

関連性行列を用いたぷよぷよの定型連鎖構成法

富沢大介^{†1} 池田心^{†1} 橋本隼一^{†1}

1991年にコンパイルから発売された対戦型パズルゲームぷよぷよにおいてコンピュータ AI の強さは 4~5 連鎖程度であり、中級者以上が満足できるレベルではなかった。本研究では、ぷよぷよの AI を強くすることを目的として、強いぷよぷよのプレイヤーに必要な要素の内の一つである、長い連鎖を効率的に組むことに着目し、囲碁の“定石”や将棋の“囲い”にあたる人智の結晶である“定型連鎖”を構成する方法に取り組む。関連性行列という形での状態表現・テンプレート表現を用いた構成法により、定型連鎖を構成することに成功した。さらに従来の探索型連鎖構成法を組み合わせることで従来法を上回る平均 11.75 連鎖を達成することに成功した。

Construction of Stylized Chain in Puyo-Puyo using Association Matrix

Daisuke Tomizawa^{†1} Kokoro Ikeda^{†1} Junichi Hashimoto^{†1}

Puyo-Puyo is a two player puzzle game, which has been released from Compile since 1991. Because a best AI for this game was able to make only 4 to 5 length chains, the strength was not sufficient for middle or higher level players. One of the lacking abilities which are essential for a strong player is to make a long chain in a compact space. We introduce 'stylized chain' which are typical strategies to place stones(puyos). These typical strategies is a grand sum of human intelligence and universally appears in other games such as Go or Shogi as 'pattern' or 'formation.' We represent a current state and favorable states in a form of 'association matrixes' in order to realize the Stylized chain. Experimental results show that our method successfully develops long chans such as 11.75 length in average.

1 はじめに

1991年にコンパイルから発売されたぷよぷよシリーズは、対戦の要素を加えたことでテトリスを凌ぐ落ち物パズルゲームの代表として認知されており、現在でも多くの上級プレイヤーがハイレベルな対戦を繰り返し広げている。詳細なルールについては後述するが、複数の色からランダムに与えられる 2 個 1 組のぷよを適切に配置し、**連鎖**と呼ばれる状況を起こして相手を攻撃することが本ゲームの中心的な課題である。双方が同時に攻撃した場合には、攻撃力の弱い側のみが被害を受けるため、与えられたぷよとフィールドサイズに合わせてできるだけ効率よく大きい連鎖を構成する能力が最も重要となる。上級者同士の戦いでは 13~14 連鎖程度が組めることが必須とされている。

ぷよぷよは分岐因子数の少なさから AI の構成が容易に見えるが、未だ人間に追いついていない。

ぷよぷよの既存研究としては、一般化ぷよぷよの NP 完全性^[3]等、数学的見地からの研究はいくつかなされているが、強くすることを目的とした研究はまだ十分に行われてきていないといえない。

本研究では、強いぷよぷよのプレイヤーに必要な要素であるとされる

- (1) 長い連鎖を効率的に組む
- (2) 相手との駆け引き
- (3) 高速な操作

の能力のうち(1)にのみ着目し、囲碁の“定石”や将棋の“囲い”にあたる人智の結晶である“**定型連鎖**”を構成する方法に取り組む。そのために本論文では関連性行列という形での状態表現・テンプレート表現を導入し、これと木探索を組み合わせた連鎖の構成法を提案する。

本論文では以降、2 章にてぷよぷよの詳細なルールと従来法の問題点、3 章にて提案手法を、4 章でテンプレート行列の作成方法について、5 章にて提案手法に関する実験と結果について述べ、最後に 6 章でまとめと今後の課題について述べる。

^{†1} 北陸先端科学技術大学院大学

Japan Advanced Institute of Science and Technology

2 従来法の問題点とアプローチ

この章ではぷよぷよのルールを述べた後、従来の AI で用いられているポテンシャル最大化法について説明し、特に長い連鎖を目指す上で従来法にどのような問題があるかを議論する。

