなプログラムが望ましい。OTL, PLG の知識集合は次のような方針のもとに構成された。

† OTL, PLG—A Knowledge Oriented Program of the Game of Othello by MORIHIKO TAJIMA (Logical Systems Section, Automatic Control Division, Electrotechnical Laboratory).

†† 電子技術総合研究所制御部論理システム研究室

| | | | | | | | | |
|---|---|---|----|----|----|---|---|---|
| | a | b | c | d | e | f | g | h |
| 1 | 隅 | | | 辺 | | | | 隅 |
| 2 | | 星 | | 中辺 | | | | 星 |
| 3 | | | 箱隅 | | | | | |
| 4 | | | | 箱 | | | | |
| 5 | | | | | 中原 | | | |
| 6 | | | | | | | | |
| 7 | | 星 | ⊙ | a | a | ⊙ | 星 | |
| 8 | 隅 | C | A | B | B | A | C | 隅 |

図 1 盤上の用語

Fig. 1 Terminology on the board.

(1) 知識を一貫した形式で書くために言語として Prolog を採用する。

(2) 候補手の評価を効率的に行う。計算は真に必要な意味があると予測できる要素に限って進めてゆく。

(3) 評価はいろいろな観点からいろいろな知識を用いて行う。評価の打ち切り、探索の起動に関する知識も用いる。

2.3 変化に関する知識

人の思考過程と同様な知識指向型の接近法を採用するにあたり、前節(2)の要請に基づき次のような性質の知識を中心とする。

[定義 1] 変化に関する知識*

局面 p^0 において候補手 m の値 v を評価するものとする。 m により p^0 は p^1 になる。 p^0 から p^1 への変化 Δp から v を評価するための知識 k を考える。一般に Δp は複数の成分に分けて考えることができる。すると k は多数の知識片の集合とみなせ、 $k = \{k_1, k_2, \dots, k_n\}$, ここで k_i は Δp の成分 $(\Delta p)_i$ の値 v_i を評価する。 k あるいは k_i を変化に関する知識と呼ぶ。

$(\Delta p)_i$ としては m により生じる各種の良形、悪形などを考えることになる (3.8 節の①~⑤より成る)。

すると m の評価は次のように行われる。

(1) 局面 p^1 の絶対的な局面評価を行いそれを m の価値とするのではなく、 Δp による相対的な局面評価で m を評価する。

(2) 局面の評価は全体的に行う必要はなく、変化した部分についてのみ行えばよい。効率的である。

(3) そのさい Δp の各成分 $(\Delta p)_i$ に対応する知識 k_i が駆動され、それらにより得られる部分評価値 v_i が総合され評価値 v が求められる。

$(\Delta p)_i$ は 2 種類に分類できる。

* プログラム的には変化により該当する知識が駆動されると考えられるので、文献 10) では「変化により駆動される知識」と呼んだ。

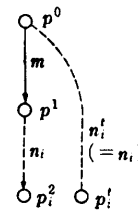


図 2 候補手と可能手

Fig. 2 Candidate move and possible moves.

A. 既知パタンの出現または消滅であり、パターンについての概念的知識を用いて v_i が直接に評価できる。

B. 既知パタンの出現、消滅でないため、 v_i が直接には評価できない。探索を起動する手続き的知識を必要とする。

概念的知識が増せば B に属する変化も A に移行する。

B に属する変化については次の方法によった。

[定義 2] 可能手の価値の変化による評価

図 2 のように、局面 p^0 あるいは p^1 のいずれかまたは両方においてありうる相手または自分 (打着者) の手 $n_i (i=1, 2, \dots, l)$ を考える。 p^1 において n_i を打つと局面 p_i^2 に、 p^0 において仮想的に n_i を打つと局面 p_i^f になるとする。後者の n_i を n_i^f と記して区別する。このときの、 n_i^f の価値と n_i の価値の差による v の評価をいう。

ただし一般には、 n_i と n_i^f の両方が必ず存在するわけではない。3 通りの場合がありうる。

(i) n_i はあるが n_i^f はない。すなわち、 m によりそれまでにはなかった新しい手が出現した。このとき n_i が敵 (打着者) の良手であれば悪い (良い) 変化である。

(ii) n_i^f はあるが n_i はない。すなわち、 m によりそれまであった手が消滅した。 n_i^f が敵 (打着者) の良手であったのならそれは良い (悪い) 変化である。

