JAIST Repository
https://dspace.jaist.ac.jp/

Title	囲碁における大局観を実現する広域パターンマッチング
Author(s)	本田，拓朗
Citation	
Issue Date	2013-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/11324
Rights	
Description	Supervisor:池田心，情報科学研究科，修士

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

修 士 論 文

囲碁における大局観を実現する広域パターンマッチング

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

本田 拓朗

2013 年 3 月

修 士 論 文

囲碁における大局観を実現する広域パターンマッチング

指導教官　池田心 准教授

審査委員主査　池田心 准教授
審査委員　飯田弘之 教授
審査委員　白井清昭 准教授

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

1010059 本田 拓朗

提出年月: 2013 年 2 月

Copyright © 2013 by Takuro Honda

概 要

チェスやオセロにおいては人間の世界チャンピオンより強いコンピュータプログラムが開発されているのに対し，囲碁においては長年にわたって棋力は伸びず，およそアマ初段ほどだった．しかし近年モンテカルロ木探索と呼ばれる新しい手法が考案され，その棋力は驚くべき向上を見せた．

モンテカルロ木探索（MCTS）を用いたコンピュータ囲碁においては，棋譜中の着手を入力とする機械学習によって行動評価関数を調整することで，木探索部で探索する枝の絞り込みやプレイアウト部での着手確率の調整を行い棋力を向上させる研究が広く行われている．学習に用いる重要な要素として，着目点の周囲の “配石パターン” があり，棋譜中に頻繁に現れるものを抽出し，重み付けすることによって着手の評価に用いられる．着手を正確に評価するためには盤上のより広い範囲をカバーする大きいパターンが必要であるが，大きいパターンは棋譜中に出現する回数が少ないため，十分に学習するためには非常に多くの棋譜が必要になるという宿命的な課題がある．

この問題に対し，本稿では，盤面を縮小することによって「広い範囲をざっと見る」新たなパターンマッチングを提案する．盤面の縮小は，盤面の勢力分布を求めること，勢力分布を分割しそれぞれの領域を平均することで縮小すること，しきい値を設けて勢力分布を 4 値化することからなる．盤上のある交点に対してパターンマッチングを行う際は，盤上の石の配置をそのままパターンとして用いるのではなく，縮小した盤上の対応する交点を中心とする 3×3 の範囲がパターンとして用いられる．

本手法の性能は盤面の区切り方に大きく影響を受けるため，発展的な手法として，異なる区切り方によって縮小された複数の盤面を用意し，それら全ての上でパターンマッチングを組み合わせるという手法も考案した．

提案手法を用いて性能評価として，機械学習時の性能，大局観が要求される典型的な局面での着手，自己対戦の 3 つの項目で従来手法との比較を行った結果，提案手法の優位性が示された．また，発展手法は機械学習時の性能においてさらに高い効果を示した．

目 次

第 1 章　はじめに **1**
1.1　背景 1
1.2　大局観とは 2
1.3　論文の構成 2

第 2 章　関連研究 **4**
2.1　モンテカルロ囲碁 4
2.1.1　プレイアウト 4
2.1.2　モンテカルロ木探索 5
2.1.3　特徴の利用 5
2.1.4　パターンマッチングのアルゴリズム 6
2.2　ファジーパターンマッチング 7
2.3　コラージュパターンマッチング 9

第 3 章　広域パターンマッチング **10**
3.1　盤面の縮小 10
3.1.1　勢力分布の計算 10
3.1.2　勢力分布の縮小 10
3.1.3　縮小された勢力分布の 4 値化 12
3.2　縮小された盤上のパターンの利用 13
3.3　位置による分類と回転による同一視 14
3.4　発展手法 CMAC 14

第 4 章　評価実験 **16**
4.1　棋譜中の着手の機械学習時の性能 16
4.2　大局観が要求される典型的な局面での着手 18
4.3　自己対戦 21

第 5 章　まとめ **23**

参考文献 **26**

図 目 次

2.1 プレイアウトの例 4
2.2 MCTS における探索中の木の様子 5
2.3 特徴の利用 6
2.4 nomitan における距離の例．数字は各点と▲との距離 7
2.5 狭すぎるパターンが悪い結果を導く例 8
2.6 ファジーパターンマッチング．パターンの周囲 8 方向それぞれが白，中間，黒のどれであるかを利用する． 8
2.7 コラージュターンマッチング．x の地点について 4 つのパターンをマッチさせる．この組み合わせとして中央にある広い領域を見ていることになる． 9

3.1 盤面の例 11
3.2 勢力分布とその分割 11
3.3 図 3.2 の左上隅の拡大図 11
3.4 盤面の縮小 11
3.5 石の数 12
3.6 4 値化された盤面 12
3.7 図 3.6 のパターンがカバーするもとの盤上の範囲 13
3.8 盤上の位置の区別 14
3.9 対称な形のパターン 14
3.10 Puddle World 15
3.11 CMAC 15
3.12 複数の区切り方を重ねた広域パターンマッチング（CMAC）の例 15

4.1 CMAC の実験で用いた 3 つの区切り方 17
4.2 大局観が要求される典型的な局面 18
4.3 図 4.2 の局面での着手の順位付け 19
4.4 相手の勢力を弱めるべき局面での着手の順位付け 20
4.5 相手の勢力を内側から荒らせる局面での着手の順位付け 20
4.6 自分に弱石がある局面での着手の順位付け 21
4.7 自己対戦方法 22

表 目 次

3.1　4 値化 13

4.1　広域パターンマッチングの汎化性能 17

第1章　はじめに

1.1　背景

人間の世界チャンピオンに勝利するほどの強さを持ったプログラムが登場しているチェスやオセロ，かつてのトッププロに勝利したプログラムが近年現れた将棋などとは違い，コンピュータ囲碁プログラムの棋力は長きにわたってアマ初段ほどに留まってきた．しかし2006年頃から，モンテカルロ法を用いたシミュレーション[1]と，UCB1(Upper Confidence Bound) 値を用いた木探索であるUCT(UCB applied to trees)[2]を組み合わせた，モンテカルロ木探索（MCTS）という手法[3]が考案されて以来，その強さは大きく向上した．

MCTSにおいては，モンテカルロ法を用いた終局までのシミュレーション（プレイアウト）での着手を完全なランダムで選択するのではなく，良さそうな手ほど高い確率で打つように着手確率を偏らせることで棋力が向上することが知られている．また，UCTにおいても同様に良さそうな手を優先的に探索を行うProgressive Widening[4]という手法が提案されている．手がどれほど良さそうかを測るためには，着手から様々な「特徴」を抽出し，プロや高段者の棋譜中の着手を再現するようにそれらを重み付けする教師あり機械学習がよく用いられており，Bradley-Terryモデルによる手法[4]やシグモイド損失関数による勾配法[5]などの手法が提案されている．

着手の特徴としては，着手位置の盤端からの距離，直前の相手の着手との距離，アタリやノビなどの囲碁独自の概念などが主に利用される．着手点の周囲の配石パターンも特徴としてしばしば利用され，着手が望ましい3×3のパターンを手作業で列挙する[3]，棋譜中に頻繁に出現するものを抽出して機械学習で重み付けする[4]などの手法が提案されている．パターンを特徴として利用する際，より正確に着手を評価するためには，着手点の周囲をより広い範囲で把握する必要があるため，マッチングする範囲を大きくする必要がある．しかし，大きいパターンはその出現回数が少なくなるため，パターンの良し悪しを正確に評価するのが困難になる．そのため，大きいパターンでも十分な回数出現するよう，非常に多くの棋譜に使う必要がある．このように，パターンマッチングには広さとサンプル数を両立できないという宿命的な課題がある．

この問題を解決するために，パターンの周囲8方向についてそれぞれどちらの色の勢力圏かという特徴を利用する手法[6]や，着目点を中心としない複数のパターンを組み合わせる手法[7]が提案されているが，棋譜中の着手との一致率はそれぞれ0.013，0.03，低下してしまっている．そこで本研究では，これらに代わる広い範囲のパターンマッチングを行う手法として，「盤面を縮小することにより，広い範囲をざっとマッチングする」パ

ターンマッチングを提案する．

1.2　大局観とは

日本棋院から出版されている用語辞典 [8] では大局観は「全体を把握して進行状態，形勢などを判断し，方針，作戦を立てる能力」であると説明されている．人によってはこれと多少異なる意味に捉えていることもあるだろうが，概ね盤面の広範囲あるいは全体を把握する点では共通すると考えられる．

本研究では「より広い範囲を」カバーするパターンマッチングを実現することを目的とするが，パターンの出現回数の低下や処理速度との兼ね合いのため，上記の説明のように「盤面全体」をカバーすることはしない．また，説明の後半部にある「方針や作戦の立案」も一般的な MCTS の枠組みのまま行われるものとする．あくまで MCTS の枠組みで利用可能な広範囲のパターンマッチングの実現と，パターンのサイズを広くすることによる問題の解決を目指す.

