

LH_002

神経力学系間の相互作用に基づく文脈依存的時空間パターン処理

Context-Dependent Processing of Spatiotemporal Patterns Based on Interaction Between Neurodynamical Systems

蓮尾 高志†
Takashi Hasuo

山根 健†
Ken Yamane

森田 昌彦†
Masahiko Morita

1. はじめに

近年、自律ダイナミクスをもった情報処理系の重要性が指摘されている [1]. 変化が微分方程式で表される現実世界を扱うには、このような系の方が記号処理系よりも適していると考えられるからである. 本論文では、その代表的なものである、連続時間で動作する神経回路網がつくる力学系 (神経力学系) を対象とし、それによる時空間パターンの処理について論じる.

一般に神経力学系では素子間の結合荷重行列 W が力学系を定めるが、学習を終えて実際に情報処理を行う際には W を一定とすることが多い. W を変化させると、目的とする情報処理を安定的に行えないからである. しかし、このような時間的に不変な力学系では、同じ状態から異なる状態へ自律的に遷移することが不可能であるため、大きな制約がある. 例えば、同じパターンが繰り返し現れるような複雑な時空間パターンを、完全に自律的に想起しようとするならば、系の一部の素子が文脈を表現し、その活動パターンに応じて異なる状態遷移を行うようにしなければならない. しかし、通常の方法では1対多対応による平均化の問題 [2, 3] が生じるために、素子数を増やして多数の時空間パターンを扱うことはできない.

この問題への解決策として、荷重行列すなわち神経力学系を複数用意しておいて状況や文脈に応じてそれらを切り替えて用いる方法 [4] や、一つの神経回路網の素子の様々な組み合わせを切り替えて用いる方法 [2] が提案されている. しかしながら、このような切り替えは離散的な操作であり、神経力学系よりも記号処理系に適したものである. 力学系による情報処理の可能性を追究するのであれば、力学系の変化も同様の神経力学系によって行われるべきであろう.

そこで、本論文では連続的に神経力学系を変化させる新たな手法を提案し、これによって神経力学系の変化を別の神経力学系によって制御できるようにする. また、この手法を用いて二つの非単調神経回路網が相互に相手の力学系を連続的に変化させながら動作する動的連想記憶モデルを構築し、これによって複雑に絡み合った時空間パターンの認識や連想が可能であることを示す.

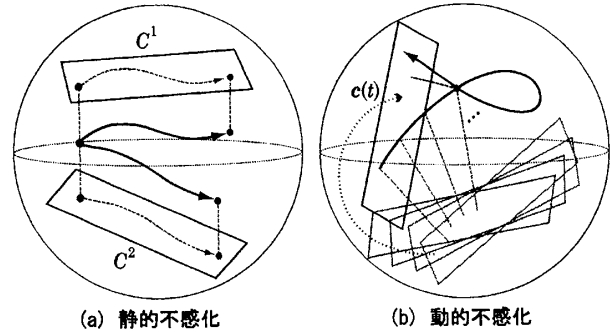


図 1: 選択的不感化による文脈依存的状態遷移の模式図

2. 原理とモデル

2.1 動的な不感化

一つの神経回路網がつくる神経力学系を離散的に変化させる仕組みとして、選択的不感化法 [2, 3] がある. これは、修飾パターンによって決まる約半数の素子を不感化する (出力を入力に関係なく中立値にする) ことによって、回路網の状態をある部分空間に射影する手法である. これにより回路網の状態はその部分空間に形成された力学系に従って動くので、修飾パターンを切り替えることによって、同一の状態から異なる状態へ遷移させることができる (図 1(a)). また、この方法を非単調神経回路網に適用することにより、任意の有限オートマトンの動作が模擬可能であることも示されている [2].

しかし、このモデルでは、ある一定の部分空間に軌道アトラクタ [5] を形成することによって安定的な状態遷移を行うため、修飾パターンは時間的に一定な空間パターンに限っていた (以下、静的な不感化と呼ぶ). これに対して本論文では、不感化する素子を連続的に入れ替える、すなわち修飾パターンを動的な時空間パターンとする手法を提案し、これを動的な不感化と呼ぶ.

