

教師あり学習を動的に補助する複合型学習システム

日大生産工(学部) ○鈴木 彩音 日大生産工 山内 ゆかり

1 まえがき

近年の人工知能に対する関心は日々高まっており、その研究も進んでいる。例えば、教師あり学習では誤差逆伝播法があり、教師なし学習では自己組織化マップ(SOM)、自己増殖型ニューラルネットワーク(SOINN)、Hopfield、深層学習(Deep Learning)があり、強化学習にはQ学習がある。以上のようにそれぞれ様々な手法が提案されてきた。しかし、実際の脳ではこれらが合わさり、複合して使われている。実際の脳のように、異なる学習システムを組み合わせることで、より高い学習能力が期待できると考えられている。

従来、笹川らは教師あり学習、教師なし学習、強化学習の三つの学習システムを利用したbrain-like学習システムを提案した。1)従来、複合型学習システムの多くが強化学習を主としているのに対し、この研究は教師あり学習を主要とするシステムであった。しかし、後の研究では組み込んだ強化学習が利用されず、局所解に対する問題も他の方法で解決に臨んでいる。

本研究では、教師あり学習を動的補助する複合型学習システムを提案し、新たな学習システムによる高速化と局所解の解決を検討する。

2 研究背景および従来手法

笹川らは自己組織化機能局在型ニューラルネットワークを提案しており、brain-like学習システムはその構成を利用し強化学習を加えたものである。2)

このシステムは教師あり学習(supervised learning)を行うSL part、教師なし学習(unsupervised learning)を行うUL part、強化学習(reinforcement learning)を行うRL partの三部構成になっている。図1にbrain-like学習システムの手順を示す。

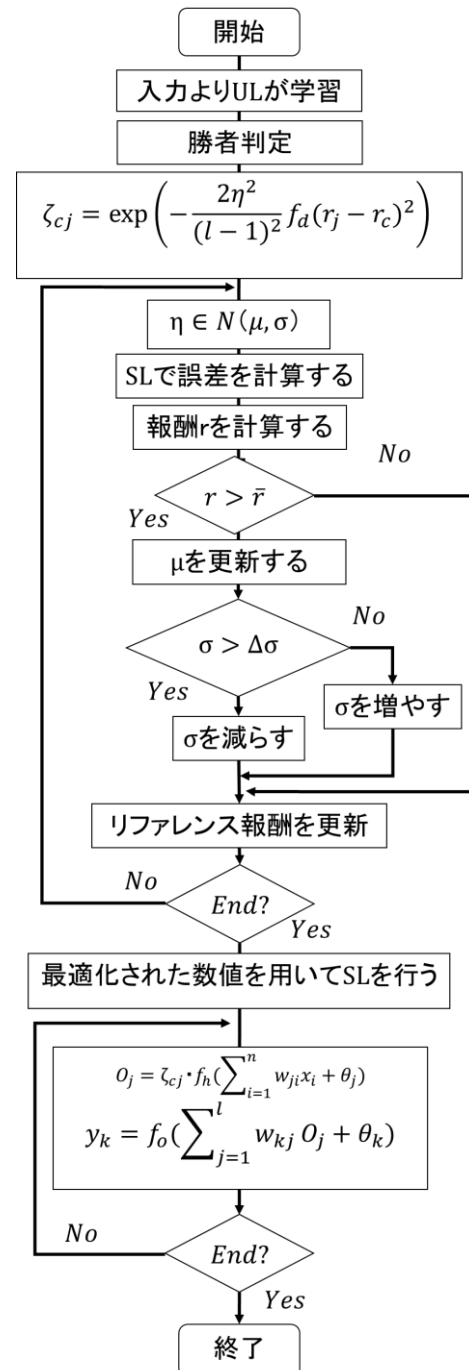


図1 brain-like学習システムのフローチャート

Integrated Brain-like Learning System

Ayane SUZUKI and Yukari YAMAUCHI

UL partはSL partと同じ数の中間ノードを持つ。これはSL partの制御を行うための発火強度を決める部であるためである。中間ノードと同じく、入力や出力も同様である。

最も入力ベクトル x に類似している参照ベクトルが m_c である。最整合ノード c を決定する条件を以下に示す。

$$c = \arg \min_j \{\|x - m_j\|\} \quad (1)$$

m_j は参照ベクトルであり、この c は決定後、ガウス関数によって計算される。制御信号の計算式を以下に示す。

$$\zeta_{cj} = \exp\left(-\frac{2\theta^2}{(l-1)^2} f_d(r_j - r_c)^2\right) \quad (2)$$

