

言語の認知科学 第4回配布資料

文字知覚，文字認知のモデル

担当：浅川伸一

2009年10月21日

来週の授業は実習をしますので 9302 教室に集まってください。

図1は様々なイメージング技法の関係を図式化したものである。横軸は対数目盛りになっており，分子レベルから脳丸ごと一つまで，空間的大きさを表している。縦軸はそれぞれの技法の時間解像度を表している。NIRS が描かれていないが板書する。

1 前回の補足，ラテラリティについて

2 パンデモニウム (伏魔殿) モデル

パンデモニウムとはセルフリッジ (Selfridge, O. G.) が考案したパターン認識のモデルである。パンデモニウムとは万魔堂のことであり，特定の処理を担うデーモン (情報処理モジュールを比喩的に示す) が階層的に多数連結されて構成されている。文字認識を行うパンデモニウムでは，イメージ・デーモンから送られた入力情報を特徴デーモンが分析し，垂直線，水平線，角などの有無がチェックされる。それらの出力を各文字に対応する認知デーモンが受け，マッチする認知デーモンの出力は大きくなる。それらの出力を評価して，最終的に決定デーモンが判定を下す。並列的であるが，基本的な情報処理の流れは，特徴抽出から解釈にいたるボトムアップ型である。

有斐閣，心理学事典より

3 パーセプトロン

パンデモニウムモデルを素直に実現したニューラルネットワークモデルとしてパーセプトロンが挙げられる。パーセプトロンは工学の分野でパターン認識との関連で長い研究の歴史がある。パーセプトロン perceptron とはローゼンブラット (Rosenblatt, 1958) によって提案された図8のような3層の階

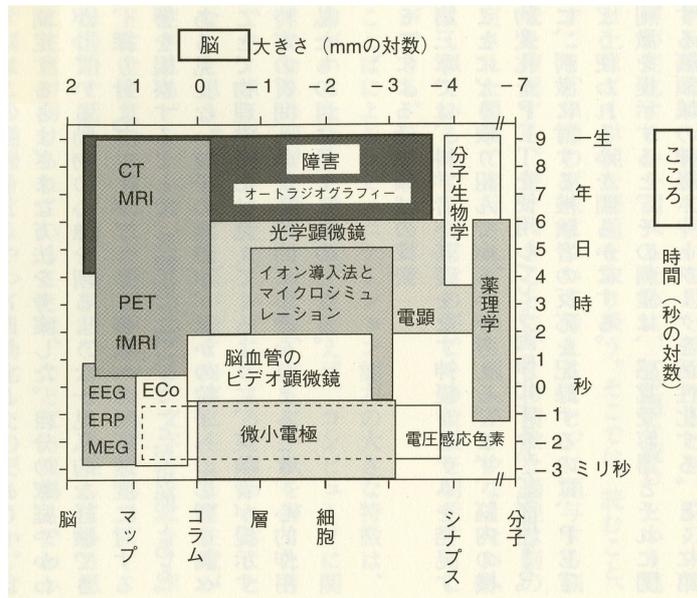


図 1: 脳のイメージング技法の特徴。脳を観る，ポズナー&レイクル著，養老，加藤，笠井訳，1997，日経サイエンス社，p.41

層型ネットワークモデルである。パーセプトロンはマッカロック・ピッツの形式ニューロンを用いて学習則にヘップ則を使ったモデルで，単純な認識能力を獲得することができる。

パーセプトロンは 3 層の階層型ネットワークでそれぞれ， S (sensory layer), A (associative layer), R (response layer) と呼ばれる層からなっている。 $S \rightarrow A \rightarrow R$ のうち パーセプトロンの本質的な部分は $A \rightarrow R$ の間の学習にある。最下層の入力層は外界からの刺激入力を表現している。中間層では入力情報の変換が行われ，最上位層である出力層で認識に到る。

パーセプトロンとは，入力パターンに現われる P^+ と P^- という 2 つのパターンを識別する問題を考えたとき， P^+ が入力されたとき 1 を， P^- のとき 0 を出力する機械である。あるいは，出力ユニットが 1 つしかない場合を考えれば，パーセプトロンは入力層に現われるパターンを 1 と 0 とに 2 分する機械であるということができる。