仮に方針や作戦の立案などを MCTS の枠組みとは別に取り扱う研究をする場合，それには広い範囲の盤面の把握が必須となるため，その点においても本研究は重要である．

1.3　論文の構成

本論文の構成は以下のようになっている．

第 2 章 関連研究

本研究で用いる nomitan[5] を含む多くの囲碁プログラムで用いられている MCTS のアルゴリズムと，棋譜からパターンなどの特徴をどのように学習して利用するかについて述べる．また，本論文が対象とする「広いパターンを用いることによるパターン出現回数の低下」という問題に対する既存の研究について述べる．

第 3 章 広域パターンマッチング

本論文が提案する広域パターンマッチング，およびその発展手法として CMAC との組み合わせについて述べる．

第 4 章 評価実験

通常のパターンマッチングだけを用いた棋譜の学習を行った場合と，それに加えて広域パターンマッチングを用いた棋譜の学習を行った場合で学習時性能および実際の着手を比較し，本手法の有効性を確認する．また，自己対戦での勝率を調べることで実際に棋力が向上するかも確認する．

第 5 章 まとめ

全体のまとめと今後の課題，展望を述べる．

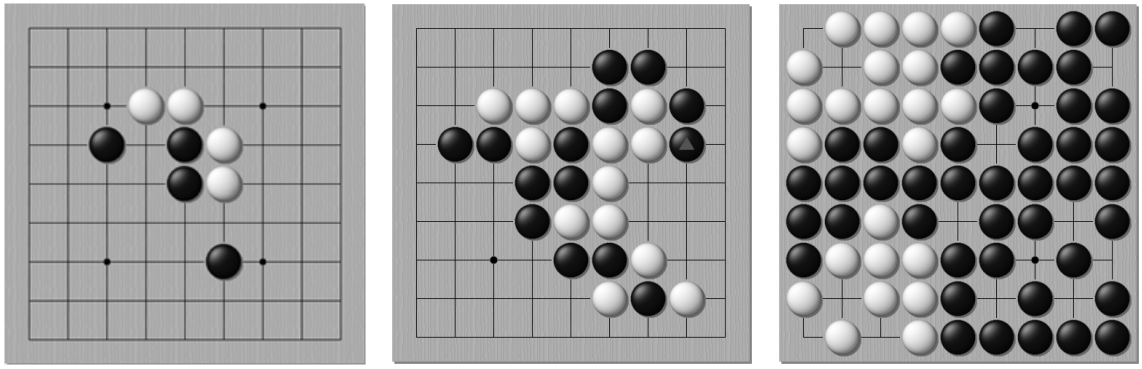
第2章　関連研究

2.1　モンテカルロ囲碁

2.1.1　プレイアウト

動きや価値の異なる駒が何種類もあるチェスや将棋と異なり，囲碁では碁石の種類が黒石と白石の 2 種類しかなく，個々の石に性質の違いは無いため，局面の静的評価関数を作るのは困難である．盤面が広いため，オセロのように盤上のある位置だけ特に価値が高いといったこともないのも，困難さを高める要因となっている．そのため，図 2.1 のように局面から終局までのランダムな対局のシミュレーションを何度も繰り返し行い，その勝率などを静的評価関数値の代わりに用いて着手を評価することが行われる．このようにモンテカルロ法を用いて終局までのシミュレーションを行うことをプレイアウトと呼ぶ．

最も単純なプレイアウトは，ルールで禁止されている着手と自分の石の眼を埋める着手以外のすべての着手を等確率に行うものである．囲碁というゲームの性質上，これだけでも終局までのシミュレーションが行え，終局すれば地を数えることで勝敗も分かる．しかし，棋力を向上させるためには，より良さそうな手ほど高い確率で打つように着手確率を調整することが有効であると知られており，そのために，棋譜からの機械学習で手の良さを推測できるようにするなど様々な研究がなされている．



(a) プレイアウト前　　(b) プレイアウト途中　　(c) プレイアウト後

図 2.1: プレイアウトの例

2.1.2　モンテカルロ木探索

多くの囲碁プログラムでは前述のプレイアウトに UCT(UCB applied to trees)[2] という木探索を組み合わせた，モンテカルロ木探索（MCTS）[3] という手法が利用されている．UCT は，式 2.1 のような UCB1(Upper Confidence Bound) という値が最も高いノードを選びながら葉まで探索していく探索法である．ここで，$\overline{X_j}$ は j 番目のノードの勝率，n_j は j 番目のノードの訪問回数，n は親ノードの訪問回数である．c は定数で 0.5～1.0 程度の値が用いられる．

$$\overline{X_j} + c\sqrt{\frac{2\log n}{n_j}} \tag{2.1}$$

モンテカルロ木探索では，一定回数以上葉ノードが訪問されたらその下に新たな子ノードを作成することで，木の成長が行われる．さらに，葉ノードからプレイアウトを行い勝敗を求め，その結果を報酬として辿ってきたノードの値の更新も行われる．探索中の木の例を図 2.2 に示す．図中の各ノードの数字はその手番のプレイヤーの勝利数とそのノードの訪問回数である．

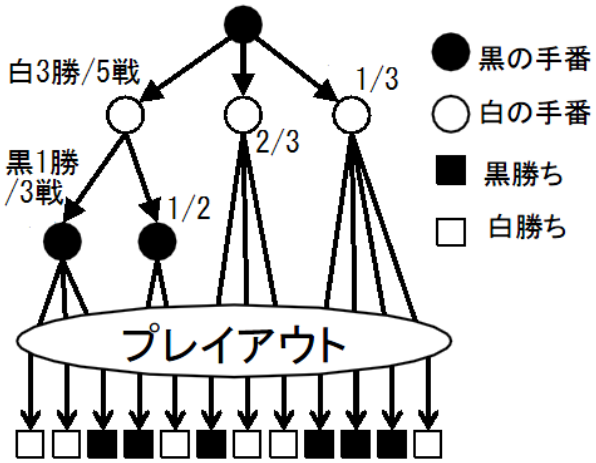
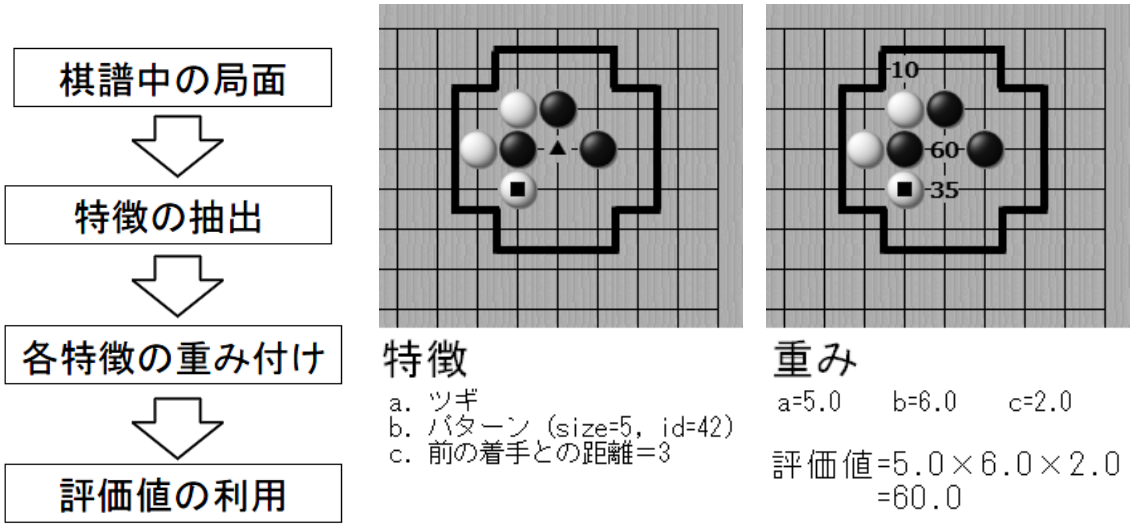


図 2.2: MCTS における探索中の木の様子

2.1.3　特徴の利用

MCTS を用いた囲碁プログラムを強くするには，何らかの方法でプレイアウトでの着手選択確率調整や UCT での探索する枝の絞込みを行うと高い効果がある．[4] その主な方法としては，着手の良さを評価する行動評価関数を用いた，プレイアウトや探索中に現れた局面の全ての合法手がどの程度打ちたい手であるかの点数付けが挙げられる．