(iii) n_i, n_i^f ともに存在する。この場合、一般に価値が変化する。敵 (打着者) 側の良手 n_i^f の価値が下がったのなら良い (悪い) 変化であり、敵 (打着者) 側の n_i の価値が上がって良手 n_i になったのなら悪い (良い) 変化である。

ここで、 n_i や n_i^f を評価するという問題が残る。これらは、いま問題にしている m の評価と同様に再帰的に行えばよい。探索はこのように進められる。

【例 2.1】 図 3(a) において白番の候補手 $m=b3$ を評価する。 $b3$ を打着後の局面 (同図(b)) には良悪が既知のパタンの出現や消滅はないので直接的な評

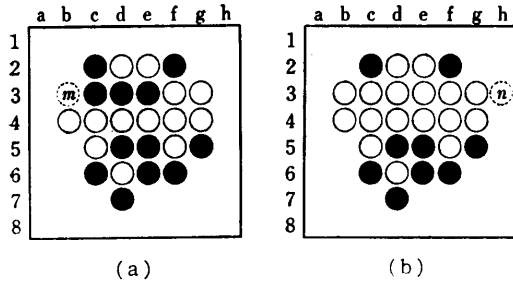


図3 敵可能手の価値の変化
Fig. 3 Change of the value of an opponent's possible move.

価はできない。ここで敵側の可能手 h3 について考える。(a)と(b)における h3 をそれぞれ n' , n とする。 n' または n 打着後の局面をそれぞれ(c), (d)に示す。 n' と n の価値を調べる。(c)にはとくに考慮すべきパタンの出現はない。一方(d)にはわくで示した部分パターンが出現するが、これは黒に権利手 h2 が残るパターンであることが直接観察できる。つまり、黒側にとり特別な手ではなかった n' が良手 n に変化したわけで、これは m によって引き起こされた悪い変化である。(例終り)

2.4 よみの深さと幅

分類Bに属する知識は手の読みを必要とする。この探索を制御するための知識を用意する。

2.4.1 幅

「変化に関する知識」により m による意味のある変化を評価するわけなので、可能手のうち m により意味のある変化を起こすと思われるもののみあらかじめ絞ることで効率を上げることが期待できる。実際、上級者ほど局面を一べつただけで読むべき手を絞っており、大変重要な種類の知識である。OTL. PLG には次の非常に初歩的な知識しかない。

選別用知識 いま、候補手 m によって影響を受けられると思われる敵の可能手 n_i, n'_i のみを選び出した。 n_i と n'_i により反転される石の集合をそれぞれ S, S'

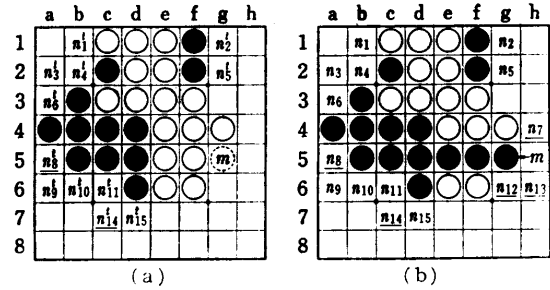


図4 探索される敵の手
Fig. 4 Opponent's moves to be searched.

とする。すると、 S と S' が一致していれば、ほとんど変化はないと考えられるから探索を省略できる。

【例 2.2】 図4(a)における黒候補手 $m=g5$ を評価中とする。 m により同図(b)の局面ができる。(b)において白の可能手は図に示すように15箇所ある。しかし選別用知識によれば次のものに絞られる。

- (1) 新たに出てきた白可能手: n_7, n_{12}, n_{13}
- (2) m の前後で反転する黒石の集合が異なる白の可能手: n_8 と n'_8, n_{14} と n'_{14} 実際には n_{14}, n'_{14} は大して調べる価値のない手だが、現 OTL. PLG はこれらをあらかじめ排除する知識をもたない。(例終り)

2.4.2 深さ

読みの深さは、各探索の必要度に応じて限界を設定しておく。状況に応じて可変にすることも考えられる。

2.5 緊急度

変化に関する知識を用いて得られる v は良悪の程度を意味しており、これがすなわち打着すべき順序を示すわけではない。 m の緊急度を考慮して最終的な m の得点 (以後 v で示す) を求める必要がある。緊急な場合には次の2種類がある。

- (1) 攻撃的緊急性: m が敵の良手の価値を減じたり打着不可にする。
 - (2) 防衛的緊急性: m が打着者の良手で敵の手 n に消されるおそれがある。 n の良悪による。
- 反対に、 m が良手だからこそ温存しておきたいという場合もある (以下温存手と呼ぶ)。

