囲碁における2点の関係性に着目した プレイアウト中の情報抽出

石田 竹至[1,a)]　中村 貞吾[2,b)]

概要：モンテカルロ木探索は囲碁などのゲームにおいて優れた探索手法であるが，深い読みを必要とする局面では十分な性能を発揮することができない．そうした局面では，ランダムプレイアウト中に妥当な着手以外を打つ確率が高く，期待する勝率とは離れてしまうため，有効な着手を発見できなくなると考えられる．パターン [2] や Criticality [1] を用いてプレイアウトの確率を制御する手法が提案されているが，十分な解決とはなっていない．本論文では，プレイアウト中の情報を用いて，2 点間の関係性を見つけ出し，前述の問題を解決する手法を提案する．

Information Extraction in Playouts forcusing on Two Points Relationship in Go

TAKESHI ISHIDA[1,a)]　TEIGO NAKAMURA[2,b)]

Abstract: Although Monte Carlo Tree Search (MCTS) is an excellent search method in game of Go, it does not work in the position that require a deep reading. In such position, the low probability of hitting reasonable move in playout. It will apart from the true winning percentage, can not be found a valid move. The method for controlling the probability of playout has been proposed (Pattern [2], Criticality [1]), not fully resolved. In this paper, by extracting statistical information focusing on the relationship between the two points from the playout, we propose a method to solve the aforementioned problems.

1. はじめに

モンテカルロ木探索は囲碁などのゲームにおいて優れた探索手法である．2016 年 3 月に人間のトッププロを 4 勝 1 敗で破った Alpha Go [4] も，モンテカルロ木探索を用いている．しかし，その対局の 4 局目は局面判断を誤って負けており，完全勝利とはならなかった．深く細い読みを要求される攻め合いや死活などにおいては，モンテカルロ木探索が有効に働かないことは既知の弱点である．そうした局面では，ランダムプレイアウト中に妥当な着手以外を打つ確率が高く，期待する勝率とは離れてしまうため，有効な着手を発見できなくなると考えられる．

この問題に対して，パターン [2] や Criticality [1] を用いてプレイアウトの確率を制御する手法が提案されており，それらしい着手のプレイアウトを実現している．しかし，攻め合いや死活においては，単純な着手の**位置**だけでなく，着手の**順序**が重要な場合があり，そうした問題に対しては十分な解決となっていない．

本論文では，プレイアウト中の情報を用いて，2 点間の関係性を見つけ出し，前述の問題を解決する手法を提案する．

2. プレイアウト中の情報抽出

通常のモンテカルロ木探索では，プレイアウトから得られる情報は，最終局面の勝敗情報のみである．これに対し，プレイアウト中の着手や，終局時に置かれている石の位置などを用いて，多くの情報を獲得する手法がある．例えば，プレイアウト中に打たれたすべての手を「最初に打たれた」と仮定して，ゲーム木中のノードを更新する AMAF [3] や，終局時の石の位置と勝敗の情報を集め，先着すべき重

[1] 九州工業大学 大学院情報工学府
[2] 九州工業大学 大学院情報工学研究院 知能情報工学研究系
[a)] t_ishida@dumbo.ai.kyutech.ac.jp
[b)] teigo@ai.kyutech.ac.jp

要な地点を求める Criticality などがある．

2.1 Criticality

囲碁において勝敗を左右する重要な地点を求める手法として Criticality が提案されている．モンテカルロ木探索によってプレイアウトを行うと，黒が勝った場合に黒が取っている点と，白が勝った場合に白が取っている点がわかってくる．プレイアウトを繰り返すことで，その点を先取することがどの程度勝敗に関与しているかを推測することができる．

N	:	プレイアウト数
B	:	黒が勝ったプレイアウト数
W	:	白が勝ったプレイアウト数
$b(x)$	:	黒が x を取っていたプレイアウト数
$w(x)$	:	白が x を取っていたプレイアウト数
$v(x)$	:	プレイアウトに勝ったほうが x を先取したプレイアウト数

としたとき，ある点 x における Criticality $c(x)$ は式 (1) のように表される．

$$c(x) = \frac{v(x)}{N} - \left(\frac{w(x)}{N} \times \frac{W}{N} + \frac{b(x)}{N} \times \frac{B}{N}\right) \tag{1}$$

