

動的バイナリーニューラルネットワークの学習と応用

上月, 良太 / KOUZUKI, Ryota

(出版者 / Publisher)

法政大学大学院理工学・工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学・工学研究科編 / 法政大学大学院紀要. 理工学・工学研究科編

(巻 / Volume)

55

(開始ページ / Start Page)

1

(終了ページ / End Page)

5

(発行年 / Year)

2014-03-24

(URL)

<https://doi.org/10.15002/00010410>

動的バイナリニューラルネットワークの学習と応用

LEARNING AND APPLICATION OF THE DYNAMIC BINARY NEURAL NETWORKS

上月 良太

Ryota KOUZUKI

指導教員 斎藤利通

法政大学大学院理工学研究科電気工学専攻修士課程

This paper studies a dynamic binary neural networks characterized by signum activation function and ternary weighting parameters. In order to store and stabilize a desired periodic orbit, we consider simple learning algorithm. In the learning algorithm, the threshold parameters are determined theoretically. The weighting parameters are adjusted based on the genetic algorithm. We use a simple return map on the lattice points in order to visualize the basic dynamics of the network, such as periodic phenomena and its domain of attraction. Performing basic numerical experiments for an example teacher signal of a matrix converter, we have confirmed storage of desired periodic patterns and their stability.

Key Words : *Dynamic Binary Neural Networks, Matrix Converter, Genetic Algorithm*

1. はじめに

動的バイナリニューラルネットワーク (DBNN) は、シグナム関数と3値重みパラメータを有する3層のフィードフォワード型のバイナリニューラルネットワーク (BNN) に遅延フィードバックを適用することによって構成される [1]。DBNN は、中間層のニューロン数が十分であれば、任意のブール関数を実現できる [2]-[4]。DBNN は、パラメータと初期値に依存して、様々な2値周期軌道 (BPO) を呈することができる。また、DBNN は様々な工学系応用が可能である。その例としては、信号処理や回路制御等が挙げられる [4] [6] [7]。しかしながら、これらのDBNN は取り扱う問題の規模により中間層ニューロン数の増加や、計算コストの上昇などの問題点を抱えている。

これらの問題を解決するために、いくつかの学習法が提案されてきた。本論文では、中間層ニューロン数を固定した状態でのDBNNの学習について考察する。このDBNNは、シグナム活性化関数を有し、荷重パラメータを3値、しきい値パラメータを整数としている [4] [5]。DBNNは2値パターンやデジタル回路による実装の実現に適している。

DBNNの動作を視覚化するために、グレイコード表示のリターンマップであるGray-code Based Return Map (Gmap) [1] を用いる。DBNNはN次元格子点上の2値ベクトルを扱うので、ダイナミクスを格子点上のマップに表すことができる。グレイコード表示を格子点に適用することで、Gmapを得ることができる。GmapはBPOやそれらの収束領域のような基本特性を把握するのに有用である。学習では、学習させる教師信号を用いて重みパラメータを理論的に決定することで、教師信号をDBNNに埋め込む。その後、教師信号への安定性の向上のために、遺伝的アルゴリズム (GA) に基づくスパース化学習を行う。学習アルゴリズムの機能を調べるために、基本的なBPOの教師信号による数値実験を行い、所望のBPOの埋め込みとそのBPOの安定性について検討する。

2. 動的バイナリニューラルネットワーク

DBNNは図1に示す3層構造のネットワークに遅延フィードバックを適用するように記述される。

$$\begin{aligned} x_i^{t+1} &= f \left(\sum_{j=1}^N w_{ij}^o \xi_j^t - T_i^o \right), \quad i = 1 \sim N \\ \xi_j^{t+1} &= f \left(\sum_{i=1}^N w_{ji} x_i^t - T_j \right), \quad j = 1 \sim M \quad (1) \\ f(x) &= \begin{cases} 1 & \text{for } x \geq 0 \\ -1 & \text{for } x < 0 \end{cases} \end{aligned}$$

ただし、 $\mathbf{x}^t = (x_1^t, \dots, x_N^t)$ 、 $x_j^t \in \{-1, 1\} \equiv B$ は離散時間 t における N 次元の2値ベクトルである。また、 $\xi^t \equiv (\xi_1^t, \dots, \xi_M^t)$ 、 $\xi_j^t \in \{-1, 1\}$ は t での M 次元中間出力ベクトルである。式 (1) を次式で略記する。

$$\mathbf{x}^{t+1} = F_D(\mathbf{x}^t), \quad F_D: B^N \rightarrow B^N \quad (2)$$

