

非単調神経素子の選択的不感化を用いた文脈依存的連想モデル

森田 昌彦[†]

松沢 浩平^{††}

諸上 茂光^{††}

A Model of Context-Dependent Association Using Selective Desensitization of Nonmonotonic Neural Elements

Masahiko MORITA[†], Kouhei MATSUZAWA^{††}, and Shigemitsu MOROKAMI^{††}

あらまし 分散情報表現を用いた神経回路網において、同じ入力から文脈に応じて異なるパターンを連想する際には、主に入力パターンに文脈を表すパターンを連結する方法が用いられてきた。しかし、この方法には本質的な問題点があり、入力及び文脈パターンの数に関して強い制約が生じていた。本論文では、これとは異なる文脈修飾の方式を非単調神経回路網に適用することによって、従来の問題点を解決する文脈依存的連想モデルを構築する。このモデルでは、入力及び文脈パターン数に関係なく、学習できる連合の数が素子数にほぼ比例して増加する。また、文脈を切り換えることによって、アトラクタ間の状態遷移を自在に制御することができる。本モデルは、素子数や学習時間の爆発を招くことなく任意の有限オートマトンの動作を模擬可能であると同時に、完全な分散表現に基づく高い汎化能力を備えており、従来の記号処理の限界を打破する可能性を秘めている。

キーワード 非単調連想モデル、軌道アトラクタ、積型文脈修飾、文脈主導型連想、有限オートマトン

1. ま え が き

一般に人間の情報処理は文脈に強く依存しており、高度に知的で柔軟な情報処理を工学的に実現するためには、文脈依存性の問題を避けて通ることはできない。人工神経回路網の場合も例外ではなく、文脈に依存した情報処理の機能は非常に重要である。

一口に「文脈」といっても様々な意味と役割があり、一般的に扱うのは困難であるが、ここでは「文脈が異なれば、同一の入力に対する望ましい出力が変わる」という基本的な観点から議論する。すなわち、本論文でいう「文脈」とは、系の入出力関係に影響を与える情報のことである。このように入力情報を修飾して出力を変化させる機能は、文脈が果たす役割の大部分に共通し、文脈処理をモデル化する際の基礎となるといえよう。

しかしながら、この観点から見たとき、従来の人工

神経回路網の文脈処理能力は極めて乏しい。分散情報表現に基づいて文脈を扱うモデルもいくつか提案されているが、ごく小規模な場合にしかうまく動作しないのが実状である。2. で詳しく論じるように、その主な原因は従来用いられてきた文脈修飾、すなわち入力パターンに文脈を表すパターンを単に連結するという方法（以下、直和型文脈修飾と呼ぶ）にある。この方法を用いる限り、ある入力パターン群を異なる多数の文脈に応じて別々の出力パターン群に対応づけることは本質的に困難なのである。

このほかに文脈に応じて入出力関係が大きく変化する場合にしばしば用いられるのが、いわゆるエキスパートネットモデル [1]、すなわち文脈ごとに特化した回路網を切り換えながら使用する方法である。確かに、この方法では後述の非直交性及び平均化の問題は生じない。しかしながら、文脈の数だけ回路網が必要であるため、素子の利用効率是非常に悪い。更に問題なのは、この方法は文脈修飾の際、実質的に文脈をシンボルとして扱っていることである。そのため、文脈情報をどのように適切にシンボル化するかという新たな問題が生じる一方で、異なる文脈の間で全く汎化が起きない。これでは神経回路網によるパターン情報処理の利点を十分に生かすことができない。

[†] 筑波大学機能工学系，つくば市

Institute of Engineering Mechanics and Systems, University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, 305-8573 Japan

^{††} 筑波大学大学院システム情報工学研究科，つくば市

Doctoral Program in Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, 305-8573 Japan

本論文では、直和型とは異なる分散文脈修飾方式を提示し、これを非単調素子を用いた連想記憶モデルに適用することによって、神経回路網による情報処理の大きな足かせとなっていた上記の問題を解決する。以下では、まず従来の文脈修飾法の問題点、及び提案するモデルのもととなる非単調神経回路網について説明する。次に、文脈パターンの積型修飾及び素子の選択的不感化について述べた上で、これを非単調神経回路網に適用することによって、規模の制約なく文脈依存的な連想を実現できることを示す。また、文脈主導型連想モデルを構築し、これによって有限オートマトンの動作が模擬できることを示す。

2. 従来手法の問題点

ここでは、文脈依存的な出力を行う代表的な神経回路網である Elman のモデル (Elman ネット) [2] を例に、従来の分散文脈表現が抱える問題点を明らかにする。

図 1 は Elman ネットの構造である。基本的には 3 層からなる層状の回路網であるが、第 1 層には入力部に加えて文脈部がある。第 2 層は状態層と呼ばれ、第 3 層は出力層である。離散時刻 t における文脈部の活動パターンが、1 時刻前における状態層のパターン $y(t-1)$ そのもの (より一般的なモデルでは、 $y(t-1)$ によって決まるあるパターン) となっているのが特徴である。第 3 層のパターン $z(t)$ は纯粹に $y(t)$ にのみ依存するから、異なる z を出力するためには第 2 層の活動パターンが異ならなければならない。そこで、第 3 層のことは無視して、入力部と文脈部を合わせた第 1 層のパターン x から y を出力する単純な 2 層の回路について考えよう。

今、この回路の入力部に $p (\geq 2)$ 通りのパターン S^1, \dots, S^p が入力され、文脈部には $q (\geq 2)$ 通りのパターン C^1, \dots, C^q が現れるものとしよう。 S^μ と C^ν はそれぞれ n_1 及び n_2 次元のパターンとすると、両者を連結した $n_1 + n_2$ 次元のパターン $x = (S^\mu, C^\nu)$ は最大 $p \times q$ 通りあることになる。これらに対して、それぞれ独立に決められたパターン $T^{1,1}, \dots, T^{p,q}$ を y として出力する回路網を構成しようとする、以下のような問題が生じる (実際には pq 通りのパターンすべてが現れるとは限らないが、ある入力 S^μ に対する出力パターンが C^ν に依存して変わり、かつ同一の文脈 C^ν における出力パターンが S^μ によって異なるのであれば、同じ議論が成り立つ)。

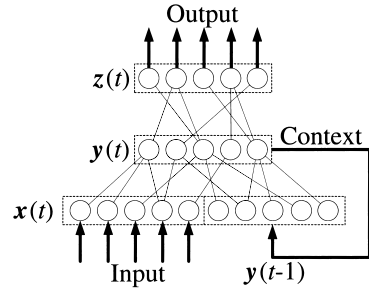


図 1 Elman ネットの構造
Fig. 1 Structure of the Elman network.

