

連想記憶の海馬モデル

正 員 森田 昌彦[†]

Hippocampus Model of Associative Memory

Masahiko MORITA[†], Member

あらまし 海馬に関する生理学的な知見と工学的な考察に基づいて構成した、連想記憶機能をもつ神経回路モデルを提案する。このモデルは、神経回路網の安定状態を“コード”として用い、入力パターンをランダムに選んだコードと連合することによって記憶するが、その過程において回路網の刺激レベルの変化とゆらぎとが重要な役割を果たす。また、互いに直交しないパターンでも記憶することができ、一対多の連想も可能である。更に時系列パターンの記憶に自然に拡張できるなど、従来の連想記憶モデルにない特徴をもっている。それと同時に、本モデルの機能や動作は、海馬の機能と記憶メカニズムに対して仮説と示唆を与えるものと考えられる。

1. ま え が き

「記憶」は、脳のもつ最も重要な機能の一つであるが、そのメカニズムの大部分は謎のままである。これを明らかにするためには、生理学や心理学だけでなく、工学的なアプローチ、特に神経回路モデルによる構成的研究が重要な役割を果たすものと考えられる。

こうした考えに従って、これまでも連想記憶機能をもつ神経回路モデルが出されているが、そのほとんどは相関行列型の連想記憶⁽¹⁾を原理として用いている。これが神経回路網における連想記憶の最も基本となる原理であることは間違いないが、記憶パターンの直交性という大きな制約条件があり、問題となっている。この問題を解決するためには、コーディングの機能をもった連想記憶モデルを構成する必要があると考えられる。しかしながら、従来のモデルではこれについてほとんど扱われていなかった。

ところで、哺乳類の脳には「海馬」と呼ばれる組織があり、記憶機能と深い関わりがあることが知られている⁽²⁾。海馬の機能はいまだに明確ではないが⁽³⁾、海馬は一種の連想記憶装置であるというかなり有力な見方が出されている。

海馬が連想記憶機能をもっているのならば、海馬に関する数多くの生理学・神経心理学的知見は、新たな

連想記憶モデルを構成するための有用な手がかりを与えてくれると期待される。また逆に、これらの知見に基づいて自然な連想記憶モデルを構成することができたならば、海馬の機能と機構を明らかにしていく上で示唆するところが大きいであろう。

本論文は、以上のような考えのもとに構成した、連想記憶の神経回路モデルを提案するものである。

以下では、まずモデルの背景となる事柄について述べる。次いで具体的なモデルについて説明し、シミュレーションの結果を示す。更に本モデルが時系列パターンの記憶モデルに自然に拡張できることを示した上で、最後にまとめと共にいくつかの考察を行う。

2. モデルの背景

2.1 相関記憶とコーディング

相関行列型の連想記憶では、 μ 番目の記憶パターンを縦ベクトル x_μ ($\mu=1, 2, \dots, k$) で表したとき、(自己)相関行列

$$M = \sum_{\mu=1}^k x_\mu x_\mu^T \quad (1)$$

を用いて記憶し (T は転置を表す)、 M の右から x_ν に似たベクトルを掛けることによって x_ν を読み出す。このとき、すべての記憶パターンが互いに直交していれば ($x_\mu^T x_\nu = 0$) もとのパターンが正確に再現されるが、直交性が悪いと記憶したものと異なるパターンが出力されやすい。これを Hopfield 型の連想記憶モデル⁽⁴⁾ で考えれば、記憶パターン以外のエネルギー極小点が多数

[†] 東京大学工学部計数工学科, 東京都
Faculty of Engineering, The University of Tokyo, Tokyo, 113
Japan

できてしまうことに相当する。

従って、相関記憶を用いる限り、すべての記憶事項は互いにほぼ直交したパターンによって表される必要がある。これを避けるために直交学習⁽¹⁾を用いる(一般逆行列を求める)方法もあるが、学習の収束が遅いなどの問題がある。

そのため、従来の連想記憶モデルのほとんどが記憶パターンの直交性を仮定している。ところが、神経回路による情報処理の利点は、パターン間の類似度が情報の類似度を表現していることによる部分が大きいから、直交性の仮定は不自然かつ強い制約である。