4 実習：パーセプトロンによる線分の認識

パーセプトロンに次のような 5×5 の受容野を考えて，線分を学習させることを考える。

縦線 5 本，横線 5 本，斜線 2 本の計 12 パターンをパーセプトロンに学習

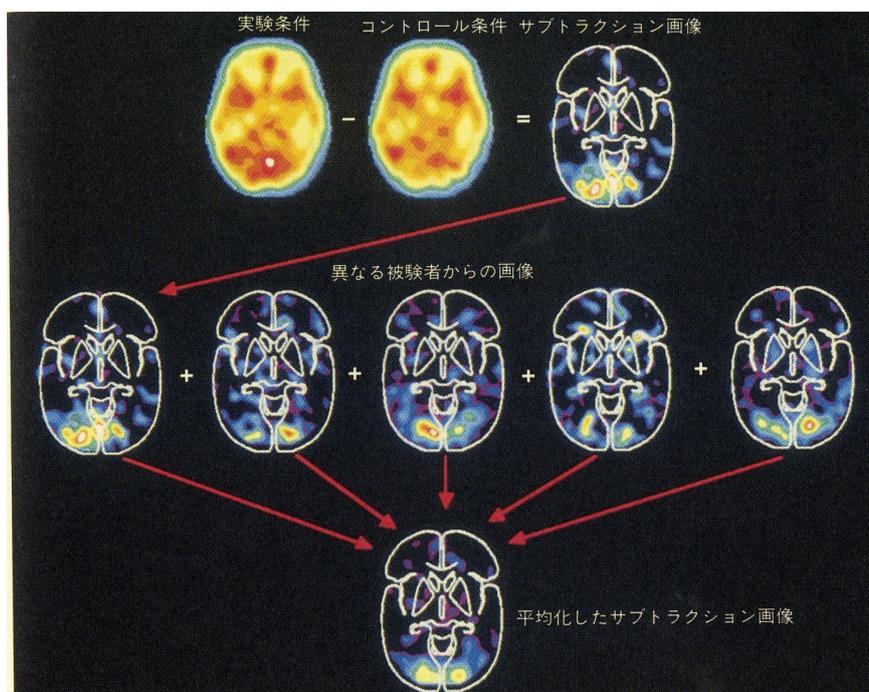


図 2: サブトラクション法の概念図。実験条件は、注視点から 5.5 度のところにある点滅するチェッカーボードを見つめることで、コントロール条件は静止した注視点を見ている状態。上段の PET 画像のように実験条件からコントロール条件を差し引くとサブトラクション画像が得られる。5 人の被験者からそれぞれの画像が得られる。これを平均化してノイズを除去すると下段の画像になる。脳を観る，ポズナー&レイクル著，養老，加藤，笠井訳，1997，日経サイエンス社，p.98

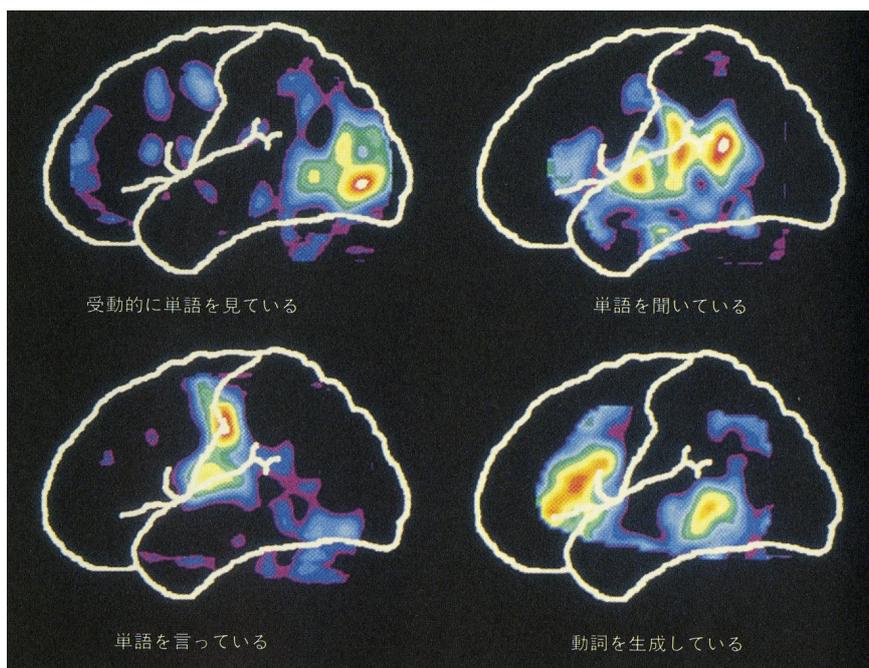


図 3: サブトラクション法によって明らかとなった認知課題との対応。脳を観る, ポズナー&レイクル著, 養老, 加藤, 笠井訳, 1997, 日経サイエンス社, p.163

