物性研究　93－1（2009－10）

修士論文（2008 年度）

多時間スケールをもつ学習過程における相空間ダイナミクス

東京大学 総合文化研究科広域科学専攻相関基礎科学系 金子研究室
栗川 知己
E-mail:kurikawa@complex.c.u-tokyo.ac.jp

目 次

1 序論

1.1 モチベーション

脳を理解するとはどのように理解されるべきだろうか?

この問いを考えるとき、外部世界をどのように解釈するかという問題は避けられないように思われる。この問題は言い替えると系の外部からの入力がどのように系の内部に埋め込めこまれるか、また埋め込まれたときに周りとの関係がどうなるかという問題である。例えば Geroge Lakoff [9] が述べているように、認知言語学では人間の思考、行動にとってカテゴリー化が極めて重要であり、概念のカテゴリ化は人間の身体性から切り離せないとして捉える。外部の対象が知覚を通して内部表象をもつとき、その関係は恣意的なものではなく身体性にもとづき何らかの意味をもつ関係であると理解するのである。カテゴリー化の問題は外部からの入力を系が内部に持っている他の構造とどのように関係づけるかという問題である。それでは関係づけられるとはどういうことか。これは入力がない状態でのダイナミクスの性質、相空間の構造と入力を印加したときの構造の変化の対応として捉えられる。

ここでは以下のように考えたい。

脳は大自由度力学系とみなすことができる。この力学系は入力と出力をもつ。そして状態変数の挙動を規定する相空間が状態変数の挙動で変わる力学系である。変数を神経細胞だとすると神経細胞のダイナミクスはシナプスの強度分布によって決まる。一方でシナプスの強度分布は神経細胞の活動によって変化する。さらに状態変数 $\leftrightarrow$ 相空間のダイナミクスに細胞の挙動を通して外部からの刺激が埋め込まれる。

イメージとしては以下のようになる。相空間を凸凹したウレタンのスポンジだと思ってその上を変数であるボールがころがる。ボールには糸がつながっていてそれを引っ張ってボールを動かせる。ただしボールはスポンジの凸凹に引っかかったり落ち込んだりするので、スポンジがどんな形をしているによって同じように糸を引いてもボールの動きは異なる。スポンジはスポンジでボールがどこに有るかによってへコんだり、復元したりでボールの動き次第で形が変わる。

このような力学系として脳を見なしたときに冒頭の問題は相空間と状態の挙動がお互いに制御する力学系にどのように入力が埋め込まれるか、どのように以前の構造と関係づけられるかという問題に言い替えられる。

この問題が本論文の大きなテーマである。

1.2 問題意識

以上のモチベーションに基づき, 本論文ではニューラルネットワークを用いた入出力関係の学習モデルを構成し、入力が相空間にどのように埋め込まれるかを議論する。そのため既存のニューラルネットワークのモデルの生物学的な観点からの問題点を指摘し、どのようなモデルを構成するべきかを明確にする。ニューラルネットワークには主に 3 つの枠組があり、それぞれ教師付き学習、強化学習、教師なし学習 [12] [14] と呼ばれる。ここではそのなかでも今問題にしている入出力関係の学習を行う枠組を中心に問題を提起する。

まず入出力関係の学習で主に用いられる教師付き学習について述べる。教師付き学習では学習により出力させたいパタンが明示的に与えられる。このパタンの事を教師信号と呼ぶ。その情報を元にシナプスを変化させることで入力とそれに対応する望ましいパタンを実際に出力するネットワークを得ることができる。ここでは最もオーソドックスな最急降下法と呼ばれるアルゴリズムを例に問題点を述べる。ニューラルネットワークの学習ではいかにシナプスを変えるかが重要な点になる。最急降下法は入力 I, シナプス $\{J_{ij}\}$(ただし j から i への結合), 細胞 $\{x_i\}$ できまる系の出力 $F(I,\{x_i\},\{J_{ij}\})$ と教師信号 $T(I)$ との誤差 $E=|T(I)-F|^2/2$ に対してシナプス J_{ij} を $\Delta J_{ij}=-\eta(\partial E/\partial J_{ij})$ という形で変化させる。これにより誤差 E を小さくするようにシナプスを変化させることができ、誤差を 0 にするまでシナプスを変化させることでネットワークの出力が教師信号と同じパタンにする方法である。良く用いられるエラーバックプロパゲー

ション法 (BP 法)[4] は複数の層があるネットワークにおいてこの ΔJ_{ij} を計算する方法である。BP 法の場合は $\Delta J_{ij} = \eta \delta_i x_j$ で書かれる。ここで x_j は細胞 j から細胞 i への入力である。細胞 i は i への入力を集め活性化関数 f(x) を通して出力する ($x_i = f(\sum_j J_{ij} x_j)$)。$\delta_i$ は出力層細胞の場合は $\delta_i = e_i f'(\sum_j J_{ij} x_j)$ で求まり ($e_i = |x_i - t_i|$)、隠れ層細胞の場合は $\delta_i = \sum_k \delta_k J_{ki}$ で求まる。隠れ層の δ_i を求めるためには、その下流の層の δ_k とその間のシナプスの値を用いる必要がある。つまり ΔJ_{ij} を計算する場合は δ_i まで 各要素の δ を出力層の誤差から出力層から入力層の方向に計算しなければならない。このように誤差情報を逆向きに計算して行くアルゴリズムなのでエラーバックプロパゲーションという名前がついている。このとき δ_i は各要素毎の情報なので学習に必要な情報は出力層の細胞各々の誤差になる。ここに教師付き学習の枠組の問題点が存在する。すなわち、学習に必要な情報量が多いという点である。今見たように教師付き学習は出力層で各要素の誤差が必要なので教師信号の次元は出力層の要素数と同じになる。これは生物的観点からみると出力細胞 (例えば筋肉細胞) にたいして個別に正しい情報が与えられるということであり妥当ではない。生物の学習としてはむしろ全体の振る舞いにたいしてそれがよいか悪いかの評価が環境から与えられ、それをもとに学習が為されるほうが自然である。

