

パ ー セ プ ト ロ ン

河 野 隆 一*

1. 序 言

“Perceptron”は1957年にCAL (Cornell Aeronautical Lab.) の F. Rosenblatt が提案した一つの“Cognitive System”に賦与した名称であって、以後アメリカ Office of Naval Research との Contract によって研究が進められている。Perceptron は工学上では図形認識装置に関する一つの試みとしての意味を持つと同時に一方では学習させる機械という点から“Self Organizing system”ひいては Brain Model の一つの考え方として Biophysics 的な面からも興味ある研究として注目されている。Perceptron は特定の pattern に対して特定の応答をうるようにならぬ方法であらかじめ構成されたものでなく、学習によってその対応を示すようになしうという特徴をもっている。Perceptron の特性を研究するには理論的考察、ディジタル計算機によるシミュレーション、あるいは実際に“hardware”を作って行なうことなどが行なわれているが、以下では主に理論的な基礎的事項に注目して Perceptron の解説を行なう。CAL では衆知のように実際に Perceptron “Mark 1”なる hardware を製作し研究に供している。

2. 3-Layer Series Coupled Perceptron

Perceptron の根本的な構成要素は S-unit, A-unit, R-unit の三つである。S-unit (Sensory unit) は外部刺激を電気信号に変換するもので、刺激されたときに出力 +1, あるいは -1 を出す。A-unit (Association unit) は幾つかの S-unit または他の unit からの入力を加え合わせ、その和がある閾値 θ を越えるときに励起されて出力 v を出す unit で、特定の強勢 (reinforcement) の方法にしたがってその出力 v は Δv なる変化をうける。R-unit (Response unit) はある閾値 θ を持ち、全入力の和が $+\theta$ より大ならば出力 +1 を、 $-\theta$ より小ならば出力 -1 (あるいは 0) を呈する。S-A-R なる三つの Layer からなり、A-unit 間に結合のないものを 3-Layer Series Coupled Perceptron

* 三菱電機鎌倉製作所

と称する。

2.1 Simple Perceptron

Simple Perceptron は 3-Layer で、しかも一つの R-unit のみをもつものであるが (図 1), 他の構成をもつ Perceptron の基本となる。S-A 間の結合はランダムであり、全 A-unit の出力が R-unit へ結合されている。

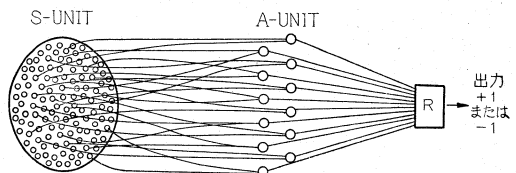


図 1 単純パーセプトロン

各種刺激は一つの R-unit によって二つのクラスに分類される。R-unit から A-unit へは強勢があるが、その方法によって Perceptron は α -system, r -system などに分類される。 $\Delta v_k(i)$ を k 番目の A-unit の出力 v_k が刺激 S_i によって受ける強勢量とすると、

$$\alpha\text{-system では } \Delta v_k(i) = \rho_i a_k(i)$$

$$r\text{-system では } \Delta v_k(i) = \rho_i \left(a_k(i) - \frac{\sum a_k(i)}{N_a} \right)$$

ただし

$$\rho_i = \begin{cases} +1 & \text{もし } S_i \text{ が第 1 のクラスに属するなら} \\ -1 & \text{第 2 } \end{cases}$$

$$a_k(i) = \begin{cases} 1 & \text{もし } k \text{ 番目の A-unit が } S_i \text{ になり} \\ & \text{励起されるなら} \\ 0 & \text{励起されないなら} \end{cases}$$

N_a は A-unit の数である。なお v_k の値に上下限のあるときを特に λ -system と称する。Perceptron の能力は、与えられた問題に対して、ある Perceptron が、ある Training system にしたがって、ある刺激列によって訓練されたときにいかなる確率で正しい応答を与えるかで示される。Training system には、訓練中、刺激に対する応答の正誤にかかわらず強勢を行なう Forced Learning System、応答の正しいときのみ正しい強勢を行なう Error Correction System、刺激に注目せず応答のみを観察して強勢を行なう Spontaneous Learning System などがある。刺激列には、

ある刺激から次のどんな刺激に移るかがある確率によって定まるランダム刺激列 (Random Sequence) とか, ある定まった刺激の列, あるいはその繰り返しからなるような固定刺激列 (Fixed Sequence) とかの種類がある。

行列による表示

ある特定の S-A 結合様式をもった Perceptron が与えられた問題に対して解を持つためには次の式を満足する \mathbf{v} および \mathbf{u} が存在すればよいことが容易に見られる。

$$A\mathbf{v}=\mathbf{u} \quad (\text{sign } \mathbf{u}=\boldsymbol{\rho}) \quad \dots\dots(1.1)$$

ただし, 行列 $A=\{a_{ki}(i)\}$, ベクトル $\mathbf{v}=[v_k]$

$$\boldsymbol{\rho}=[\rho_i] \quad \mathbf{u}=[u_i] \quad \text{である。}$$

(1.1) 式はさらに $G=\{g_{ij}\}=AA'$ を与えて

$$G\mathbf{x}=\mathbf{w} \quad (\text{sign } \mathbf{w}=\boldsymbol{\rho}) \quad \dots\dots(1.2)$$

