

ニューラルネットワークによる詰碁プログラムの研究

佐々木宣介

sasaki@bus.hiroshima-pu.ac.jp

広島県立大学

概要

囲碁は探索すべき局面数が広すぎるために、先読み探索を中心としたアルゴリズムがまだ十分に力を発揮できない。先読み探索のみに頼らない手法として、ニューラルネットワークのパターン認識能力に着目して、詰碁課題へニューラルネットワークを適用してきた。これまで、攻撃手のみの学習を行ってきたが、本稿では、 19×19 のサイズの詰碁問題に関して、攻撃手および、防御手の学習も合わせて行なった結果を報告する。約5000パターンの詰碁課題を学習したネットワークは、未知の問題において、50-60%程度の問題で、上位5位までの候補手の中に正解を含んでいた。 9×9 の狭い範囲の問題に適用した場合と比べ、正答能力の低下がみられたが、ネットワークが未知の問題に対してある程度解答能力を持っていることがわかった。また、人間のプレイヤーが直観的に解答を出した時の解答率と比較した場合、ネットワークの解答能力は、2級から初段程度に相当することがわかった。

The Study of Tsume-go Program Using Neural Networks

Nobusuke Sasaki

sasaki@bus.hiroshima-pu.ac.jp

Hiroshima Prefectural University

abstract

Go is a difficult game to make a computer program because of the space complexity. The author already reported the result of the neural network ability for Tsume-go problems' attack move. In this paper, the author applied the neural network for both attack move and defense move of 19×19 board (normal size of Go board). The network learns about 5000 patterns of Tsume-go by back-propagation method. The top five selected moves contain about 50-60% of correct answer for unknown problems.

1 はじめに

囲碁は探索すべき局面数が広すぎるため、チェスなどで有効に働いた探索中心の手法で強いプログラムを作ることが難しい。本稿では、探索のみに頼らない手法のひとつとして、ニューラルネットワークに着目し、その能力を評価した。

先に、ニューラルネットワークを 9×9 の小さな盤面の詰碁問題に適用し、その能力の評価を行ない[1]、その後、 19×19 の通常のサイズの詰碁問題に

適用した結果も報告した[2]。 9×9 の範囲の問題においては、人間のプレイヤーが直観的に解答した時の能力との比較も行なった。

これまでは攻撃手の学習のみを行ってきたが、本稿では、防御手の学習も行なった結果を加えて報告する。また、 19×19 の問題に対するニューラルネットワークの能力を人間のプレイヤーの能力と比較した結果も報告する。

なお、これまでにニューラルネットワークを囲碁プ

ログラムに応用した研究の例としては, Enzenberger の NeuroGo[3] や, Richards らの進化的アルゴリズムを用いたプログラムの研究 [4] などがある。また, 探索を中心とした詰碁プログラムについては, Wolf の GoTools がある [5]。

2 実験

2.1 ニューロンの入出力特性

人工知能研究の分野ではニューラルネットワークは広く研究対象とされており, さまざまな応用例がある。ニューラルネットワークモデルのニューロンは本物のニューロンと同様に非線型な入出力特性を持つ。

本論文では, ニューロンの入出力特性として, 式 (1) に示す連続関数を使用した。ここで, x_i は, i 番目のニューロンの出力, w_i はそのニューロンからの影響の強さを表わす結合定数である。正味の入力の総和が閾値を越えた場合にそのニューロンは活性化された状態となる。式 (1) では, 閾値は 0 となっている。式 (1) であらわされるニューロンは出力値は 0 から 1 の範囲となり, 活性化された状態では出力 y の値は 1 近く, 発火していない状態の時には出力値は 0 近くの値となる。

$$y = \frac{1}{2} \left(\tanh \left(\sum w_i x_i \right) + 1 \right) \quad (1)$$

また, 学習においては, ニューラルネットワークの教師あり学習の標準的手法である, 逆誤差伝播法 [6] を用いた。逆誤差伝播法では多層のフィードフォワード型のニューラルネットワークが使われる。

2.2 ネットワークの構造

図 1 に本論文で用いたニューラルネットワークの構造を示す。ネットワークは, 入力情報として石の配置のみを利用しており, それ以外の, 石のつながりなどの情報を明示的に入力情報として利用することはしていない。