中間層ニューロンは3値の荷重パラメータ w_{ji} 及び整数のしきい値パラメータ T_j で特徴づけられる:

$$\begin{aligned} w_{ji} &\in \{-1, 0, 1\} \\ T_j &= \text{Integer} \end{aligned} \quad (3)$$

ただし、 $i = 1 \sim N, j = 1 \sim M$ は中間層ニューロン数 M を制御する内部パラメータである。同様に出力層ニューロンは2値の荷重パラメータ w_{ji}^o としきい値パラメータ T_i^o で特徴づけられる。

DBNNのダイナミクスを視覚化するために、グレイコードに基づくリターンマップ (Gmap) を導入する。Gmapは N 次元の2値ベクトルをグレイコードで表現し、2値ベクトルの写像を格子点上の写像に変換したものである。全ての入力に対する出力から周期解数や、各周期解への収束軌道や領域を調べることができる。簡単のため、 N 次元2値ベクトルのグレイコードを C_1 から C_{2^N} で表すことにする。図2に図1のDBNNをGmapに表した例を示す。

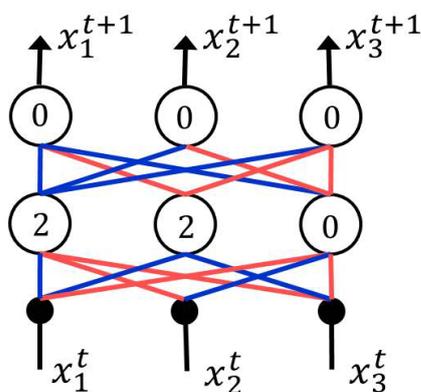


図 1 DBNN の例。青の線は $w_{ji} = 1$ 、赤の線は $w_{ji} = -1$ で繋がっていることを示す。また $w_{ji} = 0$ は線が繋がっていないことを示す。

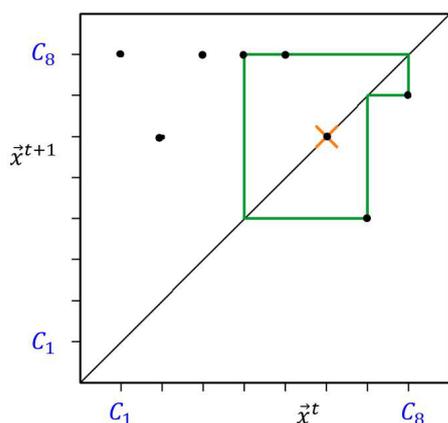


図 2 図 1 の DBNN の Gmap。黒の点は入力に対する出力を示す点であり、4 周期の BPO(緑色の軌道) が 1 つあることが解る。この BPO の軌道上にある点を Binary Periodic Point(BPP) といい、斜線上の点は 1 周期の固定点という。また、その他の点は最終的に周期軌道に収束する Eventually Periodic Point(EPP) となる。

3. 学習アルゴリズム

次に教師信号が次式で示すような周期 T の BPO である場合を対象として、中間層ニューロン数を固定した状態での DBNN の学習の問題を考える。

$$z^1, \dots, z^T, z^{T+1} = z^1, z^t = (z_1^t, \dots, z_N^t) \quad (4)$$

ここで、教師信号を構成する z_j^t と z_j^{t+1} は DBNN の変数 x_j^t と x_j^{t+1} に各々対応する。学習の目的は、教師信号を埋め込む重みパラメータとしきい値パラメータを探すことである。本論文では重みパラメータの初期値について、以下の式 (5) のように w_{ji} は教師信号の入力 z_i^j を、 w_{ij}^o は教師信号の出力 z_i^{j+1} を代入する。

$$w_{ji} = z_i^j, w_{ij}^o = \begin{cases} 1 & \text{if } z_i^{j+1} = 1 \\ 0 & \text{if } z_i^{j+1} = -1 \end{cases} \quad (5)$$

この様にして決めた重みパラメータで教師信号を埋め込む場合、しきい値パラメータを以下の式 (6) により決めることで

埋め込みが保証される。ここで、しきい値パラメータ T_j は式 (6) によって得られる値において、2 番目に大きい値とする:

$$T_j = \sum_{i=1}^N w_{ji} z_i^j + 1, j \neq t, t = 1 \sim N \quad (6)$$

$$T_i^o = 1 - \sum_{j=1}^M |w_{ji}^o|$$

これに対して、教師信号への収束率の大きい重みパラメータ w_{ji} を学習によって決定する。その方法として、スパース化を用いた安定化手法を考える。スパース化を用いた学習