$0 < \zeta_{cj} < 1$ は最整合ユニット c が選ばれた場合の j 番目のユニットの出力である。 θ はガウス関数の形状を決めるパラメータで、 $f_d(\cdot)$ はノード間の距離関数であり、 r_j 、 r_c は出力空間におけるそれぞれの位置である。

RL partでは、UL partの出力を調整する部である。ここで強化比較手法を使いSL partの性能を最適化するための発火強度制御信号、自動的に決定する。

$\eta(t)$ を決定するための式を以下に示す。

$$\eta(t) \in N(\mu(t), \sigma(t)^2) \quad (3)$$

t はステップ、 $\mu(t)$ は t の時の密度関数の平均値、 $\sigma(t)$ は密度関数の標準偏差である。RL partでは μ 及び σ の値を更新することでパラメータ η の最適値を探している。報酬の式を以下に示す。

$$r(t) = \frac{1}{E(t)} \quad (4)$$

$E(t)$ はステップ t での学習終了後のSL partの学習誤差であり、この値に応じて方策を更新する。方策の行進には強化比較手法を用いる。これまで受け取った報酬の逐次平均値であるリファレンス報酬 $\bar{r}$ という基準値を設けた。更新の式を以下に示す。

$$\bar{r}(t+1) = \bar{r}(t) + \alpha_r (r(t) - \bar{r}(t)) \quad (5)$$

$$\mu(t+1) = \mu(t) + (1 - e^{-(r(t) - \bar{r}(t))^2}) \cdot (\eta - \mu(t)) \quad (6)$$

$$\sigma(t+1) = \begin{cases} \gamma_{inc} \cdot \sigma(t) & \text{if } \sigma(t) \leq \Delta\sigma(t) \\ \gamma_{dec} \cdot \sigma(t) & \text{if } \sigma(t) > \Delta\sigma(t) \end{cases} \quad (7)$$

$\alpha_r(r(t) - \bar{r}(t))$ は $r(t) > \bar{r}(t)$ の時、 μ 及び σ が更新される。更新が終了次第、SL partの初期重みをランダムに設定し、これを繰り返す。

主となるSL partでは、RL partで求めた η を使用し、UL partで調整した教師信号を元に結

果を出力する。中間層ユニットの各出力である O_j と出力 y の式を以下に示す。

$$O_j = \zeta_{cj} \cdot f_h\left(\sum_{i=1}^n w_{ji}x_i + \theta_j\right) \quad (8)$$

$$y_k = f_o\left(\sum_{j=1}^l w_{kj}O_j + \theta_k\right) \quad (9)$$

$f_o(\cdot)$ と $f_h(\cdot)$ は出力と中間層の活性化関数であり、 w_{ji} と w_{kj} は入力から中間層間と中間から出力層間の各重みであり、 θ_j と θ_k は中間と出力層のバイアスであり、 l は中間層のユニット数である。 ζ_{cj} は式(2)の θ にRL partで求めた η を入れ、計算する。

3 提案手法

従来手法では η を適正な値に更新したが、そこで行った強化学習が膨大な計算量になっているにもかかわらず、SL partの学習が無駄になっている。そのため、 η を動的に適正な値にする。提案手法の手順を以下に示す。

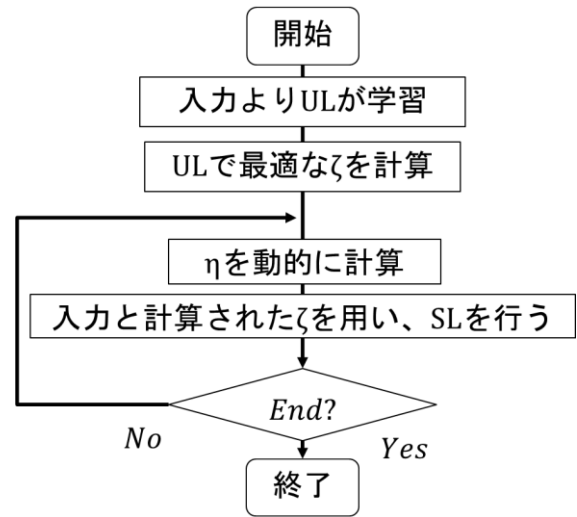


図2 提案手法のフローチャート

図2の通り、SL partを行う前に式(10)(11)により動的に $\eta(t)$ を計算する。

$$dc = \|x - m_c\| \quad (10)$$

$$\eta(t) = \eta_0 \left(\frac{1}{1 - \sqrt{dc} + \theta_s} \right)^2 \quad (11)$$