2.1.1 Criticality の式の再解釈

表 1 Criticality の分割表

	点 x を黒が先取	点 x を白が先取
黒勝ち	a	b
白勝ち	c	d

Criticality は，「点 x を獲得すること」と「勝ち負け」に対する連関係数を計算していると考えることができる．そこで，プレイアウト中の黒勝ち・白勝ちと，点 x を黒が先取・白が先取の回数を，それぞれ表 1 のように表す．a, b, c, d を用いると，式 (1) を式 (2) のように変形できる．分割表を用いることで直感的に表現できるため，本論文では以後，式 (2) を使って提案手法を表記する．

$$c(x) = \frac{2(ad - bc)}{N^2} \tag{2}$$

3. 2点間に着目したプレイアウト中の情報抽出

Criticality は，ある 1 点を先取することの重要度を測る尺度であった．囲碁では，「2 点のうちどちらかを先着すれば良い」や「特定の順番で打つ必要がある」盤面が存在する．そこで，Criticality を 2 点の関係性に拡張し，次の「見合い」と「着手順序」を測る尺度を導入する．

3.1 見合いの検出

囲碁には，2 点のうち少なくとも 1 点を先取すれば良い「見合い」という概念がある．片方を相手に打たれればすぐにもう一点を打つ必要があるが，相手に打たれない限りは自分も打つ必要はない状態である．囲碁上級者はこれを理解しており，「見合い」の考え方を用いることで，探索空間を大幅に削減している．「見合い」を測るための Criticality と同様の尺度を考える．

$b_b(x,y)$	:	x, y の両方とも黒が先取して黒が勝った回数
$b_w(x,y)$	:	x, y の両方とも黒が先取して白が勝った回数
$w_b(x,y)$	:	x, y の両方とも白が先取して黒が勝った回数
$w_w(x,y)$	:	x, y の両方とも白が先取して白が勝った回数

とした時の連結 (Joint) Criticality を式 (3) に示す．

$$\begin{aligned} JC_b(x,y) &= \frac{2((B-w_b(x,y))w_w(x,y)-w_b(x,y)(W-w_w(x,y)))}{N^2} \\ JC_w(x,y) &= \frac{2((W-b_w(x,y))b_b(x,y)-b_w(x,y)(B-b_b(x,y)))}{N^2} \end{aligned} \tag{3}$$

黒の見合い度を JC_b，白の見合い度を JC_w とした．例えば $JC_b(x,y)$ は，「黒が x, y の少なくとも一つを先取すること」と「黒が勝つこと」の相関を計算することで，黒にとっての見合いらしさを測っている (表 2)．

表 2 黒にとっての連結 Criticality の分割表

	点 x, y を 黒が 1 つ以上先取	点 x, y を 白がどちらも先取
黒勝ち	$B - w_b(x,y)$	$w_b(x,y)$
白勝ち	$W - w_w(x,y)$	$w_w(x,y)$

3.2 着手順序の検出

攻め合いなどでは，見合いとは異なり，複数の点に対して特定の順序で打たなければならない状態が存在する．例えば，図 3 のような局面（黒番，コミ 0.5 目，中国ルール）において，△印の白石を取ろうとすると，黒は内ダメである c より先に外ダメの a,b を打たなくてはならない．しかし，ランダムプレイアウトでは，a,b より先に c や，白が打つべき d,e のダメを打ってしまうこともあり，正しく攻め合いを認識できないことが多い．そこで，着手順序の重要度を測る手法として 順序 (Order) Criticality を考える．

m, n : 着手（あるプレイヤーが，ある場所に打つ）

m' : m と同じ場所に別のプレイヤーが打つ

$b(m,n)$: m, n の順に着手したときに黒が勝った回数

$w(m,n)$: m, n の順に着手したときに白が勝った回数

としたとき，表 3 のように分割表を定義する．表 1 と同じように a,b,c,d とおいた場合の, 順序 (Order) Criticality を式 (4) に示す．このとき，$a+b+c+d \neq N$ であるため，分母は $(a+b+c+d)^2$ としている．