この問題点を解決したものが次に述べる強化学習の枠組である。強化学習の枠組では、系の出力を良いか・悪いかという 1 次元情報 R(強化学習のスキームでは報酬と呼ばれる) をもとに学習をおこなう枠組である [10]。入出力関係の学習を行うモデル Associative Reinforcement Learning[17] [18](あるいは associate reward-penalty learning) と呼ばれるアルゴリズムがある。このアルゴリズムは教師付き学習での誤差 E のかわりに報酬 R の時間平均を最大化することで学習を達成する。この学習ではシナプスは R と presynaptic cell と postsynaptic cell の三者の相関により変化するので、必要な情報は系の出力が正しいかどうかという 1 次元情報であり上の必要な情報量の多さという点は問題にならない。しかしこのアルゴリズムにも問題点はある。収束性などの問題点がよく指摘されているが、ここでは以下の点に着目する。それは複数のパタンの学習を行うときの逐次性である (実は上で述べた教師付き学習にも当てはまることである)。ここでいう学習の逐次性について説明する。これらのアルゴリズムで複数の学習を行う時に一般に用いられるのは最急降下法とよばれるアルゴリズムである。最急降下法を用いてパタン A、B の二つを学習させることを考える。まずパタン A の誤差 E_A を減らすようにシナプスを $\Delta J_{ij} = -\eta(\partial E_A/\partial J_{ij})$ という形で変化させる。この時点では一度の学習では $E_A = 0$ になっておらず、学習は完了していない。次にパタン B の誤差 E_B の誤差を減らすようにシナプスを $\Delta J_{ij} = -\eta(\partial E_B/\partial J_{ij})$ にしたがい変化させる。先ほどと同様にこの時点では $E_B = 0$ になっておらず、また E_A も 0 にはなっていない。したがってまたパタン A の学習を行いシナプスを E_A を減らすように変化させ、次にパタン B を変化させ ･･･ というプロセスを繰り返すことで最終的に $E_A = E_B = 0$ になるまで[1]学習を続ける。このように、複数のパタンを学習するときには何回も繰り返し学習する必要がある。しかし実際に生物が学習を行う局面ではパタン A を学習した後に ($E_A = 0$ にした後に)、次のパタン B を学習するというように逐次的な学習の方が妥当であると考えられる。ここではこのような最急降下法の問題点を学習の非逐次性と呼ぶ。

また強化学習の他のモデルとして negative FB と winner takes all (WTA) による adaptive learning [1] がある。このモデルでは間違った出力を出したときに使われた結合は弱くする negative FB とスパースなパタンにするための WTA を導入することで学習を行うというものである。このモデルも強化学習の一種であり学習に必要な情報は出力が正しいか否かという 1 次元情報のみである。しかし離散時間でフィードフォワード結合のみの層型ネットワークでのモデルであり、力学系として考えることが難しい。この点において上記のモチベーションとは合わない。

最後に教師なし学習に関してであるが、これは入出力関係の学習というより入力をどのように埋め込むかの枠組みであるので 6 章で本論文のモデルでの埋め込まれ方との違いを議論することにする。

ここまででニューラルネットワークの既存の枠組での 2 つの問題点を指摘した。したがって本論文では

1. 出力が良いか悪いかという 1 次元情報量を用いること。

[1]オンラインラーニングでは正確に全てが 0 にはならないが、十分ちいさくなるまで学習を続ける

2. 逐次的に学習を行うこと。

という要請を行い、これを満たすモデルを構成しそのときの力学系としての性質を議論する。

1.3 構成

本論文では前節での要請を満たすモデルに入出力関係を学習させ、それによる相空間の構造の変化を解析する。

2章では本論文で用いるモデルの詳しい説明を行う。

3章では2章での説明したモデルの挙動などを解析し、実際に要請をみたすような学習ができることを示す。

3.3節では、細胞とシナプスの時間スケールの関係を変えたときに学習がどのように変化するかを解析する。このとき、記憶容量の点から最適な時間スケールの関係が存在することを示す。

以下では学習したネットワークを用いて, 最適な時間スケールにおいて本論文の学習過程がどのように相空間を形成するかについて議論を行う。

4章では入力を印加しない状況でのネットワークの相空間について解析を行う。特に学習が進むに連れて相空間がどのように変化していくかについて述べる。

5章では学習したネットワークに入力を印加し、分岐の解析を行う。学習したパタンは入力により安定な構造に分岐することで、入出力関係を為すことを述べる。

6章では以上の結果をまとめ、必要な機構と意義に関して議論を行う。

2 設定

本論文では三層ネットワークを用いて入出力関係を学習させ、それによる相空間の構造の変化を解析する。ここでいう入出力関係の学習とは入力・隠れ・出力層の三層からなるネットワークの入力層に入力パタン I^a を印加し、出力層の値が入力に対応するターゲット Ξ^a になるようにシナプスを変化させることを指す。1.2節での要請を満たす学習を行うために本モデルでは

