KPV 法に基づく前打着時評価の再利用†

田島守彦†† 実近憲昭††† 岡田義邦††††

KPV 法とは，筆者らが発表済みの知識指向型ゲームプログラムで開発・採用した候補手の評価手法である．本論文では，KPV 法を評価手段とするゲームプログラムにおける手の評価において，前打着時の評価が再利用できる方法を示す．KPV 法を概説した後，評価必要性判定用知識を用いて不要な候補手評価をカットする NV カットと称する手法を述べ，かなりの割合の候補手がカットできることをオセロゲームの例で示す．相手の思考時間を利用すればさらにカットできる候補手の割合は大きくなる．最後に，誤差が生じる問題の考慮，および NV カットの基礎について議論する．

1. はじめに

交互打着の2人零和ゲームを考える．一般に局面の絶対評価を行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，手番ごとに独立の探索木を作るため，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[1]．これはミニマクス法を基本とするプログラムでは不可避である．また，探索指向型の手法の限界を打破するために各種のアプローチによる知識指向のプログラムも開発されており[2),3)]，手の展開順序に関する記憶を利用する試みはあるが[4)]，基本的な局面評価の部分の事情は同様で，やはり再利用は難しい．

筆者は知識指向型オセロゲームプログラム OTL. PLG を開発したが[5)]，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6）で KPV (Knowledge of Position Variation) 法と呼んだ）を開発，採用した．本論文では，KPV 法で候補手を評価するときに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それにより計算時間を大きく節約できる手法，およびこの手法を利用できることが KPV 法の利点の一つであることを示す．

手はある局面におけるあるプレーヤの動作であるが，盤の同じ位置に打たれる手を，いくつかの手番にわたって異なった局面で考慮する場合があるので，ここでは局面の違いを無視する（例えばオセロゲームでは，石の色と打着マスのみに注目する）場合の手の表記を考慮している．次の記法を用いる．

- $m, m_i, m^i{}_j, \cdots$手．打着局面の違いを区別する．m はある局面の手を一般的に表す．m_i は複数個の手の第 i 番目を，$m^i{}_j$ は時点 i の第 j 番目の手を表す．
- $\boldsymbol{m}, \boldsymbol{n}, \cdots$打着局面の違いを考慮しない場合には，手 $\boldsymbol{m}, \boldsymbol{n}, \cdots$ を示すのに太字を用いる．
- $p^i, p^i{}_j, \cdots$局面．上付き，下付きは手の場合と同様．
- $v, v(m), v(\boldsymbol{m}), \cdots$手の価値

2. KPV 法[6)]

KPV 法は，手 m の評価を，m により引き起こされる局面の変化を計算することで行う一手法である．

（1） 変化に関する知識[5),6)]

局面 p^0 において候補手 m の価値 v を次のように定める．p^0 において m により得られる局面を p^1 とする．p^0 から p^1 への変化 Δp から評価値 v を得るための知識 k を考える．一般に局面の評価はいくつかの要因に基づいて行われるので，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，

$$k=\{k_1, k_2, \cdots, k_n, u\}.$$

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，u は各 v_i から v を計算するための知識である（文献5），6）の定義に u を加えている）．k を「変化に関する知識」と呼ぶ．

（2） 可能手の価値の変化[5),6)]

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 $\boldsymbol{n}_i$ $(i=1, 2, \cdots, l)$ を考える．一般に $\boldsymbol{n}_i$ の価値は候補手 m により影響を受ける．p^1 における $\boldsymbol{n}_i$ を n_i，p^0 における $\boldsymbol{n}_i$ を n'_i とする．このときの，n'_i の価

---

† Utilization of the Preceding Evaluation Based on KPV Method by MORIHIKO TAJIMA (Machine Inference Section, Machine Understanding Division, Electrotechnical Laboratory), NORIAKI SANECHIKA (AI Language Research Institute, Ltd.) and YOSHIKUNI OKADA (Distributed Systems Section, Computer Science Division, Electrotechnical Laboratory).

†† 電子技術総合研究所知能情報部推論研究室
††† (株)AI 言語研究所
†††† 電子技術総合研究所情報アーキテクチャ部分散システム研究室

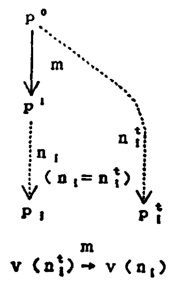


図 1 可能手の価値の変化
Fig. 1 Variation of the values of possible moves.