2.1 ぷよぷよのルール

図 1 を基にぷよぷよのルールについて説明する。ぷよぷよは二人のプレイヤーによる対戦ゲームで横 6 × 縦 13 のフィールドがそれぞれに与えられる (図では対戦相手のフィールドは省略する)。各プレイヤーは複数の色 (通常 4 色) からランダムに与えられる 2 個 1 組のぷよ (配ぷよと呼ばれる) を落下地点とその向き (上下左右 4 通り) を定めて着手する (a, b)。1 手あたりの自由度は配ぷよが異色である場合で最大 26 となる。通常、配ぷよは現在着手可能な手を含めて 3 手先まで知らされている。最終的に着手不可能となったプレイヤーが負けとなる。ぷよは 4 つ以上連結すると消滅し得点を獲得する (c)。この時、消滅により上にあったぷよが落下して再びぷよが消えると連鎖となる (d)。連鎖の過程でぷよが消滅した回数を N とするときこれを N 連鎖と呼び、およそ N の 2 乗に比例した得点が得られる。対戦では、獲得した得点に応じて発生するおじゃまぷよ (4 つ接続しても消えず、上下左右に隣接するいずれかのぷよが消えた場合にのみ消える) を対戦相手に送ることで妨害し攻撃する。双方が攻撃を行った場合には攻撃力の大きいほうはその差分だけ一方的に相手におじゃまぷよを降らせることができる。このため、大きい連鎖を作成する能力が本ゲームにおいて重要となる。

囲碁や将棋と比較すると、盤面が対戦者ごとに分かれ、自分の好きに手を進めていくことができる点が異なる。ただし連鎖によって生じるおじゃまぷよを通じて相手の盤面に干渉することはできるため、干渉のタイミングをはかることで単なる連鎖数比べとは異なる駆け引きが生じる。また、配ぷよの色にランダム性があることも大きな違いであり、どのような組み合わせの手が来ても対応できるように手広い積み方が求められる点は麻雀などに似ている面がある。

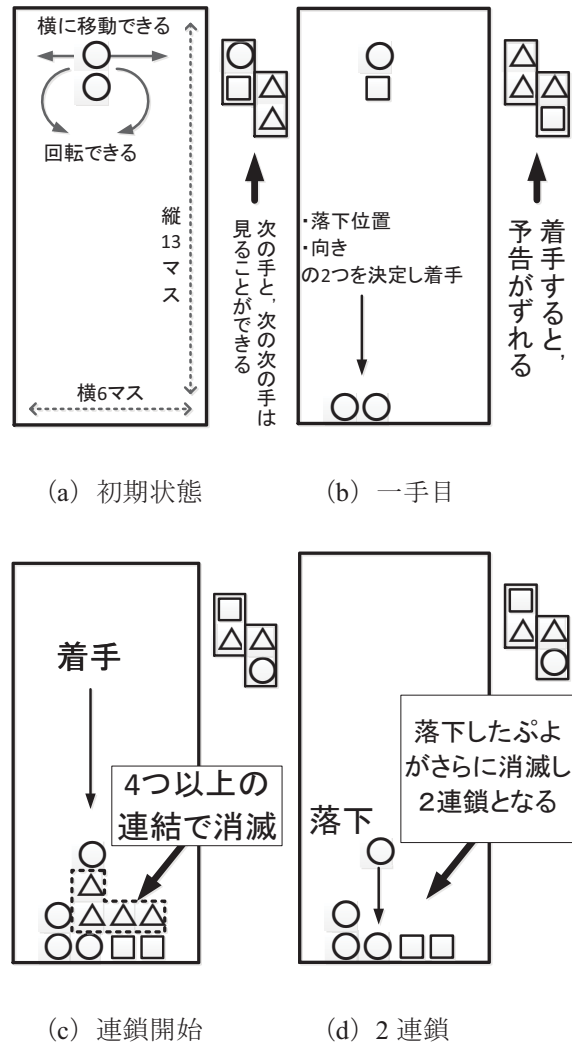


図 1：基本的なぷよぷよのルール

2.2 ポテンシャル最大化アルゴリズム

同時着手型ぷよぷよである崩珠^[1]ではポテンシャル最大化アルゴリズムが利用されており、10 連鎖程度を組むことができる。このアルゴリズムは以下の手順で、配ぷよの分かっている数手先までを木探索し、葉ノードでの攻撃力が最も大きくなる手を選ぶ。