を満足する \mathbf{x} および \mathbf{w} の存在を調べる問題と等価である。 g_{ij} は刺激 S_i および S_j によって共に励起される A-unit の数である。特定の Perceptron に対しては G が定まるのであるが, 通常 S-A 結合はランダムと考えるので, 特定の結合をランダムなあらゆる結合様式からとった一つの標本だという考えに立てば上述の g_{ij} から $Q_{ij}=E(g_{ij}/N_a)$ なる Q_{ij} を定義して行列 Q を用い

$$Q\mathbf{x}=\mathbf{w} \quad \dots\dots(1.3)$$

なる式を論ずることができる。

Q_{ij} の計算

Q_{ij} は実は刺激 S_i, S_j のいずれによっても励起される A-unit の数の割合であるが, S-A 間の結合の規則に支配される。一つの A-unit に対して (+1) の入力数を X , (-1) の入力数を Y とすれば, それらが S-unit の全集合から独立かつ一様に取られると, いわゆる “Bionominal Model” となつて

$$Q_{ij}=\sum_{\substack{E_i+E_c-I_i-I_c \geq \theta \\ E_j+E_c-I_j-I_c \geq \theta}} P_X(E_i, E_j, E_c) P_Y(I_i, I_j, I_c)$$

となる。ここで

$$P_X=\frac{X!}{E_i!E_j!E_c!(X-E_i-E_j-E_c)!} \\ \times A_i^{E_i}A_j^{E_j}C^{E_c}(1-A_i-A_j-C)^{X-E_i-E_j-E_c}$$

$$P_Y=\frac{Y!}{I_i!I_j!I_c!(Y-I_i-I_j-I_c)!} \\ \times A_i^{I_i}A_j^{I_j}C^{I_c}(1-A_i-A_j-C)^{Y-I_i-I_j-I_c}$$

ただし, A_i は刺激 S_i のみによって励起される S-unit の全体に対する比, A_j は刺激 S_j のみによって励起される S-unit の全体に対する比, C は S_i および S_j によって励起される S-unit の全体に対する比である。

Learning Curve

Q 行列がえられた場合, 次のような解析との関連において Perceptron の能力を考察することができる。今 u_x を R-unit への入力とすると, 種々の S-A 結合様式をもつ多くの Perceptron については u は確率変数と考えられて, Tschebysheff の不等式を用いると, ある標本としての Perceptron について

$$\left\{ \begin{array}{l} \text{Prob. } \{u_x > 0\} \geq 1 - \frac{1}{\mu^2(u_x)/\sigma^2(u_x)}, \quad \mu > 0 \\ \text{Prob. } \{u_x < 0\} \geq 1 - \frac{1}{\mu^2(u_x)/\sigma^2(u_x)}, \quad \mu < 0 \end{array} \right\} \quad \dots\dots(1.4)$$

したがって μ が必要な符号をもち, μ^2/σ^2 を大としうれば任意に選んだ Perceptron について u が正しい解を与える確率を 1 に近づけることができる。ただし μ は u_x の平均値, σ は標準偏差である。

[例] α -system, Forced Learning で Fixed Sequence で刺激を与えるとする。その Sequence の含む全刺激数を T とし, ある刺激 S_i の T 中に含まれる割合を P_i とすれば Sequence を示したのちにおいて刺激 S_x に対する u_x は

$$u_x=T \sum_k \sum_j \rho_j P_j a_k(i, x)$$

となる。ここで $a_k(j, x)$ は k 番目の A-unit が S_j , S_x なる刺激に対して共に励起されるなら +1, そうでなければ 0 とする。

これから

$$\mu(u_x)=T \sum_k \sum_j \rho_j P_j E\{a_k(i, x)\}=TN_a \sum_j \rho_j P_j Q_{jx}$$

同様に

$$\sigma^2(u_x)=N_a T^2 \sum_i \sum_j \rho_i \rho_j P_i P_j (Q_{ijx}-Q_{ix}Q_{jx})$$

ここで Q_{ijx} は刺激 S_i, S_j, S_x によって励起される A-unit の数の割合で, Q_{ij} と類似の方法で計算できる。したがって解の存在する条件は, S_x の応答を +1 とすると,

$$\sum_j \rho_j P_j Q_{jx} > 0 \quad \dots\dots(1.5)$$

S_x が正しい応答を与える確率は (1.4) 式を用いて

$$P \approx 1 - \frac{\sum_i \sum_j \rho_i \rho_j P_i P_j (Q_{ijx}-Q_{ix}Q_{jx})}{N_a (\sum_j \rho_j P_j Q_{jx})^2} \quad \dots\dots(1.6)$$

この確率 ρ を N_a, θ , Sequence の長さ T などを変数として求めればいわゆる learning curve がえられる。(1.6) 式によれば N_a を大にすれば P は 1 に近づき, また同じ Sequence を繰り返しても結果は変わらない。Random Sequence の場合も類似の解析ができるが, この場合 P はもちろん T に関係する。 r -system についても同様の計算によって P が計算できる。訓練の方法が Error Correction System によると

きは, learning curve は計算されていないがある方法で正しい応答をうるために必要な訓練の最大回数が計算されているが, 一般に Error Correction は Forced Learning よりも能率がよく, より一般性のある学習法であると考えられている. λ -system の場合にはその能力が A-unit の出力に限界のない場合と比較して正しい応答を与える確率は λ -system でないときに比して多少劣化することが容易に想像されよう.