値と n_i の価値の差

$$v(n_i)-v(n'_i)$$

を「可能手の価値の変化」という.

mによって新たに手が生じる場合とか，または最初にあった手が打てなくなるとかによって，いずれかが打着不可で n'_i 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.

［定義］ KPV 法

「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 KPV (Knowledge of Position Variation) 法と名づける. □

KPV 法によって上級水準（初段程度）のプレイを行う知識指向型プログラムが可能なことを実験的に証明した.

3. 新たに評価すべき手

本章では，KPV 法で候補手を評価する場合，新たに評価すべき手は，新しく打着可能性の生じた手および価値が変化する手のみであり，それが候補手全体に占める割合が小さいことがオセロゲームでは期待できることを示す.

図2を考える．現局面を p^0 とし，候補手 $m^1{}_1$～$m^1{}_s$ 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.

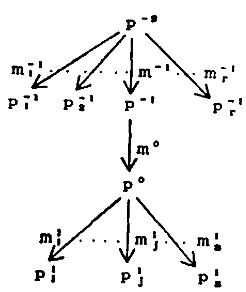


図 2 前打着と現打着
Fig. 2 Preceding moves and current moves.

m^{-1} 打着時の評価について考える．p^{-2} での候補手が $m^{-1}{}_1, m^{-1}{}_2, \cdots, m^{-1}{}_r$ であったとする．この中には m^{-1} も含まれる．また

$$M^{-1}=\{\boldsymbol{m}^{-1}{}_i \mid i=1,2,\cdots,r\}$$

$$M^1=\{\boldsymbol{m}^1{}_j \mid j=1,2,\cdots,s\}$$

とする．この時，局面 p^{-2} と局面 p^0 での $\boldsymbol{m}$ の打着可能性が n ($n\in\{m^{-1}, m^0\}$) により変化するか，あるいは $\boldsymbol{m}$ の価値 $v(\boldsymbol{m})$ がnにより変化する，すなわち

$$v(m^{-1}{}_i)\neq v(m^1{}_j) \qquad (\text{ただし } \boldsymbol{m}^{-1}{}_i=\boldsymbol{m}^1{}_j=\boldsymbol{m})$$

のとき，$\boldsymbol{m}$ はnによって影響されるという.

すると，次の原理が成り立つ.

［原理］

p^0 において新たに評価すべき手の集合 E は

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

である. □

局面 p^0 で新たに打着可能になった手はもちろん評価せねばならない．そうでない手の場合に本原理が意義をもつが，評価する前に影響の有無を容易に査定できることが本原理の有効性の鍵となる．オセロゲームにおいて手を KPV 法で評価する場合には，第**4**章で述べる「評価必要性判定用知識」を用いてこれが可能である．この知識はゲームの分野に依存した知識である.

また E が非常に大きな割合を M^1 において占めていれば，この原理の存在意義はない．しかしミニマクス法でなく KPV 法でオセロゲームの評価を行う場合には，次に説明する理由で E の M^1 に対する割合がかなり小さい.

以下では説明のため，E のうち局面 p^{-2} で既に存在していた手の集合をOで，新しく生じた手の集合をNで表すことにする．すなわち

$$E=N\cup O$$

ただし

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

オセロゲームにおいて E の M^1 に対する割合が小さい理由は，第一にNの M^1 に対する割合が通常は小さいことである．オセロゲームの戦略では，序・中盤では石を少なく取るのが良い．この戦略に従って行われるゲームでは，一手についての盤面の変化は小さいので，新しい手は余り生じない．第二の理由は，Oの M^1 に対する割合も小さいことである．これは KPV 法が局面の絶対評価でなく相対評価を行う手法である

ことから説明される．すなわち，p^{-2} から p^{-1}_i を生じる手 m^{-1}_i と p^0 から p^1_j を生じる手 m^1_j が同位置（すなわち $\boldsymbol{m}^{-1}_i=\boldsymbol{m}^1_j$）の場合，$p^{-1}_i$ の絶対的な価値が p^1_j の絶対的な価値と一致することは通常期待できないが，p^{-2} から p^{-1}_i への価値の変化と p^0 から p^1_j への価値の変化がほぼ一致することは，p^{-2} と p^0 が大きく変化していない限り大いに期待できるからである．そして第一の理由で述べたように，一手についての盤面の変化は通常小さく，したがって p^{-2} と p^0 も大きく変化しない．

あるゲームにおいて，候補手の評価を KPV 法で行うとき，そのゲームに関する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) と呼ぶことにする．

4. オセロへの適用

4.1 NV カット用知識

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．これは以下に述べるように，プレイ時の KPV 法における枝刈りに関する選別用知識[5),6)] と同一の知識である．