図 2.3 に特徴がどのように利用されるかの概念図を示す．全体としては図 2.3(a) のような流れで行われる．機械学習では，棋譜中の実際に打たれた着手を含む全ての合法手から特徴を抽出し，実際に打たれた着手から抽出された特徴の重みが大きくなるように各特徴の重み付けを行う．棋譜中の盤面から特徴が抽出される様子を図 2.3(b) に示す．この図において▲が特徴量を抽出する対象の合法手であり，■は相手が直前に打った着手である．対局時に着手の評価を行うときは図 2.3(c) のように，その着手に当てはまる特徴をすべて列挙し，その重みをすべて掛け合わせることで評価値を求める．各合法手から抽出される特徴量としては「アテ」や「ツギ」などの囲碁独自の概念や,「直前手からの距離」や「石の配置パターン」などが用いられることが多い．距離の定義は 2.1.4 節で述べる．



(a) 全体の流れ　(b) 特徴の抽出の例　(c) 着手の評価値の例

図 2.3: 特徴の利用

2.1.4 パターンマッチングのアルゴリズム

本研究で用いる囲碁プログラムである nomitan[5] では Stern ら [9] と同様，棋譜からパターンを抽出する際，複数のサイズのパターンを抽出しておき，利用時はマッチした中で最大サイズのパターンを用いている．学習では，一定回数上登場しないパターンは学習には利用しないため，頻繁に登場するパターンは大きいサイズでマッチングされ，あまり登場しないパターンは小さいサイズでマッチングされる．nomitan では図 2.4 のように着目点から (x, y) だけ離れた点の距離を $d(x,y) = x + y + \max(x,y) \leq n$ と定義し，距離が $d(x,y) \leq n$ である交点からなるパターンを利用している．

パターンマッチングにおいてパターンの**サイズ**，つまりどれだけ広い範囲を見るかは重要な要素である．サイズが小さすぎる場合，視野の狭い手を打ってしまうことにつなが

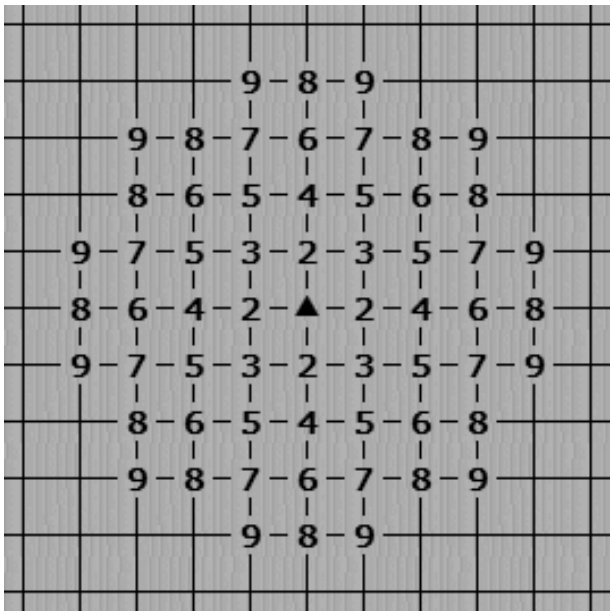


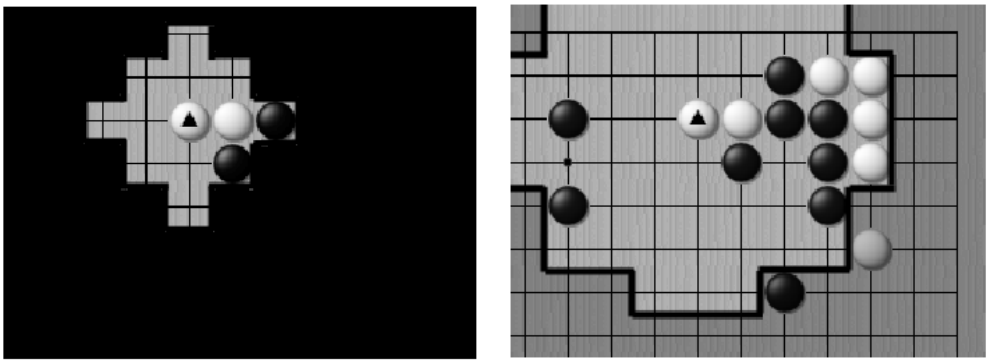
図 2.4: nomitan における距離の例．数字は各点と▲との距離

る．例えば図 2.5 のように，狭い範囲で見ると好手に思えるが，より広い範囲で見ると相手の勢力圏の中に打つため実際は悪手だという事例は多い．この場合，図 2.5(a) のような狭い範囲で見ると，黒のハネに対して白▲とノビる手は常識的な手のひとつであり，決して悪くはない．しかし，広い範囲で見た石の配置が図 2.5(b) であった場合，黒の強い勢力圏から逃げ出すわけでもなく，かといって中で生きられるほど空間があるわけでもなく，中途半端な手だということが分かる．

このようなことがあるため，正確な着手の評価のためにはサイズを大きくする必要がある．一方，サイズが大きい場合は様々な問題も発生する．まず，大きいパターンほど棋譜中に出現する頻度が低下するため，評価値を正確に決めるには学習に大量の棋譜が必要になる．また，マッチした中で最大サイズのパターンを用いるようにするためには，学習の際に最大サイズ以下の全てのパターンを抽出する必要があるため，大量の棋譜を学習に用いるとパターンがメモリに収まり切らないことも考えられる．この「大きいパターンを使いたい」「多くのサンプルがほしい」という相反する希望はパターンマッチングに宿命的な課題である．

2.2　ファジーパターンマッチング

2.1.4 の最後で述べた問題に対し，荒木ら [6] はファジーパターンマッチングという手法を提案している．これは通常のパターンマッチング（荒木らも nomitan と同じく，Stern らの研究同様に複数のサイズの中からマッチした最大のものを利用する手法を用いている）



(a) 小さいパターンの場合良い手に見える

(b) 大きいパターンの場合悪い手だと分かる

図 2.5: 狭すぎるパターンが悪い結果を導く例

に加え，そのパターンの周囲 8 方向それぞれが，黒っぽいか，白っぽいか，中間かという情報を用いるものである．

それぞれの方向の色を求めるには，まず図 2.6 のように，パターンから見てどの方向にあるかによって盤面を 8 つの領域に分ける．そして，式 2.2，式**??**のように領域内の交点ごとに算出した影響度の総和を計算する．最後に，閾値を設定することで黒，白，中間の 3 通りに領域を分類する．この式における距離の定義が何であるかは荒木らの論文には明記されていないが，本文中の図を見るにマンハッタン距離を用いているようである．

以上のような手法を用いて荒木らは棋譜の学習の実験を行ったが，棋譜中の着手との一致率で比較して， 通常のパターンマッチングが 21.0% だったのに対し，ファジーパターンマッチングでは 19.7% であり，良い結果は得られていない．

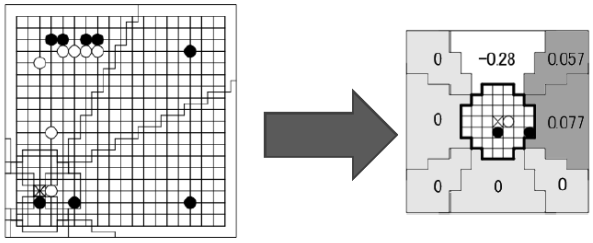


図 2.6: ファジーパターンマッチング．パターンの周囲 8 方向それぞれが白，中間，黒のどれであるかを利用する．

$$\text{領域}\,A_i\text{の値} = \sum_{s \in A_i} \frac{c(s)}{s\,\text{までの距離}} \tag{2.2}$$

$$(i = 0, ..., 7)$$

$$c(s) = \begin{cases} 1 & (s: \text{黒石}) \\ -1 & (s: \text{白石}) \\ 0 & (s: \text{その他}) \end{cases} \tag{2.3}$$

2.3 コラージュパターンマッチング

同じ問題に対し，土井 [7] はコラージュパターンマッチングという手法を提案している．これは，大きいパターンによるマッチングを行う代わりに，小さいパターンを用いたマッチングを複数組み合わせて用いる手法である．

図 2.7 のように，マッチングする交点の周囲を，左上，左下，右上，右下の 4 つに分割し，それぞれの領域に対してパターンマッチングを行う．分割によって個々のパターンのサイズは小さくなり，出現回数の増加が期待できるため，図中央のような大きな領域のパターンマッチングを用いるよりは高い効果が期待できる．しかし棋譜の学習の実験を行ったところ，棋譜中の着手との一致率で比較して，通常のパターンマッチングが 31.7% だったのに対し，コラージュパターンマッチングでは 28.6% であり，良い結果は得られなかった．ただし，目的関数をコラージュパターンマッチング用に変更して学習した実験も参考として行なっており，そちらでは 35.0% の一致率が得られている．

