

## カオスニューラルネットワークを用いた記号生成

真隅 暁, 橋本 敬

北陸先端科学技術大学院大学 知識科学研究科

抄録 本研究は, 記号生成の数理モデルを構築することを目的としている. 本論ではまず, 力学系の枠組みに基づいて記号の性質を議論する. 力学系の観点からは, アトラクタを記号として, 外部入力に応じたアトラクタ間の規則的な遷移を記号の操作と捉えられる. よってアトラクタ間の遷移が見られる力学系を用いれば, 記号の働きを表現できる可能性がある. 我々は, カオス的な振る舞いを見せる写像である NZ マップによって構成されたカオスニューラルネットワークを用いてそのような系の構築を試みる. 本論では, このシステムがマルチアトラクタ系であり, アトラクタ間の自律的な遷移が見られることを確認した. さらに系に外部入力を加えることにより, 外部入力に応じたアトラクタ間の遷移が生じ得ることを数値計算によって示した.

## Symbol formation by using a chaotic neural network

Akira Masumi, Takashi Hashimoto

School of Knowledge Science, Japan Advanced Institute of Science and Technology

**Abstract** We aim at constructing a mathematical model of symbol formation. In this paper, we discuss a concept of symbol formation on the basis of dynamical systems. From the viewpoint of dynamical systems, a symbol can be regarded as an attractor, the symbol manipulation as orderly transition among attractors. Thus, by using a dynamical system showing transition among attractors, we may be able to represent symbols. We try to construct such a system with a chaotic neural network that is a coupled NZ maps. We confirm that this model is a multi-attractor system and shows existence of transition among attractors. Further, we exemplify the transition induced by a certain external input.

### 1 はじめに

言語の使用や計算, 論理的思考などの基礎となる能力の一つとして, 記号処理能力がある. 例えば言語における記号処理とは, 文字や音声などの記号をある規則 (文法) に従って運用することである.

記号処理には次の二つの段階があると考えられる. 第一に, 外界のなんらかの対象を指し示す表象を記憶として形成する段階, 第二に, 形成された記憶をなんらかの規則に従って操作する段階である. 計算機科学などでは, 記号の操作は有限状態オートマトンなどによって記述される. すなわち, 状態を記号, 状態遷移規則を記号操作規則としたとき, 外部入力に対する状態遷移を記号の操作とする.

現在まで, 記号処理のメカニズムに関する数理的研究は, 主に人工知能やコネクショニズムなどの枠組みで行われて来た. しかしこれらの枠組みは, 前者は記号や規則が設計者によって一方的に与えられるという難点があり, 後者は記号や規則を明示的

に記述しにくいという難点がある. 従って, これらに代わる手法が必要であると考えられている [1].

近年, 脳の動的な側面を力学系の枠組で記述しようという取り組みが盛んである [2]. 特に記憶のメカニズムに関しては, 力学系におけるアトラクタを記憶とし, 記憶の想起や連想といった働きを, アトラクタへの収束やアトラクタ痕跡<sup>1</sup>の間の遷移として記述しようとしている. 例えば, 合原らは, カオス力学系により構成したニューラルネットワークモデルを用いて連想記憶のモデルを構築している [3].

以上のような力学系の性質を考慮すると, アトラクタを記号, アトラクタ間の遷移を記号の操作, アトラクタ間の遷移規則を記号の操作規則と考えることによって, 可塑的に学習が可能であり, かつアトラクタとその間の関係として記号と操作規則を表現可能な系を構築できる可能性がある.

本論ではこのような背景より, 力学系による記号

<sup>1</sup>アトラクタ痕跡とは, アトラクタのように, ある期間吸引的性質を持つが, やがてはそこから抜け出すことができるような状態空間上のある領域のことを言う.

処理の数理モデルの構築を試みる．具体的には，学習を通じてアトラクタを形成し，外部入力に対してアトラクタ間をある程度規則的に遷移するシステムの構築である．そのためには系が複数のアトラクタを持ち，その間の遷移が生じ得ることが必要である．ここでは，それが可能な系としてNZ マップの結合系を用いる．

## 2 モデル

### 2.1 NZ マップ

本研究では，上記の記号の性質を表現するためのモデルとしてNZ マップの結合系 [4] を用いる．単一のNZ マップは，連続型のホップフィールドモデルに負の自己フィードバック結合を加え，それをオイラー法によって離散化したもので，あるパラメータ領域でカオス的な振る舞いを見せることが知られている．これを相互全結合したNZ マップの結合系は，脳におけるモジュール構造をモデル化したものであり，複数のアトラクタ間を乱雑に遷移する現象を生じることが知られている [4]．