動的な不感化を行った場合、図 1(b) に模式的に示すように、回路網の状態が射影される部分空間は時間と共に連続的に変化する. そのため、静的な不感化の場合のように部分空間ごとに軌道アトラクタを形成することはできないので、目的の状態に遷移させられるという保証はない. しかし、修飾パターンの変化が比較的緩やか (回路網の状態遷移と同程度かそれ以下の速さ) ならば、静的な不感化の場合と同様な学習によって各部分空間に微少な

† 筑波大学大学院システム情報工学研究科, Doctoral Program in Systems and Information Engineering, University of Tsukuba

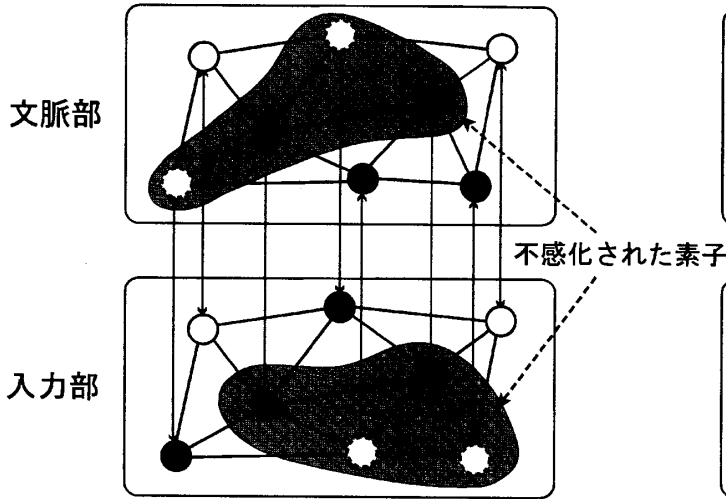


図 2: モデルの構造

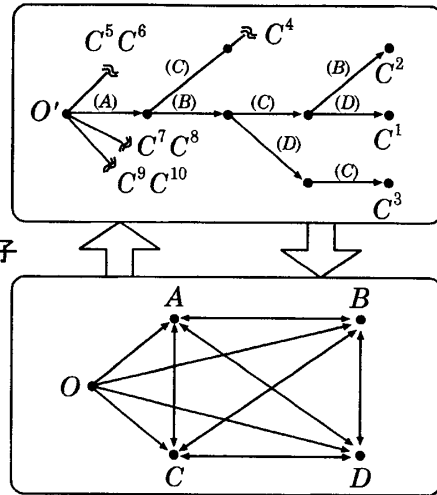


図 3: 各回路網の状態遷移図

軌道アトラクタが形成され、それらをつなぎ合わせた軌道が一つの軌道アトラクタのように振る舞う結果、目的の状態遷移が達成されると考えられる。これを計算機シミュレーションにより検証することも、本論文の目的の一つである。

2.2 モデルの構造とダイナミクス

一般に神経力学系の出力は連続的に変化する時空間パターンであるので、これを動的な不感化の際の修飾パターンとして用いることができる。このことを利用して、二つの非単調神経回路網が相互に動的な不感化することによって文脈依存の時空間パターン処理を行うモデルを構築した。

このモデルでは、図2に示すように各回路網を構成する n 個ずつの素子が1対1に対応しており、一方の内部状態に応じて他方が不感化される。二つの回路網は、構造およびダイナミクスに関しては全く同じで対称な関係であるが、一方(図の下側)の回路網は外部から入力される時空間パターンを、もう一方(上側)は入力パターンの履歴をそれぞれ扱う。前者を入力部、後者を文脈部と呼ぶ。

入力部が学習する時空間パターンは複数ある(それぞれを「系列」と呼ぶ)が、一つの系列中に同じ空間パターンが繰り返し現れたり、複数の系列が部分的に同じ時空間パターンを含んでいたりするような、複雑な系列群である(但し便宜上、初期パターンはすべて同一のパターン O とする)。一方、文脈部が学習するのは、対応する入力系列の変化に応じて同一の初期パターン O' から徐々に枝分かれしていく系列である(図3)。

いずれの回路網も、 i 番目の素子の動作は次式で表される。

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

$$y_i = g(v_i) \cdot f(u_i) \quad (2)$$

ここで、 u_i は素子の内部電位、 w_{ij} は同じ回路網の j 番目の素子からの結合荷重、 z_i は回路外部からの入力信号、 τ は時定数を表し、 $f(u_i)$ は非単調出力関数

$$f(u) = \frac{1 - e^{-cu}}{1 + e^{-cu}} \cdot \frac{1 - e^{c(|u|-h)}}{1 + e^{c(|u|-h)}} \quad (3)$$