入力層のニューロンは, 囲碁の盤面の特定の座標に黒石または白石が存在するかどうかを表わす。盤面の一点につき, 黒石用と白石用の 2 個のニューロンがあり, 盤上に石が存在する時, その座標と石の色に対応するニューロンの出力を 1 とする。したがって, 入力層のニューロン数は $722 = 19 \times 19 \times 2$ となる。ある座標に石が存在しない時には, その座

標を表わすニューロンの出力はどちらも 0 とする。出力層には, 盤面の一点につき対応するニューロンが 1 個存在する。したがって, $361 = 19 \times 19$ 個のニューロンが存在する。問題の配置を入力層に入力した時, ネットワークの出力層ニューロンの中で一番活性化している (出力値が高い) ニューロンに対応する座標が, ネットワークの求めた次の着手点であると考えられる。中間層のニューロン数は自由に設定することができるが, 中間層のニューロン数はネットワークの能力を決定する重要な要素となる。

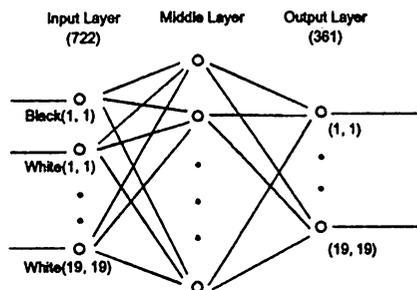


図 1: ニューラルネットワークの構造。入力層に問題配置を入力し, 出力層で一番活性化したニューロンをネットワークの解答の解釈する

2.3 学習

ニューラルネットワークは逆誤差伝播法 [6] によって詰碁のパターンの学習を行なう。学習および評価に用いた詰碁問題はすべて問題集から採取した ([7] 等)。攻撃手は黒先白死問題から, 防御手は白先白生問題からそれぞれ採取した。黒先白死問題の正答手順のすべての攻撃手, 白先白生問題の正答手順のすべての防御手を学習すべき手として採取し, その問題からさらに対称配置を生成し, それぞれを独立した学習データとした。なお, 19×19 の広いサイズの問題と言っても, 盤面全体を利用する問題はほとんど含まれず, 9×9 程度のサイズに含まれる問題も多い。 19×19 のサイズに収まる全ての問題を区別せずに学習するという意味である。

学習は用意した問題について, その問題と答えの組をネットワークに記憶させることによって行なう。攻撃手については 5000 手, 防御手は 4500 手をネットワークに学習させた。学習が終了した後に, 学習に用いた問題をどれだけ記憶したかおよび, ネットワークが学習していない, 未知の問題に対するネッ

トワークの解答能力を評価する。

ニューラルネットワークの学習は以下の手順で行なった。

1. 学習局面のすべての石の配置を入力層に入力データとして入力する。石が存在する場合、その座標と石の色に対応するニューロンの値を1とし、他のすべてのニューロンの値を0とする。
2. ネットワークの出力層において、入力した問題の答えとなる座標に対応するニューロンの教師データを0.9とし、それ以外の場所のニューロンの教師データを0.1となるように、逆誤差伝播法によってネットワーク内のニューロンの結合定数の値を更新する。
3. ネットワークが5000パターン(防御手の場合は4500パターン)の学習を1回行なったら(1ラウンド)、続いて2ラウンド目の学習を行なう。多数回の学習を経ると、出力層におけるニューロンの出力値は、正しい答えに対応する座標のニューロンは0.9、その他の座標に対応するニューロンは0.1に近づく。

なお、攻撃手の学習においては、最適な中間層ニューロン数を求めるため、中間層ニューロンの数の異なる3種類(中間層ニューロン数が400, 600, 800)のネットワークを用いて学習を行なった。

3 評価

3.1 学習した問題の解答能力

ある程度の回数の学習の後、学習に用いた問題について、どれだけのパターンをニューラルネットワークが記憶したかを調べた。

評価は以下の手順で行なった。

1. 学習に利用した問題の局面の配置を入力信号としてネットワークに与える。石が存在している場所および石の色に対応するニューロンには1を入力し、石が存在しないニューロンには0を入力する。
2. 出力層において一番高い出力を示したニューロンが対応する座標をネットワークの答えた最善手と判断する。ただし、既に石が存在する場所を答えた場合には、次に高い出力ニューロンをネットワークの出した答えとして採用する。