［評価必要性判定用知識］

打着 m^{-1}_i により反転させられる敵石の集合を S^{-1}，打着手 $m^1_j(\boldsymbol{m}^1_j=\boldsymbol{m}^{-1}_i)$ により反転させられる敵石の集合を S^1 とする．このとき，

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

ならば，$\boldsymbol{m}^{-1}_i(\neq \boldsymbol{m}^{-1})$ 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，

$v(m^{-1}_i)\fallingdotseq v(m^1_j)$　□

プレイ時の選別用知識との関係について説明しておく．KPV 法では候補手 m を評価するため，原則として全打着可能手の価値の変化を計算するが，あらかじめ m により影響を受けないと査定できる可能手については変化の計算を省略できる．このために用いられる知識が選別用知識であり，本論文の評価必要性判定用知識と同じものである．使用される場面が異なるが基本的に同一目的である．

評価必要性判定用知識は，m^{-1}_i 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．オセロゲームでは自分と相手の有効手の数が局面評価の重要な指標であるが，反転石集合が等しければ有効手数の変化もほぼ等しくなる．また石が作る盤上のパタンが将来の有効手数に影響するが，反転石集合が等しければパタンも同様に出現・消滅する．以上がこのヒューリスティクスの直観的な根拠である．この知識を用いる例を示す．

［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は，明らかな悪手である星（$b2$, $b7$, $g2$, $g7$ の 4 マス）への手を除外して全部で 12 箇所ある．

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

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 箇所である．

$M^1=\{c2, a3, h3, h4, b5, h5, h6, e7, f7, h7, d8, e8\}$

すると，新しく生じた手の集合は

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

また，$M^1\cap M^{-1}=\{c2, h3, h4, h5, h6, e7, f7, h7, e8\}$ のうち，影響を受けた手の集合 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 $S^{-1}=S^1$ 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 OTL. PLG で計算した結果を**表 1** に示す．表中の値は（a），（b），

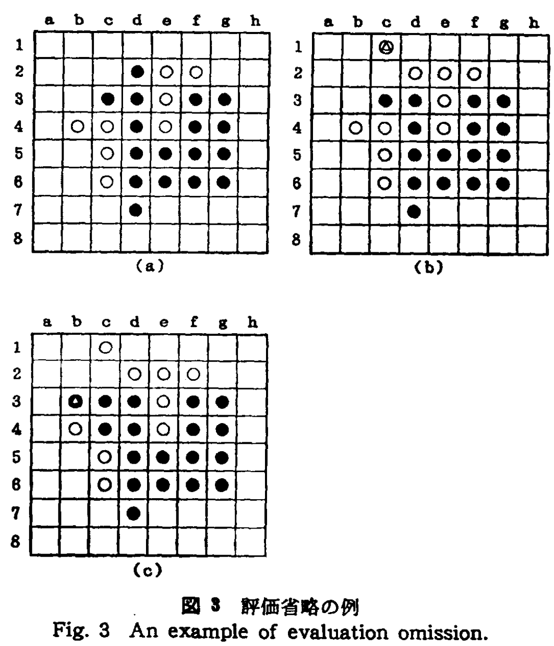


図 3　評価省略の例
Fig. 3 An example of evaluation omission.

表 1 例 1 における KPV 法による候補手の評価値
Table 1 The evaluated values of candidates by KPV method in example 1.

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

(c)ともすべて白の立場での評価値である．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，そこで評価を打ち切ってしまう．この場合を表 1 では＜−5 で示している．

(a)と(c)の値を比較してみる．$M^1 \cap M^{-1}$ 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，変化の絶対値は $e7$ の場合が 3 である．それ以外の手では評価が打ち切られている．予想通りそれらの評価値の変化は小さい（OTL. PLG では 5 以下は小さいとみなせる）か無視できる．　（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．

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］

［例 1］と同様に図 3 を例にとる．M^{-1} の手は評価済みである．

（I）(b)の黒番の局面で，黒にとって敵側の白の着手可能手を考える．星を除き次の 11 手である．（集合を M_0 と記す）

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

すると，このうち NV カットされずに残る可能手は

$\{c2, d8\}$

である．この 2 手のみを敵の思考時に評価すれば良い．

（II）敵打着後の局面(c)では，(b)と(c)の相違のみを考えれば良い．すなわち，M_0 と M^1 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 候補手の評価のみすれば良い．

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 のも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．NV カットが利用できることが KPV 法の利点の一つである．

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（例えば $e7$）．これに起因する最善手選択の誤りを防ぐには次ような後処理を行えば良い．

m^{-1} や m^0 による $m^1{}_j$ 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う．実用上十分な程度まで行

えば良い.