表 3　順序 Criticality の分割表

	m の後に n を打つ	m が打たれずに n を打つ
黒勝ち	$b(m,n)$	$b(m',n)$ + $b(n,m)$ + $b(n,m')$
白勝ち	$w(m,n)$	$w(m',n)$ + $w(n,m)$ + $w(n,m')$

$$OC(m,n) = \frac{2(ad-bc)}{(a+b+c+d)^2} \tag{4}$$

3.3　順序 Criticality のランク付け

順序 (Order) Criticality $OC(m,n)$ は，2 点間における順序の重要性であった．しかし，実際の盤面では，複数の点に対して優先すべき点と後に打つべき点があるはずである．そこで，ある着手が他に比べてどれだけ優先して打つべきかを測る尺度 $OCR(m)$ を式 (5) に示す．$OCR(m)$ は，$OC(m,n)$ と $OC(n,m)$ の差の合計を求めている．また，差が小さい場合にノイズにならないよう，差の絶対値が β 以下の場合は増減を 0 にし，差が大きすぎる場合に引っ張られないよう，増減値は α に制限している．これにより，m を先に打つ重要性を算出している．

$$OCR(m) = \sum_{n} ocd(m,n) \tag{5}$$

$$ocd(m,n) = \begin{cases} \alpha & (OC(m,n) - OC(n,m) > \beta) \\ -\alpha & (OC(m,n) - OC(n,m) < -\beta) \\ 0 & (otherwise) \end{cases}$$

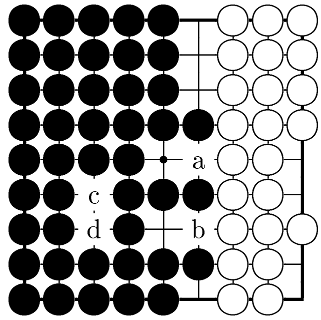

図 1　見合い局面 1

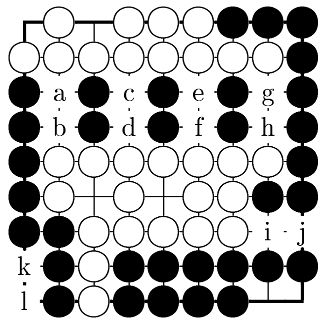

図 2　見合い局面 2

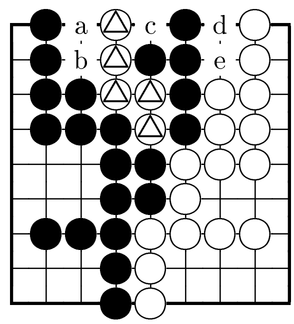

図 3　攻め合い局面 1

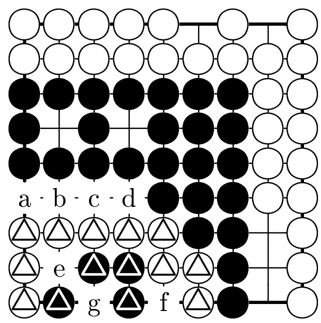

図 4　攻め合い局面 2

4.　実験

見合いが含まれた局面に対して連結 Criticality を計算することで，見合いである 2 点間のほうが，見合いではない 2 点間よりも JC_b, JC_w が高くなると予想できる．見合いが含まれる局面として，図 1,2 を用意し，連結 Criticality に差が出るか実験した．

同様に，攻め合いが含まれた局面に対して順序 Criticality

表 4　実験結果 (見合い局面 1)

x	y	$JC_b(x,y)$	$JC_w(x,y)$
a	**b**	**0.238**	0.160
a	d	0.146	0.100
b	d	0.146	0.103

表 5　実験結果 (見合い局面 2)

x	y	$JC_b(x,y)$	$JC_w(x,y)$
a	**b**	**0.0549**	-0.0377
c	**d**	**0.0545**	-0.0359
g	**h**	**0.0543**	-0.0382
e	**f**	**0.0537**	-0.0344
i	**j**	**0.0524**	-0.0157
l	j	0.0273	0.0171
k	j	0.0246	0.0248

を計算することで，着手順序が重要な 2 点間の値が大きくなるはずである．攻め合い局面として，図 3 と図 4 を用意し，順序 Criticality を計算した．その後，順序 Criticality のランク付けを行い，結果を調べた．

4.1　実験に用いた局面

図 1 は，a,b 間が黒にとって見合いである局面 (白番，コミ 0.5 目，中国ルール) である．白は a,b 両方を先取すれば黒の大石を取ることができ，大石を取れば非常に高い確率で白が勝つ．一方黒は， a,b の一方でも先取すれば（目を埋めない限りは）勝つことができる．