である (c, c', h は正の定数)。また、 v_i はこの素子を修飾するもう一方の回路網の素子の内部電位であり、 $g(v)$ は0から1の値をとる単調なシグモイド関数

$$g(v) = \frac{1}{1 + e^{-cv}} \quad (4)$$

である。

式(2)および(4)が通常の非単調神経回路網のダイナミクス [5] と異なる部分である。すなわち、 v_i が正の大きな値ならば $y_i \simeq f(u_i)$ であるが、 v_i が負の大きな値のときには $y_i \simeq 0$ 、すなわち不感化される。

なお、このモデルでは内部電位 u_i の符号が重要なので、 $x_i = \text{sgn}(u_i)$ も観測できるものとし ($\text{sgn}(u)$ は $u > 0$ のとき1、 $u \leq 0$ で -1 と取る符号関数)、ベクトル $\mathbf{x} = (x_1, \dots, x_n)$ のことを回路網の状態と呼ぶ。また、 \mathbf{x} が状態空間中を連続的に遷移していくとき、その軌跡を \mathbf{x} の軌道と呼ぶ。

2.3 学習方法

入力部と文脈部が学習する系列の組を $\mathbf{s}(t)$ と $\mathbf{c}(t)$ としよう。これらはいずれも n 次元ベクトルであり、その成分 $s_i(t)$ および $c_i(t)$ は約半数が1、残りが -1 と取るものとする。これらを動的な不感化された入力部および文脈部がつくる力学系の軌道アトラクタにすることが学習の目的である。その具体的な手順は以下の通りである。

まず、入力部と文脈部の初期状態を $\mathbf{s}(0)$ および $\mathbf{c}(0)$ に設定する。各素子への外部入力 z_i として、入力部には

$\lambda s_i(t)$, 文脈部には $\lambda c_i(t)$ を与えつつ, 式(1)~(4)で表されるダイナミクスに従って回路網を動作させる. ここで λ は学習するパターンの入力強度を表す定数である.

これと並行して, 入力部の素子間の結合荷重を

$$\tau' \frac{dw_{ij}}{dt} = -w_{ij} + \alpha s_i(t)y_j \quad (5)$$

に従って, 文脈部の素子間の結合荷重を

$$\tau' \frac{dw_{ij}}{dt} = -w_{ij} + \alpha c_i(t)y_j \quad (6)$$

に従って, それぞれ更新する. ここで τ' は学習の時定数で, $\tau' \gg \tau$ である. また α は学習係数であるが, 学習性能を向上させるため [5], 定数ではなく $\alpha = \alpha' x_i y_i$ とする (α' は正の定数).

以上をすべての系列の組について行くと, 1回の学習が終わる. これを λ を徐々に減らしながら繰り返し, $\lambda = 0$ にしても各回路網の状態が $s(t)$ および $c(t)$ に追従するようになれば, 学習の完了である.

なお, 上記の過程で入力部および文脈部においてそれぞれ $x \simeq s(t)$ および $x \simeq c(t)$ が成り立つので, 式(2)の $g(v_i)$ の代わりに $(c_i(t) + 1)/2$ (入力部の場合) または $(s_i(t) + 1)/2$ (文脈部の場合) を用いても, ほぼ同じ結果が得られる. こうすると二つの回路網を別々に学習することができるので, 計算機シミュレーションが容易になる.

3. 計算機シミュレーション

各回路網が $n = 800$ 個の素子で構成されるモデルを用いて, 計算機シミュレーションを行った.

入力部に学習させたパターン系列を表1に示す. 時間長はすべて $T = 20\tau$ とした. それぞれのアルファベットの経路点であり, いずれもランダムに選んだ800次元の2値ベクトルである. 各経路点の間では, 成分が1つずつ一定の順で変化する. また, 文脈部が学習する系列 c^1, \dots, c^{10} は, 図3に示すように, O' を始点とし, 3つの経路点を通して終点 C^1, \dots, C^{10} に達する, 時間長 $T = 20\tau$ の時空間パターンである. これらの経路点および終点も, すべてランダムに選んだ.

各パラメータは $c = 50$, $c' = 10$, $h = 0.5$, $\tau' = 5000\tau$, $\alpha' = 2$ であり, λ は学習回数が増えるにつれ0.2から0まで徐々に減少させた. 各系列の学習回数は30回である.