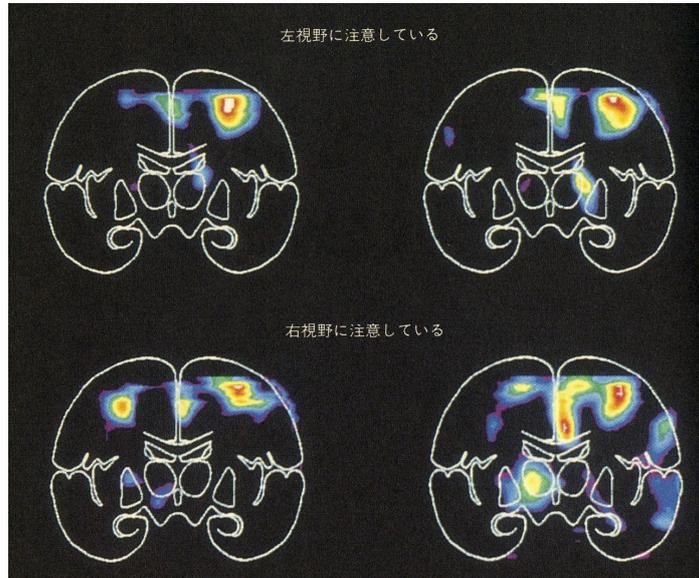


図 4: 右視野に注意を向けているときと、左視野に注意を向けているときの PET 画像

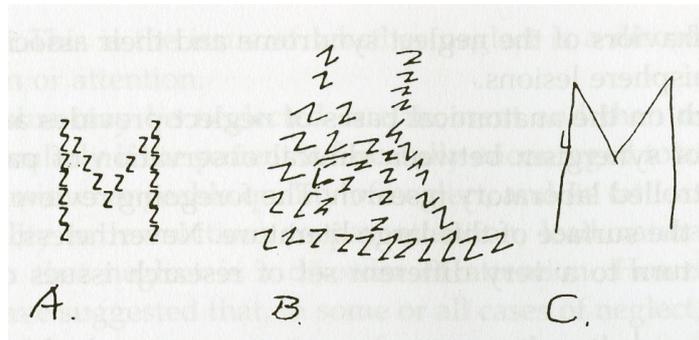


図 5: 局所情報と大域情報の分離の例。(a)z で構成されている大きな M の模写、(b) 局所情報は利用できるが、大域情報を模写できない右半球損傷患者の模写例。(c) 反対のパターンを示す左半球損傷患者の模写例。From D.C. Delis et al., "Hemispheric specialization of memory for visual hierarchical stimuli," *Neuropsychologia*, 24, 1986

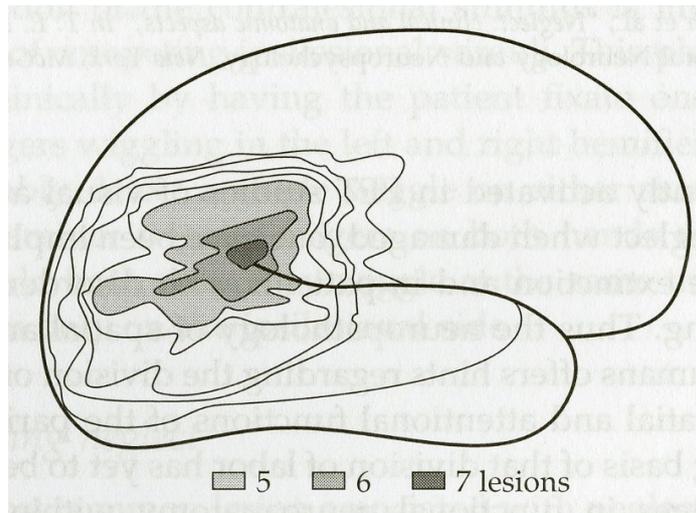


図 6: 重篤な半側空間無視患者 8 人の下後頭側頭領域の責任病巣。From G. Valzar and D. Perani, “The anatomy of unilateral neglect after right hemisphere stroke lesion: a clinical CT scan correlation study in man,” *Neuropsychologia*, 24, 1986

させることを考えよう。入力データは 5×5 の正方格子の線分とし、出力および教師信号の次元は 12 次元のベクトルとした。各々の線分に対して一つの出力層ユニットが応答するように訓練した。

結果を可視化するために出力層から入力層への結合係数の大きさを円で表したグラフを用いる。正の結合係数を \oplus で、負の結合係数を \ominus で表し、結合係数の大きさを円の直径で表したグラフをヒントングラフという。

結果はグラフから明らかである。左から 2 列目の縦線に応答する出力層ユニットから入力層ユニットへの結合係数は、左から 2 列目のみ正の結合で、その他の領域では負になっている。このことからこのユニットは左から 2 列目の縦線を検出する特性を獲得したといえることができる。

ではパーセプトロンに 5 本の縦線全てに応答するように、すなわち縦線の概念を獲得するようにさせることは可能だろうか。

5 バックプロパゲーション (誤差逆伝播法)

5.1 XOR 問題、線形分離不可能な問題

パーセプトロンでは絶対に解けない問題に排他的論理和 (XOR) 問題がある。排他的論理和とは、2 つの入力のうちいずれか一方のみが 1 のとき 1 を出力する問題である。図 11 左を見るとわかるとおり、XOR 問題は一本の判