第 1 に、第 1 層のパターン同士がかなり類似してしまうという点である。例えば、 (S^1, C^1) と (S^1, C^2) とは少なくとも n_1 個の成分が一致するし、 (S^1, C^1) と (S^2, C^1) は n_2 個以上の成分が重なる。したがって、仮に S^μ や C^ν が相互に十分離れていたとしても、 x に関する直交性は低い。特に、 n_1 と n_2 の値が大きく違うときにこれが顕著である。一般に x から y を中間層を介さずに連想する場合、入力パターンの直交性が低いほど干渉によるノイズが増え、目的のパターンが出力されるよう結合荷重を設定するのは難しくなる。 p 及び q が増えると類似したパターンの組合せが飛躍的に増大するから、結局 p と q が比較的小さく、かつ $n_1 \simeq n_2$ でなければ、上記の目的の達成は困難ということになる。これを非直交性の問題と呼ぶ。

第 2 は、1 対多対応による平均化の問題である。例えば入力パターンが S^1 のとき、状態層の活動パターンは $T^{1,1}$ から $T^{1,q}$ の q 通りをとる。このとき、入力部から状態層への結合荷重は、 S^1 を入力すると $T^{1,1}, \dots, T^{1,q}$ を平均化したパターン \bar{T}^1 を連想するような値に設定せざるを得ないし、荷重を学習によって決めるならば、どんな学習則を用いるにせよそれに近い値に近づく。同様に、 S^μ を入力したときには $T^{\mu,1}, \dots, T^{\mu,q}$ を平均化したパターン \bar{T}^μ が連想される。ところが、 q が大きければ大きいほど平均化の効果が顕著になり、 $\bar{T}^1, \dots, \bar{T}^p$ はどれも似たようなパターンとなる。これは、どの S^μ が入力されても状態層の各素子が受け取る信号には大差がないということであり、入力部から状態層への結合がほとんど無意味になるということである。とすれば状態層の出力は文脈パターン C^ν によって決めるしかないが、全く同様な議論が文脈部についても成り立つ。すなわち、ある一つの文脈 C^ν に多数の $T^{\mu,\nu}$ が対応するならば、

文脈部から状態層への結合もまた平均化によって無意味になってしまうのである。したがって、 p または q が十分小さくない限り、目的の回路網は実現不可能である。

これらの問題点のうち、後者の平均化の問題はかなり根が深く、容易には解消されない。例えば、入力部や文脈部と状態層の間に別の中間層を設けるとしても、第1層と中間層または中間層と状態層の間の問題に変わるだけで、それだけでは何ら本質的な解決にならない(Elman ネットの場合、4層の回路網になるため学習がより難しくなり、むしろ性能が低下するおそれが強い)。また、アトラクタへの引込みとといったリカレント結合によるダイナミクスを用いても、若干の量的改善はともかく、質的な改善は得られない。例えば1対多の連想モデル HASP [4] はその例であるが、記憶するパターン数が増えると、想起すべきパターンとのみ高い相関をもつパターンをバイアスとして入力しない限り連想がうまくいかない[5]。つまり、HASP は平均化の問題を解決するものではなく、むしろこの問題が生じる典型的な例である。更に、直和型以外の修飾方式(後述の種型も含む)を用いても、それだけで直ちに解決するようなものではない。

理論的には、Elman ネットは任意の有限オートマトンと同等な動作が可能である[3]。しかし、これは無限個の素子が使用可能な場合の話にすぎないし、学習によって実現できることを保証しているわけでもない。サイズが有限である現実の回路網では、上記の理由から大規模なオートマトンは実現できないし、誤差逆伝搬法による学習の計算量や収束可能性を考えると、実際に実現可能なのはごく小規模なものに限定される。有限オートマトンはシンボルを用いた基本的な文脈処理のモデルともいえるから、このことは Elman ネットの文脈処理能力の限界を示している。そして、このような限界は、Elman ネットに限らず、同様な文脈の表現及び修飾方法を用いた神経回路モデル一般に共通するものである。

3. 非単調連想モデル

上記の問題の解決法を提示する前に、非単調神経回路網による動的連想モデル(以下、非単調連想モデルと呼ぶ)について簡単に説明する。詳細については文献[6],[7]を参照されたい。

3.1 回路網のダイナミクス

非単調神経回路網[8]は、ある種の非単調な出力関

数をもつ神経素子(非単調神経素子)が相互に全結合した回路網である。ここでは、非単調出力関数として

$$f(u) = \frac{1 - e^{-cu}}{1 + e^{-cu}} \cdot \frac{1 - e^{c'(|u|-h)}}{1 + e^{c'(|u|-h)}} \quad (1)$$

を用い(c, c', h は正の定数)、連続時間のダイナミクスに従って動作するものとする。すなわち、 i 番目の素子の出力 y_i は次式で与えられる。

$$\tau \frac{du_i}{dt} = -u_i + \sum_{j=1}^n w_{ij} y_j + z_i \quad (2)$$

$$y_i = f(u_i) \quad (3)$$

ここで n は素子の数、 u_i は素子の内部電位、 w_{ij} は j 番目の細胞からの結合荷重、 z_i は外部からの入力信号、 τ は時定数を表す。

このモデルでは内部電位 u_i の符号が重要なので、 $x_i = \text{sgn}(u_i)$ を考え($\text{sgn}(u)$ は $u > 0$ のとき 1 、 $u \leq 0$ で -1 をとる符号関数)、ベクトル $x = (x_1, \dots, x_n)$ のことを回路網の状態と呼ぶ。ある時刻における回路網の状態 x は、 2^n 個の可能な状態からなる状態空間中の1点で表される。また、各 x_i は非同期的に変化するから、 x が別の状態に移移する際には、ほとんどの場合隣の点(成分が一つだけ違う状態)に移動する。したがって、状態空間には時間とともに x の連続的な軌跡が描かれることになるが、これを x の軌道と呼ぶ。

非単調神経回路網は、非常に優れた連想記憶の能力をもち、他の神経回路モデルにはない多くの特長をもつ[8]。その一つに、パターンの直交性がかなり低くても、Hebb 型の単純な学習則によって記憶が可能という性質がある[9]。これが本研究でこのモデルを採用する一つの大きな理由である。