2.2 複数の R-unit をもつ Perceptron

一般には区別すべきパターンの種類は通常 2 個以上であることが多いから, 応答を考える R-unit が複数個必要となってくる. この場合の構成には, 図 2 に示すごとく 3 種考えられるが,

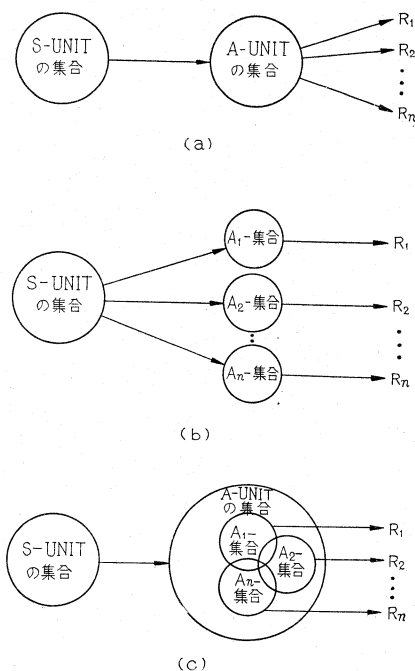


図 2 複数の R-unit をもつパーセプトロン

(a) の場合は, 各 A-unit から各 R-unit への結合があって

(b) の場合は, A-unit からの出力は各一つで各 R-unit へはランダムに選んだ A-unit が結合し,

(c) はその中間的な場合であるが, A-unit の数が大になるにしたがって (a), (b) の場合と同一になる.

ある刺激 S_x に対して正しい応答を与える確率を $P_x(R_1 \cdots R_n)$ とすると,

$$P_x(R_1 \cdots R_n) = \prod_i P_x(R_i, N_i) \cdots (1.6)$$

ただし, $P_x(R_i, N_i)$ は図 2 (b) の場合を考えたととき, N_i 個の Source A-unit をもって i 番目の R-unit

が正しい応答を与える確率である. 学習の方法として Error Correction を用いる場合には各 R-unit に対して Error Correction を行なうべきで, 全体の結果に対して強勢の可否を定める Error Correction では学習の不可能な場合がある.

複数の R-unit の場合は図形の種類にいかなる R-unit の状態を対応させるか, すなわちコード化の問題が重要である. 一つのパターンに一つの R-unit の状態を対応させるような方法 (position code) となんらかの二進数表示を対応させる方法 (configuration code) およびその混合などが考えられるが, けっきょく問題とする全刺激に対して

$$E(P_x) = \frac{1}{n} \sum_x P_x = \frac{1}{n} \sum_x \prod_i P_x(R_i)$$

を最大にするようなコードを求めることが問題となる.

2.3 3-Layer Series Coupled Perceptron による判別の例

[例 1] 20×20 S-unit, 4×20 の水平の Bar および垂直の Bar を二つのクラスと考え, Forced Learning で Fixed Sequence, α -system, $x=3$, $y=1$, $\theta=2$ なる Binomial モデルとすると,

$N_a=100$ のとき	$P=0.95$
$N_a=200$ "	$P=0.995$
$N_a=300$ "	$P \approx 1$ を得る.

Training Sequence の長さは各刺激に対して各 1 回.

[例 2] 上と同じ問題で, Random Sequence とする.

$P_j=1/40$ $N_a=100$ $p=0.88$ $T=200$
$N_a=300$ $p=0.99$ $T=1000$

[例 3] 上と同じ図形で, 各 Bar に順に番号をつけ, その奇数番目の Bar と偶数番目の Bar とを二つの class とすれば α -system, Binomial, $x=5$, $y=7$, $\theta=2$, $N_a=100$ で Fixed Sequence によって $P=0.77$

[例 4] 例 2 と同じ問題で Error Correction を用いたとき, $N_a=300$, $x=3$, $y=1$, $\theta=2$ とすれば

$$T=250 \quad \text{で} \quad p \approx 1$$

[例 5] Alphabet 26 文字の判別

Mark 1 による実験. 各文字は一種, 512 A-unit で Error Correction によって各文字につき 15 回示して $p \approx 1$.