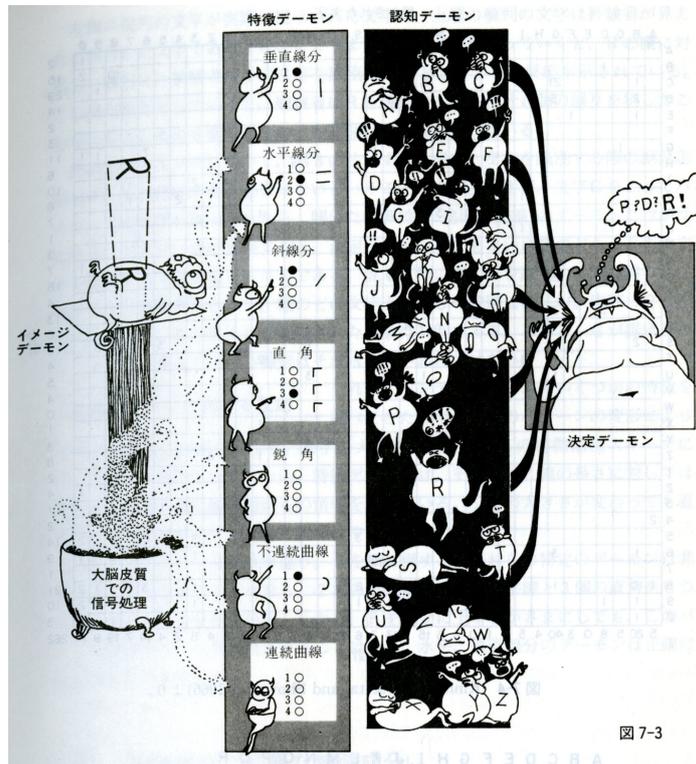


図 7-3

図 7: パンデモニウムモデル。リンゼイ&ノーマン著，中溝，箱田，近藤訳，
情報処理心理学入門 II, サイエンス社，1984，p.8

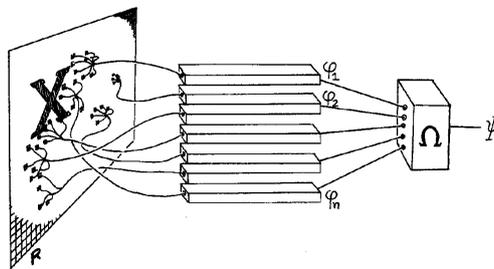


図 8: パーセプトロン。ミンスキー&パパート著，中野，阪口訳，1993，パーソ
ナルメディア

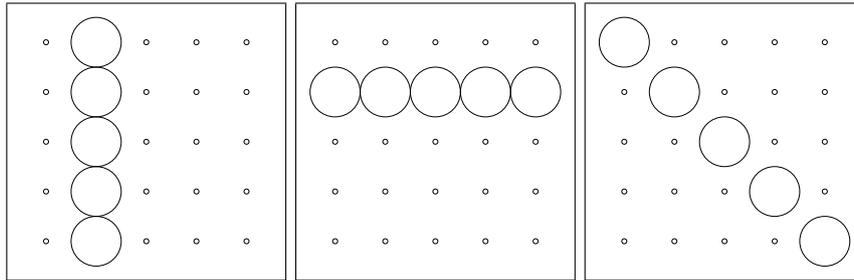


図 9: 学習パターン

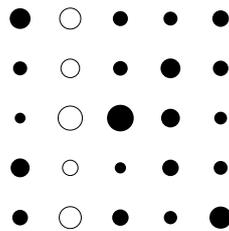


図 10: 学習結果のヒントグラフ

別直線で白マルと黒マルを分離できない、すなわち線形分離不可能な問題である。

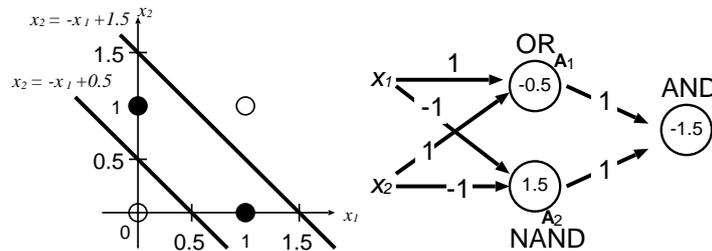


図 11: XOR 問題の幾何学的表現と XOR 問題を解くためのネットワーク

図 11 右は図 11 左の幾何学表現を対応するネットワークとして表現したものである。一番左が入力層、一番右が出力層、中間の 2 つが中間層である。ユニットの中に書かれた数値は各ユニットのしきい値を示している。中間層の 2 つのユニットのうち上の OR と書かれたユニットは、 $x_1 + x_2 - 0.5 > 0$ のとき発火する。この式を書き換えると、 $x_2 > -x_1 + 0.5$ となるので図 11 左の下斜線より上の領域に対応する。一方、中間層の下の NAND(not and) と書かれたユニットは、 $-x_1 - x_2 + 1.5 > 0$ のとき発火するので、移項して