表1, 表2に, ネットワークが学習回数によって,

どれだけのパターンを記憶していたかを示す。なお、攻撃手の学習を行なった3種類のネットワークのうち、ネットワークAが中間層ニューロン数400, ネットワークBが中間層ニューロン数600, ネットワークCが中間層ニューロン数800のネットワークである。

ニューラルネットワークは攻撃手, 防御手共に, 1000ラウンドまでにはほぼ100%の学習パターンを記憶していることがわかった。

表1: 学習したパターンに対するネットワークの正答率(攻撃手)。ネットワークA: 中間層ニューロン数400, ネットワークB: 中間層ニューロン数600, ネットワークC: 中間層ニューロン数800

学習回数 (Rounds)	正解率 (%)		
	A	B	C
1×10^1	1551(31.0)	15(0.3)	24(0.5)
2×10^1	2287(45.7)	62(1.2)	35(0.7)
5×10^1	3706(74.1)	2605(52.1)	44(0.8)
1×10^2	4619(92.4)	4432(88.6)	50(1.0)
2×10^2	4908(98.2)	4936(98.7)	59(1.2)
5×10^2	4969(99.4)	4997(99.9)	171(3.4)
1×10^3	4983(99.7)	5000(100.0)	4998(100.0)

表2: 学習したパターンに対するネットワークの正答率(防御手)。ネットワークB: 中間層ニューロン数600

学習回数 (Rounds)	正解率 (%) ネットワークB
1×10^1	22(0.4)
2×10^1	28(0.6)
5×10^1	1216(27.0)
1×10^2	3802(84.5)
2×10^2	4440(98.7)
5×10^2	4497(99.9)
1×10^3	4499(100.0)

一般的に、ネットワークの中間層ニューロン数が増えれば学習能力は向上するが、学習が収束する速度は低下することが知られている。従って、十分な学習能力を持ちつつも、できるだけ少ない数の中間層ニューロンを持つネットワークを採用する必要が

ある。表1に示したように、中間層ニューロン数が800の時には明らかに学習が終了するまでに時間がかかっており、学習回数が1000ラウンド近くになって始めて多くのパターンを記憶している。一方、中間層ニューロンが400のネットワークにおいては、学習の進行は早いものの、5000のパターンを十分に記憶することができなかった。本論文の5000パターンの学習においては、中間層ニューロン数が600程度が適当であることがわかった。

図2、図3には、それぞれ攻撃手、防御手について学習の進行にしたがって、学習したパターンに対する正答率がどのように変化したかを表わす、ニューラルネットワークの学習曲線を示す。どちらもネットワークB(中間層ニューロン数600)の結果である。

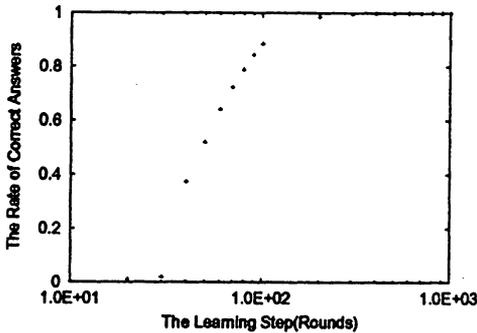


図2: 学習したパターンに対するネットワークの正答率(攻撃手)。中間層ニューロン数は600

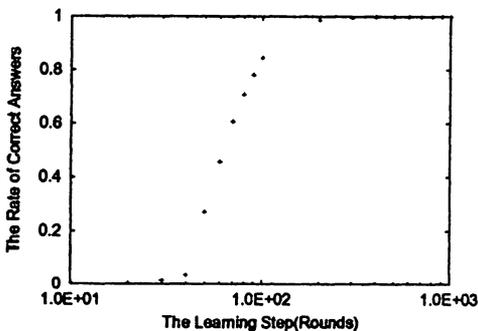


図3: 学習したパターンに対するネットワークの正答率(防御手)。中間層ニューロン数は600

3.2 未知の問題への解答能力

次に、ニューラルネットワークが学習していない、未知のパターンに対する解答能力を評価した。ネットワークは存在する詰碁のすべての局面を記憶するわけにはいかないので、未知の局面に対する解答能力が実質的にネットワークの解答能力をあらわす。

