

K P V 法に基づく前打着時評価の再利用

Utilization of the Previous Evaluation Based on K P V Method

田島 守彦 (電子技術総合研究所)

M. TAJIMA (Electrotechnical Laboratory)

1. まえがき

交互打着の2人零和ゲームを考える。一般にある局面での候補手の評価を、次の自分の手番における候補手の評価に再利用することはできない。よって通常のプログラムでは、候補手の評価を全く新しくやり直す。これは局面の絶対評価を行なうミニマクス法のような手法を基本とするプログラムでは不可避である。持ち時間を節約するための唯一の方法は、敵が考え始めるとすぐに次の自分の候補手の計算を始めることである。

筆者は知識指向型オセロゲームプログラム O T L . P L G を開発したが¹⁾、そこでは局面に及ぼす影響で候補手を評価する新手法(文献²⁾で K P V 法と呼んだ)を開発、採用した。K P V 法で候補手を評価するときの利点の一つとして、前局面での候補手の評価結果をかなりの率で再利用でき、計算時間が大きく節約できることがわかった。以下でこれを示す。

次の記法を用いる。

m, m_i, m_j, \dots

手。打着局面の違いを区別する。

m, n, \dots

打着局面の違いを考慮しない場合には、手 m, n, \dots を示すのに太字を用いる。

p^i, p_j^i, \dots

局面

$v, v(m), v(\mathbf{m}), \dots$

手の価値

2. K P V 法²⁾

K P V 法は、手 m の評価を、 m により引起こされる局面の変化を計算することで行なう一手法である。

【定義1】 変化に関する知識^{1, 2)}

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る。すなわち、

$$k = \{k_1, k_2, \dots, k_n, t\}$$

ここで k_i は Δp の成分 $(\Delta p)_i$ の値 v_i を評価するための知識片であり、 t は各 v_i から v を計算するための知識である*。 k あるいは k_i を 変化に関する知識 とよぶ。□

* 文献^{1, 2)}の定義に t を加えている。

【定義2】 可能手の価値の変化^{1, 2)}

図1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打着可能な、相手または自分の手 n_i ($i=1, 2, \dots, l$) を考える。一般に n_i の価値は m により影響を受ける。 p^1 における n_i を n_i 、 p^0 における n_i を n_i^t とする。このときの、 n_i^t の価値と n_i の価値の差を「可能手の価値の変化」という。 □

n_i^t または n_i の片方しか存在しない場合にも、広義に価値が変化したと考える。

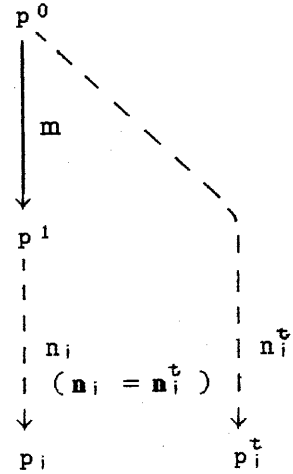


図1 可能手の価値の変化
variation of the values of possible moves

【定義3】 K P V法²⁾

「可能手の価値の変化」を含む「変化に関する知識」を利用して候補手を評価する方法をK P V (Knowledge on Position Variation) 法と名づける。 □

K P V法によれば上級水準のプレイを行なう知識指向型プログラムが可能なが実験的に証明されている。

3. 前打着時の評価と現評価

図2を考える。現局面を p^0 とし、候補手 $m_1 \sim m_s$ について評価するものとする。 m^{-1} 、 m^0 は既に打着された手である。交互打着の仮定により、 m^0 は敵側の、 m^{-1} は自分側の打着手である。

m^{-1} 打着時の評価について考える。 m^{-1} を含む候補手は m_1^{-1} 、 m_2^{-1} 、 \dots 、 m_r^{-1} であったとする。また

$$M^{-1} = \{ m_i^{-1} \mid i=1, 2, \dots, r \}$$

$$M = \{ m_j \mid j=1, 2, \dots, s \}$$

とする。

【定義4】 手 m が受ける、手 n の影響

m の打着可能性が n により変化するか、あるいは $v(m)$ が n により変化するか、すなわち $v(m_i^{-1}) \neq v(m_j)$ ($m_i^{-1} = m_j = m$)

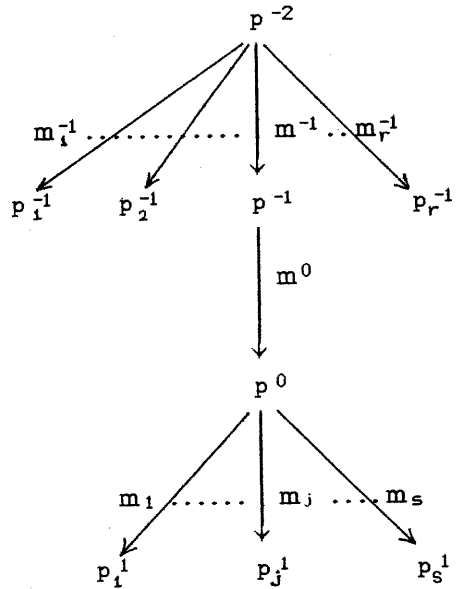


図2 前打着時と現打着時
previous move and now move

のとき、 m は n によって影響されるという。 □

次の定理は自明である。

【定理】

p^0 において新たに評価すべき手の集合 E は

$$E = \{m_j \mid m_j \text{ は } m^{-1} \text{ または } m^0 \text{ により影響される}\}$$

である。 □

E が非常に大きな割合を M において占めていれば、この定理の存在意義はない。がミニマクス法でなく KPV 法で評価を行なう場合には、以下の例に示すように $M-E$ がかなりの高率で存在し、かつ評価する前にかなり容易に判定できる。