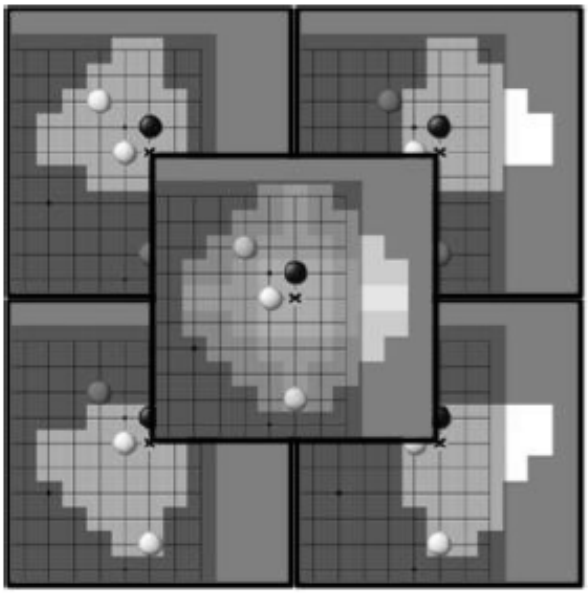


図 2.7: コラージュターンマッチング．x の地点について 4 つのパターンをマッチさせる．この組み合わせとして中央にある広い領域を見ていることになる．

第3章　広域パターンマッチング

前章の問題を解決するために，本研究では「広い範囲をざっとマッチングする」パターンマッチングを提案する．本手法では，19×19 の盤面を後述のアルゴリズムで $7\times$ に縮小し，その盤上でパターンマッチングを行う．これにより，小さいサイズのパターンマッチングでも元の盤面の広い範囲をカバーできる．また，もとの盤上で形が似通っているパターンは縮小した盤上では同一のパターンになることが期待できるため，より少ない棋譜を用いた学習でも十分なサンプル数が期待できる．

3.1　盤面の縮小

図 3.1 の盤面を例に盤面縮小のアルゴリズムを説明する．

3.1.1　勢力分布の計算

盤上の全ての交点について，それぞれが周囲にある石から受ける影響力を計算する．交点 p から距離 a 以内にある n 個の石を $s_1, s_2, ..., s_n$ とし，p とそれぞれの石の距離（図 2.4）を $d_1, d_2, ..., d_n$ とするとき，p が $s_1, s_2, ..., s_n$ から受ける影響力を求める関数 f を以下のように定義する．

$$f(p) = \sum_{i=1}^{n} \frac{g(s_i)}{max(d_i, 1)}$$

$$g(s) = \begin{cases} 1 & (s\ is\ black) \\ -1 & (s\ is\ white) \end{cases}$$

$a = 5$ として図 3.1 から求めた勢力分布を図 3.2 に示す．また，勢力の数値が潰れて見えなくなっているので，左上隅を拡大したものを図 3.3 に示す．黒の勢力圏は正の値に，白の勢力圏は負の値になり，勢力が強いほどその絶対値が大きくなる．

3.1.2　勢力分布の縮小

勢力分布を図 3.2 のように線で区切り，それぞれのマス目の平均を求めることで 7×7 の勢力分布へ縮小する．縮小後の勢力分布の交点の値 q は，マス内の m 個の交点を $r_1, r_2, ..., r_m$

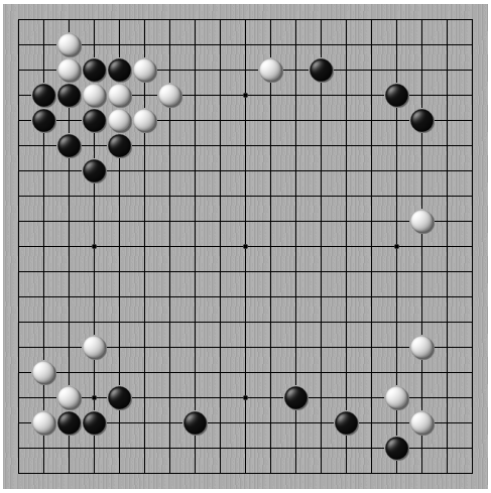


図 3.1: 盤面の例

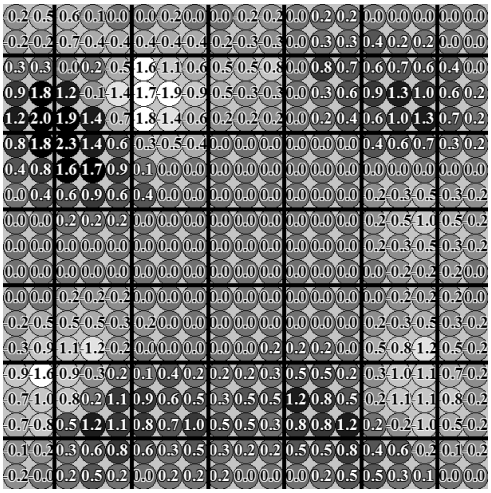


図 3.2: 勢力分布とその分割

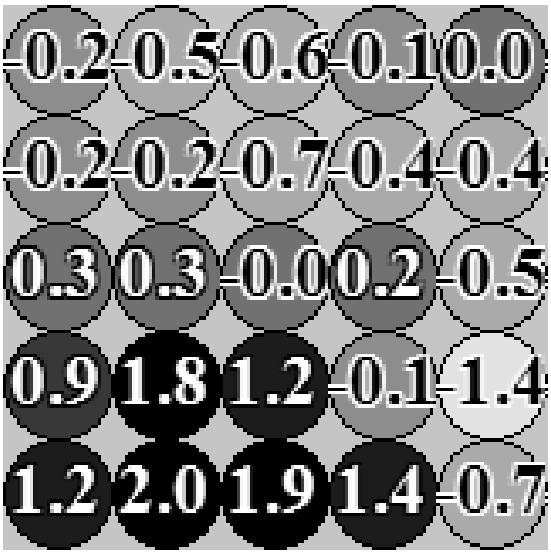


図 3.3: 図 3.2 の左上隅の拡大図

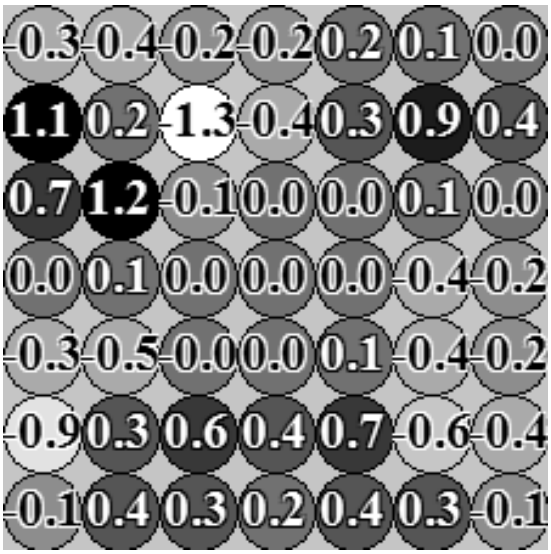


図 3.4: 盤面の縮小

として，以下の式で求められる．図 3.2 を縮小した結果として図 3.4 のような 7×7 の盤面ができる．

$$q = \frac{\sum_{i=1}^{m} f(r_i)}{m}$$

前項のように勢力分布を求めてからマス目に区切っているのは，縮小後の交点の値に出来るだけ広い範囲の盤面の状態を反映させたいからである．例えば元の盤面をそのまま用いて，単に { 黒石＝ 1，白石＝ −1，空点＝ 0} などとして平均を取ると，縮小後の交点には最大で 3×3 の範囲にある石しか影響しない．

3.1.3　縮小された勢力分布の 4 値化

縮小された盤上の各交点を表 3.1 によって { 黒の勢力圏，どちらの勢力圏でもない，両者の勢力が拮抗，白の勢力圏 } の 4 値に分類する．ここで，石数とはもとの盤面（図 3.1）を図 3.2 と同じように区切り（図 3.5），縮小された盤上の交点に対応するマスにある石の数である．4 値化された盤面は図 3.6 のようになる．