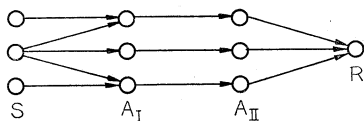
2.4 3-Layer Series Coupled Perceptron の限界

今まで見てきたように系の特性の判別に Q_{ij} が重要な役割をもっていた. すなわち (1.3)~(1.6) 式からも容易に検討されうるように一般的に二つの刺激が同じクラスに分離されるためには, その 2 刺激に

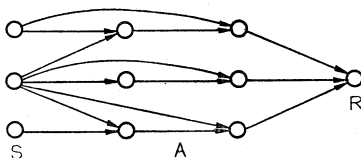
対する Q_{ij} が互いに相反するクラスに分類される二つの刺激に対する Q_{ij} よりも大きいことが望ましい。このことは Series Coupled 3-Layer においては、S-unit における図形同志の“overlap”が非常に重要となってくる。すなわちそれが大なれば、それだけ Q_{ij} を大きくする助けとなる。したがって一つの図形を学習させて、その図形の、たとえば平行移動したような図形をも学習せずに同一クラスと判断するようないわゆる“generalization”の能力は、可能ではあっても貧弱であることは容易に察せられよう。したがってあるクラスのパターンを全部学習させるのにも学習の時間を比較的長く要することになる。

3. 4-Layer および Cross Coupled Perceptron

S-A-R なる構成をもった Perceptron から一步進めると A-unit の Layer を二つ以上もつ Multi-Layer Perceptron (図 3 (a)) を考えることができる。その中には今まで述べた 3-Layer Perceptron に乏しかった Generalization の性質をより良く持つものがある。3-Layer でも A-unit 間にいわゆる Cross-coupling があるとき (図 3 (b)) は Multi-Layer と類似の取り扱いができるので同じ章で述べる。



(a) 4 層パーセプトロン



(b) A-unit 間に相互結合のあるパーセプトロン

図 3

3.1 固定の A-A 結合をもつ Perceptron (4-Layer)

S — A_I — A_{II} — R
(固定) (可変)

のような構成をもつ系では $Q_{ij}^{(1)}$ から $Q_{ij}^{(2)}$ を見出すことができればあとは 3-Layer の場合と同じ問題に帰せられる。ただし $Q_{ij}^{(2)}$ を正確に求めることは実際にはなほだ複雑であるが、A_I-unit の数が無限に多いと仮定すれば、A_I における励起された A-unit の数はちょうど $Q_{ij}^{(1)}$ などであると考えられるから、3-

Layer のときと同様の計算で $Q_{ij}^{(2)}$, $Q_{ij}^{(2)}$ などを求めることができる。この考えによれば A-unit の Layer をさらに増したときの影響を考察できるであろう。S-A_I および A_I-A_{II} 間の結合をたとえば Binominal Model と仮定すると、各 A_I-unit, A_{II}-unit への正入力力は負入力の数やその閾値が Perceptron の特性に影響を与えるが、一例として A_I-A_{II} が (-1) の結合のみでかつ A_{II}-unit の閾値を 0 としたとき、系は二つの隣り合った刺激よりも全然離れた刺激のほうが密接に関係したりするような事情が出てきて、3-Layer のときはむずかしかった前述の例 3 のような場合の判別は容易となるが、generalization の能力はかえって悪化する。同様の結合をもつ A-unit の Layer を幾つも増していくにしたがってこの傾向が助長される。

“Similarity-Constrained” Perceptron

ある“Similarity”の関係にある図形において“generalization”の能力をもたせようとして、同じ 4-Layer の Perceptron で次のような系を考える (図 4 参照)。

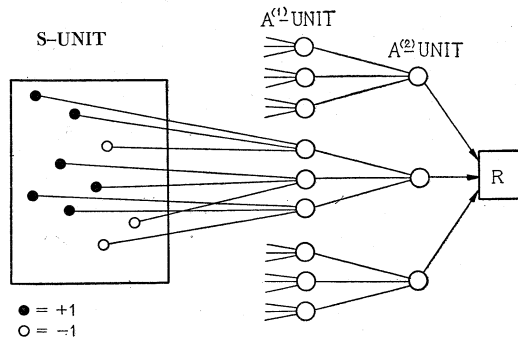


図 4 類似の関係で制約されているパーセプトロン

(1) 各 $A^{(1)}$ -unit は閾値 θ , x 個の +1, y 個の -1 入力をもち、 $A^{(2)}$ -unit の一つへ一つの出力を与える。

(2) 各 $A^{(2)}$ -unit は m 個の $A^{(2)}$ -unit からの入力を受け、閾値 +1 をもつ。

(3) $A^{(1)}$ から $A^{(2)}$ への結合の値はすべて +1 とする。

(4) ある $A^{(2)}$ -unit に結合される $A^{(1)}$ -unit はすべてある“similarity” (たとえば平行移動) の関係の下に同一と見なされる点からの結合をもつ。この系でたとえば T を平行移動の関係とすると、二つの“Random dot stimuli” S_i および $S_i' = T(S_i)$ に対して $Q_{ii'}^{(2)}$ を求めると、S-unit の数が非常に大なる場合、互いに独立な“Random dot stimuli”に対する $Q_{ij}^{(2)}$ と等しい結果が得られている。もちろん、有限の S-unit では $Q_{ii'}^{(2)} > Q_{ij}^{(2)}$ であるが差異はきわめて小で

ある。換言すれば Random dot stimuli のときは “generalization” に対してなんらの向上が見られないということである。同様の考察を、たとえば長方形のパターン間について行なった結果は、互いに “Similarity” の関係にある二つの刺激 S_i, S_i' に対する $Q_{ii'}$ は他の刺激 S_j に対する Q_{ij} よりも大となって “generalization” に対する向上がある程度認められる。