なぜなら、 KPV 法は局面の相対評価を行なう手法だからである。すなわち、 p_i^{-1} の絶対的な価値は、これに対応する p_j ($m_i^{-1} = m_j$) の絶対的な価値と一致することは全く期待できないが、 p^{-2} から p^{-1} への価値の変化と p^0 から p^1 への価値の変化がほぼ一致することは場合によっては大いに期待できる。

E のうち、新しく生じた手の集合を N 、局面 p^{-2} で既に存在していた手の集合を O で表わすことにする。すなわち

$$E = N \cup O$$

ただし

$$N = M - M^{-1}, \quad O \subseteq M \cap M^{-1} \quad .$$

4. オセロの例

オセロゲームの場合、影響の有無は、大ざっぱには反転される石の集合の変化の有無で知ることができる。これは、プレイ時の KPV 法における枝刈りに関する選別用知識^{1, 2)} と同一の知識である。以下に述べる。

4.1. NV カット

評価必要性判定用知識

m_i^{-1} が反転する敵石の集合を S^{-1} 、 m_j ($m_j = m_i^{-1}$) が反転する敵石の集合を S とする。このとき、

$$S^{-1} = S$$

ならば、 m_i^{-1} に対して m^{-1} や m^0 の影響はほとんどないと考えられる。すなわち、

$$v(m_i^{-1}) \doteq v(m_j) \quad \square$$

この知識は、 m_i^{-1} が m^{-1} や m^0 に辺上で隣接しているといった明らかに影響を受ける場合を除き、多くの場合に正しいとみなせる。例を示す。

【例1】

図3において、(a),(b),(c) というゲーム進行を考える。(a) における白番の候補手

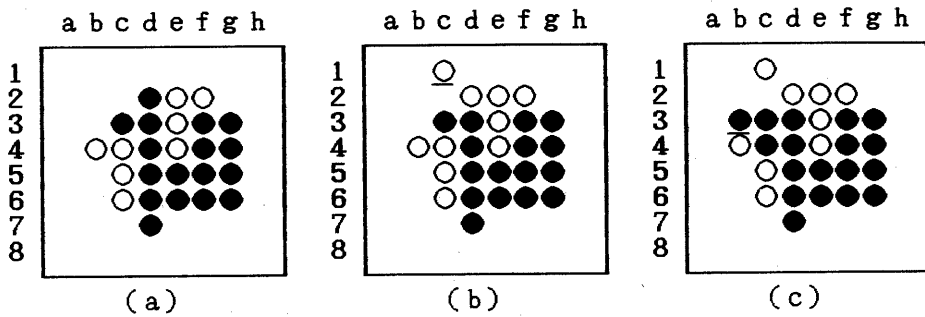


図3 評価省略の例
an example of evaluation omission

は、星 (b2,b7,g2,g7 の4マス) への手を除外して全部で12個所ある。

$$M^{-1} = \{c1, e1, c2, b3, h3, h4, h5, h6, e7, f7, h7, e8\}$$

一方、(c)における白番の候補手も星を除外して全部でやはり12個所である。

$$M = \{c2, a3, h3, h4, b5, h5, h6, e7, f7, h7, d8, e8\}$$

よって

$$N = \{a3, b5, d8\}$$

また、 $M \cap M^{-1}$ のうち、Oは、先の評価必要性判定用知識によれば、

$$O = \{c2\}$$

である。他の8候補手については $S^{-1} = S$ である。よって新たに評価すべき候補手の集合は

$$N \cup O = \{c2, a3, b5, d8\}$$

となる。

各候補手の価値をKPV法を用いた現OTL・PLGで計算した結果を表Iに示す。OTL・PLGでは、手の価値はそれに起因する局面の変化で評価される。0点が中立的な(良くも悪くもない)局面の変化を示す。(c)における白の最善手はb5である。各候補手の評価を、変化成分の良いものから始めて悪いものへという順序で見積ってゆくが、見込みがない(計算途中での評価値が-5を下回るような)手は、そこで評価を打切ってしまふ。この場合を表Iでは < -5 で示している。

表の値は(a),(b),(c)とも全て白の立場での評価値である。(a)と(c)の値を比較してみる。 $M \cap M^{-1}$ に含まれる候補手のうち、新たに評価する必要がないとされた候補手は

$$\{h3, h4, h5, h6, e7, f7, h7, e8\}$$

であるが、予想通りそれらの評価値の変化は小さい。

(例終り)