こうした問題は、記憶事項をどのようにコード化するかというコーディングの問題の一つとしてとらえることができよう。そしてこの場合、パターン間の類似度が意味をもつような表現と、すべてがほぼ直交しているような表現の二つを、相互に変換可能な形で同時にもつというのが最も自然な解答だと考えられる。

このような情報のコーディングやパターン変換のメカニズムは、これまであまり議論されていなかったが、神経回路における連想記憶にとって本質的な意味を持ち、モデルを構成する上でも重要だといえる。

2.2 海馬と連想記憶⁽⁵⁾

海馬は大脳辺縁系に属する系統発生的に古い皮質で⁽⁶⁾、主に本能的行動に関係するとされているが、記憶の機能とも深く関わっていることが知られている。例えば、海馬を損傷したヒトは特異な記憶障害を示し、ほんの少し前の出来事でも思い出すことができない^{(2),(7)}。その他、長期増強などの興味深い知見が得られており、海馬と記憶との関係をめぐり数多くの研究が進められている。

なかでも生理学者の Rolls は、海馬のニューロンの振舞いを相関行列型の連想記憶と対比して研究し、海馬の相関行列モデルを提案した⁽⁸⁾。これが海馬に関するほとんど唯一の神経回路モデルだといえるが、前述の直交性の問題をはじめとして不十分な点も多い。

ここでは、Rolls の考えも参考にしながら、後述のモデルと関係の深い知見を述べる。

図1は海馬周辺の主な線維連絡を模式的に示したものである。ここで注目すべき点の一つは、海馬(hippocampus)と大脳新皮質(cerebral neocortex)、特に連合野(association area)との間には、相互の線維連絡が広範に存在することである^{(3),(9)}。海馬のニューロンが高いシナプス可塑性をもっていること、大脳皮質へは海馬ニューロンの軸索が直接伸びていることな

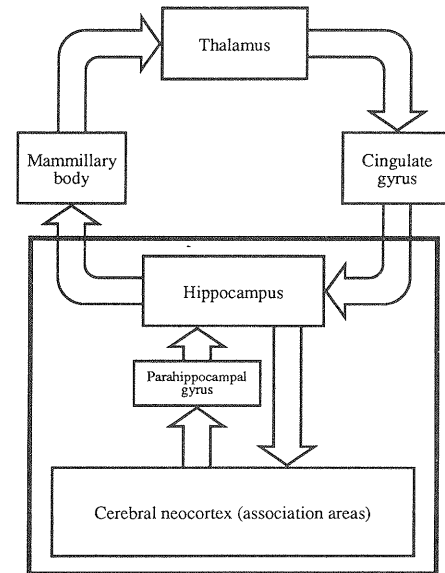


図1 海馬の主な線維連絡

Fig. 1 Schematic connections of the hippocampus.

どを併せて考えれば、海馬を大脳連合野を入出力部とする一種の連想記憶装置とみなすことが可能であろう。

海馬の線維連絡に関してもう一つ重要なのは、いわゆる Papez の回路^{(6),(10)}である。これは海馬を出て視床下部乳頭体(mammillary body)、視床(thalamus)、帯状回(cingulate gyrus)を通して海馬に戻ってくる神経の閉回路であるが、これについては 6.1 で触れる。

このほかにも数多くの知見があるが、その中で筆者が特に注目したのが、海馬 θ 波、すなわち海馬に特徴的な 4~8 Hz の規則的な脳波である^{(6),(11)}。海馬にはこれに同期して周期的に活動電位が変化する細胞(シータ細胞)があり、また海馬錐体細胞の活動も海馬 θ 波と関連があるとされている。しかし、海馬 θ 波やシータ細胞がどういう働きをしているのかは明らかでなく、記憶機能との関係も全く不明である。