$x_2 < -x_1 - 1.5$ とすれば、図 11 左の上斜線より下の領域に対応していることが分かる。さらに、AND と書かれた出力ユニットは、2 つの中間層ユニットからの情報の論理積 (AND) を計算するユニットになっている。そこで、2 つの中間層ユニットの両方が発火する場合だけ、出力ユニットも発火し、1 を出力する。これは、図 11 左では、「下の斜線より上の領域」でかつ「上の斜線より下の領域」に対応する。すなわち、図中の黒丸の領域だけが分離されることになる。このような 2 本の直線は図中にいくらかでも引けることから、XOR 問題の解も無数に存在することが分かる。

入力層		中間層		出力
x_1	x_2	a_1	a_2	r
0	0	0	1	0
0	1	1	1	1
1	0	1	1	1
1	1	1	0	0

表 1: 図 11 に対応する XOR 問題の真偽表

図 11 左にあるとおり、中間層のユニット 1 個は 1 つの線形判別関数に相当すると考えられる。中間層から出力層への結合では各出力の論理積 AND を計算していることに相当する。 n 個の中間層を用意すれば原理的には $\frac{n^2+n+2}{2}$ 個のカテゴリー分類が可能である。パーセプトロンが XOR 問題を解くことができない理由は入力層から中間層にいたる結合係数を変更する手段がないことなのである。

5.2 誤差逆伝播法 (一般化デルタルール)

XOR 問題でも見たように、パーセプトロンの問題点は学習が出力層と中間層の間だけで行われ、入力層と中間層の結合係数を更新することができない。換言すれば中間層ユニットの誤差を明示的に知る方法はない。このことは信用割り当て credit assignment 問題と呼ばれる。この信用割り当て問題を解決したのがバックプロパゲーション (誤差逆伝播法あるいは一般化デルタルール) である。

5.3 実習：排他的論理和の学習パターン

何度か XOR 問題を解かせてみると大まかに 3 パターンの学習曲線を描くことができる。

上図は横軸に学習回数、縦軸に平均二乗誤差をプロットしたものである。異なる初期値で program を動かすと、速やかに学習が終了するもの、error

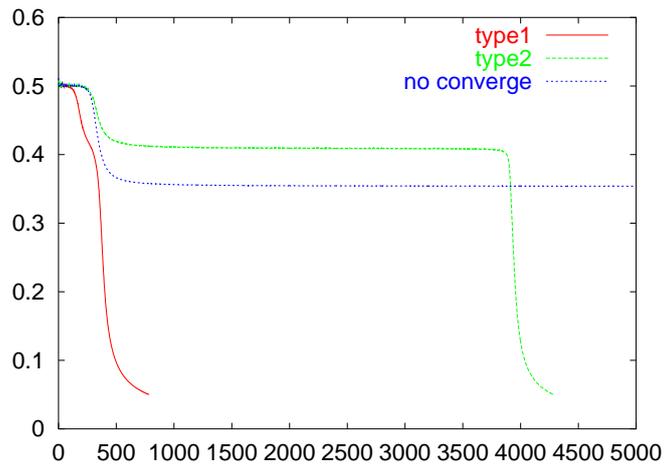


図 12: 典型的な学習パターン

が 0.4 付近まで減少したあと学習が進まないもの、error が 0.4 付近で一旦止まって、その後急激に学習が進むもの、などが観察できる。

XOR 問題が解けるか解けないか、どのパターンを示すかは初期値によって決まる。そこで、どのような初期値の時にどのパターンになるかを調べてみよ。

5.4 対称性の認識

BP 法の成功例として次に示す対称性を認識するネットワークを挙げる。図

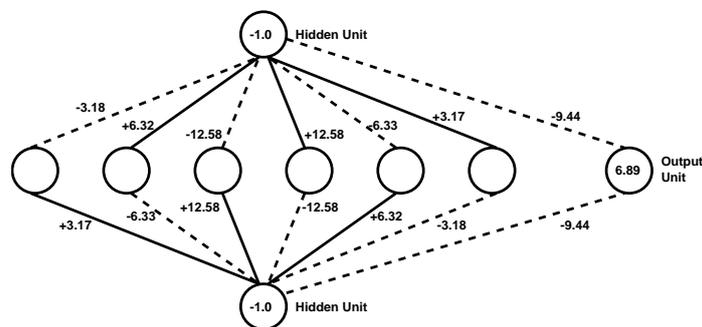


図 13: 対称性を認識するネットワーク (PDP book p.340)