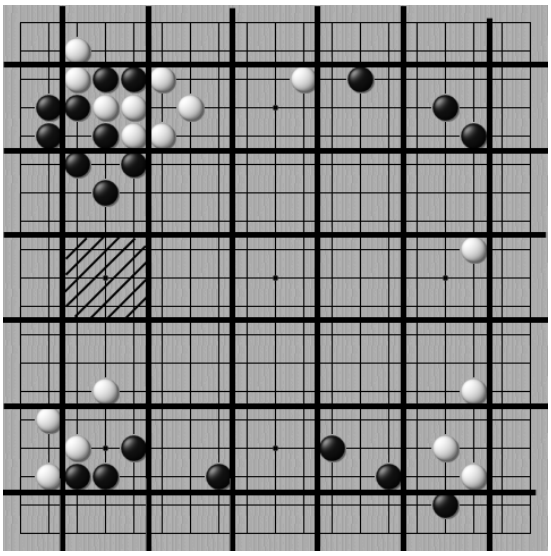


図 3.5: 石の数

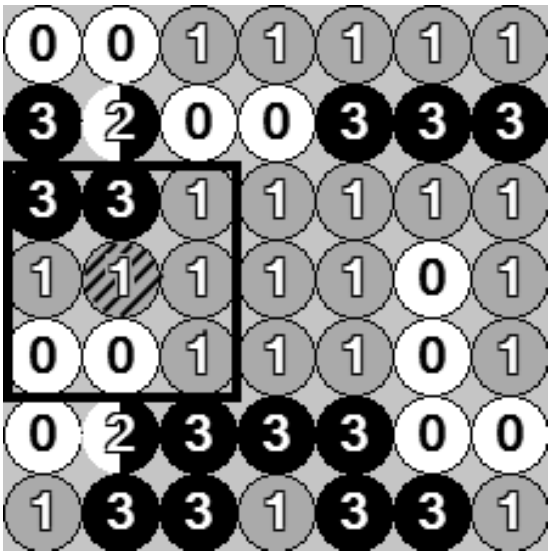


図 3.6: 4 値化された盤面

本手法では各交点を荒木ら [6] のように { 黒の勢力圏，中間，白の勢力圏 } の 3 値ではなく，前述のように 4 値に分類している．これは，特に序盤の局面においては，両者の石が近くに集まっている場所よりはどちらの石も周囲にない場所を先に占めることで勢力を築くことが重要であることが多いためである．3 値では区別できない「近くに集まっている」と「周囲にない」を 4 値では区別できるため，学習時の性能向上が期待できる．

表 3.1: 4 値化

もとの値	石数	値	意味
$-b$ 未満	-	0	白の勢力圏
$-b$ 以上 b 未満	c 未満	1	どちらの勢力圏でもない
	c 以上	2	両者の勢力が拮抗
b 以上	-	3	黒の勢力圏

3.2 縮小された盤上のパターンの利用

もとの盤上のマス内の着手可能な交点に対し，縮小された盤上のパターンを特徴として利用する．盤上のそれぞれの交点には，それを含むマスに対応する縮小した盤上の交点を中心とする 3×3 マスのパターンが対応する．例えば図 3.5 で網掛けになっているマスに含まれる交点に対応するのは，図 3.6 で網掛けになっている交点であり，それを中心とする 3×3 マス，すなわち□で囲まれた部分のパターンが特徴として利用される．

縮小した盤上では 3×3 マスという小さいサイズであっても，図 3.7 のようにもとの盤上でカバーする範囲は非常に大きくなる．この図において，実線で囲まれた範囲が縮小した盤上の 3×3 の範囲に相当し，縮小アルゴリズムのうち勢力分布計算のステップを考慮すると，その外側の点線の範囲までがパターンに影響する．

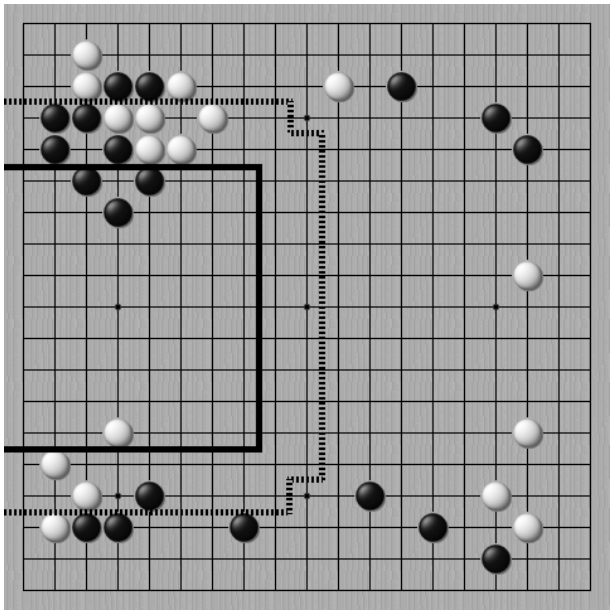


図 3.7: 図 3.6 のパターンがカバーするもとの盤上の範囲

3.3　位置による分類と回転による同一視

囲碁においては同じ配石パターンでもそれが盤の隅に現れた場合，辺に現れた場合，中央に現れた場合で着手の価値が異なることがままある．そこで図 3.8 のように，盤のどのあたりにあるかを表す値も特徴として利用することで，同じ配石でも位置によって別のパターンであると扱えるようにする．出現しうるパターンの総数は $4^9 * 8$ 通りになる．

一方，図 3.9 の左上隅と右下隅のように，本質的に同一で回転させれば完全に一致するようなパターンは同一のものとして扱う．これによって，パターンのサンプル数を増やし，より正確にパターンの良し悪しを評価できるようになる．

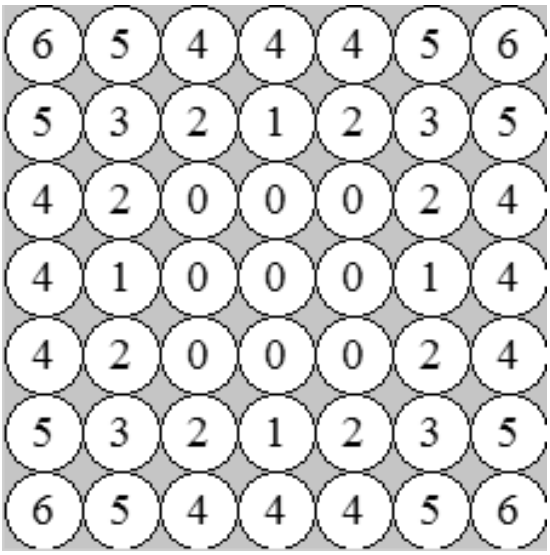


図 3.8: 盤上の位置の区別

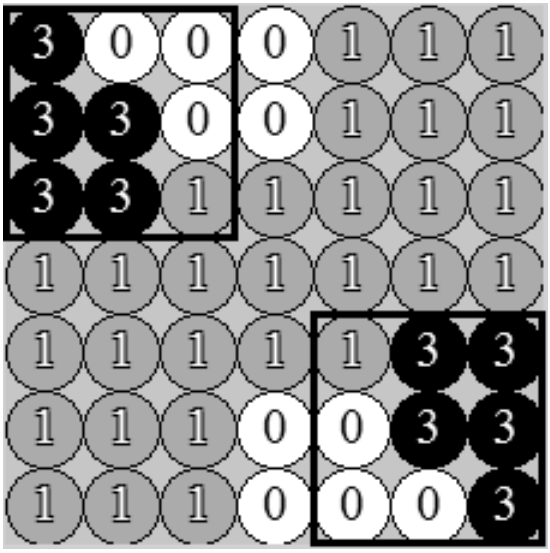


図 3.9: 対称な形のパターン

3.4　発展手法 CMAC

CMAC(Cerebellar Model Articulation Controller)[10] とは，ニューラルネットワークの一手法であり，状態空間をいくつかの「セル」に区切って学習することによって高い性能を実現できるアルゴリズムである．Sutton[11] は Puddle World という経路探索問題を定義し，それに対して CMAC を用いた教科学習を適用し高い成果を挙げた．Puddle World は図 3.10 のようなフィールドで，スタート地点にある点を 1 ステップごとに上下左右のどれかの方向に 1 単位だけ動かし，最短ステップでゴール地点まで移動させる問題である．ただし，灰色の部分に侵入するとペナルティとしてステップ数が加算される．このような問題に対し Sutton は，図 3.11 の用にフィールドを複数のセルに区切ったものを，ずらして複数重ねることで学習を行った．

本手法でも勢力分布を縮小する際に，それを分割するということを行なっている．そのため，Sutton の手法と同様，図 3.12 のように複数の区切り方を用いて縮小し，それぞれ