3. モデルの構成

以上述べた事柄を背景として構成したのが、以下に述べるモデル⁽¹²⁾である。まず、全体の構成を図2に示す。厳密ではないが、図1の太枠の内部に対応している。

本モデルの中心をなすのが P 回路網(P-network)であり、相互に結合した多数の p 細胞(p-cell)から成っている。p 細胞間の結合は対称かつランダムであって、その荷重は一定で学習による変化はない。

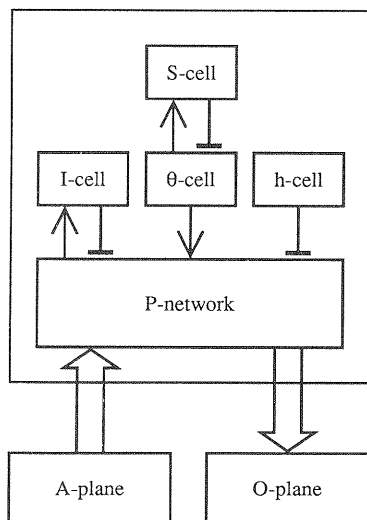


図2 モデルの構成
Fig. 2 Structure of the model.

I細胞 (I-cell) はすべての p 細胞から興奮性の結合を受けると同時に、それらに対して抑制的に結合している。この細胞には、コード (後述) の性質を良くする、システムを安定化するという働きがある。

θ細胞 (θ-cell) と S細胞 (S-cell) の間には相互に結合があり、両者の相互作用によって発振が生じる。この振動のことを、現実の海馬 θ波に倣って“θ波”と呼ぶことにする。θ細胞は、全 p 細胞に対して一様に興奮性に結合している。

A細胞面 (A-plane) と O細胞面 (O-plane) とは、それぞれ同数の a細胞 (a-cell) と o細胞 (o-cell) から成る細胞面で、この記憶システムの入出力部に相当する。便宜上、入力面と出力面とを分けたので、入力は a細胞の興奮パターンで、出力は o細胞の興奮パターンで表されることになる。a細胞から p細胞、p細胞から o細胞へはそれぞれ結合があるが、これらの結合のシナプスは可塑性をもっており、次章に述べる学習則に従って変化する。

h細胞 (h-cell) は常に 1 を出力する細胞で、p細胞および o細胞への抑制性の結合をもつ。これらの結合の荷重も学習によって変化する。

次に、各細胞の入出力関係を式で示すことにする。本モデルでは原則として連続時間一連続情報の細胞モデル⁽¹⁾を用いているが、a細胞だけは入力パターンに応じて 0, 1 の 2 値をとる。また、後述のようにモデルの動作モードには記録モードと想起モードの二つがあ

り、両者で若干の違いがある。

以下の式において、 u は細胞の平均膜電位を表し、ある上限と下限とをもつ。 τ はその変化の時定数である。また、 $w_{\alpha\beta}$ という表記は細胞 β から細胞 α への結合荷重を意味する。非線形関数 $f(u)$ としては

$$f(u) = \frac{1}{1 + e^{-cu}} \quad (2)$$

を用いた (c は正の定数)。

i 番目の p 細胞 (細胞 p_i) の出力 x_{p_i} は次式によって与えられる。

$$\begin{aligned} \tau \frac{du_{p_i}}{dt} = & -u_{p_i} + \sum_j w_{p_i p_j} x_{p_j} + w_{p_i \theta} x_{\theta} - w_{p_i} x_I \\ & - w_{p_i h} + r_{p_i} + q \sum_j w_{p_i a_j} x_{a_j} \end{aligned} \quad (3)$$

ここで、 r_i はランダム刺激であり、各細胞、各時刻においてランダムな値をとる。これが P 回路網の「ゆらぎ」を表していると考えられる。また、係数 q は、記録モードでは 0 または微小値を取り、想起モードでは 1 である。

I細胞の出力は、

$$\begin{aligned} \tau_I \frac{du_I}{dt} = & -u_I + w_{p_i} \sum_j x_{p_j} + s_I \\ x_I = & f(u_I) \end{aligned} \quad (4)$$