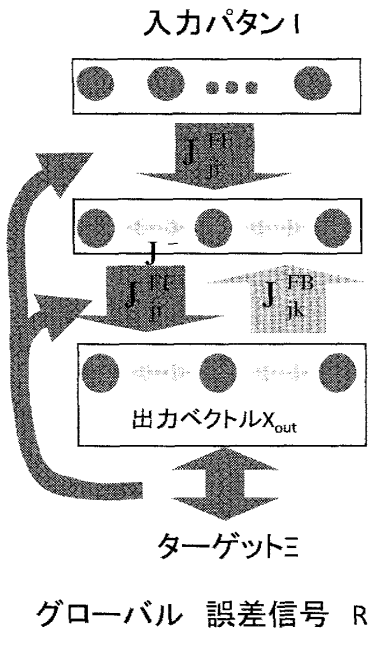
1. 1次元のターゲットと出力層の誤差情報によりシナプスの可塑性が変化すること (meta-plasticity)

2. FF・FB結合が固有の時間スケールをもっていること (multiple-timescales)

という大きな2つの特徴を有する。このモデルにより要請を満たした学習が可能になることは3章で述べることとし、まず本章では本論文で用いるモデルを2.1節で、それを用いた学習の方法を2.2節でそれぞれ定義する。

2.1 モデル

本論文で用いるネットワークは入力・隠れ・出力層からなる3層ネットワークで、3種類のシナプスをもつ。これら3つのシナプスはそれぞれ、フィードフォワード結合 J^{FF}(以下FF結合), フィードバック結合 J^{FB}(以下FB結合)、層内抑制結合 J^{-} と呼ばれる。J^{FF} は入力層細胞から隠れ層細胞、隠れ層細胞から出力層細胞への結合であり、J^{FB} は出力層細胞から隠れ層細胞への結合)、J^{-} は隠れ層内、出力層内の結合である。また各結合は全結合 (ただし抑制結合は自分との結合 J_{ii} はなし) である。従って同層内の細胞は抑制性結合だけでつながっており、異なる層の細胞は興奮性結合でのみつながっている。J^{-} はパラメタとして固定し、層間の結合 $J_{ij}^{FF,FB}$ を変化させることで学習を行う。またネットワークの全ての要素数を N、



細胞のダイナミクス

入力層　$x_i = \eta I_i$

隠れ層　$\dot{x}_j = 1/(1+\exp(-\beta u_j + \theta)) - x_j$

$$u_j = \sum_{i\in Input} J^{FF}_{ji} x_i + \sum_{k\in Outout} J^{FB}_{jk} x_k + \sum_{j\neq j'} J^{-} x_{j'}$$

出力層　$\dot{x}_k = 1/(1+\exp(-\beta u_k + \theta)) - x_k$

$$u_k = \sum_{j\in hidden} J^{FF}_{kj} x_j + \sum_{k\neq k'} J^{-} x_{k'}$$

シナプスのダイナミクス

$$\dot{J}^{p}_{ij} = R^{p}(x_i - r)x_j / \tau^{p} \quad (J_{ij} > 0) \quad (p = FF, FB)$$

$$R^{FF} = \begin{cases} +1 & for(|X_{out} - \Xi| \le \varepsilon) \\ -1 & for(|X_{out} - \Xi| > \varepsilon) \end{cases}$$

ターゲットに近いとき Hebb phase

$$R^{FB} = \begin{cases} 0 & for(|X_{out} - \Xi| \le \varepsilon) \\ -1 & for(|X_{out} - \Xi| > \varepsilon) \end{cases}$$

ターゲットに遠いとき Anti-Hebb phase

図 1: モデル イメージ図:

３層のネットワークで、３種類 (FF,FB, 抑制) のシナプスを持つ。FF 結合は入力層から隠れ層、隠れ層から出力層を全結合でつなぎ、FB 結合は出力層から隠れ層に全結合でつながる。抑制性結合は隠れ層、出力層の細胞をそれぞれ全結合で結ぶ (但し、自己結合はない。)。入力層の細胞は隠れ層に投射するだけで、値は入力パタンにより完全に規定される。ネットワークの出力とターゲットとの誤差情報 R がグローバルな情報としてシナプスを制御し anti-Hebb phase と Hebb phase を切替える。FF 結合と FB 結合は独自の時間スケール $\tau^{FF,FB}$ をもち、挙動も異なる。細胞とシナプスのダイナミクスをまとめて書いた。詳しくは本文を参照。

入力・隠れ・出力各層の要素数を $N_{in,hid,out}$ とし、$N_{in} = N_{hid} = N_{out} = N/3$ とする。以上を含め図1に本モデルについて簡単にまとめてあるので参照されたい。

以下、細胞とシナプスのダイナミクスの定義を行い、それらの定性的な振る舞いを説明する。

まず細胞のダイナミクスについて説明する。本論文で用いる細胞のダイナミクスは連続系の rate conding のモデルであり、細胞 i の状態量 x_i は発火率を表している。入力層の細胞の値は式1に従い入力によってのみきまる。隠れ層と出力層の細胞のダイナミクス $x_i(t)$ は (0,1) の値をとるシグモイド型の activation 関数 $f(x_i)$ で駆動され、同じく (0,1) の値をとる。まとめると細胞 x_i は

$$x_i = \eta^a I_i^a \quad (\text{細胞 } i \text{ は入力層}) \tag{1}$$

$$\dot{x_i} = f(\{x_i\}) - x_i = 1/(1 + exp(-\beta u_i + \theta)) - x_i \quad (\text{細胞 } i \text{ は隠れ・出力層}) \tag{2}$$