NZ マップは以下の式で与えられる．

$$p_i(n+1) = F_{q_i(n)}\{p_i(n)\} \quad (1)$$

$$q_i(n) = -\frac{1}{T_{ii}} \left\{ \sum_{j \neq i}^N T_{ij} p_j(n) + I_i \right\} \quad (2)$$

$$F_q(p) = rp + (1-r) \left[ 1 - \frac{1}{2} \left\{ 1 + \tanh \left( \frac{p-q}{2\beta} \right) \right\} \right] \quad (3)$$

ここで，式 (1) は，写像  $F_{q_i(n)}$  による  $p_i(n)$  の時間発展を表している． $p_i(n)$ ， $q_i(n)$  はそれぞれ，時刻  $n$  における  $i$  番目の要素の内部バツファと，他の要素との結合の影響を表すパラメータで，式 (2) は  $i$  番目の要素に対する他のニューロンからの結合の影響を表している． $N$  は要素の数， $T_{ij}$  は  $i$  番目と  $j$  番目の要素間の結合係数， $T_{ii}$  は自己フィードバック結合の係数， $I_i$  は閾値を表す．写像は式 (3) で定義される．

NZ マップは，パラメータ  $q_i(n)$  の変化によって写像の形が時間と共に動的に変化する．このような，写像の動的な振る舞いに起因してNZ マップの結合系は，アトラクタ間の遷移という動的な振る舞い

を見せる．本論では以下， $I_i/T_{ii} = 0.09$ ， $r = 0.7$ ， $\beta = 0.006$  とする．

### 2.2 アトラクタの埋め込み

本論で用いるのは，上述のNZ マップを相互に全結合したモデルである．そこで要素間の結合係数  $T_{ij}$  を以下の規則に従って決定する．

$$T_{ij} = \sum_s (2V_i^s - 1)(2V_j^s - 1) \quad (4)$$

ここで， $V^s = \{V_1^s, \dots, V_N^s\}$  は要素の数  $N$  を次元とするベクトルである．式 (4) に従って結合係数を決定すると，その系のアトラクタとしてベクトルパターン  $V^s$  を記録することができる [5]．本論では  $N = 16$ ， $s = 1, \dots, 3$  として，互いに直交するパターン  $(C, F, A)$  を埋め込んだ．

### 2.3 想起パターンの定義

式 (4) によって系に記録されたパターン  $V^s$  は，0, 1 の要素からなる 16 次元のベクトルである．これに対し，想起パターン  $\phi(n) = \{\phi_1(n), \dots, \phi_N(n)\}$  を以下のように定義する．

$$\phi_i(n) = \begin{cases} 1 & (q_i(n) \geq \bar{q}) \\ 0 & (q_i(n) < \bar{q}) \end{cases} \quad (5)$$

ここで，

$$\bar{q} = \lim_{n \rightarrow \infty} \frac{1}{nN} \sum_{k=0}^{n-1} \sum_{i=1}^N q_i(k) \quad (6)$$

は  $q_i(n)$  の平均値を表す [4]．

## 3 アトラクタへの収束と遷移

§1 で述べたように，記号のモデルにはアトラクタ間の遷移を生じることが必要である．NZ マップの結合系の場合，それは自己フィードバック結合  $T_{ii}$  をパラメータとして変化させることによって見ることができる．また本論では，記号の操作は外部入力に対するアトラクタ間の規則的な遷移と位置付けられる．以下では， $T_{ii}$  の変化に対する系の振る舞いと，外部入力に応じたアトラクタ間の遷移という点に関して行った数値計算の結果について述べる．

### 3.1 想起パターンの時系列

図1の二つの図は  $T_{ii} = 13.0$  とし、異なる初期値を与えた場合の  $\phi(n)$  の時系列である。トランジェントを経て、ある記銘パターンへと収束するという振る舞いが見られる。そして初期値が異なると違うパターンへ収束することから、マルチアトラクタ系になっていることが確かめられる。一方、図2は、 $T_{ii} = 15.0$  とした場合の  $\phi(n)$  の時系列である。ここでは、記銘したパターン、つまりアトラクタ間を次々に遷移するという振る舞いが見られる。