で表される。但し s_I は外部刺激である。

θ細胞と S細胞の入出力関係は次のとおりである。

$$\begin{aligned} \tau_{\theta} \frac{du_{\theta}}{dt} = & -u_{\theta} + w_{\theta \theta} x_{\theta} - w_{\theta s} x_s + s_{\theta} \\ \tau_{\theta} \frac{du_s}{dt} = & -u_s + w_{s \theta} x_{\theta} + w_{s s} x_s + s_s \\ x_{\theta} = & f(u_{\theta}), \quad x_s = f(u_s) \end{aligned} \quad (5)$$

各パラメータが適当な値であれば、約 $2\tau_{\theta}$ の周期の発振が生じる。

i 番目の o細胞の想起モードにおける出力は

$$\begin{aligned} \tau \frac{du_{o_i}}{dt} = & -u_{o_i} + \sum_j w_{o_i p_j} x_{p_j} - w_{o_i h} + s_{o_i} \\ x_{o_i} = & f(u_{o_i}) \end{aligned} \quad (6)$$

で与えられる。記録モードでは $x_{o_i} = x_{a_i}$ である。

4. モデルの動作

本モデルの動作は記録モードと想起モードとに分けられるが、どちらも θ波の 1 周期が動作の基本単位となっている。

記録モードでは、入力パターンを p細胞のある興奮パターン (“コード”と呼ぶ) と相互に連合することによ

て記憶する。一方想起モードでは、入力したキーパターンに対応するコードをP回路網上に再生し、それを通じて記憶したパターンの想起を行う。

各モードにおける動作の詳細は後で述べるとして、まずどのようにコーディングを行うかについて述べる。

4.1 コードの生成

簡単のため、I細胞を無視する($w_{pi}=0$ として考える)ことにする。そうすると、p細胞間の結合は対称($w_{p_i p_j} = w_{p_j p_i}$)なので、P回路網にはエネルギー関数⁽⁴⁾

$$E = -\frac{1}{2} \sum_i \sum_j w_{ij} x_i x_j - \sum_i x_i s_i + \sum_i g(x_i) \quad (7)$$

が定義される(添字*p*は省略した)。ここで、 s_i は細胞*p_i*への外部刺激、 $g(x)$ は $f^{-1}(x)$ の原始関数である。 s_i が一定ならば、系の状態は E が減少する方向にのみ変化する($dE/dt \leq 0$)から、回路網の安定状態はエネルギー関数 E の極小点に対応する。

この極小点は、例えば結合荷重をランダムに選ぶことによって、細胞数*n*の指数オーダーの数だけ存在するようにできる⁽¹³⁾。一方記憶システム全体としては、基本的に関連記憶を用いているため、もともと*n*の1次のオーダーしか記憶容量がない。そこで、記憶すべきある入力パターンに対し、これらの極小点(P回路網の安定状態)のうち一つをランダムに選び、“コード”として割り当てることにする。このようにしても、*n*が十分大きければ、一つのコードが複数の記憶事項に対応してしまうことはほとんどないと考えられる。

ではどうやってランダムに極小点を選ぶかが問題となるが、次に本モデルで用いているメカニズムについて定性的に説明する。

まず各細胞が受ける外部刺激 s_i を、細胞ごとに異なる成分 h_i と全細胞に共通する刺激レベル*S*とに分けて考える($s_i = h_i + S$)と、式(7)は

$$E = -\frac{1}{2} \sum_i \sum_j w_{ij} x_i x_j - \sum_i x_i h_i - S \sum_i x_i + \sum_i g(x_i) \quad (8)$$

と変形される。このとき、エネルギー関数 E の形は、刺激レベル*S*を変えることによって変化する。図3はその様子を模式的に示したものである。

S がある値 S_0 以下のときには、 $\mathbf{x}=\mathbf{0}$ の状態しか安定ではない。すなわち E の極小点はただ一つであり、これを0状態(0-state)と呼ぶことにする。 S が S_0 より大きくなると、0状態以外の極小点が存在するようになる。それと共に0状態の安定度は悪くなり、 S が