3. OTL. PLG

2章で述べた各種知識を OTL. PLG では Prolog で以下のように実現している。

3.1 評価点の設計

OTL. PLG では、候補手 m の評価を m による局面の変化を観察することで行うが、変化の種類 (良悪) およびその程度は通常の場合と同様に整数値で表

表 1 評価点と良悪
Table 1 Evaluation point vs. degree of goodness.

| 評価点 v | 良 悪 |
|------------|-------|
| >15 | 非常に良い |
| 10 | 良い |
| 5 | まあ良い |
| 0 | 良悪なし |
| -5 | まあ悪い |
| -10 | 悪い |
| ≤ -15 | 非常に悪い |

表 2 変化の価値
Table 2 Value of each variation.

| 打着マス | 変 化 | 変化の価値 |
|-------------|--------------|-------|
| 隅 | | +40* |
| 中辺 ② | 横取り | -10 |
| | a 横取り | -5 |
| | 沿い手 | +5 |
| 辺 | スカートができる | +10 |
| | C 権利手つき翼ができる | +15 |
| | 最悪の双翼ができる | -20 |
| | 双翼ができる | -15 |
| | 最悪の翼ができる | -15 |
| | 翼ができる | -10 |
| | 最悪のグライダーができる | -10 |
| | グライダーができる | -5 |
| | トーチカができる | -10 |
| | 準トーチカができる | -5 |
| | テーブルができる | -5 |
| 致命的なC打ち | | -40 |
| C 権利手に関するもの | | |
| C | 権利手そのもの | +5 |
| C | 敵の権利手を奪う | +5 |
| A | 権利手を確保する | +10 |
| A | 隣接②が空の場合 | +3 |
| B, ② | 敵に権利手を与える | -10 |

* 終盤ではこれに打着可能手数を考慮している。

す。良い方向への変化を正数、悪い方向への変化を負数、そのどちらでもない変化を0で表す。

中盤では自分と敵の可能手数の差が重要な指標であるが、手数1手が約5点になるように点数系を設計した。評価点 v と良悪の対応をあえて記せば表 1 になる。実際のゲームでは、敵が大きな間違いをしない限り、0点付近の手を打ってゆくのがやっとなのであるから、常識的には0点でも良い手ということになる。

3.2 既知パターン

2.3 節の分類 A に属する知識としては、主として辺、中辺に関するものが必要である。特定パタンの生起に対してその変化の価値を点数で与えておく (表 2)。

C 上に残る権利手は非常に安定で重要なので詳しい知識を用意する。

3.3 全体的な石の様子

序盤から中盤にかけてつねに念頭に置いておくべきことは、石がなるべく少なく内側に固まるように打つことである。この方針に基づき、壁を作るな、中を割れ等の経験的な戦術がいくつか知られている³⁾。

これらの定式化しにくい経験則に対し、同等の結果をもつ計算機向きの関数 π^* : 局面 \rightarrow 整数を用意した。石の配置から簡単な計算で概観的な評価が可能である。

3.4 可能手の価値の変化—中盤

2.3 節の分類 B に属する変化を、同一辺上の攻防の場合とそうでない場合に分けて取り扱う。

以下次の用語を用いる。

悪くない手 打着者にとって $v \geq 0$ となる手。

両空マス 候補手 m が辺上の手のとき、 m のマスを起点とした同一辺上の両方向の最初の空のマス。最大2個。

3.4.1 辺上の攻防 (m が辺上の手の場合の知識)

可能手の出現、消滅を最も読む必要があるのは辺上での攻防の場合である。各場合ごとに記す。

(1) 候補手 m が敵石に接する

まず打着者に、その辺で考えて打着可能な隅が出現するかどうかを見る (打着者の可能手の変化)。出現していれば、その隅を敵が許した場合の敵側の得点 v' と打ち返して防いだ場合の敵側の得点 v'' を計算し、 $v_i = -\max(v', v'')$ とする。

出現しない場合には(3)に示す変化を調べる。

(2) m が辺上の敵石を反転して辺をとる

敵石を反転したことで敵側のその辺上での打着可能性の意味で可能手 n が消滅する (敵の可能手の変化)。 n が敵側の悪くない手で価値が v' なら $v_i = v'$ とする。加えて(3)に示す変化も調べる。