本研究に関係するもう一つの重要な特長は、状態空間中のある連続的な軌道に沿って、任意の状態から任意の状態へ安定に移移可能ということである。このような軌道のことを軌道アトラクタと呼び、これを用いて様々な時空間パターン処理が実現できる[7],[10]。なお、通常の単調な出力関数を用いた場合、複数の素子を組み合わせると等価的に非単調素子を実現しない限り、こうした軌道アトラクタを形成することは極めて困難である[6]。

3.2 学習と連想

ここでは、ある空間パターン群 S^1, \dots, S^m から対応する別の空間パターン群 T^1, \dots, T^m を連想するこ

とを考えよう．ここで S^μ と T^μ ($\mu = 1, \dots, m$) はいずれも ± 1 を成分とする n 次元ベクトルであるが，前者を cue パターン，後者を target パターンと呼ぶことにする．つまり， S^μ を始点とし， T^μ を終点とする m 本の軌道アトラクタによって，cue から target の連想を実現しようというわけである．

このような軌道アトラクタは，以下のようにして形成する．まず， S^μ から T^μ へ至るある連続的な軌道（交差しない限り任意であるが通常は最短経路でよい）に沿って変化する時空間パターンを作成し，これを学習信号 r として回路網の学習を行う．具体的には， $x = S^\mu$ を初期状態として式 (1)~(3) に従って回路網を動作させた上で， r を $z_i = \lambda r_i(t)$ の形 (r_i は r の成分， λ は入力強度) で各素子に入力する．これと並行して，すべての結合荷重 w_{ij} を

$$\tau' \frac{dw_{ij}}{dt} = -w_{ij} + \alpha r_i y_j \quad (4)$$

に従って更新する．ここで， τ' は学習の時定数であり， $\tau' \gg \tau$ である． α は学習係数であり定数でもよいが， $|u_i|$ の増大とともに α を減少させると学習性能が高まるため，ここでは $\alpha = \alpha' x_i y_i$ とする (α' は正定数)．

学習信号 r が変化すると，回路網の状態 x は r の少し後を追従しながら，ほぼその軌道に沿って遷移する． r が終点，すなわち T^μ に達したら，しばらくの間 (x が r に十分近づくまで) $r = T^\mu$ としたまま学習を続ける．以上をすべての μ について行えば，1 回の学習が終了する．これを何回か繰り返すが，回を重ねるごとに r の入力強度 λ を弱めていくとよい．

図 2 は，学習後の回路網の状態空間における仮想的なエネルギー地形を模式的に表現したものである．平面上の各点は回路網の状態を，深さはその状態のポテンシャルエネルギー（安定度）を表す．直観的にいうならば，式 (4) の学習則は r 周辺における回路網のエネルギーを低くする．したがって， r が S^μ から T^μ まで連続的に変化することによって，図に描かれているような溝が刻まれることになる．更に， r が x よりも少し先行することによって， x から r の方向，すなわち S^μ から T^μ へ向かう緩やかな流れが溝の底に作られる．その結果，回路網の初期状態として S^μ やその近傍のパターンを与えるだけで， x は r の軌道にほぼ沿って T^μ に達し，目的とする target が連想される．

非単調回路網の性質により，このような軌道アトラクタは，交差または極端に相互に接近しない限り，近

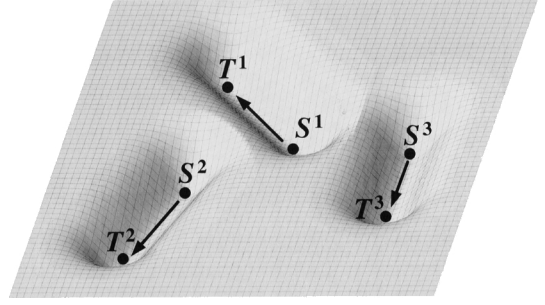


図 2 軌道アトラクタが形成された回路網のエネルギー地形の模式図

Fig.2 Schematic energy landscape of the network with trajectory attractors.

い領域に複数のものが存在可能である．また，状態空間の次元 n は十分に大きく，2 点を結ぶ軌道は多数存在するから，軌道アトラクタ同士が途中で交差する可能性はまずない．そのため，仮に二つの cue パターン S^1 と S^2 がかなり近かったとしても，それぞれから任意の target パターン T^1 と T^2 を連想することができる．

3.3 文脈修飾による軌道の分離

このように，非単調連想モデルでは，cue パターン間になんかの類似性があってもよく，前述の非直交性の問題は大幅に軽減される．しかし，完全に同一の cue から異なる target を連想するのは，原理的に無理である．これを可能にするには，文脈に応じて S^μ に修飾を加え，軌道アトラクタの始点が変わるようにする必要がある．

その手段として最も簡単なのは，Elman ネットと同様に直和型修飾を用いること，すなわち cue パターン S^μ と文脈パターン C^ν の直和ベクトル (S^μ, C^ν) を入力パターンとすることである．これによって，cue パターンの空間で重なっている軌道が回路網全体の状態空間では分離され，複雑に重なり合う軌道をもつ時空間パターンでも処理が可能になる [10]．

しかしながら，この方法を用いる限り，前述の平均化の問題は避けられない．実際 4.3 の数値実験で示すように，いくら素子数を増やしても， S^μ や C^ν の数には制限がある．そこで，本研究では直和型修飾に代わるものとして次に述べる積型文脈修飾を用い，この問題の解決を図る．

4. 文脈依存的連想モデルの原理

4.1 積型文脈修飾と選択的不感化

入力の変化に対する感度が可変な神経素子を考える。具体的には、回路網の i 番目の素子の出力が

$$y_i = g_i \cdot (f(u_i) - \bar{y}_i) + \bar{y}_i \quad (5)$$

によって与えられ、ゲイン g_i を変えることができるものとする。ここで、 u_i は素子の内部ポテンシャル（離散時間モデルの場合は入力の荷重和）、 \bar{y}_i は素子の平均出力レベルである。このとき、各素子の g_i を順に並べた感度ベクトル $G = (g_1, \dots, g_n)$ が、文脈を表現するパターン C に依存して決まる、すなわち C の関数 $G(C)$ で与えられるというのが、積型文脈修飾の基本的な概念である。