- (1) 着手可能な手を読める範囲ですべて列挙する
- (2) 1 手目に連鎖を開始する手を除外する
- (3) 2 手先、3 手先に連鎖を開始できる変化の内、攻撃力が最大であるものを選択する。

図 2 を例にして説明する。ぷよぷよでは現在着手可能な手を含めて 3 手先までの配ぷよが分かっているため、自フィールドについてだけなら 3

手先の局面をすべて列挙することが可能である。基本的に長い連鎖を目指すため、一手目で連鎖を開始する手は通常除外する。(ただし十分な連鎖数がすでにある場合や、相手に攻撃されている場合は選ぶこともある)。一方で、3手先まで読んでも連鎖が生じない場合、次の配ぶよの色は未知であり、最悪の場合はその先連鎖を起こすことができなくなってしまう可能性もある。従って4手目以降は読まない。結局、2手先・3手先に連鎖を開始できる中で最もその攻撃力が大きいものを選び、その1手目だけを選択する。2手目以降はここで選んだものにとらわれず、毎回、再度探索を行う。

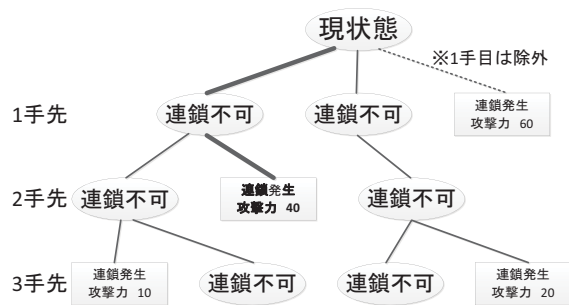


図2: ポテンシャル最大化アルゴリズム

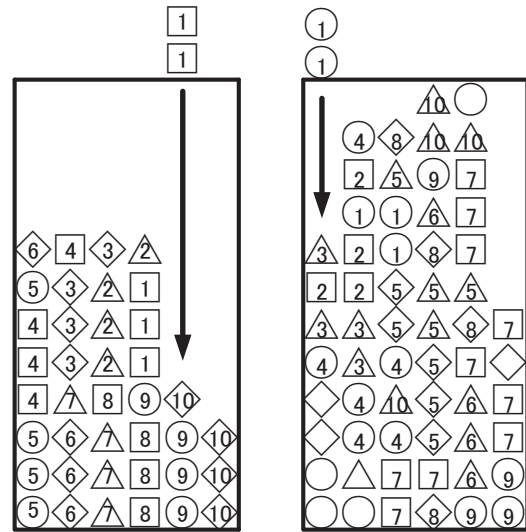
2.3 ポテンシャル最大化アルゴリズムの問題点

前節のアルゴリズムには二つの問題がある。

(1)盤面を効率よく使えない。

(2)大局的な目的に基づいて連鎖が組めない。

(1)について、ポテンシャル最大化では攻撃力が評価基準になるため連鎖数を増やせない場合には同時消し・5個以上消しなどの方法で攻撃力を補おうとする。そのため空間的な無駄を生みやすく、長い連鎖を作るのに不利な局面へと向かう傾向がある。図3(b)は従来型のAIによる10連鎖の例であるが、整然としてコンパクトな(a)に比べ、複雑・乱雑で大きなスペースを要し、これ以上の連鎖数の上昇が見込めないことが分かる。AIの組む連鎖が整然としていないことはある種の多様性・意外性による楽しさを生じさせるものの、対人戦の練習・上級者の相手をするという意味では欠点である。



(a) 定型連鎖 (b) 従来による連鎖

図3: 10連鎖の例. 左の例では、□が消えることで△が落下し、2連鎖目を生じ、続いて◇, □, ○が次々に消えることで最終的に10連鎖になる。

(2)の大局的な目的に基づいて連鎖が組めない事について、人間とAIの対戦を例に挙げて説明する。図4に、JAISTCUP^[2]における、人間とAIの対戦における序盤14手目までの一例を示す。