(3) 一般の場合 (敵可能手の変化)

m の両空マス S_1, S_2 を調べる。敵の悪くない手 n が可能で価値が v' ならば、 $v_i = -v'$ とする。そのような n がなければ $v_i = c$ とする。ここで

$$c = \begin{cases} 5 & \text{その辺が単色でないか、あるいは } m \text{ が} \\ & \text{その辺での最初の石である場合} \\ 0 & \text{それ以外の場合} \end{cases}$$

とする。

この配点は次の経験則、辺を取りすぎるのは一般に

* 文献 11) で定義した関数 f を少し調整したものである。紙面の都合で省略。

よくないが、単独の辺打ちや敵の石へのつけ手は、敵の反撃がなければまあ良手である³⁾、に基づく。

なお、(3)では m を打つ前の局面 p^0 での n' の打着可能性を調べていない。厳密にはその有無に応じて v も変わるが、第一次近似としては無視できる。 m の打たれたことによって初めて、 n の存在が意味をもってくるという状況だからである。

また、 m と同一辺上のすべての可能手を調べてはいない。他の可能手への影響は、辺上の特殊な場合ではなく、次項の一般的な場合と考える。

3.4.2 一般の敵可能手の価値の変化(全 m に関係)

3.4.1 項で考慮した可能手は考慮から除く。

(1) 敵の手 n を m で打着不可にする場合 n の価値 v' が -4 以上のときに、 m による有利な変化があったと考える。

$v_i = v' + 5$ ($v' \geq -4$ のとき) と計算する。

(2) 敵の手 n が m により打着可になる場合 n の価値 v' が 0 以上のときに、 m による不利な変化があったと考える。

$v_i = -5 - v'$ ($v' \geq 0$ のとき) と計算する。

(3) 敵の手 n の価値が m により変化する場合 n の価値が v^1 から v^2 に変化するとする。 -4 以上の範囲での値の変化に意味があると考え次のようにする。

$$w_j = \begin{cases} v^j & (v^j > -5 \text{ のとき}) \\ -5 & (v^j \leq -5 \text{ のとき}) \end{cases} \quad (j=1,2)$$

$$\Delta w = w_1 - w_2$$

$$v_i = \begin{cases} \Delta w & (\Delta w \leq -5 \text{ または } \Delta w > 0 \text{ のとき}) \\ 0 & (-5 < \Delta w \leq 0 \text{ のとき}) \end{cases}$$

(2)、(3)において値を不連続的に変化させた理由はほとんど無意味な雑音的变化を無視することにある。

3.5 可能手の価値の変化—終盤初期

3.4 節に示した知識に加え、隅や星に対する考慮が必要となる。

候補手 m が隅の手である場合に対しても可能手の変化を考慮する必要がある。辺上の場合と同様であるが、隅はそれを狭む両辺に影響する点異なる。またその隅を含む対角線上に位置するマスの列にも影響する。

(1) 斜めに接する敵の星打ち(星への打着)への影響

敵の星打ち n により、対角線上で反対側の隅を敵に渡すことになるかどうかを調べる。 n の価値を v'

とし、

$$v_i = -v' \quad (v' > 10 \text{ のとき})$$

とする。

(2) m を狭む両辺への影響

両辺上に敵の良手が生じるかどうかを調べる。隅から見て最初の空マスが両側に石のある空マスである場合にその可能性がある。それぞれの空マスへの手を n_1, n_2 とすると、

(a) n_1 または n_2 の一方が敵の悪くない手で値が v' の場合: $v_i = -v'$,

(b) n_1 および n_2 の両者が敵の悪くない手で値がそれぞれ v^1, v^2 であり、いずれか一方を打つともう一方は打てなくなるような場合: $v_i = -\max(v^1, v^2)$,

(c) (b)と同様だが、敵が両方を打てる場合:

$$v_i = -(v^1 + v^2),$$

とする。

(1)、(2)の知識は IAGO に最も欠如している種類のもの⁶⁾、これにより OTL. PLG は終盤初期で大変賢いプレイを行う。

このほか、 m が星打ちあるいは星を反転する場合、および m が自分の可能手へ影響を及ぼす場合に対する知識が終盤では追加されている。詳細は略す。

3.6 探索の量

一定の計算量を前提とすれば、考慮すべき項目の重要度に応じてこれを適当に配分しなければならない。オセロゲームでは隅、星、辺に関する読みは重要なので安定するまで見届けるが、候補手 m のそれ以外の位置への影響は比較的重要性が低いので1度しか、すなわち n の深さが2の場合にしか探索しない。つまり、3.4.1 項および3.5 節については安定するまで、3.4.2 項については深さ2までしか探索しない。普通深さ3まで読む必要があるといわれているから³⁾ 現 OTL. PLG の読みは不足であるが、後述のようになりの力を示している。