この修飾方式によって前述の平均化の問題を解決しようとするとき本質的な役割を果たすのが、一部の素子について $g_i = 0$ 、すなわち入力に関係なく $y_i \equiv \bar{y}_i$ とする操作である。これを選択的不感化と呼ぶ。このとき、文脈が違えば不感化される素子の組合せは異なっていなければならない（さもないと文脈の違いが反映されない）。また、パターン C の空間の大きさ、すなわち文脈を表現する能力を最大にするために、ある素子が不感化を受ける文脈は素子ごとに独立に選ばれるべきである。このように、文脈ごとに不感化される素子の組合せが変わり、また一つの素子が非常に多数の文脈で用いられたり不感化されたりする点で、エキスパートネットなどで用いられるゲートによる素子群の選択とは本質的に異なる。なお、不感化された素子であっても、入力信号に応じて内部ポテンシャル u_i が不感化されないときと同様に変化することに注意されたい（ただしこのことはあまり本質的ではなく、 u_i が一定であっても大きな違いはない）。

このような修飾によって、仮に同じ状態で同じパターンが入力されていたとしても、出力パターンは C によって変化することになり、異なる target の連想が可能となる。また、この過程における C の影響は間接的なものであって、 C を表現する素子群と target を表現する素子群とが直接連合されるわけではない。このことが、平均化の問題から逃れられる可能性を生むのである。ただし、不感化された素子の出力が \bar{y}_i でないと、修飾されたパターン $y(C)$ と文脈パターン C との間に相関が生じるため、 C と target との 1 対多対応による平均化の問題が残ってしまう。

式 (5) と同じまたは類似した素子はこれまでもいくつかのモデルで使われてきたし、 G が文脈を表現するという積型修飾自体はそれほど目新しいものではない。しかし、これまで平均化の問題が解決されなかったのは、一つには次節で述べる理由からであるが、それ以前にこの問題の本質が十分に認識されていなかったことや、この問題における不感化の意義がよく理解されていなかったことも大きな理由だと思われる（5.2 参照）。

以下では、議論をわかりやすくするために、次のような最も単純化した場合を扱う。まず、すべての素子について $\bar{y}_i = 0$ とする。出力関数 $f(u)$ が式 (1) のような奇関数の場合、この仮定は妥当なものであろう（厳密には成り立たなくても $\bar{y}_i \approx 0$ であればほとんど影響ない）。また、感度ベクトル G を文脈パターン C と同一視し、両者を 1 対 1 に対応させる。すなわち、 C は ± 1 の 2 値をとる n 次元のパターンとし、 $g_i = (1 + c_i)/2$ とする (c_i は C の成分)。更に、パターン中の 1 と -1 の数はほぼ均等になるようにする（この条件は、 n が十分大きければ c_i をランダムに決めても満たされるから、大きな制約とならない）。このとき、素子の約半数が不感化されて $y_i = 0$ を出力し、それ以外は通常素子と同様に $y_i = f(u_i)$ を出力することになる。

このような修飾を受けているときの回路網の状態を $x(C) \equiv (g_1 x_1, \dots, g_n x_n)$ と表すことにしよう。 $x(C)$ は成分として 0, ± 1 の 3 値をとるベクトルである。このとき、何も修飾を受けていない状態 x は、ベクトル $I \equiv (1, \dots, 1)$ による修飾を受けた状態、すなわち $x(I)$ と表現される。

同様に、cue パターン S や target パターン T への修飾を考えることができる。すなわち、文脈パターン C による積型修飾を受けた S を $S(C) \equiv (g_1 s_1, \dots, g_n s_n)$ と定義する (T に関しても同様)。 S や T の成分が ± 1 の 2 値をとるとすれば、 $S(C)$ や $T(C)$ は $x(C)$ と同じく 3 値パターンとなる。

4.2 非単調連想モデルへの適用

積型文脈修飾や素子の選択的不感化は、ほとんどの神経回路モデルに適用することができ、それによる効果も多少は期待できる。しかし、その効果を最大限に発揮するためには、適用するモデルが以下の条件を満たすことが望ましい。

第 1 に、文脈修飾を受けるパターンが回路網のア

トラクタになっていることである．例えば，ある状態 $x = S$ がアトラクタであるとしよう．この状態である文脈パターン C^1 の修飾を受けた直後には， u_i の符号が変わる可能性はほとんどないから，不感化された素子も含めて高い確率で $x_i = s_i$ が成り立つ．つまり $x(C^1) = S(C^1)$ かつ $x \simeq S$ である．この状態で修飾が外されたならば，ほとんど瞬時に $x = S$ に戻ることになる．また，文脈パターンが C^1 から別のパターン C^2 に切り換わったとしても， $x \simeq S$ のままであり， $x(C^2) \simeq S(C^2)$ が成り立つ．このように， S がアトラクタであれば，文脈修飾や文脈の切り換えによって情報が失われてしまう心配はない．

なお， S がアトラクタなら $S(C)$ もアトラクタであるとは限らないし，逆に $S(C)$ がアトラクタでも必ずしも S はアトラクタではない．しかし，十分多数の C^ν に関して $S(C^\nu)$ がアトラクタなら， S は必ずアトラクタである．したがって，いくつかの文脈において $S(C^\nu)$ がアトラクタになるよう学習すれば，文脈修飾を受けない状態で学習をしなくても， S (少なくともそれに非常に近いパターン) はアトラクタとなり，上記の条件は満たされる．

第2の条件は，非直交性の問題が生じないことである． $S(C^1)$ と $S(C^2)$ は， C^1 と C^2 が全く独立なパターンであっても比較的大きな類似性をもつ．また，文脈に関しても分散表現を採用するからには， C^1 と C^2 が類似する場合に対処できなくてはならないが，その場合 $S(C^1)$ と $S(C^2)$ の直交性は更に低い．したがって，直和型修飾の場合と同様に，非直交性の問題を解決する必要がある．

これらの条件を満足し，しかも非常に単純な構造をもつのが，3. で説明した非単調連想モデルである．これに選択的不感化を導入することにより，2. で述べた問題をすべて解消し，文脈に依存した連想を大規模に実現することができる．具体的には，モデルを以下のように修正する．

まず，回路網の動作は，式の上では式 (3) が式 (5) に置き換わるだけである．ただし，cue から target を連想する過程で，ある一定の文脈パターン C^ν を継続的に与える．学習方法も基本的に 3.2 と同じであるが， $S^\mu(C^\nu)$ から $T^{\mu,\nu}(C^\nu)$ へ変化する時空間パターンを学習信号 r とし， C^ν の修飾を加えたまま学習を行う (不感化された素子に関しては学習しない) 点が異なる．このほか必要ならば， S^μ や $T^{\mu,\nu}$ がアトラクタになるよう，文脈修飾を加えない状態 ($C = I$) で，こ

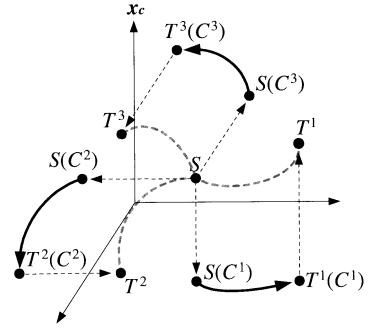


図3 文脈依存的連想過程の模式的表現
Fig.3 Schematic representation of the process of context-dependent association.