(盤面のサイズは7×11) AI側は○を消すことで現在4連鎖がいつでも開始できる形である。一方人間側は8連鎖を行う目処がついているが実際の連鎖のためには4連鎖目と8連鎖目の□2つ、6連鎖目の○1つ、2連鎖目の△1つが足りない状況となっている。人間はこのように、配ぶよに合わせて「こういう形の連鎖を組む」という想定図を思い浮かべて、それに当てはめていくことをしばしば行う。図2で示したポテンシャル最大化法によって常に連鎖を開始できるように徐々に構築を行うAIではこれではできない。図5は図4に続き、連鎖を実際に開始する局面である。人間の連鎖は初級者でも紙上で目で追うことができる程簡明であり、ほぼ全てのぷよを使って13連鎖となっている。一方でAIの連鎖は5連鎖目以降から非常に分かりにくく、最終的に10連鎖となっているのが不思議なほどである。

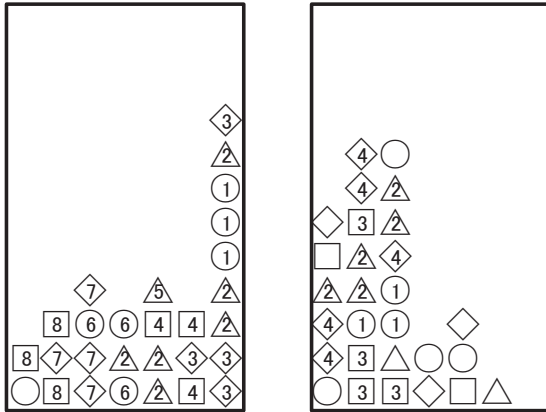


図4：人間（左）対 AI（右）序盤

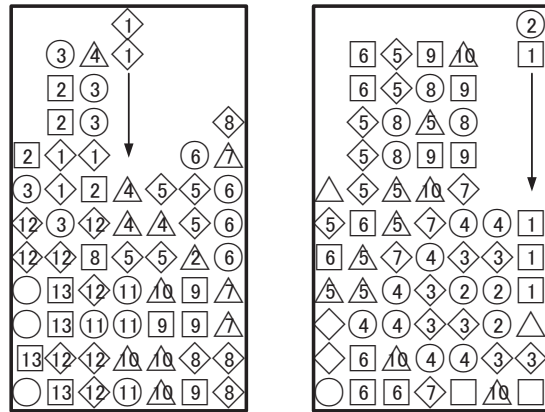
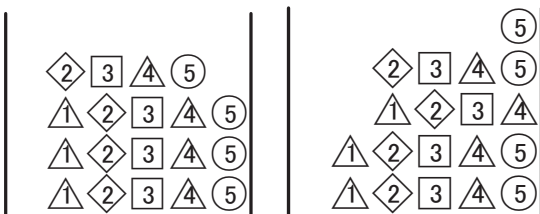


図5：人間（左）対 AI（右）終盤

3 関連性行列による連鎖生成手法

以上を踏まえて、AI の連鎖能力を改良するために、人間が連鎖を組むための思考の助けとして用いている“**定型連鎖**”という考え方を導入する。定型連鎖とは、長い連鎖を組むための土台となる部分の配置パターンであり、多くの場合規則的で最小限のぷよで構成されている。定型連鎖の例を図6に示す。将棋や囲碁と同じくこれらは長い時間をかけて進化してきた人智の結晶であり、囲碁等では相手の着手によって自分が選ぶべき定石・手が変わる一方で、ぷよぷよでは主に配ぷよによって選ぶべき定型連鎖・手が変わるという特徴を持つ。

図6(a)の定型連鎖で、例えば2連鎖目の◇色は○色であったとしても問題ない。これら全ての色の組み合わせの可能性を列挙したパターンを用意するのは効率的ではないため、本研究では、**2つのぷよの色の関連性を表した行列**を用いることでこの問題を解決する。



(a)階段積み5連鎖

(b)鍵積み5連鎖

図6：定型連鎖の例

3.1 全体の枠組み

盤上のぷよ同士の関連性を表した行列を関連性行列と呼び、これを以下で用いる。提案手法の概要を図7に示す。定型連鎖を組み合わせるための基本的な枠組みは以下の通りである。