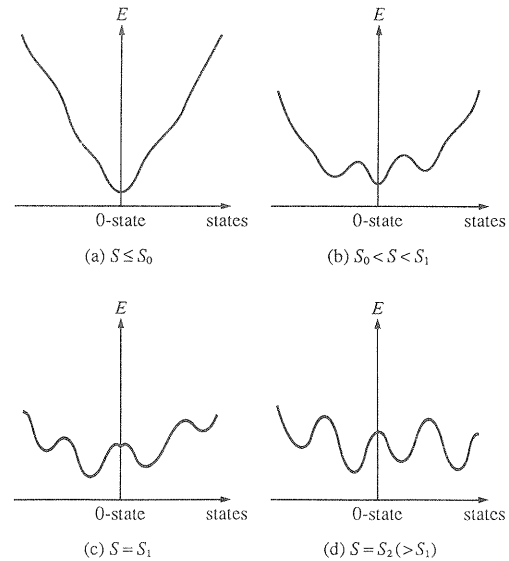


図3 エネルギー関数の刺激レベルによる変化
Fig. 3 Dependence of the energy function on the stimulus level.

ある値 S_1 を超えると不安定になってしまう。

そこで、 S を S_0 から適当な速さで大きくしていくと、 $S < S_1$ の間系の状態は0状態にとどまっているが、 S が S_1 を超えた時点で0状態から飛び出す。その後 S がある値 S_2 に保つと、系は最終的にどこかの安定状態に落ち着く。この際、どの安定状態に落ち着くかはどの安定点の引き込み領域に向かって飛び出すかに依存するわけだが、これは0状態が不安定になる瞬間に系が受ける刺激によって大きく左右される。従って、外部から与えられる刺激が小さければ、ゆらぎの影響によって落ち着き先の極小点はほぼランダムに決まることになる。

以上のように、刺激レベルを変化させることによって一時的にゆらぎの効果を拡大する方法は、初期状態を最初からランダムに与えるよりも自然であり、後述のコードの再生においても有用である。

4.2 記録モード

θ 細胞からP回路網への周期的な刺激は、刺激レベル*S*を S_0 から S_2 へ上述のごとく変化させ、ある時間一定に保ったのち再び S_0 以下に下げることに対応する。但し、 θ 細胞の出力が大きくなってp細胞のいくつか興奮するとI細胞が全体に抑制をかけるので、実際には S は S_2 まで上がった後少し下がることになる。

以上のようにして、記録モードでは θ 波の各周期ごとに異なったp細胞の興奮パターンが現れるが、これを

A 面と O 面に入力されている記憶パターンと連合する。すなわち、a 細胞から p 細胞および p 細胞から o 細胞への結合荷重を次式のように学習する。

$$\tau' \frac{dw_{pa_j}}{dt} = -w_{pa_j} + \alpha_{pa} x_{p_i} x_{a_j} \quad (9)$$

$$\tau' \frac{dw_{op_j}}{dt} = -w_{op_j} + \alpha_{op} x_{o_i} x_{p_j} \quad (10)$$

ここで τ' は学習の時定数、 α_{pa} 、 α_{op} は荷重の強化量を表す定数である。 τ' は τ に比べて十分に大きい値をとる。また、h 細胞からの抑制性の荷重も次式に従って変化する。

$$\tau' \frac{dw_{p_i h}}{dt} = -w_{p_i h} + \alpha_{ph} x_{p_i} + h_p \quad (11)$$

$$\tau' \frac{dw_{o_i h}}{dt} = -w_{o_i h} + \alpha_{oh} x_{o_i} + h_o \quad (12)$$

α_{ph} 、 α_{oh} 、 h_p 、 h_o はすべて定数である。

式(9)～(12)からわかるように、本モデルで用いている学習則自体は、普通の相関学習と全く同じである。

4.3 想起モード

想起モードでは、A 面にキーとなるパターンを入力すると O 面に想起したパターンが出力される。キーパターンは通常記憶したパターンの一部分であるが、ノイズを含んでいる場合もある。