という式にしたがって時間変化する。

ここで $u_i = \sum_{j \neq i}^{n} J_{ij} x_j$ で細胞 x_i への入力カレントである。隠れ層の細胞は入力層細胞から FF 結合、出力層細胞から FB 結合、同じ隠れ層細胞からは抑制結合を介して入力を受けるので $u_i^{hid} = \sum_j J_{ij}^{FF} x_j^{in} + \sum Jij^{FB} x_j^{out} + \sum_{j \neq i} J^- x_j^{hid}$ とかける。同様に出力層細胞は隠れ層細胞から FF 結合、同じ出力層細胞からは抑制結合を介して入力を受け, $u_i^{out} = \sum_j J_{ij}^{FF} x_j^{hid} + \sum_{j \neq i} J^- x_j^{out}$ とかける。入力がないときの細胞 i の固定点 $x_i = 1/(1 + exp(\theta))$ は自然発火率 r とみなせるので、θ は自然発火率 r を定めるパラメタである。この自然発火率 r はシナプスのダイナミクスのところで用いる (本論文では r=0.1 となるようにした)。また activation 関数の非線型性を表す β はここでは十分に大きくとり (β=42)、双安定性が現れるようにしている。同じ層の細胞からの抑制性結合 J^- の存在により、本モデルでは winner takes all (WTA) の状況ができやすい[2]。すなわち各層一つずつの細胞が1に近い値をとり (活性化)、それ以外の細胞は活性化した細胞から抑制性の入力を受けるので自然発火率より小さい値 (非活性) をとるスパースな状態が安定になりやすい。

次にシナプスのダイナミクスについて説明する。

シナプスは上で述べた通り FF 結合、FB 結合、抑制結合の3種類あるが、このうち抑制結合はパラメタであり時間変化しない ($J^- = -1.0$)。したがって入出力関係の学習は FF、FB 結合を変化させることで行われる。FF 結合、FB 結合は以下のように各々τ^{FF},τ^{FB} という固有のタイムスケールをもち、pre-synaptic cell x_j と post-synaptic cell x_i と 誤差情報 $R(X_{out}, \Xi)$ の積により変化し、

$$\tau^p \dot{J_{ij}^p} = R^p(X_{out}, \Xi)(x_i - r)x_j \qquad (J \geq 0) \qquad (p = FF, FB) \tag{3}$$

に従う。但し J は負にならない。ここで誤差情報 R はターゲット Ξ^a と出力ベクトル $X^{out}(t)$ との距離によって切り替わる関数で、

$$R^{FF} = \begin{cases} 1 & for \quad |X^{out} - \Xi| \leq \epsilon \\ -1 & for \quad |X^{out} - \Xi| > \epsilon \end{cases} \qquad R^{FB} = \begin{cases} 0 & for \quad |X^{out} - \Xi| \leq \epsilon \\ -1 & for \quad |X^{out} - \Xi| > \epsilon \end{cases} \tag{4}$$

であり、ネットワークで共通なグローバルな量である。ここで| |は N_{out} 次元のユークリッド距離である。上でも述べたように誤差情報 R によりシナプスの挙動が制御され pre-,post-synaptic cell の3者の積で表されることと、FF 結合と FB 結合が固有のタイムスケールを持ち挙動が異なることが今モデルの大きな特徴となっている。誤差情報 R はターゲット Ξ と出力ベクトル X^{out} の距離が遠いときは FF、FB 結合の R はともに負の値をとり、ターゲットと出力ベクトルの距離が近いときは FF 結合の R は正、FB 結合の R は0をとる[3]。

[2]ただし、明示的に1細胞だけが活性化するようにしていないので、他層から興奮性入力が大きいときは複数の細胞が同時に活性化することはありうる。

[3]ターゲットと出力ベクトルの距離の遠い、近いは ϵ で判定される。このパラメタの依存性は3章で議論する。

ターゲットと出力ベクトルの距離が遠いときを anti-Hebb phase, 近いときを Hebb phase と呼び分けて挙動を述べる。

anti-Hebb phase では FF、FB 結合ともに $\dot{J}_{ij} = -(x_i - r)x_j$ となる。従って presynaptic cell x_j ,postsynaptic cell x_i がともに活性化 $(x_i > r)$ している場合、J_{ij} は小さくなり J_{ij} を介した興奮性入力が減少するので活性化している postsynaptic cell x_i は非活性方向に圧力を受ける。presynaptic cell x_j が非活性の場合、postsynaptic cell x_i の値には関係なく J_{ij} は変化しない。逆に presynaptic cell x_j が活性で、postsynaptic cell x_i が非活性の場合は J_{ij} は正の方向に変化するので, 非活性な postsynaptic cell x_i を活性化する圧力が働く。以上のように anti-Hebb phase では現在の postsynaptic cell の活性・不活性の状態を反転させる方向に J_{ij} は変化するので、現在の活性・非活性のパタンに対して不安定化力として働くと考えることができる[4]。

Hebb phase では FF 結合は R が正の値をとるので、anti-Hebb phase とは逆に活性化している隠れ層細胞から出力層細胞の間の結合が強化される。FB 結合は R=0 なので変化しない[5]。したがってターゲットの近くのパタンにネットワークの状態が遷移すると安定方向に相空間は変化する。このように R がシナプスの挙動を安定化力と不安定化力の間で切替える役割を担っておりこれがターゲットに収束する機構に重要な役割を果たす。この点は 3 章で議論する。

またもう一つの特徴である FF 結合と FB 結合が独立な時間スケール τ^{FF}, τ^{FB} をもつことは、序論でも述べた学習に置ける相空間のタイムスケールと細胞のタイムスケールの関係を議論するとき重要な点であり 3.3 節で議論する。

以上が本論文で用いるモデルである。このモデルをもちいて以下では学習が可能かどうか、学習過程の時間スケール依存性 (3 章), 学習後の相空間の構造 (4 章)、その分岐 (5 章) について述べる。