教師信号への収束率の大きい重みパラメータ w_{ji} の決定のために、GA に基づく学習を使ってスパース化を行う。同アルゴリズムでは、3 値重みパラメータ w_{ji} で構成される N 個の要素からなるベクトルを染色体とする。

Step1:(初期遺伝子の決定)

k をある染色体の番号とし、 t 番目の初期染色体を手動で生成した重み値パラメータ行列 $w_{ji}(k)$ とする。初期染色体の値は教師信号を重みパラメータに代入したものとす。このような染色体を G_e 個用意する。進化過程では、評価の高い上位 1 割の染色体は保存することとする。便宜上、その染色体を 1 番目 ($k = 1$) の染色体とする。また、 g を現在の世代数とし、最大世代数を G_{max} とする。

Step2:(探索対象の決定)

現在の各染色体の持つ重みパラメータ行列 w_{ji} から得られる BPO の中に教師信号の点を多く含む染色体を探索対象とする。

Step3:(評価・並び替え)

G_e 個の染色体それぞれについて、染色体の w_{ji} から得られた DBNN が持つ BPO に含まれる教師信号へ 1 離散時間で落ち込む初期値の割合 (CR) が多い順に染色体を並べ替える。CR は以下の式 (7) で定義される:

$$CR = \frac{\#(1 \text{ 離散時間で教師信号に落ち込む初期値})}{2^N} \quad (7)$$

Step4:(淘汰・増殖)

上位 1 割の染色体を次世代へ残す。また、DBNN の持つ BPO に教師信号 BPO が含まれない染色体は淘汰し、上位 1 割の染色体を淘汰した数だけ増殖する。

Step5:(GA に基づく探索)

G_e 個の染色体について、GA を用いて全ての入力教師信号により多く収束するようなネットワークとなる染色体を決定する。下位 9 割の染色体の全ての要素 w_{ji} について突然変異確率 P_m に従い、1 または -1 の要素には 0 の値を代入し、0 の要素には式 (5) の値を代入する。

尚、この学習では交叉を用いないものとする。

Step6:(終了判定)

$g = g + 1$ とし、Step3 へ戻る。 g が最大世代数となったら学習終了となる。

4. 数値実験

GA の学習性能を検証するために、 N 次元 M 周期の教師信号を用いて学習を行った。この実験では各パラメータを以

下のように設定する。

$$(G_{max}, G_e, P_m) = (100, 100, 0.1) \quad (8)$$

この実験で用いた教師信号は、3 相交流電圧を別な周波数、大きさの3 相交流電圧に変換する、マトリクスコンバータのスイッチングパターンを用いる。である BPO(表 1) である。