中の点線は負の結合を、実線は正の結合係数を表す。また、円の中に書かれた数字はしきい値である。図を見ると分かるとおり中間層への結合係数が綺

麗に対称になっており二つの中間層ユニットで正負が逆転していることが分かる。

5.5 実習：対称性の認識

対象性を学習するバックプロパゲーションのプログラムで、学習セットをいくつか取り除いて学習させよ。学習成立後、取り除いたパターンを用いて一般化能力を試せ。

全 64 個の学習パターン中、対称なパターンを 1 個、非対称なパターンを 8 個取り除き学習させよ。学習後に取り除いた 9 個のパターンを入力すると正解が得られるかどうかを調べよ。

同様に、対称なパターンを 2 個、非対称なパターンを 16 個取り除き学習させよ。学習後に取り除いた 18 個のパターンを入力すると正解が得られるかどうかを調べよ。

対称性を学習するネットワークでは、最低何個の学習パターンが必要かを検討せよ。

5.6 砂時計モデル

バックプロパゲーションの学習アルゴリズムにの応用例として、恒等写像の学習による情報圧縮がある。恒等写像の学習とは、入力データと教師信号を同じくして学習させることである。中間層のユニット数が入力層ユニットや出力層ユニットに比べて少ないとき、中間層ユニットにおいてデータ圧縮表現が得られる。

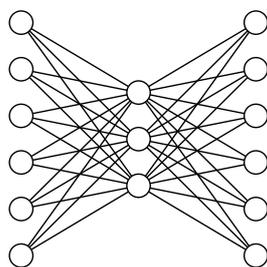


図 14: 砂時計モデル

5.7 実習：データ圧縮

128 × 128 ピクセルの一枚の画像を 8 × 8 の小領域に分割し、恒等写像を学習させた。恒等写像なので入力ベクトルも出力ベクトルも同じ 64 次元で

ある。ここでは中間層のユニット数を 16 とした。すなわち $16/64 = 1/4$ のデータ圧縮を行なったことになる。

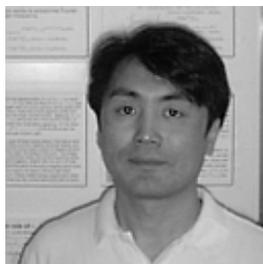


図 15: 恒等写像を行なった原画像

一ピクセルあたり 256 階調のグレースケールのデータを 0 から 1 までの値に変換し、 $8 \times 8 = 64$ 次元のデータを 256 個使って訓練した。

学習係数は 0.01 とし、 $MSE = 0.03$ に達するまでに要した学習回数は 1460 回であった。



図 16: 恒等写像により復元された画像

6 ログジェンモデル

7 Interactive Activation Model

8 実習 : PMSP96 英単語の読み

Plaut ら (Plaut, McClelland, Seidenberg & Patterson, 1996) は単音節の英単語 2998 語を音韻に変換するモデルを提出している。彼らの用いたのは 3 層のバックプロパゲーションである。

Plaut らのモデルをトライアングルモデルという。トライアングルモデルにおける読みの説明は以下のとおりである。書記素層から音韻層への直接経

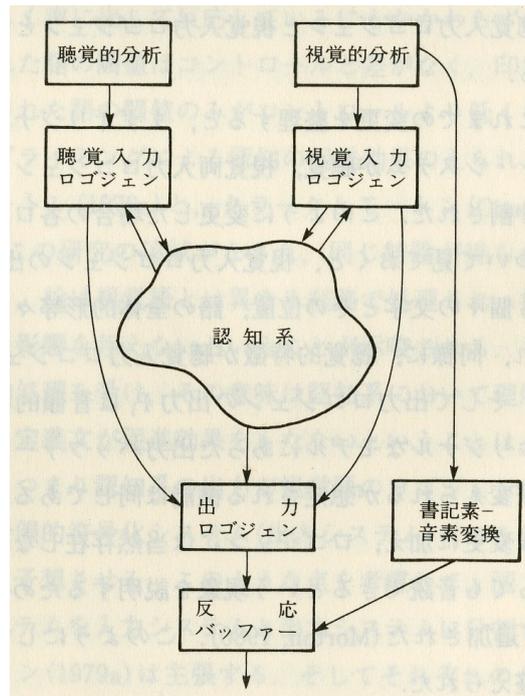


図 17: ログジェンモデルの概要

路では、多くの単語と発音規則が一致する規則語と高頻度の不規則語が学習される。一方、低頻度の不規則語は意味系に依存すると仮定される。すなわちトライアングルモデルにおける直接経路では規則語と高頻度例外語が学習され、従って直接経路は単語の頻度効果に、すなわち単語の統計情報（生起確率）に敏感である。規則語および高頻度例外語と低頻度例外語との処理の違いには労働の分割と呼ばれる作用が関与する。

単語の入出力表現にはオンセット、母音、コーダという表現を持っている。単音節の単語なので、母音については一つだけのコーディングが必要である。加えて母音の前後に子音のクラスターが必要である。母音の前の子音をオンセット、母音の後の子音をコーダという。