3.2 可変 A-A 結合をもつ 4-Layer Perceptron

これは “generalization” の立場から非常に有力と考えられる系である。A⁽¹⁾-A⁽²⁾ 間の結合を可変にした系において、特に系が次のような規則にしたがうとする (図 5)。

- (1) A⁽¹⁾ の各 unit a_r ⁽¹⁾ は閾値 θ をもつ。
- (2) A⁽²⁾-A⁽²⁾ は $v=\theta$ なる固定値および可変の結合 u_{rs} をもつ。
- (3) 刺激は $t, t+Dt, t+2Dt, \dots$ に与えられる。
- (4) もし A⁽¹⁾-unit a_r ⁽¹⁾ がある時刻 t_1 に励起され、 $t+Dt$ に A⁽²⁾-unit a_s ⁽²⁾ が励起されるならば、 u_{rs} は強勢をうける。
- (5) 全 u_{rs} は常に $(\delta Dt)u_{rs}$ なる減衰をうけている。

このような系に対する解析はまず刺激 S_i に対する s 番目の A⁽²⁾-unit への全入力信号 α_s ⁽ⁱ⁾(t) の成り行きを求めてのち Q_{ij} ⁽²⁾ を求める。

さて、

$$\begin{aligned}\alpha_s(i)(t) &= \theta a_s(i) + \sum_{r=1}^{N_a} u_{rs}(t) a_r(i) \\ &\equiv \beta_s(i) + r_s(i)(t) \quad \dots\dots (3.1)\end{aligned}$$

時刻 t_0 に刺激 S_j , t_0+Dt に S_k が加えられたとき、上記規則 (4), (5) により、

$$\begin{aligned}u_{rs}(t_0+2Dt) - u_{rs}(t_0+Dt) \\ = (\eta Dt) a_r(j) \phi\{\alpha_s^{(k)}(t_0+Dt)\} \\ - (\delta Dt) u_{rs}(t_0+Dt) \quad \dots\dots (3.2)\end{aligned}$$

ただし、 $\phi(x) = \begin{cases} 0 & x < \theta \text{ のとき} \\ 1 & x \geq \theta \text{ のとき} \end{cases}$

$a_s(i)$ などは 2.1 で定義した量である。

(3.1), (3.2) 式から

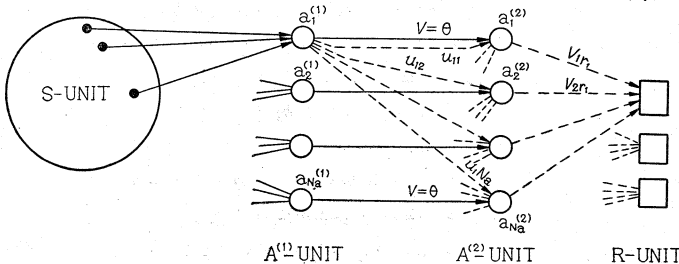


図 5 4 層パーセプトロン可変 (A-A)

$$\begin{aligned}r_s(i)(t_0+2Dt) - r_s(i)(t_0+Dt) \\ = \sum_r [u_{rs}(t_0+2Dt) - u_{rs}(t_0+Dt)] a_r(i) \\ = (\eta Dt) \phi\{\alpha_s^{(k)}(t_0+Dt)\} \sum_r a_r(j) a_r(i) \\ - (\delta Dt) \sum_r u_{rs}(t_0+Dt) a_r(i)\end{aligned}$$

この式から $a_r(j) a_r(i) = Q_{ij}$ ⁽¹⁾ の関係を用い、 $\Delta t \rightarrow 0$ と考えてけっきょく r に関する方程式

$$\frac{dr_s(i)}{dt} = (N_a \eta) \sum_{j=1}^n C_{ij} \phi(\beta_s(j) + r_s(j)(t)) - \delta r_s(i)(t) \quad \dots\dots (3.3)$$

$$C_{ij} = \sum_{k=1}^n Q_{ij}^{(1)} f_{kj} \quad (n: \text{パターン数}) \quad \dots\dots (3.4)$$

がえられ r に関して解くことによって $\lim_{t \rightarrow \infty} r_s(i)(t)$ の状況を知りうる。ただし上式で f_{kj} は刺激対 $S_k S_j$ の起こる割合である。

平衡状態では

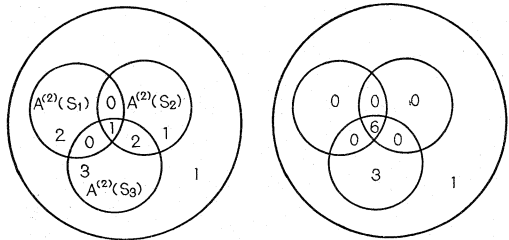
$$r_s(i) = \frac{N_a \eta}{\delta} \sum_{j=1}^n C_{ij} \phi(\beta_s(j) + r_s(j)) \quad \dots\dots (3.5)$$

が成立する。実際に求めるには

$$r_{s,m}(i) = \frac{N_a \eta}{\delta} \sum_{j=1}^n C_{ij} \phi(\beta_s(j) + r_{s,m-1}(j)) \quad \dots\dots (3.6)$$