- (1) 定型連鎖を表現するテンプレートを、関連性行列の形で**テンプレート行列**として複数用意しておく。
- (2) 対局時は、配ぷよの見える数手先までの全ての局面を探索し、その葉ノードを関連性行列の形で**状態行列**として表現する。
- (3) 各葉ノードと各テンプレートを比較して合致度スコアを計算する。
- (4) スコアに応じて最も有望なノードを選択する。実際に着手されるのはルートノードからの1手のエッジとなる。

(1)のテンプレートの作成のみ現在は手作業で行っているが、今後は棋譜からの自動生成も視野に入れている。提案手法によって定型連鎖を生成したのちに、従来手法の木探索アルゴリズムに切り替えることによってAIの連鎖能力向上を図る。

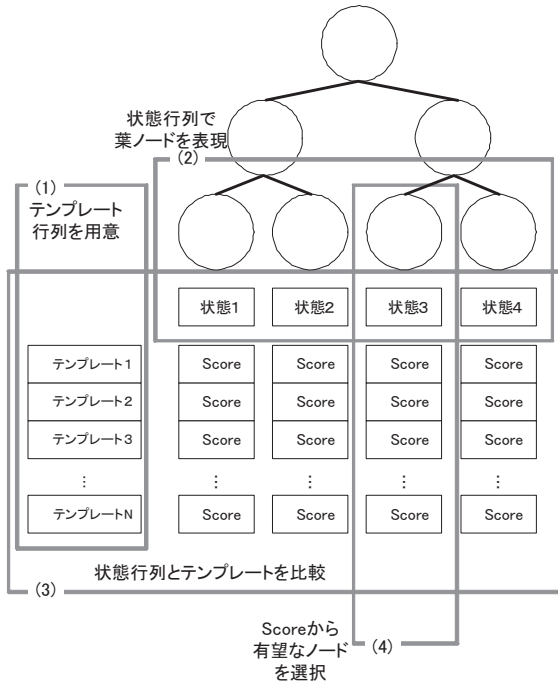


図7：関連性行列による連鎖生成手法

3.3 テンプレート行列

テンプレート行列 T は定型連鎖を表したものである。盘面のある注目するマス i のふよの色に対して他のマス j のふよの色が同色であるべきか、異色であるべきか、どちらでもよいか、によって評価値を与える。同色がふさわしければ正の評価値を、異色がふさわしければ負の評価値を与える。さらに、評価値は優先的に着手すべき箇所（例えば色の制限が厳しい箇所）の評価を重くするように設定する。状態行列 T は対称行列で、対角成分は正値又は0になり、実際に必要なのは右上三角成分のみとなる。

$$T = \begin{pmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & & \vdots \\ t_{1n} & \cdots & t_{nn} \end{pmatrix}$$

3.2 状態行列

状態行列 S は盘面のある注目するマス i のふよの色に対して他のマス j のふよの色が同色か異色かを表す。 n は盘面のサイズである（通常 $6 \times 13 = 78$ ）。同色は $+1$ 、異色は -1 、ふよが存在しない場合は 0 で表す。ふよの色を 6×13 の行列で表したわけではないことに注意されたい。状態行列 S は対称行列で、実際に必要なのは右上三角成分のみとなる。

$$S = \begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & & \vdots \\ S_{1n} & \cdots & S_{nn} \end{pmatrix}$$

$$s_{ij} = \begin{cases} +1 & \text{両方にふよがあり同色の場合} \\ -1 & \text{両方にふよがあり異色の場合} \\ 0 & \text{どちらかのマスが空きの場合} \end{cases}$$

$$\begin{cases} t_{ij} < 0 & \text{異色であるべき場合} \\ t_{ij} > 0 & \text{同色であるべき場合} \\ t_{ij} = 0 & \text{どちらでもよい場合} \end{cases}$$

$$\begin{pmatrix} +1500 & -1500 & -1500 & 0 & \cdots & 0 \\ -1100 & +1100 & -1100 & 0 & \cdots & 0 \\ -1100 & -1100 & +1100 & 0 & \cdots & 0 \\ 0 & 0 & -1000 & +1000 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