まず、入力表現である orthography は、オンセットが Y, S, P, T, K, Q, C, B, D, G, F, V, J, Z, L, M, N, R, W, H, CH, GH, GN, PH, PS, RH, TH, TS, WH の 30 とおりであり、母音は E, I, O, U, A_i, Y, AI, AU, AW, AY, EA, EE, EI, EU, EW, EY, IE, OA, OE, OI, OO, OU, OW, OY, OW, OY, UE, UI, UY の 27 とおり、コーダは H, R, L, M, N, B, D, G, C, X, F, V, J, S, Z, P, T, K, Q, BB, CH, CK, DD, DG, FF, GG, GH, GN, KS, LL, NG, NN, PH, PP, PS, RR, SH, SL, SS, TCH, TH, TS, TT, ZZ, U, E, ES, ED の 48 とおりであった。すなわち入力表現は計 105 次元のベクトルとして表

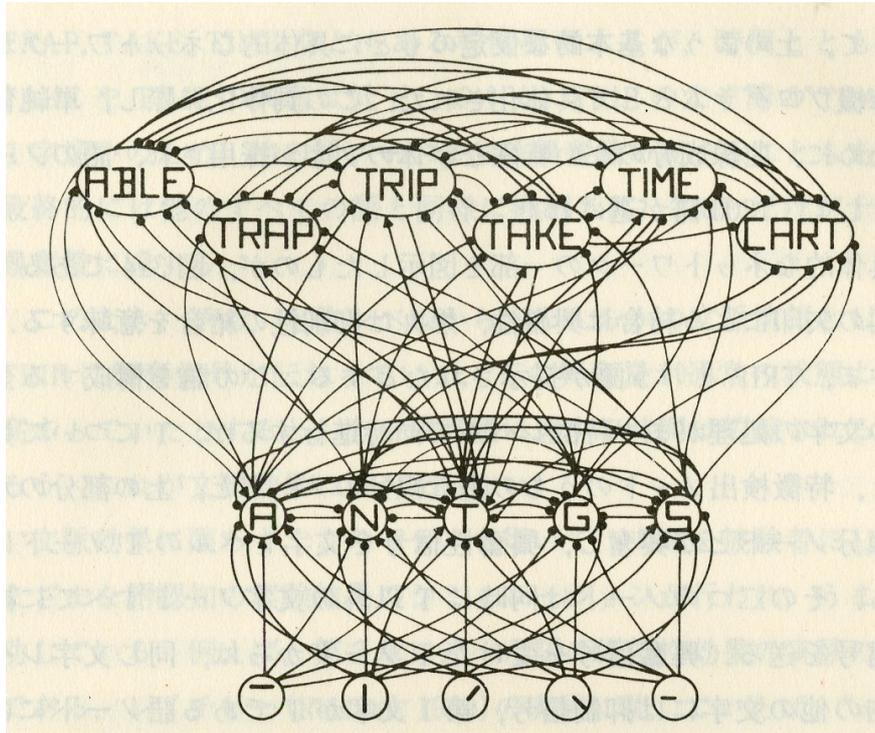


図 18: IA モデルの結線図

現された。

出力表現である phonology の表現としては、オンセットが (s, S, C), (z, Z, j, f, v, T, D, p, b, t, d, k, g, m, n, h), (l, r, w, y) の 23 次元、母音が a, e, i, o, u, @, ^, A, E, I, O, U, W, Y の 14 次元、コードが (r), (l), (m, n, N), (b, g, d), (ps, ks, ts), (s, z), (f, v, p, k), (t), (S, Z, T, D, C, j) の 24 次元の計 61 次元のベクトルとして表現された。出力表現は彼ら読字の表記方法であり、いわゆる発音記号とは関係がない。表記は次のようなものである。/a/ は POT の、/@/ は CAT の、/e/ は BED の、/i/ は HIT の、/o/ は DOG の、/u/ は GOOD の、/A/ は MAKE の、/E/ は KEEP の、/I/ は BIKE の、/O/ は HOPE の、/U/ は BOOT の、/W/ は NOW の、/Y/ は BOY の、/^/ は CUP の、/N/ は RING の、/S/ は SHE の、/C/ は CHIN の、/Z/ は BEIGE の、/T/ は THIN の、/D/ は THIS の音を各々表現している。

母音の前後にある子音には順序関係についての制約がある。例えばオンセットクラスターにおける /s/, /t/, /r/ は順序が /str/ でなければならない。出力表現である phonology のオンセットとコードにあるカッコ内の音が相互に排他的な表現であることを意味している。この制約によって子音の順序が一意的に定まるようになる。子音は必ず上に表記した順序で音声化さ