学習後のモデルに, 様々な時空間パターンを入力し (入力強度 $\lambda = 0.4$), その挙動を調べた. 動作例を図4に示す. 図の各グラフは, 下から順に入力パターン, 入力部の状態, 文脈部の状態を表しており, 縦軸はそれらと経路点等との類似度 (パターンベクトル間の方向余弦で定義する), 横軸は時定数 τ を単位とする時間である. なお, 文脈部に関しては, 最も近い経路点または終点との類似度のみを表示している.

表 1: 入力部が学習する時空間パターン

$s^1 = \{OABCD\}_T$	$s^2 = \{OABCB\}_T$
$s^3 = \{OABDC\}_T$	$s^4 = \{OACBD\}_T$
$s^5 = \{OBCBD\}_T$	$s^6 = \{OBDAC\}_T$
$s^7 = \{OCADB\}_T$	$s^8 = \{OCADA\}_T$
$s^9 = \{ODADC\}_T$	$s^{10} = \{ODCDA\}_T$

(a)は, 系列 s^2 に対して, 10%の成分を反転した上で, 一部の区間で時間伸縮を施した時空間パターンを入力した場合である. 入力部の状態がほぼ入力パターンからノイズを除いたものになっていること, 文脈部の状態はほぼ c^2 の軌道に沿って遷移し, 最終的に C^2 に到達していることがわかる. このことから, 入力した時空間パターンは2番目の系列として認識されたと言えよう. 同様の実験をすべての系列について行ったところ, すべて正しく認識されることが確認された.

次に, (b)は系列の一部を断続的に入力した場合である. 最初, O から D に変化する途中のパターン D' まで入力した後, しばらく何も入力しないと, 入力部の状態は D の近くまで自律的に遷移した後, A および C にわずかに寄った地点で停留した. また, このとき文脈部の状態も最初の経路点付近で停留した. これは, 複数の軌道アトラクタが拮抗した状態と考えられる.

その後, D' から A の方向へ変化するパターンを短時間に入力したところ, 入力部は $A \rightarrow D \rightarrow C$ と状態遷移し, 同時に文脈部は c^9 の軌道に沿って C^9 まで状態遷移した. このことは, モデルがこの入力パターンを9番目の系列と認識したことを意味するが, 「ある系列を認識するとは, その系列の続きを自律的に再生できることである」と考えるならば, 文脈部の状態が C^9 に達したときではなく, 最後の分岐を過ぎた時点 ($t = 15\tau$ 付近) で認識したと見なせる.

ここで, 入力部の状態が s^{10} の軌道から離れてしばらくしてから, 文脈部の状態が c^{10} の軌道から離れていることに注意されたい. これは, 入力部による動的不感化によって, 文脈部の状態が c^9 に対応する軌道アトラクタを含む部分空間に射影される結果, 文脈部が軌道アトラクタ間の拮抗状態から抜け出すことを反映していると考えられる. 逆に, 入力部が A から D へ状態遷移する際には, 文脈部の状態が追いつくまで入力部が少し待っており, 文脈部による動的不感化が入力部にとっても重要な働きをしていることがわかる.

以上の結果から, 動的不感化がうまく機能することが確認されると共に, 本モデルにおいて二つの回路網が自律的かつ協調的に動作し, 複雑な時空間パターンの認識および想起が可能であることが示された.