想起を行うためには、まず対応するコードを再生することが必要であるが、その過程はコード生成の場合とほぼ同様である。すなわち、 θ 細胞の出力が増加すると 0 状態が不安定になり、別の安定状態に落ち着くというダイナミクスを利用する。但し、a 細胞からの刺激があるため、落ち着き先の安定状態はランダムには決まらない。つまり、対応するコードが埋め込まれている安定点へ向かう可能性の方が、一般にそれ以外の安定点へ向かう可能性よりも高い。

ここで問題となる点は、記憶パターン自体の直交性は全く保証されていないため、入力パターンからコードへの対応を記憶している相関行列 ($w_{p_i a_j}$ を成分とする行列) の性質はよいとは限らないことである。そのため、うまく目的のコードが再生されないことがしばしばありうる。

しかしながら、p 細胞間の結合は変化せず、また a 細胞から受ける刺激は p 細胞相互がやり取りする刺激の総量 (式(8)右辺の第 1 項に相当する) に比べて小さくてもかまわないため、P 回路網の安定点の分布はほとんど変化することがない。つまり、コードは常に安定点のままであり、それ自体はいつまでも記憶として残って

いることになる。従って、1 度コードの再生に失敗した場合でも、 θ 波の各周期ごとに上述の過程を繰り返すうちにうまく再生されることもある。

このようにして正しいコードの再生に成功しさえすれば、コード相互の直交性はよいため、 $w_{o_i p_j}$ を成分とする相関行列によって比較的容易に記憶したパターンを想起できる。

5. シミュレーション

以上示したモデルの動作を確認するために、コンピュータシミュレーションを行った。本モデルは実用性を目指したものではないので、パラメータに関する定量的な議論等にはあまり意味がない。従って、以下ではいくつかの例を示し、定性的な性質を述べることにする。

まず、記録モードにおけるコード生成過程をシミュレートした。図 4 がその結果の一例であり、上から順に θ 細胞、I 細胞、および 60 個の p 細胞の出力の時間変化を表している。この例では、p 細胞間の結合荷重は 1、-1 のどちらかをそれぞれ 1/3、2/3 の確率でとるよう乱数で決めた。また h 細胞からの結合の荷重は、適当な初期値から始めたのち、式(11)による修正がある程度落ち着いた時点での値を用いた。

図に示されているように、 θ 細胞の出力が大きくなるたびに異なる p 細胞の興奮パターンが現れ、ランダムにコードが生成されることが確かめられた。但し、p 細胞数 $n=60$ というのは必ずしも十分な数ではないため、長く動作させていると同じようなパターンが現れてしまうことがある。

次に、モデル全体をシミュレートし、実際にいくつかのパターンを記憶させる実験を行った。

まず、記録モードにおいて、図 5 に示す五つのパターン A～E をそれぞれ 1 度ずつ入力した。そののち、想起モードでいくつかのキーパターンを入力し、想起を行わせた。その結果の例を図 6 に示す。図の上半分が A 細胞面 (キー入力)、下半分が O 細胞面 (想起出力) を表しており、それぞれ各 a 細胞および o 細胞の出力が時間を横軸としてプロットされている。

(a) は、A～E の各記憶パターンの一部をキー入力とした場合である。すべて 1 度で正しいパターンが想起されているが、このことは P 回路網上にコードがうまく再生されたことを意味している。