なる m についての逐次繰返しを用いる。

さて、 $\beta_s(i) + r_s(i) \geq \theta$ ならば a_s ⁽²⁾ は励起されるから、 S_i に対して励起される A⁽²⁾-unit の集合 $A^{(2)}(S_i)$ 、したがって $A^{(2)}(S_i) \cap A^{(2)}(S_j) \equiv Q_{ij}$ ⁽²⁾ / N_a を、 β すなわち $Q^{(1)}$ を出発点として求めることができる。



(a) 最初の A-unit (b) 平衡時の A-unit

図 6 4 層パーセプトロンの例

[例] 三つの Stimuli S_1, S_2, S_3 があって $N_a=10$ とし、 Q_{ij} ⁽¹⁾ (Q 行列) を次のごとくする (図 6)。

$$\begin{pmatrix} 0.3 & 0.1 & 0.1 \\ 0.1 & 0.4 & 0.3 \\ 0.1 & 0.3 & 0.6 \end{pmatrix}$$

Training Sequence を $S_1 S_2 S_2 S_1 S_2 S_3 S_1 S_3 S_1 S_2$ の繰返しとする。そうすれば

f_{ij} の行列は

$$\begin{pmatrix} 0 & 0.3 & 0.1 \\ 0.2 & 0.1 & 0.1 \\ 0.2 & 0 & 0 \end{pmatrix}$$

ゆえに

$$C_{ij} = \begin{pmatrix} 0.04 & 0.10 & 0.04 \\ 0.14 & 0.07 & 0.05 \\ 0.18 & 0.06 & 0.04 \end{pmatrix}$$

(3.5) 式は

$$\begin{pmatrix} r^1 \\ r^2 \\ r^3 \end{pmatrix} = \frac{\eta}{\delta} \begin{pmatrix} 0.4 & 1.0 & 0.4 \\ 1.4 & 0.7 & 0.5 \\ 1.8 & 0.6 & 0.4 \end{pmatrix} \begin{pmatrix} \phi(\beta^1 + r^1) \\ \phi(\beta^2 + r^2) \\ \phi(\beta^3 + r^3) \end{pmatrix}$$

最初 $\beta = \theta(1, 0, 0)$ から出発すると, $\eta/\delta = \theta = 1$ とおき (3.6) 式を繰り返してけっきょく

$$\begin{pmatrix} r^1 \\ r^2 \\ r^3 \end{pmatrix} = \begin{pmatrix} 1.8 \\ 2.6 \\ 2.8 \end{pmatrix}$$

をうる. すなわち $\beta = (1, 0, 0)$ に対応した A-unit は S_1, S_2, S_3 のいずれによっても励起されることになる. 同様にしてけっきょく次の Q を得る (図 6 (b)).

$$Q_{ij}^{(2)} = \begin{pmatrix} 0.6 & 0.6 & 0.6 \\ 0.6 & 0.6 & 0.6 \\ 0.6 & 0.6 & 0.9 \end{pmatrix}$$

種々の問題に対して (3.5) 式を考慮すればその場合の状況が求められるが次の一例は興味がある.

[例] 刺激を $\{S_1, S_2, \dots, S_k\}$ および $\{S_{k+1}, \dots, S_n\}$ なる二つのクラスに別け, ランダムな Training Sequence において, 一つのクラス内における刺激間の遷移確率を, 互いに異なる刺激間のそれよりはるかに大にしておく. 今一つのクラスを一つの図形の種々の変換図形, たとえば平行移動図形群, 第 2 のクラスを他の図形の同じ変換図形群と考えると, ある条件下においてそれぞれの図形のあらゆるその変換図形はそれぞれのクラスに正しく分類される.

[例] 刺激 S_1, S_2, \dots, S_n があり, そのある変換を $T(S_1), T(S_2), \dots, T(S_n)$ とする. 訓練を $\{S_{k1}, T(S_{k1}), S_{k2}, T(S_{k2}), \dots\}$ なる sequence にて行なう. k_1, k_2, \dots は n 個の刺激からランダムに抜きとったものである. 次にさらに $S_1 \sim S_n$ 中になかった S_x, S_y をとってきて互いに二つの異なるクラスになるように学習させる. そこで $T(S_x), T(S_y)$ なる刺激を与えるといかなる応答を示すか? 結果だけを示すとある条件下で S_x および S_y とそれぞれ同一の応答を与えることができる. すなわち一つの変換に対する generalization が行なわれたことになる.

その他この型の Perceptron においてはその学習の方法や Training の Sequence を適当にくふうすることによって種々広範な図形認識の問題が解かれる.

3.3 Cross Coupled Perceptron

これは図 7 に示すような系で 3.2 でのべた 4-Layer ときわめて類似している. R と可変の結合をもつもの

は A''-unit のみで各 A'-unit から 4-Layer の $A^{(1)}$ – $A^{(2)}$ と類似の可変の結合がある. 4-Layer の場合の (3.4), (3.5) 式がこの場合も成立する. β -Component はこれが $A^{(1)}$ からでなく S-unit から直接与えられることだけが異なる. ある例について 4-Layer の場合と比較した結果, cross coupled のときのほうが generalization の能力を示しうる $N_a\eta/\delta$ の値の範囲が広いということが知られている.