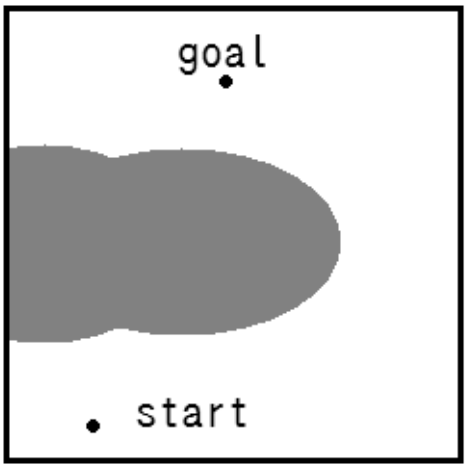


図 3.10: Puddle World

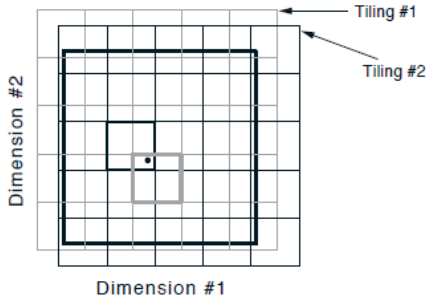


図 3.11: CMAC

の縮小盤面上でパターンマッチングを行うことで高い効果が期待できる．そこで，次章ではこれまで説明した広域パターンマッチングを使った実験に加えて，この複数の区切り方を用いた広域パターンマッチングを使った実験も行った．1 つの天に対して複数のマッチングを行うという意味では土井の手法 [7] とも関連性が高い．本稿ではこれ以降，特に断らない限りは，Albus や Sutton の手法ではなくこの手法を CMAC と呼称する．

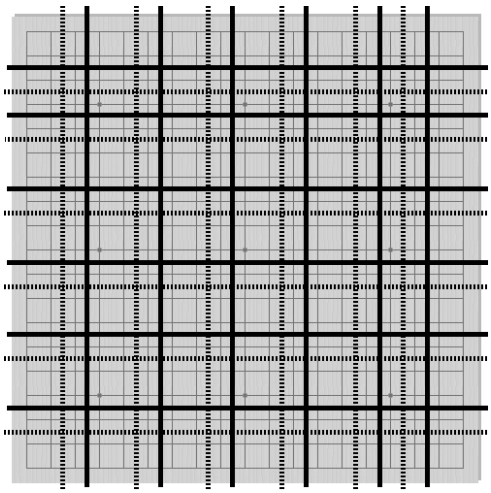


図 3.12: 複数の区切り方を重ねた広域パターンマッチング（CMAC）の例

第4章　評価実験

4.1　棋譜中の着手の機械学習時の性能

囲碁では序盤ほど大きいサイズのパターンがマッチしやすい傾向がある [12]．そのため，より大きいパターンマッチングを実現する本手法も**序盤〜中盤**に効果があると考えられる．そこで，棋譜の 50 手目までについて本手法と従来のパターンマッチングを組み合わせて，木探索に用いる行動評価関数の機械学習を行い，従来のパターンマッチングのみを用いた場合と性能を比較した．従来のパターンマッチングでは距離 9 までの範囲（図 2.4）の配石に対してマッチングを行った．パターン以外の特徴としては，着手によるダメの変化と， 直前の着手との距離を用いている．学習のモデルには Bradley-Terry モデルを，最適化には勾配法を用いている．棋譜は，学習用に 29 万枚，性能評価用に 1 万枚の計 30 万枚を用いた．対局時に学習に用いた棋譜中の局面ばかり登場するということは無いので，性能評価には学習に用いたものとは別の棋譜を用いている．このように別の棋譜を用いて測った性能を本稿では汎化性能と呼ぶ．

本手法のアルゴリズムには前章で示したように，影響力の範囲 a，4 値化のしきい値 b，石数のしきい値 c という 3 つのパラメータがある．この実験では $a = 5, b = 0.01, c = 7$ という値を設定して行った．これらの値は本章でこれから述べるのと同様の実験を，学習用に 9 万枚，テスト用に 1 万枚の計 10 万枚の棋譜を用いて，a, b, c の値を様々に変えて行なって得たものである．b の値は非常に小さく，0 にしても良いように思われるが，実際 0 にして実験をしてみると 0.01 の場合と比べて性能は大きく低下した．

学習した係数の評価は以下に示す一致率，平均ランク，20 位以内率，平均対数尤度の 4 つで行った．

1. **一致率**
   教師信号（棋譜中の着手）の評価値が最も高くなった割合を一致率と呼ぶ．高いほど良い．

2. **平均ランク**
   教師信号が何番目に良い手だと評価されたかの平均を平均ランクと呼ぶ．順位なので数字が若いほど良い．

3. **20 位以内率**
   教師信号が 20 番目以内に良い手だと評価された確率を 20 位以内率と呼ぶ．21 番目以降の手は探索中に全く考慮されない可能性が高い．高いほど良い．

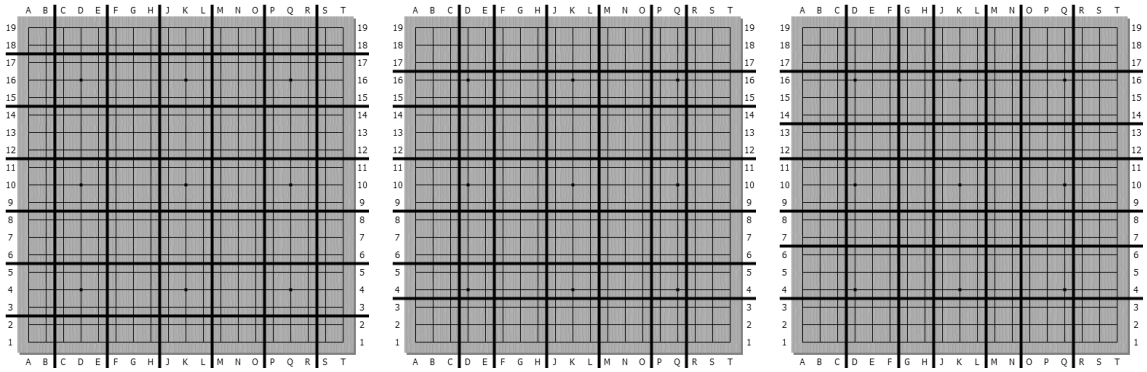
4. **平均対数尤度**

教師信号の選択確率の対数を取り全棋譜について平均を取ったものを平均対数尤度（Mean Log Evidence）と呼び，他論文 [4][9] でも評価に用いられている．もし教師信号の確率が 100%であれば 0 となるので，0 に近いほど良いといえる．これは学習時の目的関数（最大化）でもある．

以上 3 つの指標を用いて，テスト用の汎化性能を比較した結果を表 4.1 に示す．いずれの指標においても，広域パターンマッチングを通常のパターンマッチングと組み合わせたものの方が，通常のパターンマッチングのみを用いたものより，**良い性能を発揮した**ことが分かった．また，4.1 に示すような 3 つの区切り方を使った CMAC を用いて学習を行った結果，より良い性能を発揮した．参考として，$b = 0.00$ にして縮小後の盤面を事実上 3 値化した場合の性能も掲載する．通常のパターンマッチングのみの場合よりは性能が向上したものの，4 値の広域パターンマッチングを組み合わせたものと比べると大きく劣った．

表 4.1: 広域パターンマッチングの汎化性能

	通常のみ	通常＋広域	通常＋ CMAC	通常＋ 3 値広域
一致率	0.405	**0.420**	**0.423**	0.409
平均順位	8.587	**7.922**	**7.778**	8.276
20 位以内率	0.888	**0.906**	**0.909**	0.895
平均対数尤度	-2.432	**-2.366**	**-2.345**	-2.402



(a) 2333332^2　　(b) 3233323^2　　(c) 3323233^2

図 4.1: CMAC の実験で用いた 3 つの区切り方

4.2　大局観が要求される典型的な局面での着手

大局観は序盤において特に重要な要素であり，好手を打つために大局観が必要な局面はたびたび出現する．例えば図 4.2 のような局面（白番）では▲付近の点にワリウチ打つのが好手である．ワリウチとは主に相手の勢力圏の形成を妨げる目的で三線あるいは稀に四線に打たれる手で，両側に二間ビラキ（4.2 では■の点）を打つ余裕があるものをいう．逆に黒から▲に打たれた場合，右辺のほとんどが黒の大きな勢力圏になる．プログラムでこのことを検知するためには，右上隅の小ゲイマジマリと右下隅の小目定石によって，両方の隅に黒の強い勢力があることを検知する必要がある．