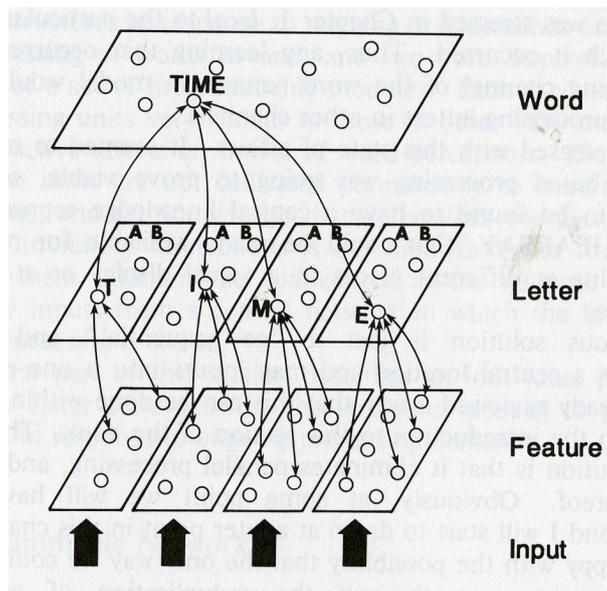


図 19: Interactive Activation Model

れるという制約がある。

これに加えて単語 CLASP と LASPE とでは /p/ と /s/ との順序関係を表現できないため /ps/ というユニットが加えられている。同じ理由により /ks/, /ts/ というユニットが加えられた。

英語はアルファベットを表記記号とする言語であるが、単語の書記形態の一部が音韻形態に対応しているに過ぎない。そこで orthography のユニットとしては単一文字からなるユニットの他に 2 つの文字の組み合わせからなるユニットも用いられた。

これらの表現の詳細については Plaut ら (Plaut, McClelland, Seidenberg & Patterson, 1996) の原典を参照して欲しい。

3 層のバックプロパゲーションネットワークを用いて Plaut らの学習させた 2998 単音節単語を学習させた。中間層のユニット数は彼らのシミュレーションと同じ 100 個にした。MSE=0.0305 程度にまで学習が進行し、このときの正解率はおよそ 94.46 % であった。同じ学習セットをパーセプトロンで学習させる (MSE=0.05) と 87.26 % ほどの正解率になる。

ちなみに中間層のユニット数を 30 にしても学習が成立する。中間層のユニット数 30 のときの MSE=0.05 の場合正解率は 89.26 % であった。

学習の成立したネットワークを用いて、Plaut らの論文にあるような非単語を入力した結果が表 2 である。

人間の被検者が読んだ場合 93.8 % の正解率の一貫語が Plaut らのモデルでは 97.7 % と読めているのに対して、bp3 では 90.7 % であり、パーセプト

表 2: Glushko(1979) の非単語を読ませた結果

	一貫語	非一貫語
人間	93.8	78.3
PMSP96	97.7	72.1
bp3(中間層 100,MSE=0.03)	90.7	53.5
bp3(中間層 100,MSE=0.05)	95.3	58.1
bp3(中間層 30,MSE=0.05)	88.4	58.1
perceptron(MSE=0.05)	93.0	67.4

ロンでは 93.0 % であった。非一貫語についても bp3 が 53.5 % であるのに対してパーセプトロンでは 67.4 % であった。このことは訓練した 2998 単語の正解率ではパーセプトロンの成績は最も悪かったにもかかわらず、非単語の読みに対してはパーセプトロンは人間の読みの成績に近いものになっていることが分かる。このことは逆説的ではあるが、より能力の高いバックプロパゲーションを使うよりも、よりシンプルなパーセプトロンを使った方が、一見すると矛盾するような結果になっている。これは 2998 単語の読みを学習させるためにバックプロパゲーションによる学習では平均二乗誤差が 0.03 になるまで学習させた結果かも知れない。すなわち訓練のしすぎによる過学習が起きてしまったため、非単語の読みにおける一般的な能力が低下してしまっただけかも知れない。平均二乗誤差を 0.05 で打ち切ると訓練データである 2998 単語の正解率は 88.4 % と落ちるものの glushko の非単語リストの成績は平均二乗誤差を 0.03 とした場合より向上した (表 2)。

このように学習をどこで打ち切るかと言う問題は注意を要するものである。また、Plaut らの結果とやや異なる結果が得られたことは興味深く、示唆に富んでいる。

引用文献

- Behrmann, M. (1994). Neglect dyslexia: Attention and word recognition. In M. J. Farah & G. Ratcliff (Eds.), *The Neuropsychology of High-Level Vision* (pp. 173-214). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Lindsey, P. & Norman, D. (1983). 情報処理心理学入門 2 注意と記憶, 中澤幸夫 & 箱田裕司 & 近藤倫明・訳. サイエンス社.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S. & Patterson, K. (1996).

Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56–115.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65, 386–408. In J. A. Anderson and E. Rosenfeld (Eds.) *Neurocomputing* (1988), MIT Press.