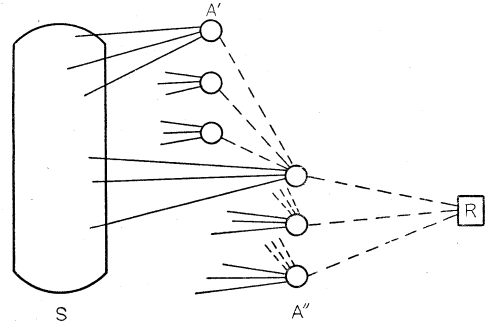


図 7 A-unit 間に相互結合のあるパーセプトロン

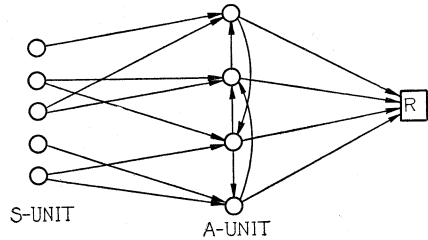


図 8 A-unit 間に相互結合のあるパーセプトロン (閉ループ)

図 8 に示すように各 A-unit が他の A-unit の入力を受けかつ他の A-unit へ出力を与えるような場合 “closed-loop” と称する. open-loop の場合と異なる点は時刻 t における状態が現在の S-unit からの input のみならず, それ以前の A-unit の状態にも関する点で, 系の安定問題が重要となる. したがってここでは種々の刺激列 S_i に対して Q 関数を定義する必要がある. A-A 結合が固定値をもつとき, $Q_{i\mu}$ や $Q_{i\mu, i\mu}$ を求めるには, 最初の Q_{i0} を出発点として順次計算せねばならない. N_a を非常に大とするときは, これに比較的, 簡単な計算となる. ただし, ここでたとえば i_μ は $t-\mu_i$ に始まる刺激列 S_i を指す (τ は各刺激の時間々隔).

与えられた問題に対して解の安定性をうるにはパラメータの選択に条件が課せられる.

A-A 間に可変な結合をもつ場合にはちょうど可変な $A^{(1)}$ – $A^{(2)}$ 結合をもつ 4-Layer の場合と類似な解析

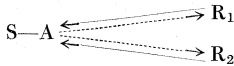
が、各刺激列に対してなされる。

Closed-Loop Cross Coupled Perceptron は Sequence すなわち時間的なターンの識別能力を有するから音声パターンの識別などにも応用が考えられる。

4. Back Coupled Perceptron

R-unit から A-unit の入力への back coupling がある系を考える。特にここでは $R \rightarrow A$ 結合は固定値とする。

たとえば



のようなものである。

ここで back coupling は各 R-unit から各 A-unit に対して存在し、一つの R-unit が励起されているとき、それから全 A-unit 中ランダムに選んだ半分には $+1$ 、あとの半分には -1 なる値がフィードバックされると仮定する。Back Coupled System のもつ特殊の性質を次の例によって見る。

【例 1】 S_1, S_2 なる刺激があって

S_1 は応答 $(r_1, r_2) = (1, 0)$

S_2 は応答 $(r_1, r_2) = (0, 1)$

を与えるようにあらかじめ独立に学習させてあったとする。そこへ今 S_1, S_2 を同時に与えたときの状況を考えると、図 9 を参照してわかるように、そのとき (r_1, r_2) は $(0, 0)$ または $(1, 1)$ なる状態をとらずに必ず $(0, 1)$ または $(1, 0)$ のどちらかを安定状態としてとる。

図 9 中、各円はそれぞれの刺激と R-unit の組み合わせ状態に対して励起される A-unit を指す。

この例で言えることは、この系には二つの図形が同時に与えられても必ずそのいずれかの一つの図形に対する応答を示し、その混合を示す中間的な応答を与えないといういわゆる“Selective Attention”の性質を

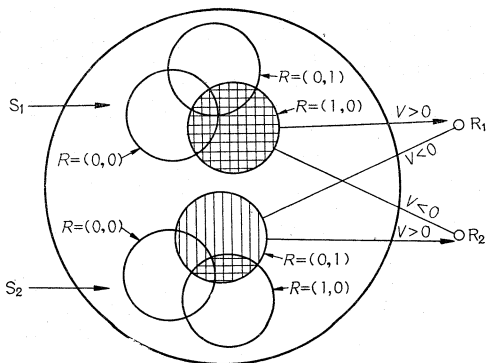


図 9 “Selective Attention” の例

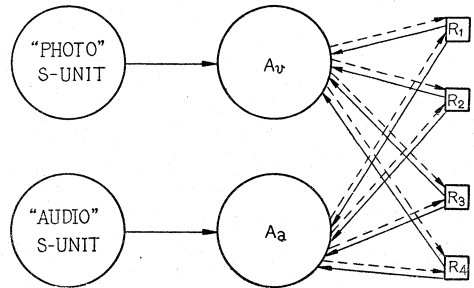


図 10 二種の入力をもつ系の例

示す。

図 10 に示すように Input が Dual Mode の系で次の例を与える。

【例 2】 S-unit に図形が三角であるか、四角であるか、それが上部にあるか、下部にあるかということに対して、四つの R-unit を考えておき、図 11 中において、photo. のほうから、上部に三角、下部に四角などの独立パターンを与えて、学習させておく。