最後に今モデルの特徴の一つである誤差情報 R によるシナプスの制御の生理学的な対応について少し述べる。この制御はターゲットから遠いときの不安定化方向の変化と近いときの安定化方向の変化の 2 種類を切替えることで行われる。これらの生理学的な直接的な根拠は現在のところないと思われるが、いくつかの知見はそのことを示唆していると考えられる。

1) 誤差情報の表現

本論文の冒頭に述べたように誤差情報は各部位・要素ではなく個体全体に与えられるべき、ということから 1 次元の誤差情報を学習するにあたっての要請とした。実際に個体においての誤差情報はどのように表現されているのかは、難しい問題であるが答えの 1 つとして大脳基底核での情動系があると考えられる。たとえばこの分野ではドーパミンニューロンの活動と報酬予測などとの関連が最近活発に研究されている [20]。

2)neuronal modulator によるシナプスの可塑性の変化

ドーパミン、アドレナリンなど neuronal modulator と呼ばれる一群のシグナル存在により LTP や LTD などのシナプスの可塑性自体が変化する事はよく知られている。またマクロのレベルでもドーパミンがないときでの working memory 課題の performance 低下などが見られる [15]。これらの neuronal modulator は基底核等から大脳新皮質を含めた大域的な領域に投射回路が存在することは解剖学的に確かめられている。

以上の二つをつなぎ合わせて考えると、基底核での活動として個体の行動が評価され、その活動が大域的な投射回路を通してシナプスの可塑性をコントロールしている可能性がある。このことが、今モデルの誤差情報によるシナプスの制御のような機能を果たしていることはありえないことではないとかんがえられる。

[4]ここで $(x_i - r)x_j$ の部分は covariance rule[19] と呼ばれる。ここの部分が他の形例えば、$x_i x_j$ のときは anti-Hebb では常に負の値しかならず活性化細胞の遷移が起きない。同様に $(x_i - \theta)(x_j - \theta)$ のときは presynaptic cell が非活性のときもシナプスが変化してしまい、全体の結合が大きくなり過ぎてしまう。従って $(x_i - r)x_j$ のような非対称な形が必要となっている。

[5]FB 結合も FF 結合と同様に Hebb phase で正の値をとると、パタンの安定化が強過ぎて周りの構造を消してしまう。これは学習ができなくなる要因となるので今モデルでは Hebb phase では R=0 とした。

2.2 学習方法

ここでは前節で説明したモデルを用いてどのように学習を行うのかについて説明する。2 章の冒頭でも述べたが、本論文では入出力関係の学習は入力パタン I^a を印加しときにターゲット Ξ^a を出力するようなネットワークを形成する事を指す。N_{pat} セットの入出力関係の学習を以下のように行う。

0. 初期のネットワークは特に断らない限り、FF・FB 結合ともに 0 のネットワークを用意する。

1. 入力 I^a を印加し、それに対応するターゲット Ξ^a を提示する

   入力層の細胞を式 1 に従い入力パタン $I^a = (I^a_0, I^a_1, \ldots, I^a_{N_{in}})$ に固定する。このとき学習過程を通して入力強度 η は 1.0 に固定する。また誤差信号 R はネットワークの出力とターゲット Ξ^a の距離で変化する。ここで入力パタン・ターゲットは各要素が 0, 1 の値をもつ 2 値ベクトルであり、本論文で用いるのは特にスパースなパタン、すなわち 1 の値をもつ要素は一つだけでの残りは全て 0 というパタンを用いる。

2. anti-Hebb phase と Hebb phase の切替えによりターゲット Ξ^i を探索する

   前節で見たようにターゲットから遠い時 (anti-Hebb phase) ではシナプスは細胞の状態が不安定になるように変化するので、細胞の状態は次々に変化する。その結果ターゲットに近づくと Hebb phase になりシナプスは細胞の状態を安定化するように変化するので細胞の状態は変化しない。この二つの phase により細胞の状態はターゲットに収束することが期待される。実際にそのように収束する事を 3 章で示す。適当な時間 T_{dyn} 内にターゲットに収束できなければ失敗とみなす。しかし 3.1 節で見るようにほとんどの場合でターゲットに収束するので失敗した場合は事実上考えなくてよい。

3. 探索によりターゲットを探し細胞の状態がターゲットに収束したのち (Hebb phase) に適当な時間 (T_{mem})、学習セット I^a, Ξ^a を提示し続ける

   2 のプロセスで出力層のパタンがターゲットに収束したのちもシナプスは式 3 に従い変化を続ける。従って適当な時間 T_{mem} をパラメタとして定めておき、収束してから T_{mem} までは変化を続けさせる。

4. 1-3 のプロセスを N_{pat} 回繰り返す

   3→1 のプロセスの切り替わりでは、新たなターゲット Ξ^b を提示することによりターゲットとの距離が ϵ を越えて大きくなり、anti-Hebb phase に変わる。anti-Hebb phase に変わったことでシナプスが現在のパタンを不安定化し、細胞の状態は相空間を探索し始める。