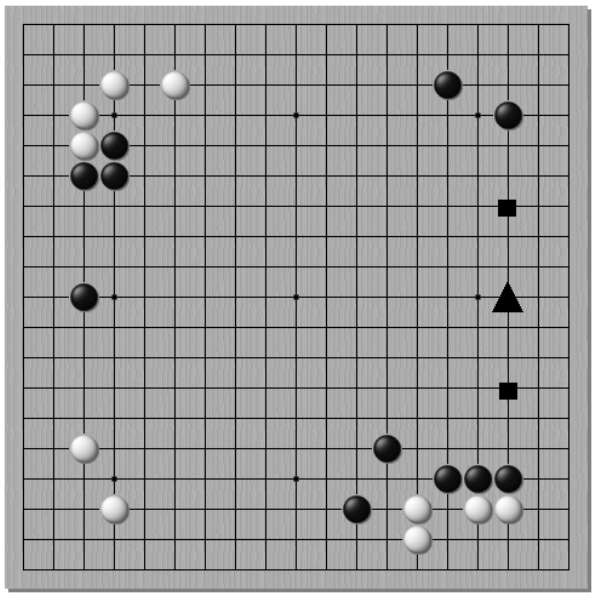
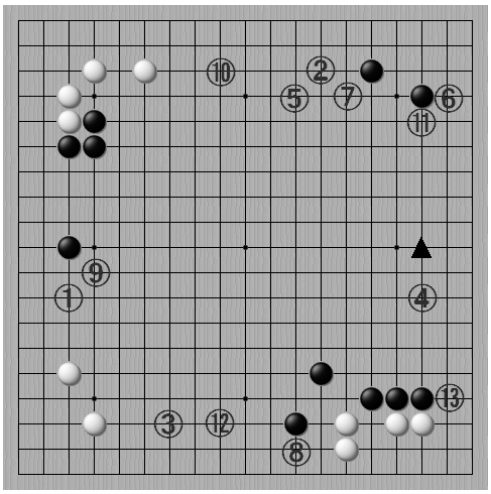


図 4.2: 大局観が要求される典型的な局面

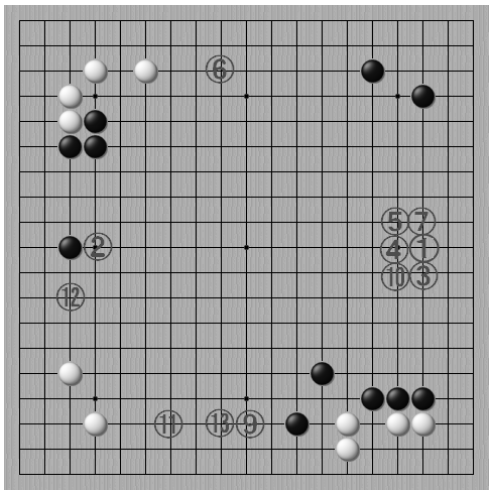
本手法の導入によって，この局面での着手が従来の nomitan と比べてどのように変わったかを調べた．図 4.3 は図 4.2 の局面において，本手法を用いない nomitan と本手法を用いた nomitan それぞれに，評価値が高い順に着手を順位づけさせたものである．図 4.3(a) に本手法を用いない場合の着手の順位を示す．1 位から 3 位までの手は悪い手ではないがいずれも右辺に見向きもしない手である．4 位は右辺に打つ手ではあるが，▲に打たれるとその反対側へ二間ビラキする余裕が無いためワリウチではない．むしろ自分から窮地へ向かおうとする悪手である．

本手法を用いた場合の着手の順位を図 4.3(b) に示す．図 4.2 で示したワリウチが 1 位になっていることが分かる．また，3 位，4 位，5 位，7 位，10 位の手もワリウチであるが，このようにほぼ同等の手が高順位に固まると，探索でもこれらの手にリソースが割かれるため効率が悪い．なぜこのようなことが起きるかというと，盤面分割の際にこれら全ての交点が一つのマスの中に入るため，最終的に同じ特徴が割り当てられるためだと考えられ

る．そこで，複数の区切り方を組み合わせる CMAC を用いれば，この問題は緩和されると考えられる．



(a) 本手法を用いない場合（4 は▲と打たれ不利）



(b) 本手法を用いた場合（1 が正答．その周囲も良い手）

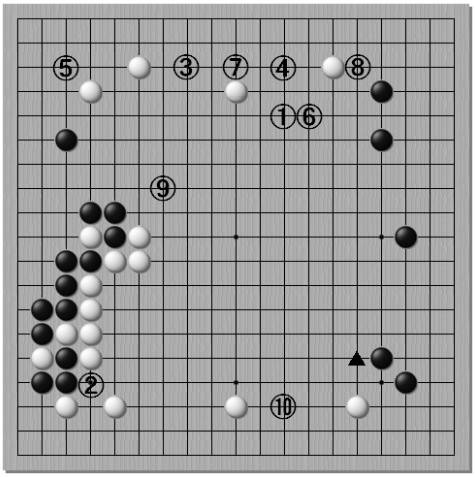
図 4.3: 図 4.2 の局面での着手の順位付け

図 4.3 以外にもいくつか序盤の典型的な局面で同様の順位付けを行った．これらの局面は全て市販の問題集 [13] より，正答を導き出すのに大局観が必要だと考えれられるものを選択した．

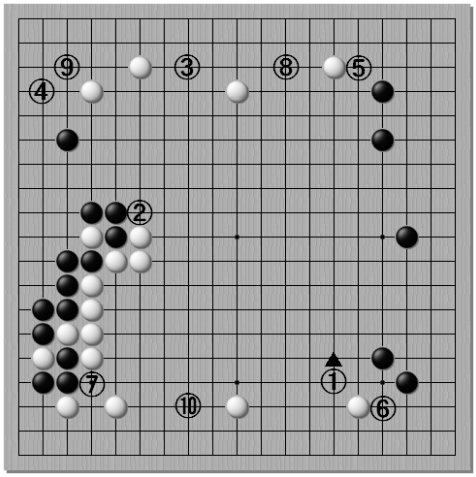
図 4.4（黒番）は左辺の下半分に白の大きな壁があり，これを背景に白が大きな勢力圏を築きつつあるため，それを阻止する必要がある局面である．本手法を用いない場合の着手の順位を図 4.4(a) に示す．1 位の着手は右辺の黒の勢力圏を広げる点では悪い手ではないが白の勢力圏を見ていない．9 位と 10 位の着手は白の勢力圏拡大を阻止するという点では良い手であるが，10 位の着手は深入りしすぎであり，白から▲などの手で右辺の黒と分断して攻められると不利になるため，総合的に見るとあまりよい手ではない．

本手法を用いた場合の着手の順位を図 4.4(b) に示す．1 位の着手が正答であり，白から▲に打たれた場合と比べると，白の勢力圏拡大をうまく阻止していることが分かる．2 位の着手も白の勢力圏を阻止しする点では良い手であるが，白の側から勢力するのに最も良い手は▲であるため，これを阻止できない点で少し劣る．1 位，2 位とも白の勢力圏の拡大を阻止する手であることからも本手法の効果がうかがえる．

図 4.5（黒番）は左上の白と右下の黒が大きな勢力を張り合っている局面である．左上の白の勢力には隙があり，黒は▲（正答）に打つことで白の勢力を大きく荒らすことができる．本手法を用いない場合の着手の順位を図 4.5(a) に示す．1 位の手は局所的には悪い手ではないものの，ほぼすべての着手が中央の白と黒との境界線付近に集中している．本手法を用いた場合は図 4.5(b) のとおり，本問の正答が 1 位となった．