変動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程度の誤差をあらかじめ見込んでおけば良いようである．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. NV カットの基礎に関する議論

前章までで，評価必要性判定用知識を用いる評価省略手法「NV カット」を導入し，これが有効なことが KPV 法の一利点であることを述べた．NV カットをオセロゲームの例で説明してきたが，本章では KPV 法と NV カットの重要性を示すため，これらが基礎的な問題と関連していることに言及する．

NV カットのアイデアは広義のインクリメント更新の枠組みから見ることができる．すなわち，時間と共に変化する世界の記述におけるデフォールト性の役割である．

このアイデアを成立させている根拠は次の二つの条件である．条件1は，変化が局所的，すなわち変化しない部分が大部分であること．ここでいう「変化」とは，見かけだけの変化でなく本質的な変化である．条件2は，本質的な変化であることを見分ける，比較的単純な知識の存在である．条件1は，適用分野によって定まる．しかし場合によっては局所性の仮定が崩れることがあり，その場合の検出法と対応策が必要である．条件2は，オセロの「反転石集合」のように，うまく見つかる場合もある．しかし一般には難しく，いくつかの知識片を組み合わせて近似的に合成したり，うまい知識がないときは局所的な探索で代用（実際には知識片と局所的探索の併用）したりすることが行われる．

7. おわりに

NV カットの手法の有効性は適切な評価必要性判定用知識の有無によって決まるため，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」の計算の省略と同様の処理であり，間に別の手が1手しか入らない場合，すなわち本論文の場合に置き換えて述べれば，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 2/3 がカットできる[5),6)]．NV カットの効果は著しい．

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（評価必要性判定用知識としては，例えば出入り計算による勢力圏の変化の有無に関する知識を利用できる）．また，一般の意思決定問題においても，選択肢の評価に KPV 法や NV カットを利用することで，効率的な問題解決が期待できる．いずれにしても，適用分野に応じた適切な評価必要性判定用知識を見いだすことが鍵となる．

謝辞　本研究の機会を与えて下さった当所若松清司元制御部長，日頃御指導御討論頂いていた論理システム研究室の各氏に感謝の意を表する．

参 考 文 献

1) 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).
2) Wilkins, D.: Using Patterns and Plans in Chess, *Artif. Intell.*, Vol. 14, pp. 165-203 (1980).
3) Berliner, H. J. and Ebeling, C.: The SUPREM Architecture: A New Intelligent Paradigm, *Artif. Intell.*, Vol. 28, pp. 3-8 (1986).
4) 鳥居　稔，小谷善行：ゲームにおける知識処理と探索の融合，情報処理学会記号処理研究会報告，Vol. 87，No. 79，43-1 (1987).
5) 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，Vol. 25，No. 4，pp. 597-605 (1984).
6) 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，Vol. 48，No. 12，pp. 951-965 (1984).

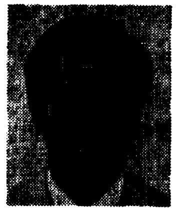
(昭和63年11月 9 日受付)
(平成元年 4 月 11 日採録)

田島　守彦（正会員）

昭和22年生．昭和44年東京大学工学部計数工学科卒業．同年電気試験所（現電子技術総合研究所）入所．現在知能情報部推論研究室主任研究官．入所以来，拡張閾値論理，並列計算機システム，ゲームの研究を行う．現在機械学習を中心とする人工知能の研究に従事．著訳書，マイクロコンピュータ教科書Ⅰ～Ⅳ(丸善，Ⅰ～Ⅲ共訳)ほか．電子情報通信学会，人工知能学会各会員．

実近　憲昭（正会員）

昭和14年生．昭和40年東京大学工学部電気工学科卒業．同年電子技術総合研究所入所，以来論理回路，システム設計の研究に従事．また，ゲームを題材とした人工知能問題に興味を持つ．昭和63年(株) AI 言語研究所に所長として出向．CESP 言語の開発に従事．



岡田　義邦（正会員）

昭和17年生．昭和40年名古屋大学工学部電子工学科卒業，昭和42年同大学院工学研究科電気電子専攻修士課程修了，昭和45年同博士課程修了，同年電気試験所（現在電子技術総合研究所）勤務．現在同研究所情報アーキテクチャ部分散システム研究室長．分散・並列計算機システム，光電子コンピュータの研究に従事．電子情報通信学会，計測自動制御学会各会員．