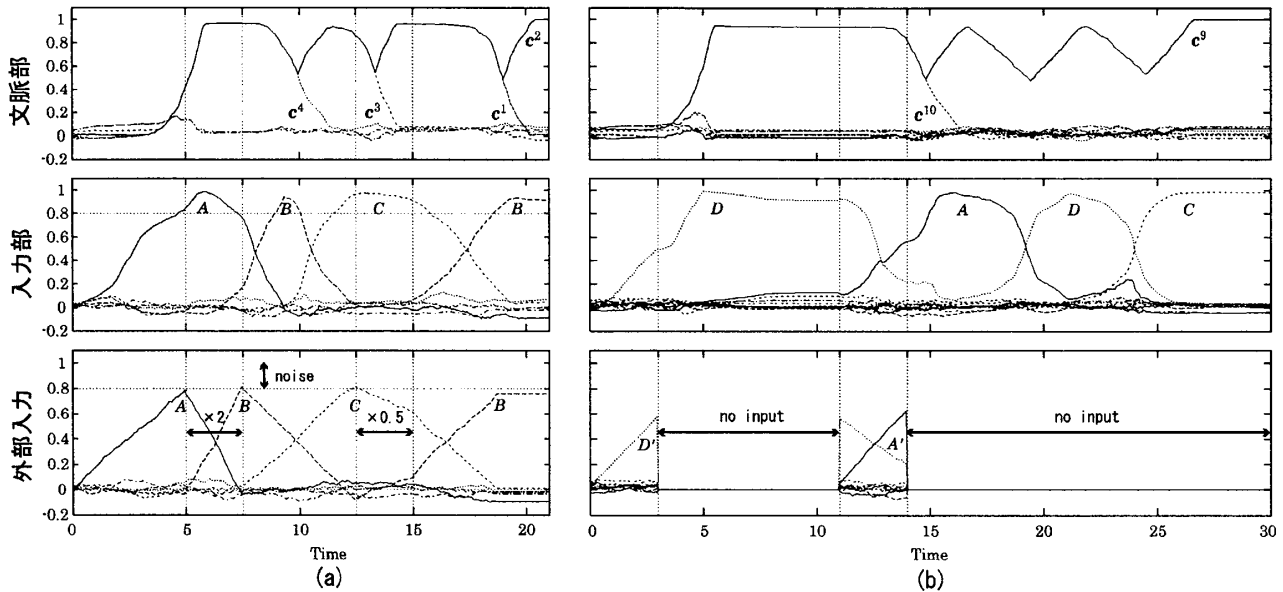


図4: 複雑な時空間パターンの認識と想起の例

4. おわりに

動的不感化によって神経力学系を連続的に変化させる方法を提案し、これを用いて神経力学系同士が相互作用しながら文脈依存的な時空間パターン処理を行うモデルを構築した。このモデルは、複雑な時空間パターンを認識できるだけでなく、ノイズや時間方向の伸縮に強い、系列の途中まで入力しただけでそれ以降を想起できる、といった優れた能力をもつ。

本モデルを構成する二つの非単調神経回路網は、全く同じ構造とダイナミクスをもつ、独立に学習できる、両者を結ぶのは素子数と同数の固定結合のみであるなど、モジュール性が高い。更に重要なのは、二つの回路網を従来の方法で連結すると、ダイナミクスはそのままに素子数が2倍の神経力学系ができるだけであるのに対し、両者を動的に不感化によって結合するのはそれとは全く異なるということである。そのため、単一の神経力学系ではできない処理が可能であり、また多数の回路網を同様に結合すれば、より複雑で高度な情報処理が実現できると考えられる。このように、本研究の結果は、神経力学系による情報処理、ひいては自律ダイナミクスをもつ情報処理系全般の可能性を拓けるものと言えよう。

今後、回路網の数を増やすなど本モデルを発展させると共に、これを動画像の認識やロボットの行動制御などへ応用したいと考えている。そのためには、モデル自体の研究以外に、現実世界の情報を本モデルで扱うのに適した時空間パターンに変換する方法についても検討が必要である。

謝辞

本研究の一部は、文部科学省科学研究費補助金基盤研究(B)(No.15300068)および特定領域研究「情報爆発」(No.18049008)の補助を受けた。

参考文献

- [1] 松山隆司, 杉本晃宏, 佐藤洋一, 川嶋宏彰, “人間と共生する情報システムの実現を目指して”, 人工知能学会誌, vol.19, no.2, pp.257-266, 2004.
- [2] 森田昌彦, 松沢浩平, 諸上茂光, “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” 信学論(D-II), vol.J85-D-II, no.10, pp.1602-1612, October 2002.
- [3] 森田昌彦, 村田和彦, 諸上茂光, 末光厚夫, “選択的不感化法を適用した層状ニューラルネットの情報統合能力” 信学論(D-II), vol.J87-D-II, no.12, pp.2242-2252, December 2004.
- [4] J. Tani and S. Nolfi, “Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems”. *Neural Networks*, vol.12, no.7-8, pp.1131-1141, 1999.
- [5] M. Morita, “Memory and learning of sequential patterns by nonmonotone neural networks”. *Neural Networks*, vol.9, no.1, pp.1477-1489, 1996.