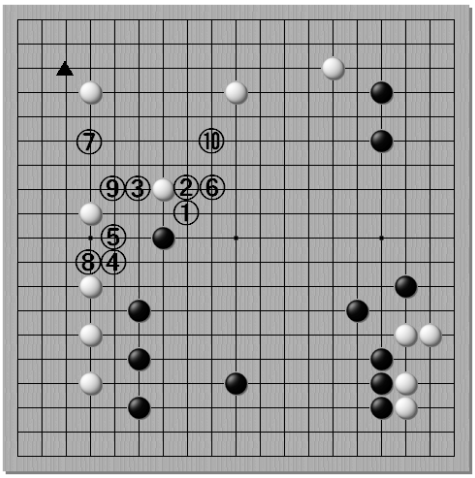


(a) 本手法を用いない場合（9 は良い手．10 は▲と打たれて不利）

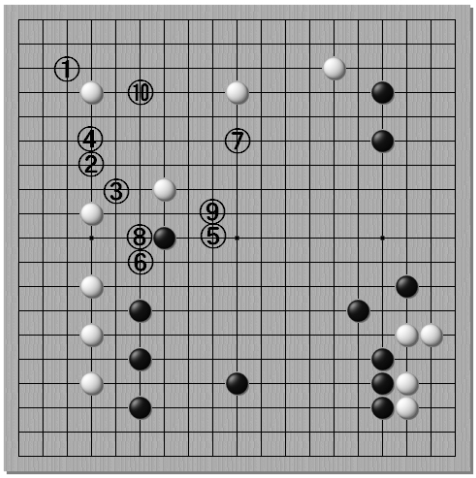


(b) 本手法を用いた場合（1 が正答．打たなければ白▲が良い手に）

図 4.4: 相手の勢力を弱めるべき局面での着手の順位付け



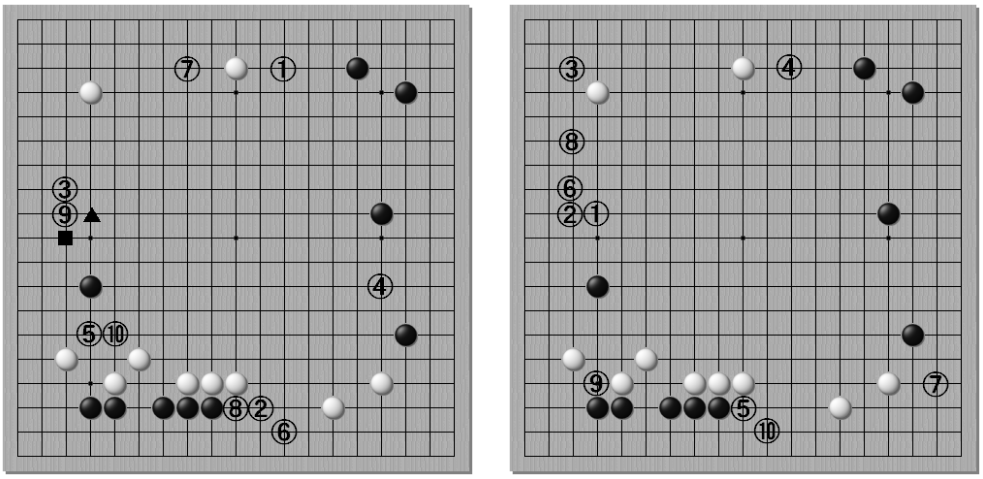
(a) 本手法を用いない場合（▲に打つべき局面）



(b) 本手法を用いた場合（1 が正答）

図 4.5: 相手の勢力を内側から荒らせる局面での着手の順位付け

図 4.6（黒番）は下辺に白の壁，その近くに黒の孤立した石があり，それを守る必要がある．正答は▲である．本手法を用いない場合の順位を 4.6(a) に示す．1 位の手こそそっぽを向いているものの 3 位や 9 位の手もほぼ正答である．しかし，3 位の着手は少し隙があり，■に打たれると攻められる可能性がある．本手法を用いた場合の順位を 4.6(b) に示す．1 位の着手が正答であり，2 位の着手も同じくらい良い手である．図 4.6(a) では 9 位だったこの着手が 2 位になっていることからも本手法の効果が伺える．



(a) 本手法を用いない場合（▲が正答．3，9 も良い手．ただし 3 は■に打たれる危険）

(b) 本手法を用いた場合（1 が正答．2 も良い手）

図 4.6: 自分に弱石がある局面での着手の順位付け

4.3 自己対戦

本手法によりどれだけ nomitan 棋力が向上したかを調べるために，本手法を用いない nomitan との自己対戦の勝率を測定した．使用した特徴量の重みは 4.1 節で学習したものを用いた．この係数は棋譜の 50 手目までしか学習ていないものであるため，2 対 2 という変則的な方法をとった．まず，本手法なしで 50 手目まで学習した nomitan(A) と，同じく本手法なしで 51 手目以降から学習した nomitan(B) を用意する．そして，対戦時はまず本手法を用いた nomitan(C) と A を対局させ，51 手目以降は B 同士を対局させる（図 4.7）．このようにして C と B のペアが A と B のペアに有意に勝ち越せば，本手法の効果があるものとする．

対局に用いたコンピュータの性能はメモリが 24Gb，CPU が Intel Xeon X5680(3.33GHz, 12 スレッド) を 2 つである．対局は 1 手 2 秒で 600 局行い，1 局ごとに手番を交代させた．結果は本手法を用いた B と C のペアが 334 勝し，勝率は 55.7%となり，有意水準検定 (5%)

により，本手法を用いた nomitan が有意に強いことが示された．以上により，本手法で実際に棋力が向上したといえる．

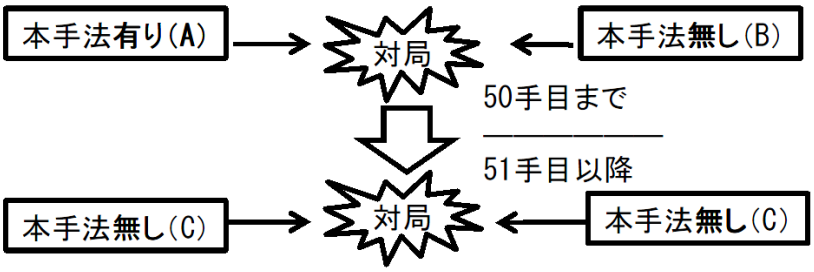


図 4.7: 自己対戦方法

第5章　まとめ

本稿では，囲碁で大きいサイズのパターンマッチングを行う既存手法の問題点を指摘し，その解決策として,「盤面を縮小することで広い範囲をざっとマッチングする」パターンマッチングを提案し，機械学習時の性能，大局観が要求される典型的な局面での着手,自己対戦の勝率の 3 つの指標でその有効性を確認した.

今後は，より良いパラメータの探索とアルゴリズムの改良を行なっていく予定である.石の勢力範囲を求める処理では，単に距離に比例して影響力を計算するのではなく，相手の石の向こう側には影響が及ばないようにしたり，必要以上に凝り固まっている石の集団の影響力を小さくしたりするなどによってより正確な勢力分布を求められることを目指す.

本研究はゲームプログラミングワークショップ (GPW2012) においても発表しており，研究奨励賞を受賞した．本手法をさらに改良したものを情報処理学会のジャーナルへの投稿予定である.

謝辞

本研究を進めるにあたりご指導頂きました池田心准教授と飯田弘之教授に深い感謝の意を表します．また，池田研究室の Simon Viennot 氏や元飯田研究室の橋本隼一氏をはじめとする，池田・飯田研究室の皆様にも様々なご協力を頂き，大変感謝しております.

参考文献

[1] Bernd Brügmann. Monte Carlo Go, 1993.

[2] Levente Kocsis and Csaba Szepesvári. Bandit based Monte-Carlo Planning. In *ECML*, pages 282–293, 2006.

[3] Sylvain Gelly, Yizao Wang, Rémi Munos, and Olivier Teytaud. Modification of UCT with Patterns in Monte-Carlo Go. Technical Report 6062, INRIA, 2006.

[4] Rémi Coulom. Computing Elo Ratings of Move Patterns in the Game of Go. *ICGA Journal*, 30(4):198–208, 2007.

[5] 松井利樹, 野口陽来, 土井祐紀, and 橋本 剛. 囲碁における勾配法を用いた確率関数の学習. In **情報処理学会論文誌**, volume 51, pages 2031–2039, 2010.

[6] 荒木伸夫, 吉田和弘, 鶴岡慶雅, and 辻井潤一. 囲碁における正確な着手予測のためのファジーパターンマッチング. In *The 20th Annual Conference of the Japanese Society for Artifcial Intelligence*, 2006.

[7] 土井 佑紀. 局所化と汎化を両立させる囲碁パターンマッチング. Master's thesis, 北陸先端科学技術大学院大学, 2011.

[8] 日本棋院. **新・早わかり 用語小事典—読んで調べる囲碁知職**. 日本棋院, 1997.

[9] David Stern, Ralf Herbrich, and Thore Graepel. Bayesian Pattern Ranking for Move Prediction in the Game of Go. In *Proc. of the International Conference of Machine Learning*, pages 873–880, 2006.

[10] James Sacra Albus. A new approach to manipulator control: The cerebellar model articulation controller (CMAC). *Journal of Dynamic Systems, Measurement, and Control*, 97(3):220–227, 1975.

[11] Richard S. Sutton. Generalization in reinforcement learning: Successful examples using sparse coarse coding. In *Advances in Neural Information Processing Systems 8*, volume 8, pages 1038–1044, 1996.

[12] 長谷川敦史, 池田 心, 飯田弘之, and 橋本隼一. プレイアウト局面の乱雑さを考慮した学習用棋譜作成. In **第 *16* 回ゲームプログラミングワークショップ** *2011*, pages 33–40, 2011.

[13] 趙治勲. **ひと目の総合問題**. 毎日コミュニケーションズ, 2010.