神経回路網特論
情報システム工学科　　2005 年 6 月 2 日
伊達 章　　更新：2005 年 6 月 2 日

自己組織化神経回路モデル (3)

- **論文紹介**:
  P. Foldiak
  Forming sparse representations by local anti-Hebbian learning,
  Biological Cybernetics, vol. 64, pp. 165-170, 1990.
- キーワード
  - 局所表現，分散表現，スパース表現
  - 情報分離，情報圧縮，独立成分，教師なし学習
- いろいろな情報表現（発火率の違いに着目）
  - ニューロン数 n，活動度 $p, 0 \leq p \leq 1$,（興奮しているニューロンの割合）
  - 局所表現 ($p = 1/n$)
    - 「おばあさん細胞」表現とも呼ばれる
    - ヘブ学習と競合学習で実現
    - 表現形成の最も単純なモデルが SOM
  - 完全分散表現 ($p = 1/2$)
  - スパース表現 (例： $p = \log n$)
  - それぞれの利点，容量，耐ノイズ性，$\cdots$
- ネットワークの構造
  - 入力 $x_1, x_2, \cdots, x_m \in \{0, 1\}$
  - 出力 $y_1, y_2, \cdots, y_n \in \{0, 1\}$
- 学習アルゴリズム
  - 入力素子と出力素子間：
    ヘブ学習 $\Delta w_{ij} = \alpha(y_i y_j - p^2) \rightarrow$ 学習の平衡状態で $E[y_i y_j] = p^2$
  - 出力層内：反ヘブ学習 $\Delta q_{ij} = \beta y_i (x_j - q_{ij})$
  - 出力素子の閾値：$\Delta t_i = \gamma(y_i - p) \rightarrow$ 平均活動度を p にする]
- 例 1：縦線，横線の学習
  - 図 1： ネットワークの構造（入力層 64 個，出力層 16 個）
  - 図 2： 入力信号の例 ($8 \times 8 = 64$ ピクセル)
    入力は縦線，横線の組合わせ．各縦線，横線が確率 1/8 で独立に on/off
    入力層には 64 素子 $\rightarrow$ 表現可能なパターンの総数 2^{64} 個
    そのうち縦線と横線の組合わせのパターンの総数 2^{16} 通り
  - 図 3： 自己組織化の様子（$t = 0, 400, 800, 1200$）
- 例 2：文字の学習

– 図 1： ネットワークの構造（入力層 120 個，出力層 16 個）

– 図 4： 自己組織化の様子（$t = 0, 4000, 8000, 16000$）
入力は Sun-3 ワークステーションのフォント．各文字を入力する頻度は，ある英文で使われている頻度に設定．→入力層には 120 素子あるので表現可能なパターンの総数は 2^{120} 個．そのうち出現するのは 82 個．出力層には 16 素子あるので $2^{16} = 256$ までは 1 対 1 で表現できる構造にはなっている（実際には 256 パターンのうち，48 パターンが使われるようになった）．

– 表 1： 学習後，各文字がどういう出力（16 ビット）に変換されるようになったか示した一覧表．多対 1 の写像になっていることに注意．例：k, B, R, F を入力すると，どの場合も出力は 0000010000010000 になる．

– 表 2：符合の性質

  * 入力
    $A_j \cdots j$ 番目の文字の出現頻度．$j = 1, \cdots, 82$.
    $E(A) = -\sum_j A_j \log_2 A_j = 4.34$
    j 番目の文字の入力パターン $(a_{1j}, a_{2j}, \cdots, a_{120j})$
    $q_i = \sum A_j a_{ij} \cdots i$ 番目のビットの平均値（ビットが 1 になる確率）
    $e(A, a) = -\sum_i [q_i \log q_i + (1 - q_i) \log(1 - q_i)] = 24.14$

  * 出力
    A_j を入力したときに得られる出力 $\boldsymbol{b}_k = (b_{1k}, b_{2k}, \cdots, b_{16k})$
    $B_k \cdots \boldsymbol{b}_k$ を観測する確率
    $E(B) = -\sum_k B_k \log_2 B_k = 4.22$
    各文字パターンを 16 ビットの列に変換したところ，出力パターンの総数は 82 個より小さく，48 個になっていた（$k = 48$）．そのため，もとのエントロピーより小さくなっている．
    $p_i = \sum B_k b_{ik} \cdots i$ 番目のビットが 1 になる確率
    $e(B, b) = -\sum_i [p_i \log p_i + (1 - p_i) \log(1 - p_i)] = 5.86$

    スパース性：16 ビットのうち，1 の個数は $0 \sim 4$ になっている．
    ※出現頻度が高い文字ほど 1 の数が少ないパターンに符合化されている．

  * 各ビットのエントロピーの和は何を意味するか．
    $H(X), H(Y) \leq H(X, Y) \leq H(X) + H(Y)$ の関係を思い出す．

    $H(p_i), \leq H(p_i, p_2, \cdots, p_{16}) \leq H(p_1) + H(p_2) + \cdots + H(p_{16})$
    16 個のビットが互いに独立な場合に $=$ が成立する．
    $H(p_i, p_2, \cdots, p_{16}) = 4.22$ [bits]
    $H(p_1) + H(p_2) + \cdots + H(p_{16}) = \sum_{i=1}^{16} H(p_i) = 5.86$ [bits]

    冗長性:$[e(B, b) - E(B)]/E[B] = (5.86 - 4.22)/4.22 = 0.388$
    この意味：すべてのビットが独立だと仮定すると 5.86 ビットの情報を得られる．実際には 4.22 ビットしか得られなかった．これをどうして著者らは冗長性と定義したか．→ 4.22 ビットの情報を表現するのに 5.86 ビット使用，と考

える．入力パターンについても同様に考える：
冗長性:$[e(A,a)-E(A)]/E[A]=(24.14-4.34)/4.34=4.562$

- パラメータ $\alpha,\beta,\gamma,\lambda$ などは，どう決定したか.
- 速修 確率と情報理論
  - 情報の量をいかに定めるか
  - エントロピーとは
  - 確率変数の依存性，確率変数が互いに独立であるとは