dc により最整合ノードとの距離を求める。 $\eta(t)$ は学習回数 t の時の η であり、 η_0 は η の初期値である。この式により、 η は最整合ノードに近ければ発火強度を強め、遠ければ発火強度を弱める効果が期待される。

4 実験結果および検討

従来研究と同様に、ベンチマーク問題としてTwo-spirals問題を用い、提案システムの性能が自動的に最適化されるかのシミュレーションを行う。図3は使用した螺旋を示したものである。

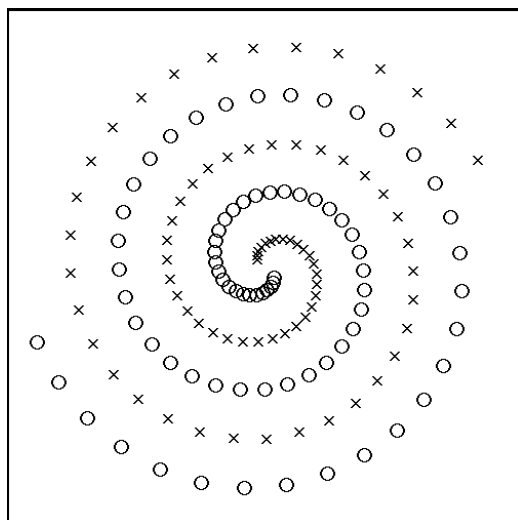


図3 Two-spirals問題

この問題は描かれた赤と青の二つの入れ子状の螺旋を判別するものである。図3では○が赤、×が青となっている。螺旋の点の数は76ずつの計152点である。

表1 実験環境

BP	$\eta = 0$
BP+SOM	$\eta(t) = \eta_0 \in \{1.0, 3.0, 6.0\}$
Brain-like	$\eta(t) \in N(\mu, \sigma)$ $\mu_0 = \eta_0 \in \{1.0, 3.0, 6.0\}$
提案手法	$\eta(t) = \eta_0 \left(\frac{1}{1 - \sqrt{dc} + \theta_s} \right)^2$ $\eta_0 \in \{1.0, 3.0, 6.0\}$ $\theta_s = 0.0002$
UL part	Learning steps 10,000 $\theta = 6.0$ $A_{turn} = 0.002$
RL part	SL Learning steps 10,000 Learning steps 5,000 $\gamma_{inc} = 1.03$ $\gamma_{dec} = 0.99$ $\alpha_r = 0.1$
SL part	Learning steps 500,000

入力とは各点の座標の2つとなり、中間ノードは従来と同じく各システム同じ9つとする。出

力は表示する色の数値に使用した。その教師信号は赤が0.1、青が0.9である。

使用したアルゴリズムはSOM、Backpropagation(以下BPと示す)である。BPは η を0にすることによりUL partの影響を受けなくなることを利用し、0にして比較する。brain-like学習システムでは、比較するため適切な η を決定した後のSL partを500000回行った。