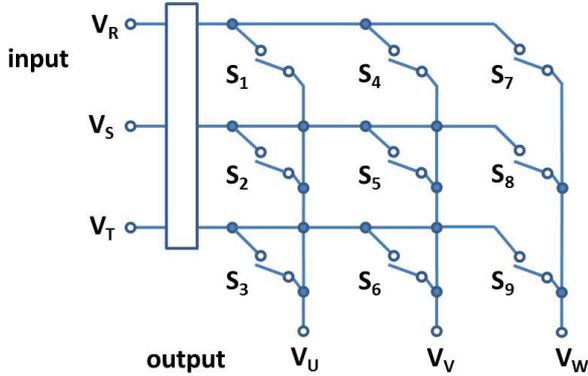


図 3 回路図：マトリクスコンバータ

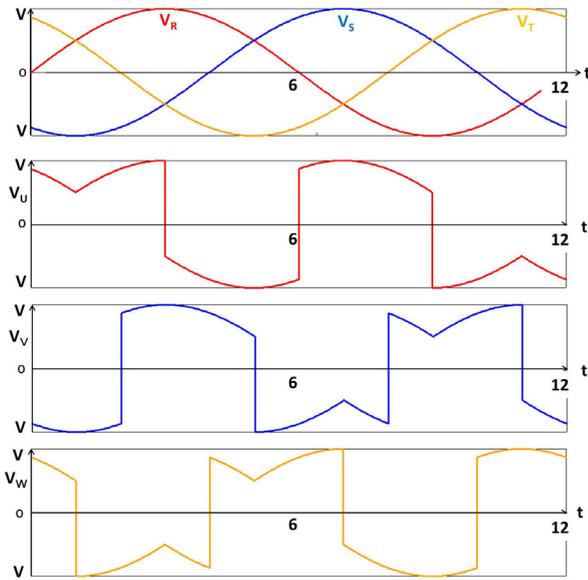


図 4 出力波形：マトリクスコンバータ

この教師信号について学習を行った。図 5 に基づく学習を行う前の DBNN に対応する GMAP を示す。図?? において、CR は初期状態において 14% となった。またこの図におけるその他の入力は全て不動点 0 に収束していることがわかった。続いて、GA に基づく学習結果を示す。表 3 は GA に基づくスパース化学習の結果の重みパラメータ w_{ji} である。この学習結果の DBNN に対応する GMAP を図 6 に、図 7 に GA に基づくスパース化学習の学習過程を示す。GA に基づく学習を行った結果、例 1 において図 6 では CR が 78% となり、図 7 からわかる通り CR の上昇が起きていることからこの学習が機能していることがわかる。また、この学習の結果から得られた軌道では不動点 0 に収束する初

表 1 教師信号 BPO

$z(1)$	$(-1, -1, +1, -1, +1, -1, -1, -1, +1)$
$z(2)$	$(+1, -1, -1, -1, +1, -1, -1, +1, -1)$
$z(3)$	$(+1, -1, -1, +1, -1, -1, -1, +1, -1)$
$z(4)$	$(-1, -1, +1, +1, -1, -1, -1, -1, +1)$
$z(5)$	$(-1, -1, +1, +1, -1, -1, +1, -1, -1)$
$z(6)$	$(-1, -1, +1, -1, -1, +1, -1, +1, -1)$
$z(7)$	$(-1, +1, -1, -1, -1, +1, -1, +1, -1)$
$z(8)$	$(-1, +1, -1, +1, -1, -1, +1, -1, -1)$
$z(9)$	$(-1, +1, -1, -1, -1, +1, -1, +1, -1)$
$z(10)$	$(+1, -1, -1, -1, -1, +1, +1, -1, -1)$
$z(11)$	$(+1, -1, -1, -1, -1, +1, -1, -1, +1)$
$z(12)$	$(+1, -1, -1, -1, -1, +1, -1, -1, +1)$

表 2 例 1：スパース化前の重みパラメータ w_{ji}

j	w_{j1}	w_{j2}	w_{j3}	w_{j4}	w_{j5}	w_{j6}	w_{j7}	w_{j8}	w_{j9}	T_j
1	-1	-1	+1	-1	+1	-1	-1	-1	+1	6
2	+1	-1	-1	-1	+1	-1	-1	+1	-1	6
3	+1	-1	-1	+1	-1	-1	-1	+1	-1	6
4	-1	-1	+1	+1	-1	-1	-1	-1	+1	6
5	-1	-1	+1	+1	-1	-1	+1	-1	-1	6
6	-1	-1	+1	-1	-1	+1	-1	+1	-1	6
7	-1	+1	-1	-1	-1	+1	-1	+1	-1	6
8	-1	+1	-1	+1	-1	-1	+1	-1	-1	6
9	-1	+1	-1	-1	+1	-1	+1	-1	-1	6
10	+1	-1	-1	-1	-1	+1	+1	-1	-1	6
11	+1	-1	-1	-1	-1	+1	-1	-1	+1	6
12	-1	+1	-1	-1	+1	-1	-1	-1	+1	6

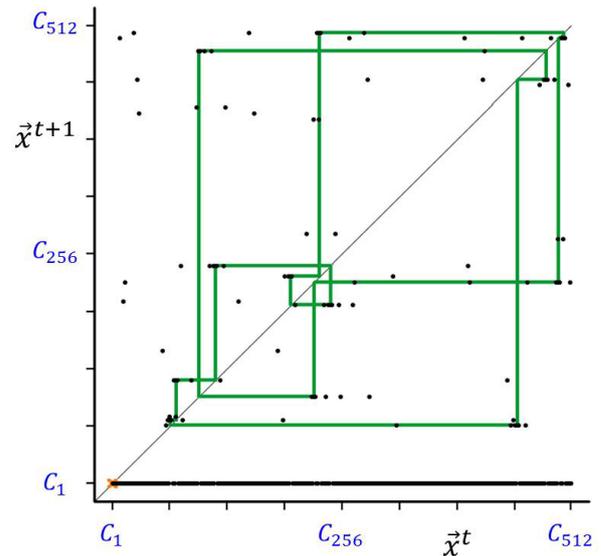


図 5 教師信号 BPO のスパース化前の Gmap:g=0,CR=14%

期値が 22%残っていることがわかる。ここでもう一つの学習結果の例として例 2 を以下に示す。表 4 は GA に基づくスパース化学習の結果の重みパラメータである。この学習結果の DBNN に対応する GMAP を図 8 に、図 9 に GA に基づくスパース化学習の学習過程を示す。