攻撃手については5000パターン、防御手については1000パターンの未知の局面について、ネットワークの解答を評価した。ただし、未知の局面に対する評価では、ネットワークの出力する候補手を上位5番目まで調べ、正解と一致するか調べた。

表3、表4に、未知のパターンに対するネットワークの正答率の評価結果を示す。

表3: 未知のパターンに対するネットワークの正答率(攻撃手)。ネットワークA: 中間層ニューロン数400, 学習回数500ラウンド, ネットワークB: 中間層ニューロン数600, 学習回数500ラウンド, ネットワークC: 中間層ニューロン数800, 学習回数1000ラウンド

候補手	正解率(%)		
	A	B	C
第1	1576(31.5)	1613(32.3)	1616(32.3)
計	1576(31.5)	1613(32.3)	1616(32.3)
第2	646(12.9)	647(12.9)	701(14.0)
計	2222(44.4)	2260(45.2)	2317(46.3)
第3	331(6.6)	337(6.7)	331(6.6)
計	2553(51.1)	2597(51.9)	2648(53.0)
第4	186(3.7)	205(4.1)	202(4.0)
計	2739(54.8)	2802(56.0)	2850(57.0)
第5	141(2.8)	132(2.6)	144(2.9)
計	2880(57.6)	2934(58.7)	2994(59.9)
誤り	2120(42.4)	2066(41.3)	2006(40.1)

ネットワークの解答の上位3位までの中に、攻撃手については、50%程度の正解が含まれ、防御手については、40%程度の正解が含まれていた。また、上位5位までの中には、攻撃手は60%程度、防御手は50%程度の正解が含まれていた。

図4、図5に、それぞれ攻撃手、防御手について学習の進行にしたがって、未知のパターンに対する正答率がどのように変化したかを示す。どちらも、ネットワークB(中間層ニューロン数600)の結果である。

表 4: 未知のパターンに対するネットワークの正答率(防御手). ネットワーク B: 中間層ニューロン数 600, 学習回数 500 ラウンド

候補手	正解率 (%) ネットワーク B
第 1 計	271(27.1)
第 2 計	93(9.3)
第 3 計	48(4.8)
第 4 計	36(3.6)
第 5 合計	473(47.3)
mistake	527(52.7)

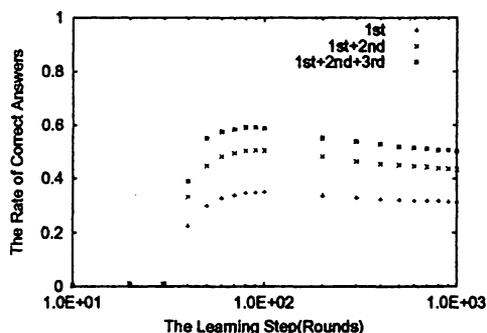


図 4: 未知のパターンに対するネットワークの正答率(攻撃手). 中間層ニューロン数は 600

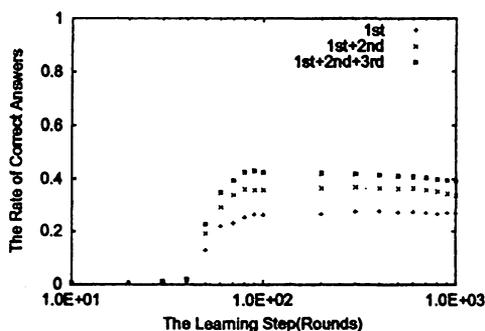


図 5: 未知のパターンに対するネットワークの正答率(防御手). 中間層ニューロン数は 600

先行研究において 9×9 の範囲の攻撃手の問題に適用した場合には, 上位 3 番目までの候補手に約 55-65%の正解が含まれていた。したがって, 9×9 の狭い範囲の問題に適用した場合と比べ, 正答能力の低下が見られる。ただ, 学習するパターンと, 評価に使う問題の組によって, ある程度は正答率が変化することが考えられるため, さらに多くの学習問題と評価問題についての実験を行ない, 比較すべきと考えられる。