3.7 各変化の総合

Δp の値 v は、各変化 $(\Delta p)_i$ に対する v_i を総合的に考慮して計算する。現 OTL. PLG では単純にこれらの総和を求めている。すなわち

$$v = \sum_i v_i$$

より高度なプログラムを目的とするならこのような単純な計算では不十分であろうが、十分実用的であることは確かめられている。

3.8 評価要素の評価順と打切り

v_i を順に評価し累積してゆくに際し、その順序を

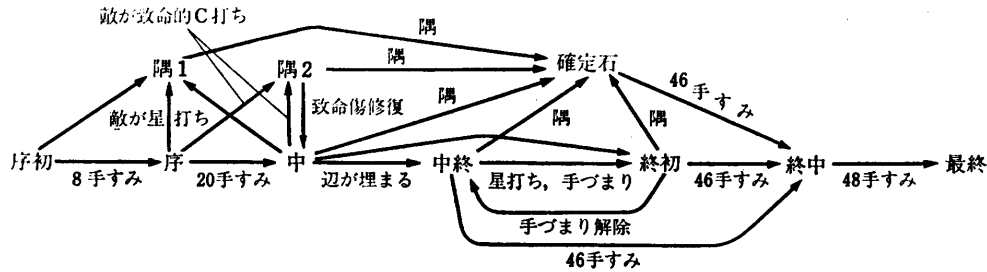


図5 状態の遷移

Fig. 5 Transition of states.

適当に設定しておく。ここで人が通常行うように、見込みがないことが明白になった時点で評価を打ち切り効率を上げることを考慮する。

既知パタンの出現等速く計算できる要素を先にすれば能率が上がる。しかし順序によっては打ち切るべきでない手を打ち切ってしまうことになる。正確さと速度の間にはトレードオフが存在する。OTL, PLGは正確さを重視し、原則として良い変化を先に計算している。

評価順は次のようである。

- ① 良形の出現
- ② m 打着前後での関数 π の計算: $\pi(p^1) - \pi(p^0)$
- ③ 可能手の自分側にとって有利な変化
- ④ 可能手の自分側にとっておもに不利な変化
- ⑤ 悪形の出現

④が不利な変化のみならば最後にすべきだが、プログラミング上有利な変化の一部が混入しているのでこの位置にある。①②⑤は計算が容易である。③④は探索なので時間がかかる。

全候補手が打ち切られると打ち切り点を下げて評価をやり直す。打ち切り点は1回目 -5, 2回目 -10 にした。こうすると1回目の場合、平均して約1/2が②までで、約1/4が④までで打ち切られ、かなり能率が向上した。

【例 3.1】 図3(a)における白番の候補の評価は次のようになる。星をあらかじめ除外すると候補手は13手ある。このうち、e7, f7, g6, c1, c8, e8, b1 の7手が②までで、b3, e1, h5 の3手が④までで打ち切られる。h6, d1, d8 のみが完全に評価される。

(例終り)

打ち切り点は形勢に応じてその都度設定するのが望ましいが、現 OTL, PLG では固定値にしてある。

3.9 緊急度の得点化

V は v と緊急度から次のように計算される。

- (1) $V=v: v \leq 0$ または m が緊急な手の場合
- (2) $V=v/c: v > 0$ かつ m が緊急でも温存手でもない場合. c は定数で5に設定した。
- (3) $V=-v/c-d: v > 0$ かつ m が温存手の場合. c, d は定数で $c=5, d=2$ に設定した。これらの定数およびその値は経験的に定めた。

3.10 OTL, PLG の全体構成

OTL, PLG は図5のような状態遷移図をもっており、ゲームの進行中つねに現状態を把握する。1手終了するたびに遷移を起こす条件を検討し、満足していれば状態を遷移させる。10の状態を区別し、それぞれに使用する戦略を対応させてある。中と中終では同じ戦略を用いるが遷移が異なる。盤上の石数が増大するに従って偶の価値が絶対でなくなるからである。

序初, 偶1, 偶2の各状態では簡単な知識を用いてプレイを行う。序, 終中, 最終, 確定石, の各状態では通常の α - β 探索による先読みを行う¹²⁾。

4. プログラミング

OTL, PLG の主要部分は Prolog で書いたが、関数 π や α - β 探索の部分等は FORTRAN によった。このため、Marseille 大学版 Prolog に FORTRAN プログラムとのインタフェース機能を付与し、必要な組み込み述語を FORTRAN および Prolog で作成し組み込んだ。このようにして、論理プログラミング機能と高効率なプログラムを使い分けている。Prolog を用いることには以下の利点があった¹³⁾。