図8 テンプレート行列の例

3.4 合致度スコアの算出手法

状態行列 S とテンプレート行列 T の比較は、基本的に要素ごとの積を加算することで行う。行列の要素が同符号ならばスコアに正の値として加算されるが、一つでも異符号つまり異色（同色）であるべき箇所が同色（異色）の場合はスコアを負値とし採用しない。分母はテンプレート間の数値のばらつきを無視するための正規化項である。

$$score_T(S) = \begin{cases} -\infty & \text{if } \exists i, j \quad s_{ij}t_{ij} < 0 \\ \frac{\sum_{i=1}^n \sum_{j=1}^n s_{ij}t_{ij}}{\sum_{i=1}^n \sum_{j=1}^n |t_{ij}|} & \text{otherwise} \end{cases}$$

3.5 最善手の選択

一つの状態行列に対して複数の合致度スコアが得られるので、最善手の選択手法には様々な方法が考えられる。(A) 単純に最大のスコアを持つものを最善手とする手法, (B) スコアが正のものの数が最大となるものを最善手とする手法, (C) スコアが負のものを除外した上での平均スコア等である。本論文では最大のスコアを持つものを最善手とする手法を用いる。

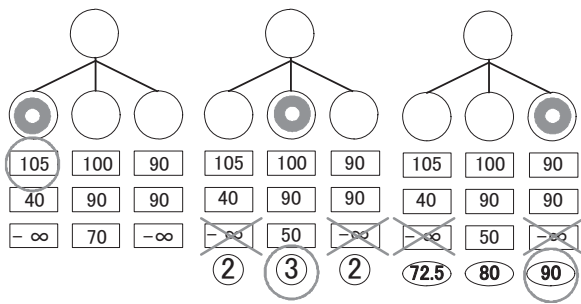


図9：各最善手選択手法(A)(B)(C)の例

3.6 計算量

提案手法を実装した場合、計算量は $O(\text{合法手}^3 \cdot \text{フィールドマス数}^2 \cdot \text{テンプレート数})$ が必要となるが、実際には、1手目で一部でも不一致があればその先はそのテンプレートを計算する必要が無いなどの理由で大幅に削減することができる。

4 テンプレートの作成手法

テンプレートの作成は次の手順で行う。

- (1)ラベリング表の作成
- (2)真理値表と重み付け表の作成
- (3)テンプレート行列の作成

このうち手順(2)の一部と(3)は自動化できており、本論文の実験で用いたテンプレート (24 個) を

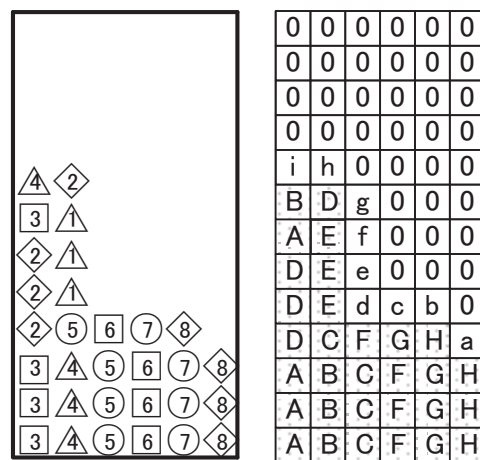
作るのに必要な時間は2時間程であった。

まず、作成したい定型連鎖からラベリングした表を作成する(図10)。このとき、同じ文字である箇所は、一回の連鎖で消えるぶよの箇所になるようにし、定型連鎖に隣接する箇所はすべて異なる文字にし、それ以外の連鎖に関係のない領域はまた異なる文字で埋める。今回の例では、定型連鎖に関する箇所は英字の大文字で、定型連鎖に隣接する箇所を英字の小文字で、関係のない箇所を0で埋めた。(図10(b))