今モデルではネットワークの出力とターゲットとの距離に応じて anti-Hebb phase と Hebb phase を切替えることで、ネットワークはターゲットを出力できる固定点に収束すると述べてきた[6]。しかしこの固定点は正確にいうと細胞の状態をみたときの固定点であり、シナプスのダイナミクスは式 (3) から明らかなように固定点ではない[7]。したがって (細胞の状態が) 収束したのちに、適当なパラメタ T_{mem} でシナプスのダイナミクスをとめる必要がある。そのために上の手順 3 のように収束後 T_{mem} たった後に学習パタンを切替えることで一パタンの学習を終了させている。ターゲットに収束した後のシナプスのダイナミクスは Hebb phase であり、活性化している細胞同士を結ぶ FF 結合は強化される。したがって手順 3 は収束したネットワークの状態 (ターゲットパタン) を安定化させる過程である。以上のように本論文において学習過程はターゲット探索過程 (手順 1、2) とその後のターゲット安定化過程 (手順 3) を繰り返すことで行われる。また今モデルでは、入力パタンは常に提示し続けるので、入力は系にとって境界条件のパラメタとみなせる。異なる入力・ターゲットを学習する過程では、系はそれぞれ異なる境界条件のもとでシナプスを変化させており、その意味で異なる力学系とみなすことができる。

[6]実際にターゲットに収束できるかは次章で詳しく解析している。

[7]もちろん、シナプスが変化している以上細胞のダイナミクスも実際には収束していないが、シグモイド型の activation 関数の効果で 0,1 付近に急速に近付くと後はシナプスが強化されてもほとんど変化しない。

次に N_{pat} セットの学習を行ったネットワークに対して'記憶'を定義し、それを調べる手順について述べる[8]。今モデルでは学習後のネットワーク (シナプスを固定したネットワーク) は入力層細胞は入力により値がきまるので、隠れ・出力層細胞からなる $N_{hid}+N_{out}$ 次元の力学系とみなすことができる。この力学系において入力 I^a とターゲット Ξ^a という入出力関係の記憶を、I^a を強度 $\eta=1.0$[9]で印加したネットワークの相空間の中においてターゲットパタン ξ^a の basin volume が $\theta^{mem}(=0.5)$ より大きくなる事として定義する。すなわち、入力をいれたことによる分岐を力学系の出力・応答とみなし、記憶を定義する。今後の解析では次の手順を行うことで記憶を調べる。

1. 学習により形成されたネットワーク J^{FF}, J^{FB} を以下では用いる。入力パタン I^a を印加した相空間での各アトラクタの basin volume を調べる。

   記憶を調べるときは学習時と違いシナプス J^{FF}, F^{FB} は変化させず、細胞のダイナミクスのみを動かす。入力パタン I^a を印加した状態で初期状態 (隠れ・出力層細胞の初期状態) を多数とり、そこからの軌道がどのアトラクタに収束するかでアトラクタの basin を調べる。basin volume は全体で 1 になるように規格化しておく。

2. ターゲットパタン ξ^a が固定点アトラクタの basin volume が 0.5(全体の basin volume が 1 となるように規格化) より大きいときに記憶ができているとする。

3. 1 セット学習するごとに今まで学習した全セットについて 1-2 を行い、記憶できていたセット数を各学習段階での記憶数とする。

このように本論文において N_{pat} セットの記憶しているネットワークの力学系は、入力のない状態の相空間に N_{pat} の入力パタンを印加したときに、対応する N_{pat} のターゲットパタンが大きな basin をもつ安定固定点として分岐する構造をもつ力学系であるといえる。

最後にいくつか用語の説明をしておく。・学習段階・学習セット数

今モデルでは N_{pat} の入力パタンとターゲットのセットを逐次学習していく。この学習過程を解析するとき学習セットを 1 つ与えて学習を行う度にネットワークを取り出し、そのネットワークの相空間の性質を調べるということを行う。このように入力・ターゲットのセットを 1 セット学習したネットワーク、2 セット学習したネットワーク … N_{pat} セット学習したネットワークを調べることで、N_{pat} セットの学習過程でネットワークがどのように変化していくかを記述する。このときの各段階のネットワークを学習段階 1(学習セット数 1) のネットワーク、学習段階 2(学習セット数 2) のネットワーク … というように以下では呼ぶことにする。また学習段階毎の平均とは初期条件・学習パタンの組合せなどが異なる多数の学習過程に渡って同じ学習段階のネットワークで平均をとる事を指す。

・ターゲット Ξ とターゲットパタン ξ についての区別

ターゲット Ξ は提示された学習すべき出力層のパタンを意味し次元は N_{out} である。前節で述べたようにターゲット Ξ が提示され出力層がそのパタンをとるように細胞の状態は収束する。このとき隠れ層のどの細胞が活性化するかについては学習セットからは情報はあたえられない。そこでターゲット Ξ を提示され収束したときの隠れ層の活性化パタンも含めたパタンは Ξ とは異なるので、これをあらためてターゲットパタン ξ とよぶことにする。ターゲットパタン ξ の次元は $N_{hid}+N_{out}$ である。今までの議論では細胞の状態がターゲット Ξ に収束する、という表現は正確にはターゲットパタン ξ に収束すると表現すべきであるが混乱をさけるためにそのように表記してきた。以下では探索過程ではターゲット Ξ と学習後はターゲットパタン ξ として別のものとして表記する。

[8]上で述べたが今モデルでは学習したパタン=記憶したパタンは自明ではなく、学習セット数より記憶数は一般的に小さくなることに注意。

[9]学習で用いた入力強度である。

3 ターゲットの学習

今章では 2 章で定義した学習過程を用いて、3.1 節では探索過程について、3.2 節で学習過程全体について、3.3 節においては時間スケールに依存性についての結果を述べる。

3.1 探索過程:ターゲットへの収束

ネットワークが提示されたターゲットに収束できるかどうかは今モデルにとって重要な問題である。そこで今節ではターゲットを探索する機構について述べ、適当なパラメタのもとではほとんどがターゲットに収束することをしめす。