これらのパターンを学習信号として短時間学習する．

学習後の回路網では，文脈によって異なる部分状態空間に軌道アトラクタが形成されており，これを用いて文脈依存的連想が実現される．その様子を模式的に表したのが図3である．

この図では，回路網の n 次元状態空間が3次元的に表現されている．すなわち， $x = (x_a, x_b, x_c)$ である．説明の都合上，文脈パターン C^1, C^2, C^3 によってそれぞれ x_c, x_b, x_a に対応する素子群が不感化されて0を出力するとしよう (実際には，異なる二つの文脈で不感化される素子は約半数が重なるし，それ以上重なってもかまわない)．このとき，回路網に例えば C^1 による修飾を加えることは， x を部分空間 $(x_a, x_b, 0)$ に射影し，そこでのダイナミクスに従って回路網を動作させることに相当する (0 は0ベクトル)．したがって，cue パターン S が初期状態として与えられたならば，この部分空間に形成された軌道アトラクタによって $x(C^1)$ は $S(C^1)$ から $T^1(C^1)$ に遷移する．前述のように， T^1 がアトラクタならばこの時点で $x \simeq T^1$ であるし， $C = I$ とすれば正確に T^1 を想起することができる．同様に， C^2 や C^3 による修飾を加えると，それぞれ別の部分空間に形成された起動アトラクタに沿って状態遷移し， T^2 及び T^3 が連想される．

ここで重要なのは，同一及び異なる部分空間に形成された軌道アトラクタ間の干渉を避けることである．例えば x を (x_a, x_b, X_c) に射影する (X_c は ± 1 を成分とするあるパターン) のでは，平均化の問題と同根の強い干渉が生じる．これに対して，不感化された素子の出力をその平均値に近づければ，このような干渉を最小限に抑えることができるのである．

4.3 数値実験

非単調連想モデルに積型文脈修飾を適用したモデル（以下、積型モデルと略称する）について計算機シミュレーションを行い、その性質を実験的に調べた。また、比較のため、直和型の文脈修飾を適用した場合（以下、直和型モデル）についても実験した。

実験に用いたのは、 p 個の cue パターン S^1, \dots, S^p と q 個の文脈パターン C^1, \dots, C^q のすべての組合せ（ただし $p, q \geq 2$ ）に対して、それぞれ与えられた target パターン $T^{\mu, \nu}$ ($\mu = 1, \dots, p, \nu = 1, \dots, q$) を連想するという課題である。これらのパターンは、各成分が $1/2$ の確率で 1 または -1 をとるようランダムに作成した。積型モデルの場合、 S^μ と C^ν の次元はいずれも回路網の素子数 n と等しい。一方、直和型の場合、 n 個の素子のうち cue 及び target を表現するものを n_1 個、文脈を表現するものを n_2 個としたとき、 S^μ 及び $T^{\mu, \nu}$ は n_1 次元、 C^ν は $n_2 (= n - n_1)$ 次元のパターンとなる。

まず、 $p = q$ かつ $T^{\mu, \nu}$ はすべて異なるという条件の下で、正しく連想できる p の範囲を調べた。なお、 $m = pq$ 通りすべての連合について、正しい target との類似度（パターンベクトル間の方向余弦で定義する）が 0.75 以上のパターンを想起するとき（実際にはほとんどの場合、類似度は 1 に達するか小さな値にとどまるかのどちらかである）、連想可能と判断した。また、直和型モデルでは、 n_1 と n_2 の比を変えながら実験を行い、最適と思われる値を用いた。両モデルで共通のパラメータは、以下のとおりである。

$$c = 50.0, c' = 10.0, h = 0.5, \alpha' = 2.0$$

学習の際の学習信号 r の時間長（cue から target まで変化するのに要する時間）は 5τ としたが、target パターンを学習信号とする学習を 1τ の期間追加した。学習は必要な回数だけ繰り返したが、各連合につき $10 \sim 30$ 回程度で十分であり、それ以上増やしてもほとんど変化は見られなかった。

図 4 は、3 通りの n について、学習可能な m の最大値（連想の容量）をプロットしたグラフである。積型モデル（実線）の方が直和型モデル（破線）より容量がかなり大きいことがわかる。更に重要な点は、前者では容量がほぼ n に比例して増加しているのに対し、後者では n に対する容量の比が n の増加とともに減少していることである。直和型でも、一つの S^μ が q 個の C^ν のうちごく少数による修飾しか受けない

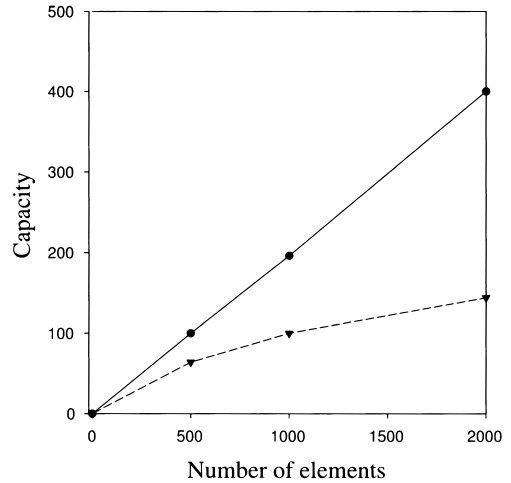


図 4 target パターンに重複のない場合の連想の容量
Fig. 4 Capacity of association when the targets have no duplication.

