確率的非巡回神経回路網のための並列相関学習法

5H－3

三谷　光照　　大堀　隆文　　渡辺　一央

北海道工業大学　電気工学科

1　まえがき

ゆらぎをともなう線形しきい値ユニット群からなる，確率的非巡回神経回路網 (PFN　$Fig.1$) のモデルとそのための並列相関学習法を提案し，理論的検討を行ったので報告する．

2　ニューロンモデル

PFNを構成する各ニューロンk($Fig.2$) は,(1) ゆらぎ発生源 $n_k(t)$ を有し,(2) 情報伝達諸量が入力パターン $\mathbf{x}(p)$ によって変化するが，一定入力に対しても期間的ゆらぎが存在する，ことをのぞいては McCulloch-Pitts のニューロンモデルと同じである．ニューロン kの $\mathbf{x}(p)$ に対する情報伝達量は以下のように記述できる．ただし,$f\{\cdot\}$ は線形しきい値関数を表す．

シナプス前ニューロン群出力：

$\mathbf{q}_k(p,t) \in \{0,1\} = \mathbf{q}_k\{\mathbf{x}(p), \mathbf{w}, \mathbf{n}(t)\}$　(1)

入力総和　$: u_k(p,t) = \mathbf{w}_k^T \mathbf{q}_k(p,t)$　(2)

瞬時活性値　$: v_k(p,t) = u_k(p,t) + n_k(t)$　(3)

瞬時出力値　$: r_k(p,t) = f\{v_k(p,t)\}$　(4)

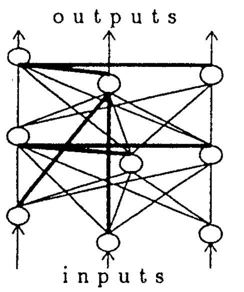

Fig.1
Probabilictic Feedforward NN

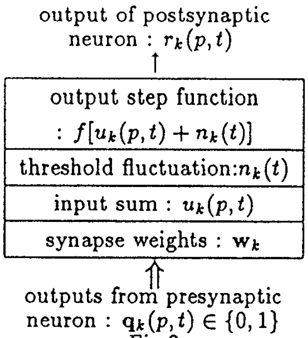

Fig.2
Model of a neuron k for PFN

3　確率的情報伝達機構

PFNの情報伝達量について離散時間確率系で考察する．あるパターンの入力期間中 (入力継続期間) におけるニューロンkの伝達諸量 $\mathbf{q}_k(p,t)$, $u_k(p,t)$, $n_k(t)$, $v_k(p,t)$, $r_k(p,t)$ を確率変数として $\mathbf{q}_k$, u_k, n_k, v_k, r_kで表し，各確率変数の分布を $Q(\mathbf{q}_k)$, $U(u_k)$, $N(n_k)$, $V(v_k)$, $R(r_k)$ で表す．ゆらぎに

条件 1〕パターン入力継続期間 Tp 内の小区間 Ts のゆらぎ分布 $N'(n_k)$ は，微小な偏り（活性値バイアス）ν_kを有していることを除いては,Tp 内のゆらぎ分布と等しい．

すなわち，$\nu_k \simeq 0$, $\langle\nu_k\rangle_{Tp} = 0$, $N'(n_k) = N(n_k - \nu_k)$

を設定する．ここで，$\langle\rangle_{Tp}$は Tp 期間の期待値を表す．式 (1)〜(4) より，伝達諸量の確率分布は，以下の特徴，関係を持つ．以下，時間変数 t，パターン変数 p，期待値の期間 Tp を省略する．

$V(v_k) = \sum_{q_k} Q(\mathbf{q}_k) N(v_k - \mathbf{w}_k^T \mathbf{q}_k - \nu_k)$　(5)

$R(r_k = 1) = \sum_{q_k} Q(\mathbf{q}_k) \int_0^\infty N(v_k - \mathbf{w}_k^T \mathbf{q}_k - \nu_k)\, dv_k$　(6)

$R(r_k = 0) = 1 - R(r_k = 1)$　(7)

ニューロン出力 r_kは後続のニューロン jの入力ベクトル q_jを構成し，$S(j)$ を jのシナプス前ニューロン群の集合とすると，その確率分布 $Q(\mathbf{q}_j)$ は，次式で表される．

$Q(\mathbf{q}_j) = Pr\{r_k : k \in S(j)\}$　(8)

よって,PFNは式 (5)〜(8) に示した確率分布変換を繰り返し，情報伝達を行っていると解釈できる．

一方，可視ニューロン群 S_vに対する平均出力誤差 $\langle e\rangle$ は，ζ_jを教師信号として，次式で表せる．

$\langle e\rangle = \sum_{j\in S_v} \{R(r_j = 1) - \zeta_j\}^2 = \sum_{j\in S_v} \{\langle r_j\rangle - \zeta_j\}^2$　(9)