(b) は、パターン B と C との共通部分をキーとして入力した場合である。キーからコードへの対応が一意でないためコードの再生は容易ではないが、何度か繰り返

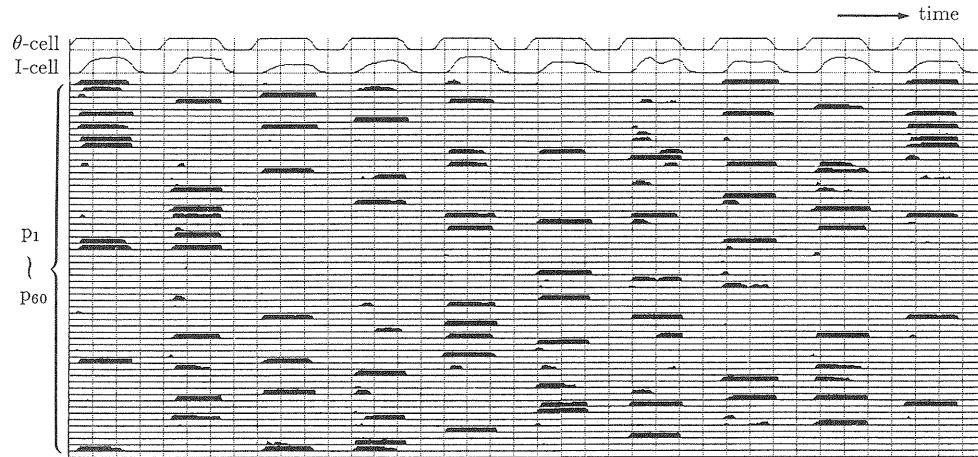


図4 コード生成過程のシミュレーション例
Fig. 4 Simulation results of the code production process.

返すことによって、両方のパターンをなんとか想起することができた。この結果から、本モデルが一対多の

連想能力をある程度備えていることがわかる。

6. 時系列パターンの記憶モデル

6.1 モデルの概要

海馬はある一連の出来事を記憶するのに重要であると考えられている。そしてこのような記憶には、海馬だけでなく Papez の回路全体が関係することが示唆されている^{(2),(5)}。例えば、ヒトの乳頭体や視床の損傷は特に記憶の文脈に影響を与えるといわれているし、ラットなどでは帯状回の損傷により一連の秩序だった行動が阻害される⁽⁶⁾という。

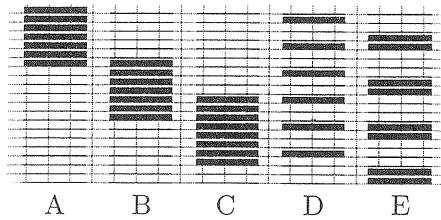


図5 実験に用いた記憶パターン
Fig. 5 Memory patterns used in the experiment.

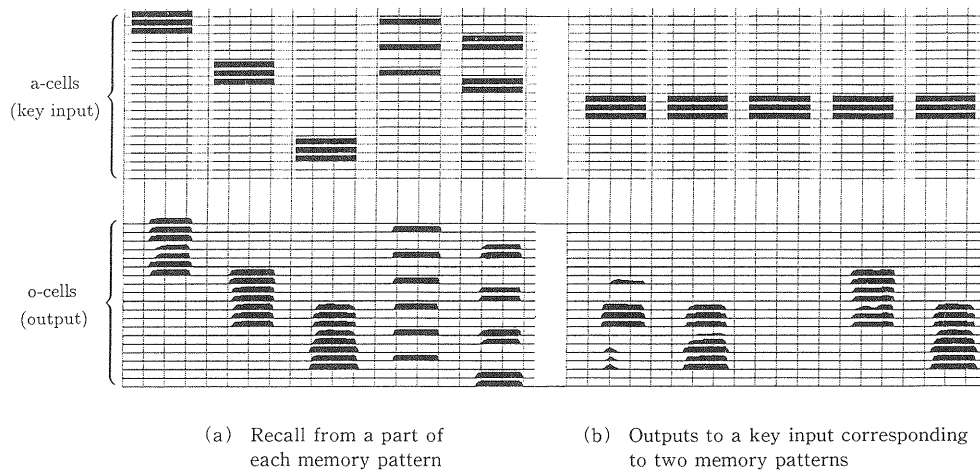


図6 パターンの想起実験
Fig. 6 Output patterns in the recalling mode.