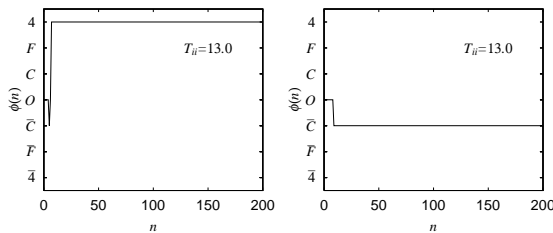


図1: 想起パターン  $\phi(n)$  の時系列。横軸は時間、縦軸は想起パターン。C-bar F-bar 4-barはそれぞれのパターンの反転を表し、O は記銘していないパターンを指す。左右の図はそれぞれ初期値が異なる。

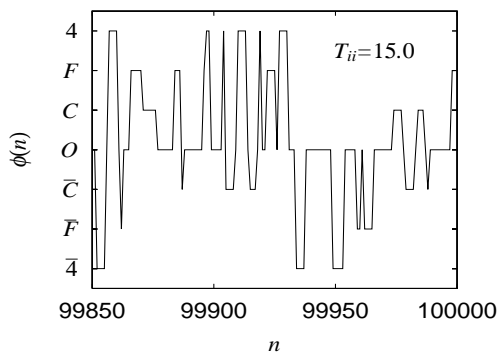


図2: 想起パターン  $\phi(n)$  の時系列。

### 3.2 アトラクタと軌道との距離

時刻  $n$  における系の軌道とアトラクタとして埋め込んだ記銘パターン  $V^{s'}$  との間の距離  $Dist^{s'}(n)$  を以下のように定義する<sup>2</sup>。

$$Dist^{s'}(n) = \sqrt{\sum_{i=1}^N (V_i^{s'} - q_i(n))^2} \quad (7)$$

図1, 2の想起パターンの時系列は、系の軌道を、埋め込んだパターンに至っているか否かという観点で

<sup>2</sup>ここで  $s'$  は、反転も含めた場合の記銘パターンについての添字であり、 $s' = 1, \dots, 6$  である。

粗視化しているのに、埋め込んでいないパターンOに至っているときには具体的にどのような状態にあるのかが不明である。そこで、アトラクタと軌道の距離  $Dist^{s'}(n)$  の時系列を見ることによって、より詳しく系の状態を知ることができる。

### 3.3 外部入力

記号の操作は、本論においては外部入力に応じたアトラクタ間の規則的な遷移と位置付けられる。そこで、まずは系に対して外部入力を与えたときの系の振る舞いを見る必要がある。

考えられる入力の与え方は様々である。ここでは最も単純なものとして定数入力と、正弦波による周期的入力を与え、それに対する系の振る舞いを見る。外部入力  $\delta$  は、式(2)において

$$q_i(n) = -\frac{1}{T_{ii}} \left\{ \sum_{j \neq i}^N T_{ij} p_j(n) + \delta(n) + I_i \right\} \quad (8)$$

のように与えた。

まず、定数入力として、 $\delta(n) = 0.3$  を  $n = 10000 \sim 20000$  の期間与えた。このとき図3に示すように、軌道があるアトラクタ付近に固定されることがある。また、どのアトラクタにも固定されない場合も存在する。これは今、 $T_{ii} = 15.0$  とし、図2に見られるような状態にある系に対して入力を与えているので、入力し始めた時点 ( $n = 10000$ ) での系の状態は初期値によって異なる。これにより、上記の違いが生じる。

次に、周期的な入力として

$$\delta(n) = A \sin(2\pi\omega n) \quad (9)$$

を  $n = 10000 \sim 20000$  の期間与えた ( $A = 0.7$ ,  $\omega = 0.001$ )。この場合、図4に見られるように、正弦波の周期性によってアトラクタ間の遷移が誘発される。それぞれのアトラクタにおける滞在時間は、周期  $1/\omega$  に応じて変化する。

定数入力を与えた場合のアトラクタ付近への軌道の固定は、図3に示すように  $q_i(n)$  がクラスターを形成し、それによって式(5)で定義した  $\phi_i(n)$  が時間的に変化しないことによって生じる。これは、どのアトラクタに近接した場合でも成り立つ性質であり、定数入力を与えた際に常に見られる。一方、正

弦波を入力した場合, 正弦波の振幅がゼロ付近で系をトランジェント状態に至らせる. その後, 振幅が増加するにつれてその時の状態によってあるアトラクタへ遷移し, それが  $1/\omega$  の期間持続する.