図 2 も黒の見合いの含まれる局面 (白番，コミ 0.5 目，中国ルール) である．a,b や c,d などは「タケフ」と呼ばれる形であり，白に一方を打たれても，黒はどちらか一方を繋げばよいため見合いである．黒がどこか 1 箇所でもタケフを繋ぎ損ねると，白が勝つ局面になっている．

図 3 は前に述べた通り，外ダメと内ダメの着手順序がある．黒は a,b から攻めなくてはならない．

図 4 も着手順序が重要な局面 (黒番，コミ 0.5 目，中国ルール) である．△印の石を取り合う形であり，正しい順序で打てば黒が勝つ．基本的には外ダメの a,b,c,d を詰めてから e か f に打てばよいが，途中で白が e か f に打ってきた場合は黒が g に打って 5 目ナカデにしなければならない．

4.2　実験方法

オープンソースのモンテカルロ碁プログラムである libEGO をベースとし，JC, OC, OCR を計算できるように実装した．UCT やパータンは用いずに 10,000 回プレイアウトした結果の JC, OC, OCR を求めた．

OCR 計算時の α, β はどちらも 0.01 とした．

4.3　連結 Criticality の実験結果

見合い局面である図 1,2 に対する JC_b, JC_w の一部を表 4, 5 に示す．表は JC_b で降順にソーティングしている．**太字** で示した部分が見合いである．

図 1 の局面では，a,b 間の JC_b が一番高くなった．値が大きいほど，「見合いらしさ」が高い．2 番目の a,d 間とは大きく差があり，見合いを認識できていると言える．

図 2 の局面でも，a,b 間，c,d 間，e,f 間，g,h 間，i,j 間のすべてのタケフの JC_b が最も高い値を示した．いずれも絶対値はそれほど高くなかったが，続く見合いでない着手の l,j 間とは約 2 倍の差があるため，見合いを認識できていると言える．

表 6　実験結果 (攻め合い局面 1)

m(プレイヤー 場所)	n	$OC(m,n)$
B a	B c	0.0918
W a	B c	0.0915
B b	B c	0.0910
W b	B c	0.0885
B a	W e	0.0794
B b	W d	0.0778
B a	W d	0.0772
B b	W e	0.0713

表 7　実験結果 (攻め合い局面 2)

m(プレイヤー 場所)	n	$OC(m,n)$
B g	B e	0.0743
B d	B f	0.0320
W e	B g	0.0319
B a	B f	0.0307
B b	B f	0.0289
B c	B f	0.0289
W e	B f	0.0284
W c	B f	0.0273
W a	B f	0.0268
W b	B f	0.0256
B b	B g	0.0249
W f	B g	0.0247

4.4　順序 Criticality の実験結果

攻め合いである図 3,4 に対する OC の一部を表 6, 7 に示す．値が大きいほど黒に有利な着手順序，値が小さいほど白に有利な着手順序である．

図 3 の局面の OC 上位 8 つを表 6 に示した．上の 4 つは，黒は a,b よりも後に c を打ったほうが，OC が高い（黒に有利）となった．次の 4 つは，白の外ダメ d,e よりも先に a,b を打ったほうが OC が高い結果となっている．黒が優先して打つべき場所，後に打つべき場所が，高い数値として現れている．

図 4 の局面の OC 上位 12 個を表 7 に示した．2 番目と 4〜6,8〜10 番目までは，黒は外ダメである a,b,c,d よりも後に，内ダメの f を打ったほうが OC が高くなることを表している．3 番目，12 番目は g を後に打つ手順であり，特に「白 e の後に黒 g」と「白 f の後に黒 g」はナカデに

表 8 実験結果 (攻め合い局面 1 のランク)

m(プレイヤー 場所)	$OCR(m)$
B a	0.71
B b	0.71
B c	-0.65
B d	-0.75
B e	-0.75

表 9 実験結果 (攻め合い局面 2 のランク)

m(プレイヤー 場所)	$OCR(m)$
B a	0.06
B b	0.06
B c	0.07
B d	0.07
B e	-0.12
B f	-0.18
B g	-0.07