ここでは、その他さまざまな知見を背景としてモデルを拡張し、時系列パターンの記憶を扱うことにする。その基本的アイデアは、入力されたパターン系列そのものでなく、対応するコードの系列を記憶することにある。

拡張したモデル⁽⁵⁾の構成図を図7に示す。現実との対応や今後の発展も考えてこのような構成にしたが、特に必然性があるわけではないので、細かい式等は省略することにする。

M回路網(M-network)とC回路網(C-network)はそれぞれp細胞と同数のm細胞(m-cell)とc細胞(c-cell)から成り、回路網内部の結合荷重は τ_θ より短い時定数 τ'' で変化する。両回路網の動作はほとんど同じであり、共に一種のラッチとして機能する。すなわ

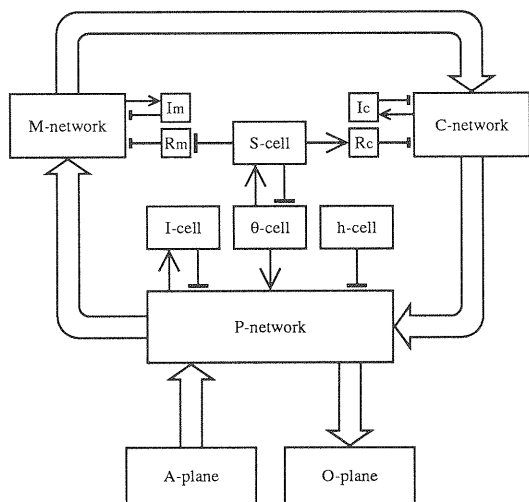


図7 時系列パターンの記憶モデルの構成図
Fig. 7 Structure of the extended model for temporal pattern memory.

ち、m(c)細胞はp(m)細胞から入力を受けて興奮するが、ある期間入力パターンが一定であれば、入力刺激がなくなった後もm(c)細胞相互の刺激によってパターンを維持することができる。また、I_m(I_c)細胞が興奮していない細胞を抑制するため、R_m(R_c)細胞が全体をリセットするまで、別の入力パターンが入ってきても前のパターンを保持し続ける。

R_m細胞とR_c細胞は共にS細胞の出力に従って出力を出す。その位相が θ 波の半周期だけずれている。そのため、P回路網に現れたコードはM回路網からC回路網に順次保持され、この間に次のコードがP回路

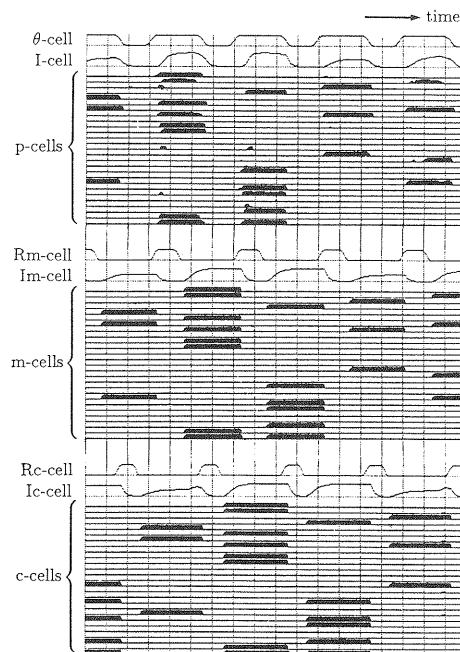


図8 拡張部分の動作シミュレーション
Fig. 8 Simulated behavior of the extended parts.

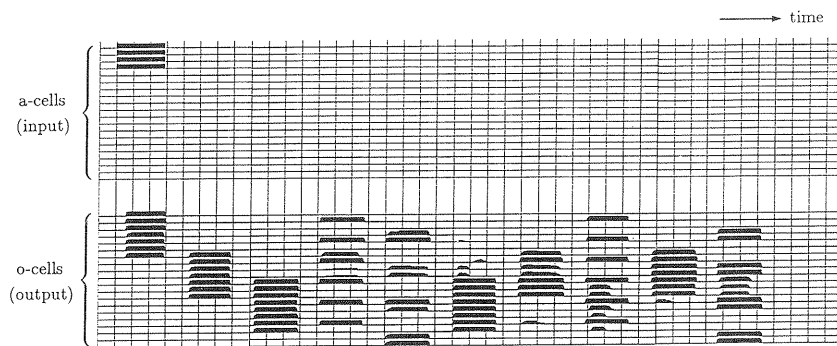


図9 時系列パターンの想起例
Fig. 9 Recall of a temporal pattern.