今モデルは 2.1 節で述べたように隠れ・出力層それぞれ 1 つの細胞が活性化しているスパースなパタンが安定になりやすい性質をもつ。このスパースなパタンを Θ^a とし、各層で活性化している細胞を隠れ層の活性化細胞 x_a^{hid} と 出力層の活性化細胞 x_a^{out} とする。また学習すべき入力を I^a(このとき活性化している細胞を x_a^{in}), ターゲットを Ξ^a とする。本論文ではネットワークのパタンを Θ^a と表記した時に、パタンの index a に対応して隠れ・出力各層で活性化している細胞をそれぞれ $x_a^{hid,out}$ と記述する。この記法は Θ^a に限らず ξ^a でも同様であり、とくにターゲットの場合 $x_a^{hid,out}$ を隠れ・出力層ターゲット細胞と呼ぶ。以下では出力層のパタンがターゲット Ξ^a から遠いとき (anti-Hebb phase) と近いとき (Hebb phase) に分けてその挙動を述べる。

1)Θ^a の出力層のパタンがターゲットとことなる時:anti-Hebb phase

anti-Hebb phase なので x_a^{in}、x_a^{hid}、x_a^{out} をつなぐシナプスの強度は FF,FB 結合共に減少する。逆に他のシナプスのうち $x_a^{in,hid,out}$ から投射されるシナプスの強度は強化され、$x_a^{in,hid,out}$ 以外から投射されるシナプスの強度は変化しない。このようにシナプスの変化が分かれるのは presynaptic cell が非活性の場合シナプスは変化しないことが効いている。強化されるシナプスの postsynaptic cell に着目する。この細胞への入力は $x_a^{in,hid,out}$ のうち他層の細胞からの興奮性入力と同じ層の細胞からの抑制性入力の和で表せる。他層からの興奮性入力はその間のシナプスの強度が強化されることで増加する一方で、同層からの抑制性入力は抑制性のシナプス強度が固定されているので変化しない[10]。この結果ある時点で興奮性入力が抑制性入力をこえ postsynaptic cell が活性化し、逆に元々活性化している細胞は抑制され非活性化する。このようにして活性化する細胞は遷移し、この遷移は anti-Hebb phase であるかぎりつづく[11]。

どの細胞があらたに活性化するかはそのときの結合強度の分布に依存し、この分布の形成はそれまでの探索に依存している。例えば presynaptic cell が何回も活性化しているにもかかわらず、ある postsynaptic cell x_a が活性化しない場合はその間のシナプス強度は単調に増加し大きな値をもつはずである。その結果次に presynaptic cell が活性化する時、この postsynaptic cell x_a が活性化する可能性は一度活性化したことの有る細胞に比べ高くなる。このように anti-Hebb phase の遷移はランダムではなく、それまでの探索により形成された構造に依存して遷移が行われる。また FF 結合と FB 結合は固有の時間スケールをもっているので、隠れ層と出力層の活性化細胞の遷移の時間はこの時間スケールに依存し一般に同じではない。

2)Θ^a がターゲット近くの時:Hebb phase

この場合上の anti-Hebb phase とは逆の現象が起きる。すなわち活性化した細胞 $x_a^{in,hid,out}$ の間の FF 結合が強化され、現在の状態が安定する方向にシナプスの強度は変化する[12]。この結果パタン Θ^a が安定になり細胞のダイナミクスは止まる。

以上の機構を力学系の相空間の変化として捉え直すと以下のように考えられる。もともと各層 1 細胞が活性化しているパタンは安定になりやすい。したがって相空間上にはいくつかのスパースなパタンが安定な

[10]同層の活性化細胞も入射するシナプスの強度の減少により弱くなっているので、抑制性入力も徐々に小さくなる

[11]ただし以下で見るように、ϵ の値次第では anti-Hebb phase でも遷移が止まり系が収束する場合がある。

[12]ただしあくまで"安定する方向に変化する"なので、この状態が元々不安定であれば、安定になる前に細胞の状態が遷移してしまうことはある。これは後にタイムスケールの依存性のところで議論する。

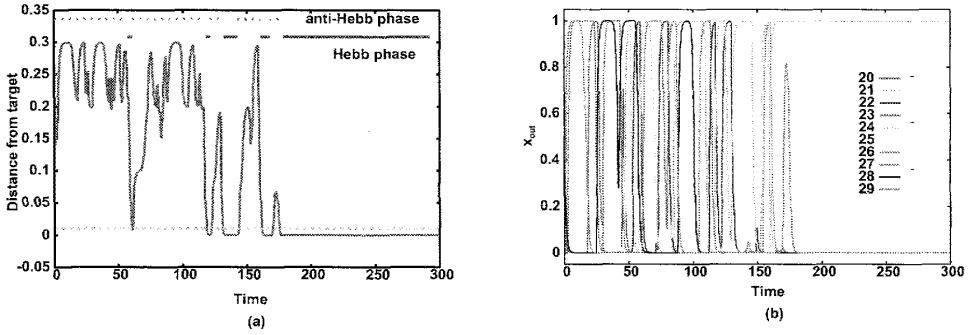


図 2: ターゲット探索過程:
探索前のネットワークとしてシナプスの値を $(0,1)$ からランダムに選んだネットワークをとり (但し $J^- = -1.0$)、細胞の初期状態として $(0,1)^N$ からランダムにえらんだ状態にとった。また入力パタンとターゲットは入力層細胞、出力層細胞それぞれ 1 細胞だけ活性化したスパースなパタンからランダムに選んだ。以上の初期条件でターゲットの探査をおこないその時系列をプロットした。(パラメタは $N = 30, N_{out} = 10, \epsilon = 0.01, \tau^{FB} = 8, \tau^{FF} = 64$)
(a) ターゲット Ξ と出力 X^{out} の距離の時系列:
ただしここでは距離はユークリッド距離の 2 乗を 1 に規格化したものとした。green line は $d = \epsilon$ である。この line よりターゲットとの大きいときは anti-Hebb phase (図上部の点線) であり、小さいときは Hebb phase (図上部の実線) である。
(b) 出力層細胞 N_{out} 個の時系列の重ね書き:
出力層の各細胞 N_{out} 個の軌道を重ねてプロットした。