- (1) 自動後戻りの機能により非決定的なプログラムが可能で、①次の手番における全候補手の展開、②全 ($4p$) の検査および評価、が複雑な制御なしで実現した。
- (2) また同機能はマスの順次走査や定石走査時にも有用で、それらは列挙しておくのみでよい。
- (3) パターン照合機能が定石の照合等で有用であ

った。

(4) 引数への入出力方向の自由さにより、多くの箇所で同一プログラムを多目的に使用できた。知識ベースの多様な利用法が可能である。

(5) (1)~(4)は、宣言的記述を許す Prolog の性質によっている。これは知識の追加・削除の容易さを意味する。事実、筆者は序盤の学習の簡単な実験から、学習プログラムの記述にも Prolog が適していることを示している¹³⁾。

5. OTL. PLG の動作

5.1 計算内容の解析

OTL. PLG の用いている中盤の戦略の性質を観察し、それが通常の α - β 探索と非常に異なることを示すため、一つの典型的なゲームの計算内容を解析する。

図6の局面は状態が「中」に移行したところである。ここから黒白ともに OTL. PLG に打たせる。「中」の状態が続く32手目までの各種統計を表3に示す。28, 29手目は打ち切り点を下げて再度評価を行っている。各記号は次を示す。

- N : 候補手の数。ただし星は初めから考慮されない。
- G : 全計算中に照合された良形の数。
- B : 全計算中に照合された悪形の数。
- D : 全計算中に良い変化を起こした敵可能手の数。
- Z : 全計算中に悪い変化を起こした敵可能手の数。
- S_p^o : 探索された敵側の n_i のうち m により消滅す

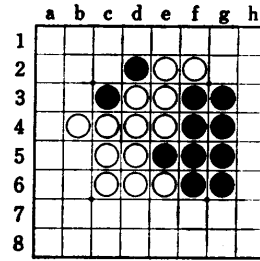


図6 中盤に移ったところ

Fig. 6 When middle game opens.

る手の総数。

S_s^o : 探索された敵側の n_i の総数。 S_p^o を含む。

P_s^o : 敵側の n_i のうち探索するかどうかを考慮されたものの総数。2.4節で述べたように、探索に至らない n_i もある。

H : 辺上の攻防に関する探索の総数。ただし交互の打着のみ (すなわち n_i は含まない)。辺上では深さが3以上になるが、これらすべて。

H_1 : H と同様だが、第1回目 (すなわち深さ2) のものみの総数。 H の一部。

S^o : 探索された深さ2の敵側の n_i の総数。 H_1 を含む。

P^o : 敵側の n_i のうち探索するかどうかを考慮されたものの総数。探索されない n_i もある。

S^p : 探索された打着者側の n_i の総数。

表3から以下が観察される。

(1) 手を決定するための通常の先読みの回数 $N+S^o-H_1+H$ は通常の α - β 法等に比して非常に

表3 中盤の計算内容
Table 3 Contents of computation of a middle game.

| 手番号 | 手 | N | G | B | D | Z | S_p^o | S_s^o | P_s^o | H | H_1 | S^o | P^o | S^p |
|-----|----|-----|-----|-----|-----|-----|---------|---------|---------|-----|-------|-------|-------|-------|
| 21 | d7 | 11 | 13 | 5 | 3 | 3 | 6 | 22 | 90 | 61 | 3 | 27 | 81 | 3 |
| 22 | c1 | 12 | 12 | 2 | 5 | 2 | 5 | 10 | 65 | 33 | 3 | 12 | 29 | 3 |
| 23 | b3 | 11 | 3 | 1 | 2 | 1 | 3 | 8 | 24 | 9 | 0 | 7 | 24 | 0 |
| 24 | b5 | 12 | 3 | 3 | 2 | 1 | 4 | 17 | 44 | 17 | 0 | 18 | 35 | 1 |
| 25 | c7 | 10 | 3 | 1 | 3 | 2 | 5 | 14 | 42 | 11 | 1 | 11 | 26 | 0 |
| 26 | b6 | 11 | 7 | 3 | 1 | 1 | 2 | 6 | 20 | 19 | 1 | 6 | 19 | 2 |
| 27 | c2 | 8 | 3 | 2 | 2 | 2 | 6 | 10 | 48 | 16 | 3 | 9 | 26 | 1 |
| 28 | | 10 | 3 | 1 | 0 | 2 | 1 | 2 | 18 | 10 | 0 | 4 | 7 | 0 |
| | a3 | 10 | 21 | 4 | 2 | 5 | 2 | 14 | 72 | 77 | 9 | 27 | 47 | 1 |
| 29 | | 7 | 7 | 3 | 0 | 1 | 3 | 3 | 28 | 17 | 2 | 3 | 3 | 0 |
| | d1 | 7 | 24 | 12 | 0 | 7 | 9 | 22 | 98 | 76 | 8 | 23 | 34 | 3 |
| 30 | f1 | 11 | 12 | 4 | 0 | 2 | 5 | 10 | 55 | 28 | 4 | 12 | 38 | 4 |
| 31 | b1 | 8 | 12 | 1 | 1 | 0 | 3 | 9 | 56 | 31 | 4 | 10 | 26 | 0 |
| 32 | e1 | 11 | 4 | 0 | 0 | 1 | 1 | 3 | 18 | 6 | 1 | 5 | 12 | 1 |
| 合計 | | 139 | 127 | 42 | 21 | 30 | 55 | 150 | 678 | 411 | 39 | 174 | 407 | 19 |