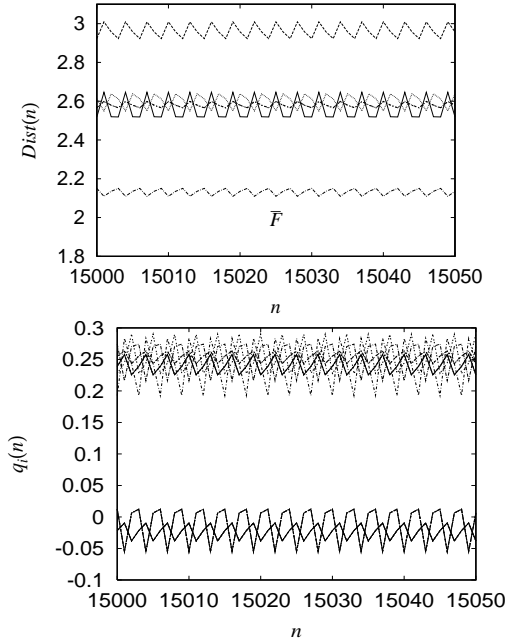


図 3:  $n = 15000 \sim 15050$  における  $Dist^{s'}(n)$  の時系列 (上段) とそのときの  $q_i(n)$  の時系列の一部 (下段).  $T_{ii} = 15.0$ .  $Dist^{s'}(n)$  については反転も含め 6 個のパターン,  $q_i(n)$  については 16 個の要素を重ね書きしている. 横軸は時間を表す. 上段の図中の記号は, この時に想起しているパターンを示す.

#### 4 おわりに

本論ではまず, 記号には表象形成の段階と, 操作対象としての段階があることを論じた. 次にこれを数理的に表現するため, 力学系におけるアトラクタと, その間の遷移という性質が適当であることを論じた. そして, NZ マップの結合系にアトラクタ間の遷移が見られることを確認し, 外部入力に応じて, アトラクタ間の遷移が生じ得ることを数値計算によって示した. これより NZ マップの結合系は, 記号の性質を表現するモデルとして有効であると言えるだろう. NZ マップの結合系は, 大域結合写像 (GCM) の一種と捉えることができる. GCM に関しては, 系に摂動を加えたときのアトラクタ間の遷移が調べられている [6]. 今後は, これらを参考にアトラクタ間の遷移に関して調べるとともに,

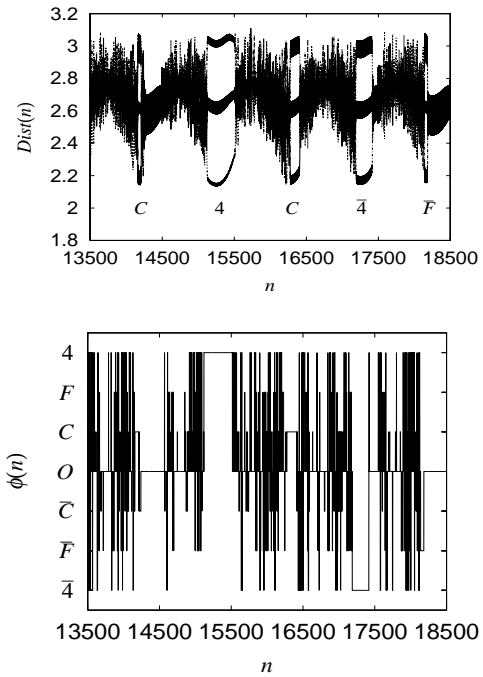


図 4:  $n = 13500 \sim 18500$  における  $Dist^{s'}(n)$  の時系列 (上段) と, そのときの想起パターンの時系列 (下段).  $T_{ii} = 15.0$ . 横軸, 縦軸, 図中の記号は既出の図と同様.

アトラクタをトップダウン的に埋め込む以外のやり方で形成することが課題である.

#### 参考文献

- [1] Harnad, S. The symbol grounding problem. *Physica D* **42** 335-346 (1990).
- [2] 津田 一郎. 複雑系脳理論「動的脳観」による脳の理解.(サイエンス社) (2002).
- [3] Adachi, K., and Aihara, K. Associative dynamics in a chaotic neural network. *Neural Networks* **10** 83-98 (1997).
- [4] Nozawa, H. A neural network model as a globally coupled map and applications based on chaos. *CHAOS* **2** (3) 377-386 (1992).
- [5] Hopfield, J.J. Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. USA* **81** 3088-3092 (1984).
- [6] Kaneko, K. Dominance of Milnor Attractor and Noise-Induced Selection System. *Physical Review Letters* **78** 14 (1997).