それから audio のほうで三角、四角、上部、下部のパターンを考えて各 R-unit をそのおのおのに対応するように学習させておく。それから photo. のほうに上部に三角、下部に四角のパターンを同時に示し、audio のほうに、三角、四角、上部、下部のいずれかを示す。

例 1 のごとく、photo. の入力だけでは、上部に三角か、下部に四角のどちらかに対する応答のみが安定であることが見られたが、さらにそのどちらかに audio のほうからバイアスをつけることになる。すなわち audio 入力が photo. の系を所定の応答に向けることができる。

Back Coupled System ではさらに不変の R-A 結合をもつものも考えられよう。

5. そ の 他

5.1 Program Learning Perceptron

今まで扱ってきたような比較的簡単な図形や刺激列の認識よりもっと複雑な物体の詳しい認識やあるいは、抽象的な関係の認識問題を今までのような Perceptron で一挙に実現してしまうのは経済的にとてもむずかしいことであると考えられる。

この場合には一時にはパターンの一部に注目してプログラムによって一連の観察を行なった結果に基づき判断をなすような方法が一つの解決になろう。このような方向において、今まで述べてきた種々の Perceptron のそれぞれの特徴を巧みに利用して総合された系を種々考えることもできよう。図 11 はその一例で

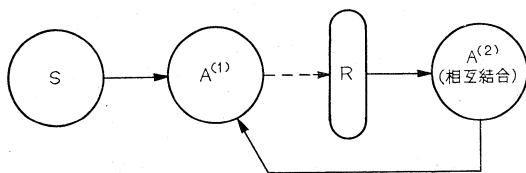


図 11 組み合わせパーセプトロンの例

ある。

これは種々の刺激に、それぞれの特定の応答の例を対応させるような能力のあることが予想されるであろう。

5.2 Sensory Analyzing Mechanism

複雑な対象の認識を助けるための他の一つの有力な方法は図形の“Sensing”の段階において、特殊の性質や構造を検知する機構を用いることである。

検知すべき性質、構造は対象によって実に多様であるが、二次元的な図形ではたとえば、局所的な性質を検知する“Local Property Detector”が有用である。その一例を図 12 に示す。この図では $A^{(1)}$ が Local Property Detector となり $A^{(2)}$ が性質の組み合わせを検知する部分となる。図 12 で上段は線の終点を検知するため、中、下段は水平の線を検知するものであるが、さらに、中、下段を組み合わせ、それぞれの入力のくる時間を図のように加減すれば、与えられた問題、たとえばある大きさの \sqcup と \sqcap を区別するような問題とか、 \top と \perp を区別する問題とかそれぞれに対して適している property detector は存在するはずで

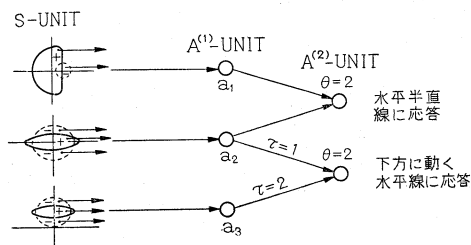


図 12 局所的な性質を見出す系

ある。確率の大小によって判定できる。

6. 結 言

以上 CAL における Perceptron の研究の一部を概観した。研究はさらに複雑な対象に関する認識の問題を扱うために前章で少しふれたような考え方にしたがって種々のモデルや変形が考えられている。

それらは“qualitative”な検討に加えてさらに、デジタル計算機によるシミュレーションや理論的な計算によって量的に評価がなされるべきであろう。

その点からは、4 章までに述べた perceptron についても種々の検討が残されている。

Perceptron を実際の装置として製作するためには、技術的に廉価な A-unit を考えることが問題であろう。この面では electrochemical な技術や半導体の技術が研究されているということである。

Perceptron の原理は空間的および時間的なパターンに対して広く応用され、実際、文字読取のみならず、音声の識別にも応用が考えられている。

最後に Perceptron の紹介を許可していただいた ONR の Dr. Yovits, CAL の Dr. Rosenblatt, 御討議、御援助下さった三谷尚正氏に謝意を表する。

参 考 文 献

- (1) F. Rosenblatt: “The Perceptron: A Probabilistic Model for Information Storage & Organization in the Brain”, Psych. Review, vol. 65, No. 6, 1958.
- (2) R. D. Joseph: “On Predicting Perceptron Performance”, 1959 IRE, Int'l Conv. Rec., Part 4.
- (3) A. E. Murray: “A Review of the Perceptron Program”, Presented at 15th Annual National Electronics Conf., 1959, Oct.
- (4) F. Rosenblatt: “The Perceptron, A Theory of Statistical Separability in Cognitive System”, CAL Rep., No. VG-1196-G-1, 1958.
- (5) R. Joseph: “Contribution to Perceptron Theory”, CAL Rep., No. VG-1196-G-7, 1960.
- (6) F. Rosenblatt: “Perceptron & the Theory of Brain Mechanisms”, CAL Rep., No. 1196-G-8, 1961.