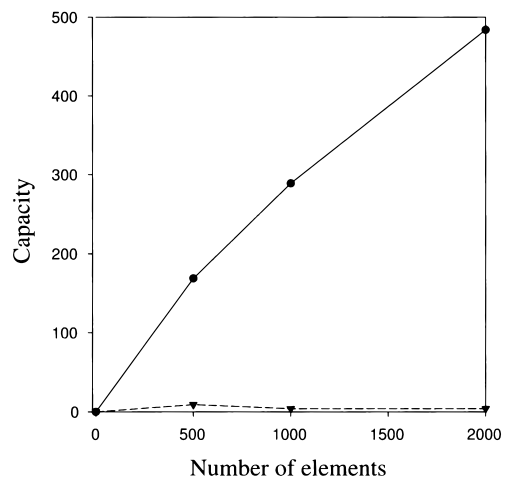


図 5 多数の target パターンが重複している場合の連想容量
Fig. 5 Capacity of association when the same pattern is repeatedly used for many targets.

という条件下では容量が n に比例して増加するから、このような効率の低下は 1 対多対応による平均化の問題が原因であると考えられる。

この傾向は、 $T^{\mu, \nu}$ に重複があって、target が S^μ や C^ν と多対多に対応するとき、より顕著に現れると考えられる。そこで、表 1 に示すように、target を p 個のパターンに限定して同じ実験を行った。表の μ 行 ν 列が $T^{\mu, \nu}$ を表し、 $\mu + \nu$ の値に応じて T^1, \dots, T^p のいずれかのパターンとなっている。各行及び列に關

表 1 平均化効果が顕著に現れる連合の例
Table 1 Example of associations causing a strong averaging effect.

	C^1	C^2	C^3	C^4	C^5	C^6	C^7	C^8	C^9	C^{10}
S^1	T^1	T^2	T^3	T^4	T^5	T^6	T^7	T^8	T^9	T^{10}
S^2	T^{10}	T^1	T^2	T^3	T^4	T^5	T^6	T^7	T^8	T^9
S^3	T^9	T^{10}	T^1	T^2	T^3	T^4	T^5	T^6	T^7	T^8
S^4	T^8	T^9	T^{10}	T^1	T^2	T^3	T^4	T^5	T^6	T^7
S^5	T^7	T^8	T^9	T^{10}	T^1	T^2	T^3	T^4	T^5	T^6
S^6	T^6	T^7	T^8	T^9	T^{10}	T^1	T^2	T^3	T^4	T^5
S^7	T^5	T^6	T^7	T^8	T^9	T^{10}	T^1	T^2	T^3	T^4
S^8	T^4	T^5	T^6	T^7	T^8	T^9	T^{10}	T^1	T^2	T^3
S^9	T^3	T^4	T^5	T^6	T^7	T^8	T^9	T^{10}	T^1	T^2
S^{10}	T^2	T^3	T^4	T^5	T^6	T^7	T^8	T^9	T^{10}	T^1

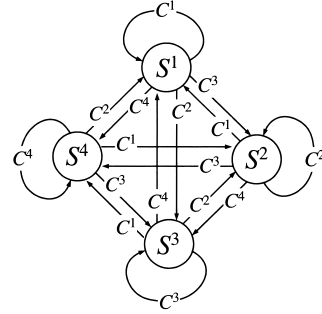


図 6 モデルの状態遷移図の例
Fig. 6 Example of the state transition diagram of the model.

る target パターンの平均はすべて等しいから、このとき前述の平均化効果が非常に強く生じるはずである。

結果を図 5 に示す。直和型モデル（破線）では、target に重複のない場合（図 4）に比べて容量が極端に小さいことがわかる。更に、 n の増加に対して容量が増えるどころか、かえって減少している。このような性能の低下は、一つには強い平均化効果によって生じるが、もう一つの理由として、図 4 の実験では target パターンが強いアトラクタであればあるほど target の連想が容易になった（各 target パターンは一つの cue 及び文脈パターンに対応するので、仮に target が非常に強いアトラクタであればそれだけで連想が可能である）のに対し、この場合一つの target が複数の S^μ 及び C^ν に対応しているため、そのような効果が生じないことが挙げられる。

一方、積型モデルでは、容量が n にほぼ比例して増加しているだけでなく、その値は図 4 よりも大きい。このことから、積型の場合 target に重複があっても全く問題なく、むしろ target パターン数の減少により容量が増えることがわかる。

以上は $p = q$ という条件での実験であるが、 $p \neq q$ の場合も同様である。様々な p と q の組合せで実験したところ、直和型では規模が大きくなったり target に重複があると性能が低下するのに対し、積型ではどんな場合でも $0.16n$ 以上の容量があった。

5. 文脈主導型連想

前節の数値実験では、target パターン $T^{\mu,\nu}$ と cue パターン S^μ とは全く別々に選んでいた。しかし、実際には両者に共通するものがあってもかまわない。特に、すべての target パターンが cue パターンのいずれかであるとき、連想した target パターンを cue と

表 2 実験に用いた状態遷移規則

Table 2 State transition rules used for the experiment.

	C^1	C^2	C^3	C^4	C^5	C^6	C^7	C^8	C^9	C^{10}
S^1	S^1	S^2	S^3	S^4	S^5	S^6	S^7	S^8	S^9	S^{10}
S^2	S^{10}	S^2	S^1	S^3	S^4	S^5	S^6	S^7	S^8	S^9
S^3	S^9	S^{10}	S^3	S^1	S^2	S^4	S^5	S^6	S^7	S^8
S^4	S^8	S^9	S^{10}	S^4	S^1	S^2	S^3	S^5	S^6	S^7
S^5	S^7	S^8	S^9	S^{10}	S^5	S^1	S^2	S^3	S^4	S^6
S^6	S^5	S^7	S^8	S^9	S^{10}	S^6	S^1	S^2	S^3	S^4
S^7	S^4	S^5	S^6	S^8	S^9	S^{10}	S^7	S^1	S^2	S^3
S^8	S^3	S^4	S^5	S^6	S^7	S^9	S^{10}	S^8	S^1	S^2
S^9	S^2	S^3	S^4	S^5	S^6	S^7	S^8	S^{10}	S^9	S^1
S^{10}	S^1	S^6	S^7	S^2	S^8	S^3	S^9	S^4	S^5	S^{10}

して次々とパターンを連想できる。このとき文脈を能動的に切り換えることによって、連想するパターンの系列を制御することも可能である。このように、cue パターン S は初期状態においてのみ与え、その後は文脈パターン C だけを入力して行う連鎖的連想のことを、文脈主導型連想と呼ぶことにする。