4　並列相関学習法

4.1　誤差最小化基本原則

パターン入力継続期間 Tp の平均出力誤差 $\langle e\rangle$ を，各ユニットkへの結合係数 $\mathbf{w}_k$に関する最急降下法によって，最小化する．結合係数の修正量$\Delta_p \mathbf{w}_k$は，次式で与えられる．

$$\Delta_p \mathbf{w}_k^T = -\epsilon \frac{\partial\langle e\rangle}{\partial \mathbf{w}_k} = -\epsilon \frac{\partial\langle e\rangle}{\partial \nu_k} \cdot \frac{\partial \nu_k}{\partial R(r_j = 1)} \cdot \frac{\partial R(r_j = 1)}{\partial \mathbf{w}_k} \quad (10)$$

4.2　誤差駆動項 $\mathrm{ED}(k) = \frac{\partial\langle e\rangle}{\partial\nu_k}$ の推定

Ts 区間の平均出力誤差 $\langle e\rangle_{Ts}$をνに関してマクロリン展開すると次式を得る．

$$\langle e\rangle_{T_s} \fallingdotseq \langle e\rangle_{T_s}|_{\nu=0} + \frac{\partial\langle e\rangle_{T_s}}{\partial\nu}|_{\nu=0} \cdot \nu \quad (11)$$

条件 1 より，$\langle e\rangle_{T_s}|_{\nu=0}$は $\langle e\rangle_{Tp}$ と等しくなる．ここで，式 (11) の両辺に，右からν^Tを乗じ，期間 Tp の期待値を求めると，

$$\langle\langle e\rangle_{T_s} \cdot \nu^T\rangle \fallingdotseq \frac{\partial\langle e\rangle}{\partial\nu} \langle\nu\nu^T\rangle \quad (12)$$

となる．ゆらぎに，

条件 2〕PFNでは，各ニューロン内ゆらぎ n_kは，互いに独立．

を設定すると，$\langle\nu\nu^T\rangle = \mathrm{diag}\{\langle\nu_k^2\rangle :^\forall k\}$ であり，式 (12) は次式となる．

$$\mathrm{ED}(k) = \frac{\partial\langle e\rangle}{\partial\nu_k} \fallingdotseq \frac{\langle\langle e\rangle_{T_s} \cdot \nu_k\rangle}{\langle\nu_k^2\rangle} \quad (13)$$

4.3　局部可塑項 $\mathbf{LP}(k) = \frac{\partial\nu_k}{\partial R(r_j=1)} \cdot \frac{\partial R(r_j=1)}{\partial \mathbf{w}_k}$ の推定

$\mathbf{LP}(k)$ は式 (6) から次式となる．

$$\mathbf{LP}(k) = \frac{\sum_{q_k} Q(\mathbf{q}_k) N(-\mathbf{w}_k^T \mathbf{q}_k) \mathbf{q}_k^T}{\sum_{q_k} Q(\mathbf{q}_k) N(-\mathbf{w}_k^T \mathbf{q}_k)} \quad (14)$$

ここで,$N(\cdot)$ が分散σ^2を有する正規分布のとき，

$$N(x_0) = \frac{2}{\sigma^2} \int_0^\infty (x + x_0) \cdot N(x + x_0)\, dx \quad (15)$$

の関係を用いて式 (14) を変形すると，次式となる．

$$\mathbf{LP}(k) = \frac{\langle r_k n_k \mathbf{q}_k^T\rangle_{Tp}}{\langle r_k n_k\rangle_{Tp}} \quad (16)$$

4.4　学習則 (結合係数修正式)

式 (10),(13),(16) から以下の簡潔な学習則を得る．

$$\Delta_p \mathbf{w}_k^T = -\epsilon \frac{\langle\langle e\rangle_{T_s}\nu_k\rangle_{Tp}}{\langle\nu_k^2\rangle_{Tp}} \cdot \frac{\langle r_k n_k \mathbf{q}_k^T\rangle_{Tp}}{\langle r_k n_k\rangle_{Tp}} \quad (17)$$

5　むすび

線形しきい値ユニット群からなる PFNモデルと，その学習法として，平均出力誤差と活性値バイアスとの相関，および，局部可塑項からなる並列相関学習法を提案した．理論的検討から，本学習法は，

1. 線形しきい値ユニット群からなる PFNに所期の入出力特性を満たさせることが可能,

2. すべての結合係数を並列かつ同時に修正可能,

などの特徴を有することが推論される．今後，性能支配要因や構造化学習について検討する．

Probabilistic Feedforward Neural Network
and its Parallel Correlation Learning
Mitsuaki MITANI, Takahumi OOHORI, Kazuhisa WATANABE
Hokkaido Institute of Technology
Teine-ku,Sapporo,JAPAN