表 10 実験結果 (攻め合い局面 2 から e を打った後のランク)

m(プレイヤー 場所)	$OCR(m)$
B a	0.09
B b	0.09
B c	0.08
B d	0.07
B f	-0.16
B g	**0.1**

表 11 実験結果 (攻め合い局面 2 から f を打った後のランク)

m(プレイヤー 場所)	$OCR(m)$
B a	0.07
B b	0.08
B c	0.09
B d	0.08
B e	-0.15
B g	**0.15**

する形である．しかし，特に順序関係が無いように思える $OC(Bg, Be)$ が高い値を示した．これは，表 3 の「m が打たれずに n を打つ」場合が以下のようになったためだと思われる．

- (m', n)：「白が g に打ってから黒が e に打つ」状態だが，白が g に打つためには e と f が埋まってないといけないため起こり得ない．ほぼ 0.
- (n, m)：「黒が e に打ってから黒が g に打つ」状態だが，自分で目を埋めることはほぼありえないようになっている．ほぼ 0.
- (n, m')：「黒が e に打ってから白が g に打つ」状態で，この場合ほぼ確実に白が勝つため，$w(n, m')$ が高くなる．

以上によると，「m が打たれずに n を打つ」場合では白勝ちがほぼ全てとなってしまう．式 (4) で言えば $d \gg b$ となってしまうため，$OC(Bg, Be)$ が大きな値になったのだと考えられる．

このように，打つ順序が重要な並びは浮かび上がってきているが，単純にこれらの値を使うには無理がある．次に示す順序 Criticality のランク付けが必要である．

4.5 順序 Criticality のランク付け

図 3,4 の OC を用いて OCR を計算した結果を，表 8,9 に示す．

図 3 のランク付け (表 8) では，外ダメ a,b にプラスの大きな値がついており，内ダメや白の外ダメの c,d,e にはマイナスの大きな値が出ている．

同じく，図 4 のランク付け (表 9) でも，外ダメ a,b,c,d にプラスの値がつき，内ダメ e,f にはマイナスの値がついた．g は内ダメではないが，白が e か f に打つまでは必要のない手なので，マイナスの値がついた．

また，図 4 の局面から 1 手進めた盤面を考えてみる．白が e または f に打った場合，黒はナカデにする g の優先度が上がる．白が e を打った場合の OCR を表 10 に，白が f を打った場合の OCR を表 11 に示す．どちらとも g にプラスの大きな値がつき，正しく認識できていると言える．

以上から，順序 Criticality の順序付けを行うことで，実用的な着手の優先度が取得できたと言える．

5. おわりに

深い読みを必要とする局面に対して最適な着手の探索が難しいというモンテカルロ碁の弱点を克服するために，2 点間の関係性に着目したプレイアウト中の情報抽出を提案した．提案手法を実装したプログラムで，いくつかの局面を計算させる実験を行った結果,「見合い」や「着手順序」を考慮した着手の優劣を測ることができた．

順序 Criticality においては，局面によって JC の最大値は異なるため，絶対値を見て「ある値以上を見合いである」という認識の仕方はできなかった．しかし，明らかに見合いである手は，その局面の JC の中では最大になり，見合いでない手との間には割合的なギャップを含むことが見て取れた．そのギャップを見つけることで，局面中の見合いを機械的に認識できそうである．

今後の計画として，提案手法を用いて，実際にプレイアウトの確率を制御することが挙げられる.「見合い」であれば，プレイアウト中の直前の手を参照し，順序 Criticality が高い手に打つ確率を上げることで，より自然なプレイアウトが行えるはずである.「着手順序」を考慮した順序 Criticality のランクを用いれば，内ダメ外ダメがある局面でも，外ダメを優先して打つことで，自ら不利になる手の確率を下げることができる．以上の手法を用いることで，探索の効率が上がり，複雑な局面でも有効な着手の発見ができると考える．

謝辞

本研究は，JSPS 科研費 15K00506 の助成を受けた．ここに感謝する．

参考文献

[1] Rèmi Coulom.: "Criticality: a Monte-Carlo Heuristic for Go Programs", Invited talk at the University of Electro-Communications, Tokyo, Japan (January 2009).

[2] Rèmi Coulom.: "Computing Elo Ratings of Move Patterns in the Game of Go", ICGA Journal Vol.30 pp.198-208, 2007.

[3] Sylvain Gelly, David Silver.: "Combining Online and Offline Knowledge in UCT", ICML 2007.

[4] David Silver *et al.*, "Mastering the game of Go with deep neural networks and tree search", Nature 529, 484-489 (28 January 2016).