ネットワークの未知の問題に対する解答能力は, 学習が完全に収束する前に最高点に達しているが, その後はわずかながら減少する傾向が見られる。これは教えられたパターンを完全に記憶した結果, ニューラルネットワークがいわゆる過学習の状態になって, 未知の問題への適応能力が若干低下したと推測される。

以上の結果から, 19×19 の広いサイズの盤面においても, ネットワークが未知の問題に対して一定の解答能力を持っていることがわかった。

3.3 人間の能力との比較

次に, ニューラルネットワークの能力を人間のプレイヤーの能力との比較を行なった。人間のプレイヤーは, 先読みも行なって正解に辿りつく。一方, ニューラルネットワークは, そのような先読みを行わずに, 瞬時に答えを返す。したがって, ニューラルネットワークの機能は, 人間の直観的な解答能力と比較すべきと考えられる。

小島らによって行なわれた, 人間のプレイヤーが詰碁問題を先読みせずに直観的に答える実験 [8] によって得られた人間の正答率を表 5 に示す。これは, ある詰碁の問題を人間のプレイヤーに与え, 人間のプレイヤーが数秒以内に直観的に解答した結果をまとめたものである。プレイヤーは複数の解答を答えることが許されており, その場合には, $(1/\text{解答数})$ が得られるポイントとなる。例えば, ある問題に 2 つの解答を答え, そのうち一方が正解だった場合には, プレイヤーは 0.5 ポイントを得る。正解率は, $(\text{正解率} = \text{獲得ポイント}/\text{問題数})$ によって求められている。

表 5 から, プレイヤーの強さによって, 正答率に相関があることがわかる。ニューラルネットワークの正答率をこの結果と比較する。

ニューラルネットワークが学習したのは, 黒先白死問題の攻撃手と白先白生問題の防御手で, コウのような問題には対応できない。表 5 の実験において

表 5: 人間のプレイヤーが数秒間で答えた場合の正答率. すべての結果は [8] より引用

難易度 (問題数)	基本 (100)	三段用 (100)	五段用 (100)	平均
強さ	正解率 (%)			
六段 a	63	68	43	58.0
六段 b	53.3	59.5	38.3	50.4
四段	58	68	38	54.7
三段	27.0	30.5	21.0	26.2
初段 a	40.2	36.8	20.3	32.4
初段 b	30.8	40.3	22.3	31.1
2 級 a	21.5	18.3	16.0	18.6
2 級 b	12.3	15.8	11.8	13.3
4 級	13.5	12.5	10.8	12.3

用いられた問題のうち、ネットワークに適用可能な問題は、基本問題が 82 問、三段用が 65 問、五段用が 66 問である。

ニューラルネットワークの解答結果を表 6 に示す。小島らの実験では、人間が複数の解答を答えることが許されているため、ネットワークの解答で複数解答を選択した場合についても得点の計算を行なった。正解率 a は、ネットワークが答えを 1 つだけ解答した場合の得点を元に計算した正解率を表わし、正解率 b は、ネットワークが上位 2 つの候補手を解答として答えた場合に得られる正解率である。

ネットワークの解答結果を人間のプレイヤーの結果と比較すると、非常に粗い見積りではあるが、2 級から初段のプレイヤーの結果に相当する正答率であることがわかった。

4 まとめと今後の課題

4.1 まとめ

本論文においては、先行研究として行っていた、ニューラルネットワークの攻撃手のみへの適用に加え、 19×19 の通常の囲碁の盤面の大きさの問題に対して、攻撃手と防御手をネットワークにそれぞれ学習させた結果を報告した。

3 層構造のフィードフォワード型ニューラルネットワークにおいて、逆誤差伝播法により、ニューラルネットワークに 5000 パターン (防御手については 4500 パターン) の問題と正解手の組を教師データと

表 6: ニューラルネットワークの正答率. 正解率 a は、ニューラルネットワークが上位 1 つの手のみを答えた時に得られる得点から求めた値で、正解率 b は、ネットワークがすべての問題に上位 2 つの手を答えたと想定した時に得られる値である