例 2 において、図 8 では CR が 61% となり、図 9 から収束率の上昇が起きていることからこの学習が機能しているこ

とがわかる。また、例2の結果から得られた軌道では不動点が教師信号BPOとは別に2つ現れており、そちらへ収束する割合は39%ということがわかる。

考察

例1、例2より重みパラメータをスパース化することで安定性の向上が見られた。このことから、教師信号を初期重みパラメータとした場合、不要な結合が多く存在し、それらを切り離すことで安定化したと考えられる。

5. むすび

本論文では、DBNNにおける中間層数を固定した状態でのGAに基づく学習について考察した。それらによって得られるDBNNのダイナミクスを系統的に解析するために、GMAPを導入した。GMAPを用いることで、DBNNのダイナミクスを調べることができた。典型的な結果として、その教師信号BPOを銘記することができ、教師信号以外の初期値から出発する信号の収束特性を確認することができた。GAに基づく学習により、教師信号への収束安定性の向上が確認することができた。

今後の課題として、DBNNの学習過程の解析、ダイナミクスの解析、安定化のメカニズムの解明等が挙げられる。

謝辞: 本論文は著者が法政大学大学院工学研究科電気工学専攻に在学中の2年間、同大学理工学部電気電子工学科教授斎藤利通博士の指導下で行ったものである。研究活動を遂行するにあたり、同教授から懇切に御指導、御鞭撻下さった同博士に深謝致します。

最後に法政大学理工学部電気電子工学科斎藤利通研究室の皆様にはいろいろな有益な御討論・ご助言を戴きました。ここに深謝致します。

参考文献

- 1) R. Ito, Y. Nakayama and T. Saito, Analysis and Learning of Periodic Orbits in Dynamic Binary Neural Networks, Proc. IEEE-INNS/IJCNN, to appear, 2012.
- 2) D. L. Gray and A. N. Michel, A training algorithm for binary feed forward neural networks. IEEE Trans. Neural Networks, 3, 2, pp. 176-194, 1992.
- 3) J. H. Kim and S. K. Park, The geometrical learning of binary neural networks. IEEE Trans. Neural Networks, 6, 1, pp. 237-247, 1995.
- 4) R. Kouzuki, and T. Saito, Learning of Simple Dynamic Binary Neural Networks, Trans. IEICE, E96-A, 8, pp. 1775-1782, 2013
- 5) K. Nowara and T. Saito, Guaranteed storing of limit cycles into a discrete-time asynchronous neural network, Trans. IEICE, E75-A, 11, pp. 1579-1582, 1992.
- 6) P. L. Rosin, Training cellular automata for image processing, IEEE Trans. Image Process., 15, 7, pp. 2076-2087, 2006.
- 7) W. Wada, J. Kuroiwa, S. Nara, Completely reproducible description of digital sound data with cellular automata, Physics Letters A 306, pp. 110-115, 2002.

表 3 例 1 : 学習後の重みパラメータ w_{ji}

j	w_{j1}	w_{j2}	w_{j3}	w_{j4}	w_{j5}	w_{j6}	w_{j7}	w_{j8}	w_{j9}	T_j
1	0	0	+1	0	+1	0	0	0	0	1
2	0	-1	-1	0	+1	0	0	0	0	2
3	+1	0	0	+1	-1	-1	0	+1	-1	3
4	0	0	+1	0	-1	0	0	0	+1	2
5	-1	-1	0	0	-1	0	0	-1	-1	5
6	0	0	+1	-1	-1	0	+1	0	-1	3
7	-1	0	-1	0	-1	0	0	+1	0	3
8	0	+1	0	1	0	0	0	-1	-1	3
9	-1	+1	-1	-1	0	-1	0	-1	-1	7
10	+1	0	-1	-1	0	+1	+1	0	0	4
11	0	0	-1	0	-1	0	0	-1	+1	3
12	0	+1	-1	0	+1	0	-1	0	0	3