少ない。この例では1手平均57回で人の場合と比較しうる。

(2) α - β 法では読まれない読みがある。 S^g, S^p がそれらである。これらを含めても読みの回数は少ない。

(3) 辺上の攻防は安定するまで読むため深さが最大7になるが、枝数が1か2なので総数 H は巨大でない。

(4) m 打撃後の局面における、辺の攻防以外の読むべき手の選択率 r は $r=(S^o-H_i)/(P^o-H_i)$ で計算される。全体を通じて $r=0.37$ となり、ほぼ2/3の手は初めから m による変化を受けないとして除外できている。2.4節に述べた選別用知識は大変大雑把なものなので、人なみの判定知識を導入すれば0.2以下にすることは容易と思われる。

(5) 辺上の攻防以外の探索数に対する可能手の変化の発見率 r_v は $r_v=(D+Z)/(S^g+S^o-H_i)$ となる。 m の前後での重複する読みは1回と数える。全体を通じて $r_v=0.27$ となり、あまりよくない。無駄な探索がかなり多く、改善の余地が大きい。

(6) 探索総数 $S=(N+S^g+H+S^o-H_i+S^p)$ は全体で854回である。同じ例を評価打ち切りなしで行ってみると $S=5,739$ となった。打ち切りは深さ2以上でも行われるから計算量の節約効果は大きい。

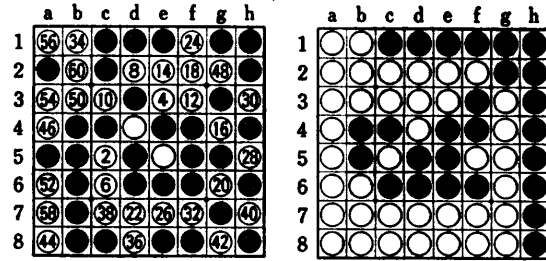
5.2 候補手の評価

中盤～終盤初期において OTL. PLG は全候補手について順次評価を行い最善手を選び出す。各評価過程は、予想されたように、人が通常行うような思考過程と大変類似したものとなった^{11), 13)}。見込みのない手は早々に切り捨てられる。悪くない手は詳細に調べられ、必要な探索が次々に駆動される。とくに辺上の攻防が追跡される。複数の良い、あるいは悪い要素が総合され、最後に緊急性が考慮されて候補手の値が定まる。これらの実行過程は人が十分追跡できる程度のものである。

5.3 OTL. PLG の実力

強力なことで知られる探索型の市販ゲームマシン Reversi Challenger[®] (RC と略す) を相手にした結果は以下のものであった。

(1) 序盤の読みの深さが5のとき、ほぼ RC のエキスパートレベル5と同程度の序盤の戦いとなる。このとき、中盤以後の戦いも同程度の力をもつことがわかった。十分強力であり上級者に匹敵する力をもつ。



(a) 棋譜 the record (b) 結果 the result

図7 Reversi Challenger との対戦例

Fig. 7 An example of a game against Reversi Challenger[®]. White: OTL. PLG.