難易度 (問題数)	基本 (82)	三段用 (65)	五段用 (66)
候補手	正解数		
1 位 計	29 29	19 19	14 14
2 位 計	11 40	14 33	2 16
正解率 a	35.4(%)	29.2(%)	21.2(%)
正解率 b	24.4(%)	25.4(%)	12.1(%)

して記憶させた。その結果、ほぼ全てを正確に記憶することができた。次に未知の局面 5000 パターン (防御手については 1000 パターン) をネットワークに入力し、その正答能力を評価した。ネットワークの解答上位 5 位までの中に正解手が含まれる割合は 50-60% 程度であった。

9×9 の狭い範囲の問題に適用した場合と比べ、正答能力の低下がみられたが、ネットワークが未知の問題に対してある程度解答能力を持っていることがわかった。

また、人間のプレイヤーが直観的に解答を出した時の解答率と比較した場合、ニューラルネットワークの解答能力は、2 級から初段程度に相当することがわかった。

4.2 今後の課題

本論文で得られたニューラルネットワークの解答能力は、詰碁の一連の手順を正確に導き出すということを考えて十分とは言えない。今後はさらに学習させる問題数を増やすこと、中間層ニューロン数を始めとするネットワークの構造の最適化を行なうことなどにより、未知の問題に対するネットワークの正答能力の向上を目指す。

また、本手法においては、盤面の石の配置のみをニューラルネットワークへの入力情報として与えている。一方、Enzenberger の NeuroGo のように、連、空点、端などの情報に変換してネットワークへの入力として利用している研究も多い。吉川らによ

り、高段者は石の形そのもののみを知識として用いるのではなく、その他の概念も利用した複合的な知識を用いて正解を見つけていることが示唆されている [9]。今後は、ネットワークの最適化を行なった上で、単に石の配置のみの情報に留まらず、石のつながりなどの概念も入力情報として用いていく必要があると考えられる。

ニューラルネットワークによるシステムはその解答が間違いなく正解であるという厳密な証明を行なうことが困難である。ニューラルネットワークによって何種類かの候補手をリストアップし、その候補手を探索などにより詳しく検討して厳密な答えを求めるといった形で、探索等の手法と組み合わせていくことも必要になると考えられる。

また、詰碁だけでなく、囲碁のゲーム全体に適用可能であるようにシステムを拡張することも今後の課題である。

謝辞

本研究は科学研究費補助金（若手研究 (B) [旧奨励研究 (A)] #13780299）による助成を受けた。

参考文献

- [1] N. Sasaki, Y. Sawada, and J. Yoshimura, “A Neural Network Program of Tsume-Go”, *Lecture Notes in Computer Science 1558* (eds. Jaap van den Herik and Hiroyuki Iida), Springer-Verlag, pp.167–182, 1999.
- [2] 佐々木宣介, “ニューラルネットワークによる詰碁解答能力”, 情報処理学会研究報告, vol. 2002, No.69, pp. 45-51, 2002.
- [3] M. Enzenberger, “The Integration of A Priori Knowledge into a Go Playing Neural Network”, available from <http://www.markus-enzenberger.de/neurogo.html>.
- [4] N. Richards, D. Moriarty, and R. Miiikkulainen, “Evolving Neural Networks to Play Go” *Proceedings of the 7th International Conference on Genetic Algorithms*, East Lansing, MI, 1997.

- [5] T. Wolf, “About problems in generalizing a tsumego program to open positions”, *Proceedings of the 3rd Game Programming Workshop*, Hakone, pp.20-26, 1996.
- [6] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, “Learning representations by back-propagating errors”, *Nature*, vol. 323, pp. 533-536, 1986.
- [7] 石田芳夫, “基本詰め碁 100 題”, 日本文芸社, 1997.
- [8] T. Kojima and A. Yoshikawa, “A Two-Step Model of Pattern Acquisition: Application to Tsume-Go”, *Lecture Notes in Computer Science 1558* (eds. Jaap van den Herik and Hiroyuki Iida), Springer-Verlag, pp.146–166, 1999.
- [9] A. Yoshikawa and Y. Saito, “The Difference of the Knowledge for Solving Tsume-Go Problem According to the Skill”, ゲーム・プログラミングワークショップ '97, pp.87–95, 1997.