表 4 例 2 : 学習後の重みパラメータ w_{ji}

j	w_{j1}	w_{j2}	w_{j3}	w_{j4}	w_{j5}	w_{j6}	w_{j7}	w_{j8}	w_{j9}	T_j
1	-1	-1	0	-1	+1	0	0	-1	0	4
2	+1	0	-1	-1	0	-1	0	0	0	3
3	+1	-1	-1	+1	-1	-1	-1	+1	-1	6
4	0	0	0	+1	-1	-1	-1	-1	+1	3
5	0	-1	0	+1	0	-1	+1	-1	0	4
6	0	-1	+1	0	0	+1	0	+1	0	1
7	0	+1	-1	0	-1	+1	0	0	0	3
8	-1	0	-1	+1	0	-1	+1	0	-1	5
9	0	+1	0	0	+1	0	0	-1	-1	3
10	0	-1	-1	-1	0	0	+1	-1	-1	5
11	0	0	-1	0	0	+1	0	-1	+1	3
12	-1	0	-1	-1	0	-1	0	0	+1	4

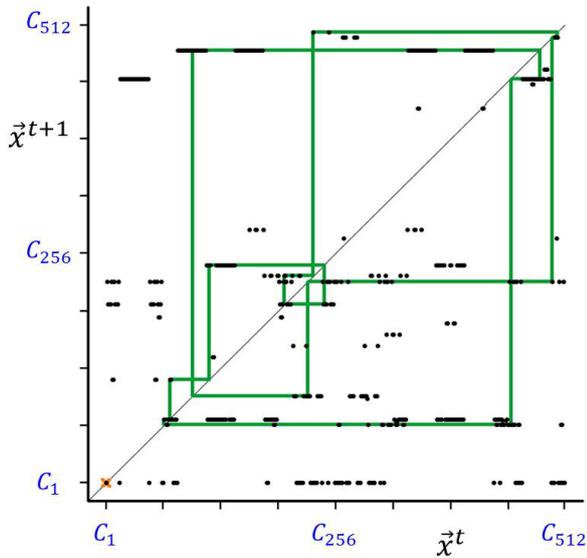


図 6 例 1 : 教師信号 BPO の GA に基づく学習後の Gmap:g=100,CR=78%

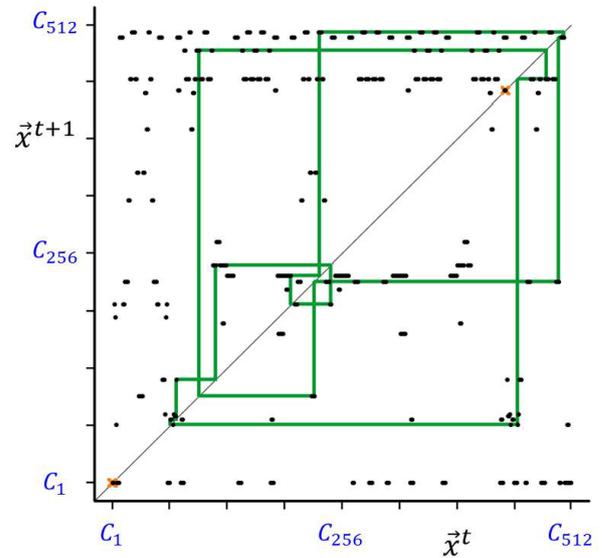


図 8 例 2 : 教師信号 BPO の GA に基づく学習後の Gmap:g=100,CR=61%

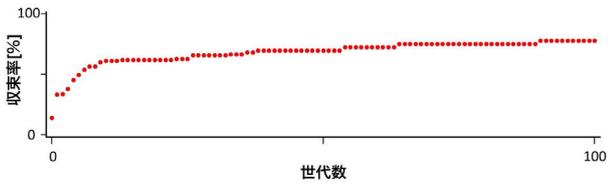


図 7 例 1 : 教師信号 BPO の GA に基づく学習後の学習過程

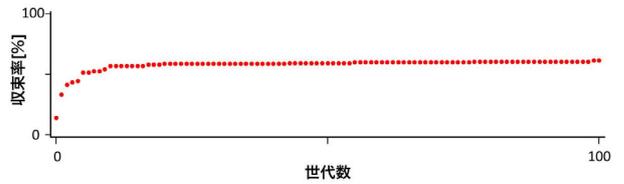


図 9 例 2 : 教師信号 BPO の GA に基づく学習後の学習過程