(2) 打ち方はかなり異なる。OTL. PLG は形を重視するのに対し、RC は手数先の可能手数を重視している。このため OTL. PLG が中盤で手詰りになることもあるが、中盤を乗り切ると全体的な形の良さから最終的に勝つことが多い。

OTL. PLG の序盤の読みの深さを6としたときの、RC エキスパートレベル8との対戦例を図7に示す。RC は可能手数を重視するあまり黒39等で悪形を作り、結局 OTL. PLG を手詰りにしても不利になる様子がわかる。RC との対戦における典型例である。

5.4 ミニマクス法との比較

変化に関する知識を用いた手法は以下の点でミニマクス法と大きく異なっている。

(1) 候補手 m により生じる変化を観察することで、その手のもつ意味が直接理解される。ミニマクス法では全探索の結果間接的に求められるにすぎない。

(2) とくに m によって消滅させられる敵可能手を調べることができる利点をもつ。

(3) m による影響を多面的に評価できる。ミニマクス法で直接得られるのは、次の敵側の最善手のみ。

(4) 結果的に、非常に少ない探索で大きな効果をあげることができる。

5.5 一般的考察

OTL. PLG は AI プログラム一般に対して次のような示唆を与えている。

(1) いわゆる機械的な計算の不利

プログラムは一般的で単純な算法によるのが通常望ましい。ミニマクス法はこの原則に沿っている。しかし実際的な AI プログラムでは、計算の多くの段階で、場合場合に依じた細かい判断を伴う算法が要求される。

(2) 各種の知識の必要性

(1)を実現するには、計算に直接用いられる知識に

加え、必要な計算のみを選択するための多様な制御用知識をもつ必要がある。探索の爆発を防止できる。

(3) 意味のある作業の発見

必要な計算の判定には、その計算の意味を把握しておく必要がある。ある動作の意味は、それが及ぼす影響の大きさで理解されることが多い。

6. 結 び

通常の探索型のものとは異なり、より人の思考に近い形の知識を主体としたゲームプレイプログラムを作成し、かなりの実力をもつものが可能であることを示した。探索は非常に浅く知識量も上級者に比べてはるかに少ないが、探索型の強力なプログラムに対抗できることを示し、知識指向型のゲームプログラムの可能性を確かめた。能率向上のための、あるいは手詰りの予測に関するより高度な知識を導入することで棋力をより向上させることも可能である。この方法の原理は大変汎用性があるので、膨大な探索に依存している他の分野にも応用できよう。今後はさらに進んで、知識の自動的な学習を目標としたい。

謝辞 本論文作成に当たり、本研究の機会を与えて下さった当所若松清司制御部長、Prolog の利用について指導していただいた田村浩一郎数理情報研究室長、ご指導ご討論いただいた論理システム研究室の各位に感謝する。

参 考 文 献

1) Berliner, H. J.: Backgammon Computer Program Beats World Champion, *Artif. Intell.*,

Vol. 14, No. 2, pp. 205-220 (1980).

- 2) Frey, P. W.: The Santa Cruz Open Othello Tournament for Computers, *BYTE*, Vol. 6, No. 7, pp. 26-37 (1981).
- 3) 井上 博: 逆転の発見, (株)企画室ネコ, 東京 (1977).
- 4) Kowalski, R.: *Logic for Problem Solving*, North-Holland, Amsterdam (1979).
- 5) Nilson, N. J.: *Problem-Solving Methods in Artificial Intelligence*, McGraw-Hill, New York (1971).
- 6) Rosenbloom, P. S.: A World-Championship-Level Othello Program, *Artif. Intell.*, Vol. 19, No. 3, pp. 279-320 (1982).
- 7) 実近憲昭: ゲームプレイングプログラムの近年の成果, 情報処理, Vol. 20, No. 7, pp. 601-611 (1979).
- 8) 実近憲昭: ゲームと探索, 電子通信学会誌, Vol. 65, No. 4, pp. 405-412 (1982).
- 9) N. Sanechika and M. Tajima: Relationship between Evaluation Functions of Moves and Positions in Othello Game, 電総研彙報, Vol. 43, No. 4, pp. 1-17 (1979).
- 10) 田島守彦: ゲームにおける知識とその利用, 情報処理学会第24回全国大会 1B-7 (1982).
- 11) 田島守彦: ゲーム中盤の論理型プログラム, 信学技法 PRL 81-52 (1981).
- 12) 田島守彦: OTL. PLG 第一版—知識指向型ゲームプログラム, 情報処理学会第26回全国大会 5D-9 (1983).
- 13) 田島守彦: ゲーム・プログラムにおける PROLOG の使用, Prolog Conference (1982).

(昭和58年7月21日受付)

(昭和59年2月14日採録)