次に、ラベリングした表から真理値表と重みづけの表を自動生成する(図11)。真理値表は、ある記号の箇所に隣接した記号の箇所は、異色であると定義し、消えるべき箇所では同色と定義し、それ以外の箇所はどちらでもよい箇所と定義する。重みづけの表は連鎖に関する箇所を1000、それに隣接する箇所を100、関係ない箇所を0の評価を自動的に与える。

こうして自動生成した表をもとに、自動生成で補えない部分を人の手で修正を行う。例えば階段折り返しのタブーと呼ばれるラベリング表(図12中)で、AとCの箇所が同色であるとAが消滅するときにCが暴発してしまうため、「異色」と修正を入れないといけない。こうした箇所を元に真理値表を人の手で修正を行う。さらに、それをもとに重みづけの表を、条件の制約が厳しい所や、優先的に積むべき所を重点的に重みづけの修正をする。

最後にラベリングした表と真理値表、重みづけの表を基にテンプレートの行列を自動生成して図8のような行列を得る。



(a)定型連鎖 (b)ラベリング表

図10：ラベリング表の作成

	A	B	C	D	E	F	G	H	a	b	c	d	e	f	g	h	i	0	
A	S	X	X	X															
B	X	S	X	X													X		
C	X	S	X	X	X														
D	X	X	S	X											X	X			
E	X	X	X	S					X			X	X						
F		X			S	X			X										
G				X	S	X			X										
H					X	S	X		X										
a							X												
b				X	X														
c					X														
d						X													
e				X															
f				X															
g			X																
h	X																		
i			X																
0																		0	

S : 同色
X : 異色
無 : どちらでも

重み	
A	1000
B	1000
C	1000
D	1000
E	1000
F	1000
G	1000
H	1000
a~i	100
0	0

図 11 : 生成された真理値表と重みづけ表

	A	B	C	D	E	F	G	H	a	b	c	d	e	f	g	h	i	0	
A	S	X	X	X															
B	X	S	X	X															
C	X	S	X	X	X														
D	X	X	S	X															
E	X	X	X	S															
F		X			S	X													
G				X	S	X													
H					X	S	X												
a							X												
b				X	X														
c					X														
d						X													
e				X															
f			X																
g			X																
h	X																		
i			X																
0																			0

A と C が同色だと暴発するので、A と C の関係は異色にする必要がある

⊙ A ⊙ C
⊙ A ⊙ B ⊙ C
⊙ A ⊙ B ⊙ C
⊙ A ⊙ B ⊙ C

重み	
A	1500
B	1100
C	1100
D	1300
E	1300
F	1000
G	1000
H	1000
a~i	100
0	0

図 12 : 修正した真理値表と重みづけ表

5 実験

5.1 提案手法の有効性の検証

提案手法の有効性を示すために、崩珠クライアント^[2]上に AI を実装し、AI が定型連鎖を完成させるまでの手数と上級者が定型連鎖を完成させるまでの手数を比較した。ここでの定型連鎖の完成とは合致度のスコアが 1 になることを示す。実験の条件として、盤面のサイズを 6×13、配ぶよは 4 色、テンプレートは 8 連鎖のものを 24 種類用意した。連鎖完成までの手数を昇順でソートしたグラフを図 13 に示す。