文脈主導型連想では、 C を回路網への制御入力とみなし、これによって回路網の状態 x が S^{μ_1} から S^{μ_2} へ遷移するという見方ができる（ S と C の関係が Elman ネットとは逆になることを注意しておく）。ここでは特に、任意の S^{μ_1} から任意の S^{μ_2} へ直接遷移できる場合を考えよう。例えば $p = 4$ のとき、最低四つの文脈パターン $C^1 \sim C^4$ があれば、図 6 のようにすべての状態遷移が実現できる。同様に、 $p = q = 10$ の場合、表 2 に例示するような連合関係があれば、任意の状態間の遷移が可能である。

5.1 実現例

実際に表 2 の連合を $n = 1000$ のモデルに学習させた。モデルのパラメータや学習方法は 4.3 とほとんど同じだが、 $r = S^\mu(I)$ の状態での学習は必要ないので

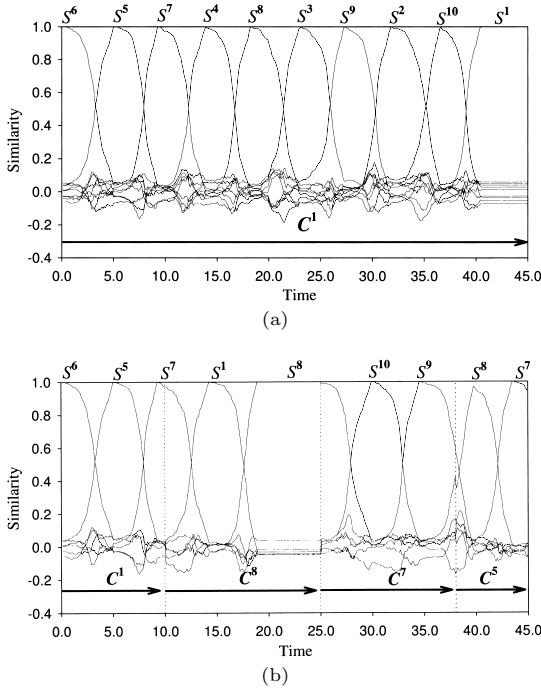


図 7 文脈主導型連想モデルの動作

Fig. 7 Behavior of the context-directed association model.

行っていない。学習回数は各連合につき 22 回である。学習後のモデルの挙動を図 7 に示す。これは、回路網の状態 x とパターン $S^1 \sim S^{10}$ との類似度

$$\frac{1}{n} \sum_{i=1}^n x_i s_i^\mu$$

の時間変化をプロットしたグラフである。

(a) は、 $x = S^6$ を初期状態とした上で、文脈パターンとして C^1 を入力し続けた場合である。 x は S^6 から S^5, S^7 などの状態を次々に経て S^1 に達し、その後は同じ状態にとどまっていることがわかるが、これは表 2 の 1 列目に示されているとおりの状態遷移である。一方、(b) は $t = 10\tau$ までは (a) と全く同じであるが、その後 C を切り換え、 $10\tau < t \leq 25\tau$ では C^8 、 $25\tau < t \leq 38\tau$ では C^7 、 $38\tau < t \leq 45\tau$ では C^5 をそれぞれ入力した場合である。この例が示すように、任意の順に S^μ を連想したり、一時的に S^μ を保持し続けたりすることが可能である。また、 $t = 38\tau$ で C を C^7 から C^5 に切り換えているが、このように状態遷移の途中で文脈パターンが変化しても、たいいていの場合問題なく連想を続けることができる。

そのほか、すべての (S^μ, C^ν) の組合せについて正しいパターンが連想されることが確かめられた。また、 $x \simeq S^\mu$ の状態であれば、 C をどのように切り換えても意図した状態遷移が生じることも確認された。このことは、適当な C の入力によって、任意の S^μ の系列を想起可能であることを意味する。

5.2 考 察

図 6 はある 4 状態 4 入力の有限オートマトンの状態遷移図にほかならない。同様に、表 2 は 10 状態 10 入力有限オートマトンの状態遷移規則の一例である。したがって、前節の回路網が表 2 のとおりに状態遷移するということは、この表で表される有限オートマトンが回路網上に実現されたといつてよい。

しかも、このオートマトンはすべての状態間を直接遷移する経路をもつから、これが実現できるならば、同じ大きさの有限オートマトンはすべて実現できるはずである。更に、 $p = q$ である必要は全くないこと、素子数を増やせばそれに比例して p や q をいくらかでも増やせることを考慮すると、このモデルは任意の p 状態 q 入力オートマトンの動作を模擬する能力をもつといえる。また、このとき必要な素子数は $n = 10m$ 程度 (m は遷移の経路数で $m \leq pq$) にすぎない。更に、軌道アトラクタの形成に必要な学習回数は n によらずほぼ一定であり、かつ原理的には全素子が完全に並列動作するから、学習に要する時間は m に比例して増加する程度である。

ただし、以上のことは本モデルが p 状態 q 入力有限オートマトンと全く等価であることを意味するものではない。まず、回路網がとり得る状態の数は p よりもはるかに多く、 S^μ やその間を結ぶ状態だけでなく、未知の入力に対してその他の状態をとることもある。また、 S^μ, C^ν ともに分散表現に基づいているため、汎化が生じることも大きな違いである。汎化にもいろいろなレベルがあるが、少なくとも C^ν に十分近いパターンを入力したとき C^ν の入力時と同じ状態遷移が行われることは、アトラクタの性質から保証される。また、いくつかの異なる文脈で同一の連合を学習したとき、学習していない全く新しい文脈下でもそれと同じ連想をすることが予備の実験によって確かめられている (もちろん別の連合を学習することも可能である)。このような複数の文脈に渡る汎化は、エキスパートネットモデルにはないものである。更に、本モデルにはより高度な汎化の能力があることを示唆する結果も得られているが、これに関する詳細な検討は今

後の課題としたい。

ところで、文脈依存的連想を行いオートマトンと同様の状態遷移が可能な神経回路モデルには、望月と大森 [11] が提案し PATON と名づけたモデルがある。このモデルは、分散表現以外にシンボルに対応する局所表現を用いている点で、分散表現の枠組みの中で可能性を広げようとする本研究とは方向性が異なるが、文脈に応じて一部の素子の出力ゲインを 0 にするという点で本モデルと類似している。しかしながら、PATON に関しては、数十個の素子からなる小規模な回路網に簡単な例題を実行させたシミュレーション結果が示されているだけで、平均化の問題などモデルの大規模化に伴う問題点が解決されているわけではない。実際、以下の理由から、PATON は大規模化が困難だと考えられる。