この判定用知識は大変初歩的なものであり、現実にはより高度な知識で人は判定を行なっている。それらを実現すればより能率的なプログラムが可能となる。

【定義5】 NVカット

候補手の評価をKPV法で行なうとき、評価必要性判定用知識を利用し、価値の変化が生じないと予測できる手の評価を省略することをNVカット (NV-cut, No Variation cut) とよぶ。 □

4.2. 敵の思考時間の利用

自分の手 m^{-1} により生ずる変化は、敵の思考時間中に計算しておくことができる。この場合には、2手打着後の変化分ではなく、1手打着後の変化分をその都度計算する方式になる。

【例2】

やはり図3を例にとる。

(I) (b) の黒番の局面で、黒にとって敵側の白の打着可能手を考える。星を除き次の11手である。(集合を M_0 と記す)

$$M_0 = \{c2, b3, h3, h4, h5, h6, e7, f7, h7, d8, e8\}$$

このうち、NVカットされずに残る可能手は

$$\{c2, d8\}$$

表I KPV法による候補手の評価値

候補手	(a)	(b)	(c)
e7	0	-4	-3
b5			7
c2	<-5	-5	1
f7	<-5	<-5	<-5
b3	<-5	-3	
c1	7		
a3			<-5
h3	<-5	<-5	<-5
h6	<-5	<-5	<-5
e1	<-5		
d8		<-5	<-5
e8	<-5	<-5	<-5
h4	<-5	<-5	<-5
h5	<-5	<-5	<-5
h7	<-5	<-5	<-5

である。この2手のみを敵の思考時に評価すればよい。

(II) 敵打着後の局面(c)では、(b)と(c)の相違のみを考えればよい。すなわち、 M_0 とMの間で価値の変化し得る手のみを評価する。NVカットされずに残る候補手は
{c2,a3,b5}

の3手である。(I)の過程が敵打着以前に終了していれば、自分の持時間にはこの3候補手の評価をしさえすればよい。

やはり現OTL・PLGによる評価値を調べてみる。(I)の段階では、NVカットされる候補手は

{b3,h3,h4,h5,h6,e7,f7,h7,e8}

である。主な値の変化はe7の-4である。全体に変化は大きくない。(II)の段階では、NVカットされる候補手は

{h3,h4,h5,h6,e7,f7,h7,d8,e8}

である。これらには大きな値の変化はない。(例終り)

OTL・PLGの持つ知識も不完全であるから正確な数値ではないが、NVカットが、概して不必要な評価をうまく省略できる手法であることが示された。例2では、評価の必要な候補手数を12から3に減らすことができた。このような計算量の節約は、KPV法による評価で初めて可能になる。

5. 実際上の考慮

例2で見たように、実際にはNVカットされる候補手の価値も多少変動している場合がある。これに起因する最善手選択の誤りを防ぐには次のような後処理を行なえばよい。

m^{-1} や m^0 による m_j の価値の変動の上限と下限をあらかじめ見積っておく。それらの変動を考慮すると最善手の可能性をもつ候補手が複数になる場合には、それらの最善手候補でNVカットされているものを、良い順に1手ずつ再評価するという過程を、最善手が確定されるまで行なう。実用上十分な程度まで行なえばよい。

変化の量をどの程度に見積るかにより、新たに計算する必要の生じる候補手の数が決まる。実際よりも小さく見積れば誤りは増加し、大きく見積れば不要な計算が増加する。OTL・PLGでは±5程度の誤差を予め見込んでおけばよいようである。例1、例2においては、±5の変動の可能性を考慮しても、図3(c)における「最善手=b5」という結論が変わる可能性はないので、そのような再評価の必要はない。より精密な方法としては、見積られる変化に応じて複数水準の評価必要性判定用知識を用意することが考えられる。

6. 結び

NVカットの手法の有効性は、対象とするゲームの種類に大きく依存する。オセロゲームの場合、これは局面の変化の評価時における「可能手の価値の変化」の計算の省略^{1,2)}と同様の処理であり(これもNVカットと呼ぶことにしたい)、間に別の手は1手しか入らない場合、すなわち p^{-1} における各候補手の価値が既知の場合、初歩的知識のみでも辺

以外の位置への候補手の約率がカットできる可能性があることが分っている^{1, 2)}。

しかし、本来NVカットは、より大規模な盤上でのゲームに有用と考えられる。特に碁では、その局所戦の多さ等を考慮すれば、劇的な割合のカットが期待できる。

謝辞

本研究の機会を与えて下さった当所若松清司制御部長、日頃御指導御討論頂いている論理システム研究室の各位に感謝する。

文献

- 1) 田島守彦、知識指向型オセロゲームプログラムOTL . PLG、情報処理学会論文誌、25、4、pp.597-605 (1984)。
- 2) 田島守彦、KPV法による知識指向型ゲームプログラムOTL . PLG、電総研彙報、48、12、pp.951-965 (1984)。