結果から、上級者は平均 23.55 手を完成に費やしているのに対して AI は平均 24.45 手と上級者とほぼ遜色ない性能となった。図 14 は AI の作成

した連鎖の例であるが、無駄は殆どなく、ここからポテンシャル最大化法を行うことで 13~14 連鎖が十分期待できる。

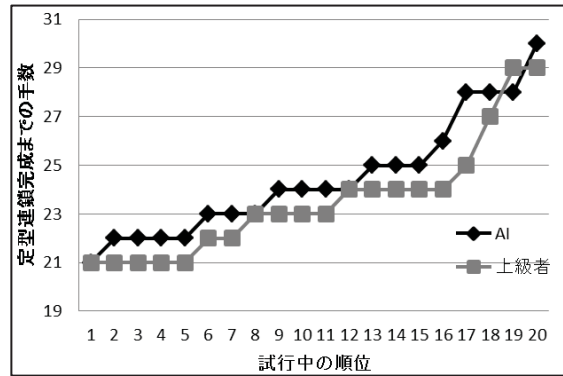


図 13: 定型連鎖完成までの手数

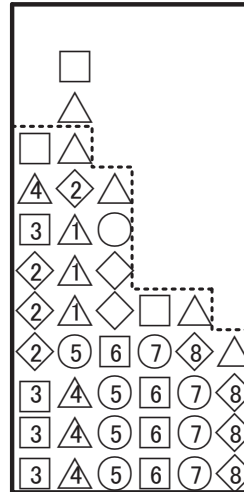


図 14: 定型連鎖完成の例

5.2 ポテンシャル最大化法との組み合わせによる連鎖生成能力の検証

ポテンシャル最大化法との組み合わせによる連鎖能力の向上を検証するため、関連性行列とポテンシャル最大化法を組み合わせた AI と、ポテンシャル最大化法のみ AI の連鎖数の比較を行った。関連性行列とポテンシャル最大化法の組み合わせ方法は、合致度スコアが 0.85 を超えたときにポテンシャル最大化法にスイッチするという至ってシンプルなものである。実験の条件は 5.1 の実験と同様に、盤面のサイズを 6×13、配

ぷよは 4 色，テンプレートは 8 連鎖のものが 24 種類である．実験結果を図 15 に示す．グラフと平均連鎖数を比較すると 1.6 連鎖分組み合わせによる手法が上回っていることから，ポテンシャル最大化法との組み合わせで連鎖数の向上が確認できた．実際に生成された連鎖の 1 例を図 16 に示す．

表 1 各 AI の平均連鎖数

	平均連鎖数
関連性行列 +ポテンシャル最大化法	11.75
ポテンシャル最大化法	10.15

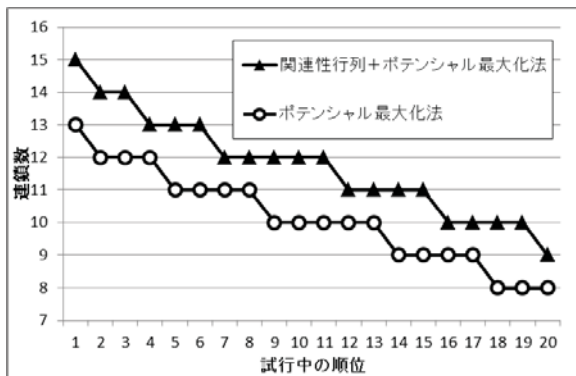


図 15 : 各 AI による連鎖数の比較

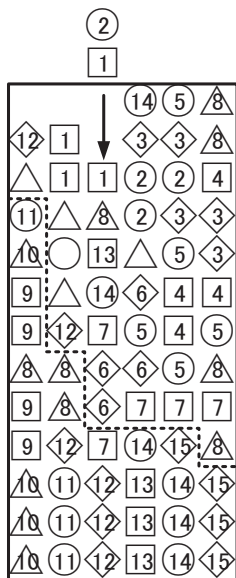


図 16 : 生成された連鎖の 1 例

6 まとめと今後の予定

本論文では，ぷよぷよにおける定型連鎖を AI が生成するための手法を提案し，上級者が定型連鎖を作る場合とほぼ遜色ない性能を持つことを示した．さらに，ポテンシャル最大化法と組み合わせることによって，ポテンシャル最大化法を上回る連鎖数を生成することに成功した．今後は，多くのテンプレート行列を効率的に構成する方法，棋譜からのテンプレート自動抽出，ロバストな最善手選択法の考案，機械学習による評価関数の調整等によってさらなる AI の性能向上を実現したい．

また，ぷよぷよに限らず，麻雀やコラムス，パネルでポン等といったランダム性を持ち，複数人ゲームで自分のフィールドの独立性が強い，形を構成するタイプのゲームに対する拡張性についても検討していきたい．

参考文献

- [1]. PojeMaster, <http://www.vector.co.jp/soft/win95/game/se095051.html>
- [2]. JAIST CUP 2011, http://www.jaist.ac.jp/jaistcup2011/poje_details.html
- [3]. 松金, 輝久; 武永, 康彦. 一般化ぷよぷよの NP 完全性. 数理解析研究所講究録 2005.