第 1 に、PATON では素子の平均出力は 0 ではない（出力は 0 から 1 の値をとる）にもかかわらず、ゲインが 0 の素子は出力も 0 である。これは不活性化というべきものであって、本モデルの不感化とは意味が違う。前述のように、このような不活性化では平均化の問題が完全には解消されないはずである。第 2 に、PATON における不活性化は各属性に対応する素子群の単位で行われるから、同一属性の素子は同一の文脈修飾を受ける。そのため、ある属性を表現する素子群に注目したとき、そこに不活性化されない状態で同じ特徴パターンが繰り返し現れ、それが多数の target と対応する場合に、平均化の問題が生じる。第 3 に、PATON ではいくつかの素子群を不活性化することによってエネルギー関数を動的に変化させ、あるアトラクタから目的のアトラクタへの状態遷移を実現しているが、この方法には大きな問題点がある。それは、分散表現を用いた場合、アトラクタの形成しようとした状態以外に多数のエネルギー極小点ができてしまうことである。この現象は出力関数が単調な素子を用いた場合に顕著であり、その数は素子数の指数オーダーで増加するから、素子数が数百（形成するアトラクタの数にもよる）以上になると無数のエネルギー極小状態が目的のアトラクタを取り囲むことになる。素子のゲインを変えてもこの状況は変わらないため、このときエネルギー極小状態にトラップされることなく目標状態まで遷移させるのは容易でない。

以上の議論から、本モデルと PATON モデルとは、構造や情報表現が大きく異なるだけでなく、文脈修飾の方法やダイナミックスの点でも本質的な違いがあり、

両者の類似性は表面的なものにすぎないことがわかる。このように、文脈依存的連想における平均化の問題を完全に解決し、完全な分散表現に基づいて大規模な有限オートマトンを模擬可能な神経回路モデルは、今のところ本モデル以外に知られていない。もちろんこのことは、そのようなモデルがほかに存在しないことを意味するものではないが、少なくとも本モデルより単純なものはないと思われる。

6. む す び

分散情報表現を用いた神経回路網における従来の文脈修飾方式が抱える問題点について考察し、選択的不感化による文脈修飾を非単調神経回路網に適用したモデルによってこの問題が解決されることを明らかにした。また、このモデルが多数の文脈に応じてそれぞれ多数の異なるパターンを連想できること、更に任意の有限オートマトンを模擬可能であることを示した。

以上の結果は、神経回路網による情報処理の可能性を大きく広げるとともに、分散表現のみを用いて記号処理と同等以上の文脈依存的情報処理を実現できる可能性を示唆するものであり、いわゆる symbol grounding 問題や frame 問題に起因する従来の記号処理の限界を打破する糸口を与えるかもしれない。この可能性を検証するためには、表現に何らかの構造（パターン間の類似関係）を導入してモデルの類推能力を調べることなどが必要であるが、これは今後の研究課題である。

また、本論文では脳との関連性について全く触れなかったが、筆者らは脳の神経回路においても実際に選択的不感化が用いられていると考えている。これには十分な計算論的及び生理学的な根拠があるのであるが、これについてはまた別の機会に論じたい。

謝辞 本研究は、文部科学省科学技術振興調整費目的達成型脳研究「文脈主導型、認識・判断・行動機能実現のための動的記憶システムの研究」の補助を受けて行われた。

文 献

- [1] R.A. Jacobs, M.I. Jordan, S.J. Nowlan, and G.E. Hinton, "Adaptive mixture of local experts," *Neural Computation*, vol.3, pp.79-87, 1991.
- [2] J.L. Elman, "Finding structure in time," *Cognitive Science*, vol.14, pp.179-211, 1990.
- [3] A. Cleeremans, D. Servan-Schreiber, and J.L. McClelland, "Finite state automata and simple recurrent networks," *Neural Computation*, vol.1, pp.372-381, 1989.

- [4] Y. Hirai, "A model of human associative processor (HASP)," IEEE Trans. Syst. Man. & Cybern., vol.SMC-13, no.5, pp.851-857, 1983.
- [5] 川村正樹, 岡田真人, 平井有三, "1 対多の連合を記憶する相関型連想記憶モデルの解析," 信学論 (D-II), vol.J81-D-II, no.6, pp.1336-1344, June 1998.
- [6] 森田昌彦, "非単調ダイナミクスを用いた時系列パターンの連想記憶," 信学論 (D-II), vol.J78-D-II, no.4, pp.679-688, April 1995.
- [7] M. Morita, "Memory and learning of sequential patterns by nonmonotone neural networks," Neural Networks, vol.9, no.8, pp.1477-1489, 1996.
- [8] M. Morita, "Associative memory with nonmonotone dynamics," Neural Networks, vol.6, no.1, pp.115-126, 1993.
- [9] 森田昌彦, 吉澤修治, 中野 馨, "非単調ダイナミクスを用いた構造をもつパターンの連想記憶," 信学論 (D-II), vol.J75-D-II, no.11, pp.1884-1891, Nov. 1992.
- [10] 森田昌彦, 村上 聡, "非単調神経回路網による時系列パターンの認識," 信学論 (D-II), vol.J81-D-II, no.7, pp.1679-1688, July 1998.
- [11] 望月彰子, 大森隆司, "PATON: 文脈依存性を表現する動的神経回路網モデル," 日本神経回路学会論文誌, vol.3, no.3, pp.81-89, Sept. 1996.
- [12] T. Omori, A. Mochizuki, K. Mizutani, and M. Nishizaki, "Emergence of symbolic behavior from brain like memory with dynamic attention," Neural Networks, vol.12, no.7-8, pp.1157-1172, 1999.

(平成 14 年 2 月 20 日受付, 5 月 15 日再受付)



諸上 茂光 (学生員)

平 12 筑波大・工学システム学類卒。現在、同大学院システム情報工学研究科在学中。神経回路モデルの研究に従事。



森田 昌彦 (正員)

昭 61 東大・工・計数卒。平 3 同大学院博士課程了。日本学術振興会特別研究員、東京大学工学部助手を経て、平 4 筑波大学電子・情報工学系講師。現在、同大学機能工学系助教授。生体の情報処理機構及び神経回路網の研究に従事。平 5 日本神経回路学会研究賞, 平 6 同学会論文賞, 平 11 日本心理学会研究奨励賞受賞。



松沢 浩平

平 12 筑波大・工学システム学類卒。現在、同大学院システム情報工学研究科在学中。神経回路モデルの研究に従事。