教師あり学習 (BP)のみ、教師あり学習と教師なし学習(BP+SOM)、Brain-like学習システム、提案手法の誤差をグラフにした。

実験の結果は以下に示す。

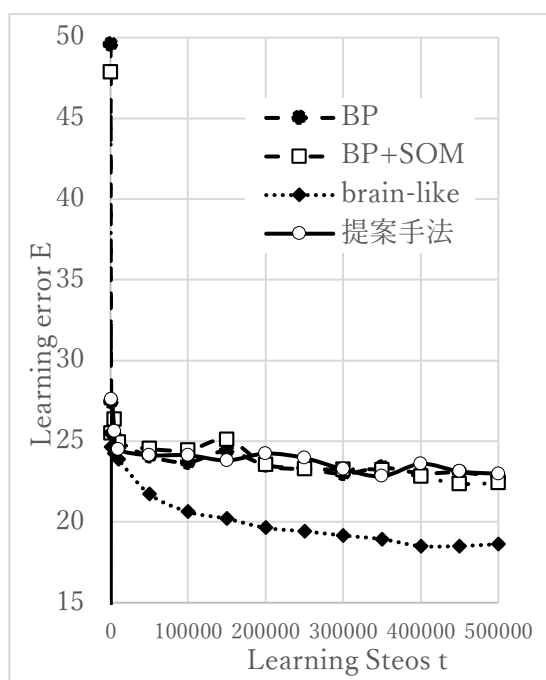


図4 初期値 $\eta_0 = 1.0$ の学習誤差

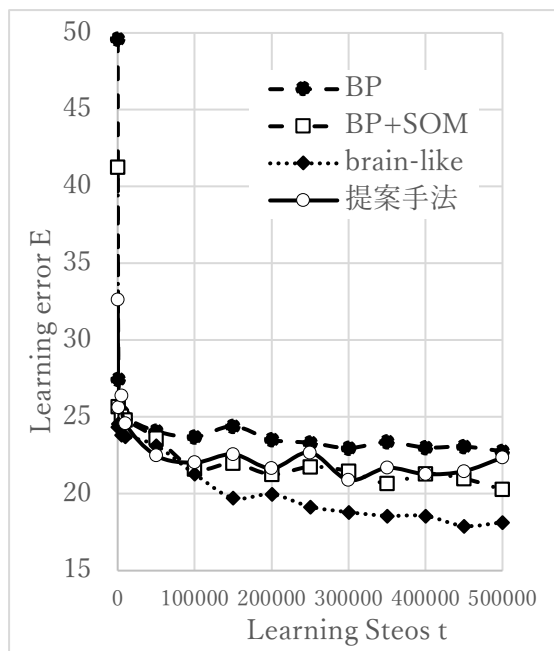


図5 初期値 $\eta_0 = 3.0$ の学習誤差

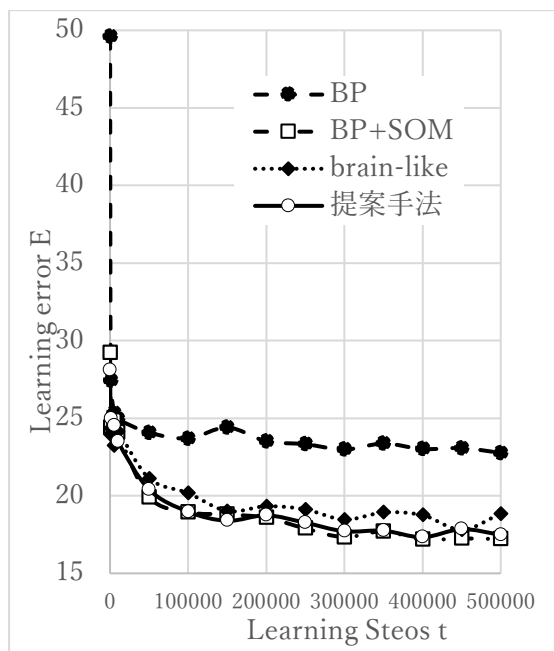


図6 初期値 $\eta_0 = 6.0$ の学習誤差

学習誤差を見ると、動的に η の値が変わることによってBPのみやBP+SOMよりも少ない誤差になることもあるが、Brain-like学習システムを上回る誤差にはならなかった。

また、図6のように η が適切な $\eta_0 = 6.0$ で計算するとかなり良い結果になるが学習誤差が安定しない結果となった。

5 まとめ

機能局在型ニューラルネットワークの η を動的に適正な値にする学習システムを提案し、そのシステムの実験としてTwo-spirals問題を用いた。

従来のBrain-like学習システムの強化学習の計算をせずに従来の結果を出すことを試みたが、よい結果を得ることが出来なかった。

勝者ノードとの近傍を利用し、 η を動的に変化させることで、従来手法には及ばずとも多少の良い結果を出すことは出来たが、安定しないという課題が残った。

【参考文献】

- 1) 笹川 隆史, 古月 敬之, 平澤 宏太郎, 教師あり学習・教師なし学習・強化学習を複合したbrain-like学習システム, 電気学会論文誌C, Vol. 126, No.9 (2006) p.1165-1172.
- 2) 笹川 隆史, 古月 敬之, 平澤 宏太郎, 自己組織化機能局在型ニューラルネットワーク, 計測自動制御学会論文集, Vol. 41, No.1 (2005) p.67-74.