活性化した細胞が次々と遷移し、最終的にターゲットとの距離が 0 の状態に収束しているのがわかる。

点として存在し、初期状態に依存して細胞の状態は適当なパタン Θ^a へと向かう。しかしパタン Θ^a がターゲット Ξ から遠いと相空間はパタン Θ^a が不安定化するように変化する[13]。この変化は同時に他のパタン Θ^b が安定化する変化でもある。この結果 Θ^a から Θ^b に向かう軌道が形成されパタンが遷移する。パタン Θ^b がどれになるかはそれまでの探索で作られた相空間の構造に依存する。このように探索により相空間を形成し、またその相空間に依存しながら探索が行われる。遷移を繰り返した結果、ターゲット Ξ に遷移すると安定化し、ターゲット Ξ がアトラクタとなるような相空間が形成される[14]。

このように、anti-Hebb phase では 1. 安定なパタンに状態が遷移、2. パタンが不安定化することで別の安定なパタンに遷移、というサイクルを繰り返す。一方ターゲットの近くにあるパタンに遷移すると Hebb phase になり安定化し収束する。以上のようにターゲットとの距離に応じて、安定化と不安定化を切替えることでターゲットの近くのパタンに収束する。この探索過程の一例が図 2 である。パタンが次々と活性化する細胞が変化し (anti-Hebb))、最終的にターゲットに収束しているのが分かる (Hebb phase)。この図でターゲットに近付き Hebb phase になってもターゲットが安定しない場合がいくつか見られる。これはターゲットに近付いた時に相空間はターゲットを安定化するように変化するが、ターゲットが実際に安定な構造に変化する前に細胞の状態がターゲットの近傍から脱出してしまうからである。

[13]注意すべきなのは不安定化したパタンも全ての方向にわたって不安定になるわけではない点である。パタン Θ^a で活性化している細胞 $x_a^{in,hid,out}$ から投射しているシナプスは上述のように不安定方向に変化するが、他の細胞から $x_a^{hid,out}$ に入射するシナプスは presynaptic cell が非活性なので変化しない。このことにより不安定化したパタンも一部の方向にたいしてのみ不安定化するのであり、多くの方向からは依然として安定である。
[14]ただし、この相空間構造は入力 I^a が印加されているときの構造であり、入力が印加されていないときの構造とは異なる。入力が印加されていないときの構造については 4 章で議論する

3.1.1 探索過程のパラメタ依存性

以上のように今モデルではanti-Hebb phaseの不安定化とHebb phaseによる安定化がターゲットと出力の距離に依存して切り替わることで最終的にターゲット近くのパタンに収束する。ここでターゲットと出力との距離の情報はϵより大きいか小さいに応じたシナプスのHebb phaseとanti-Hebb phaseの切替えに使われる。したがって〝近い〟かどうかはターゲットとの距離がϵより小さいかどうかで決まる。つまりϵはネットワークのターゲット認識の解像度のパラメタと考えることができる。以下今小節ではこの解像度ϵがターゲット探索にどのような影響を与えるのかを述べる。

ϵを変えたときにネットワークが収束するパタンがどのように変化するかを調べた。以下では探索前のネットワークとしてシナプスの値を$(0,1)$からランダムに選んだネットワークを用いた。また細胞の初期値として同じく$(0,1)^N$のN次元パタンからランダムに選び、入力・ターゲットもそれぞれスパースなパタンからランダムに選んだ。ここでターゲットと出力との距離dはN_{out}次元上のユークリッド距離の2乗を規格化し最大値を1にしたものとした。今モデルにおいて出力細胞の値は0,1付近をとるので、出力パタンは近似的に$\{0,1\}^{N_{out}}$に属するとみなせる。したがって出力パタン間の距離はi/N_{out}になる。

さまざまなϵに対して収束したときの出力とターゲットとの距離の分布を計ってみると距離d=ϵ,$i/N_{out}(i=0,1,2)$の値にピークが立つ分布になっている。但し$N_{out}=N/3$である。以下で$d=\epsilon$と$d=i/N_{out}$のふたつ場合に分けて解析する。

1)$d=i/N_{out}$に収束する場合。

上で述べたようにϵはネットワークのパタン認識の解像度とみなすことができる。したがってϵが大きく解像度が粗いとターゲットパタンとそうでないパタンの区別がつかなくなるので、ターゲットパタン以外のパタンに収束してしまう[15]。異なるパタン間の規格化距離はi/N_{out}なので$\epsilon > i/N_{out}$の時はターゲットパタンとi/N_{out}内にあるその他のパタンを区別がつかず他のパタンに収束する場合がある。このときの分布が$d=i/N_{out}$のピークに現れていると考えられる。

2)$d=\epsilon$にも収束する場合。

今モデルではJのダイナミクス(式3)により、Jの固定点は$x_i=r, x_j=0$しかないが、細胞の状態がターゲットとの距離ϵをはさんで激しく振動することで、平均的にR=0の解をつくる場合がある。図4に$d=\epsilon$に収束するときのターゲットとの距離dの時系列をプロットした。ターゲットとの距離dが$d=\epsilon$をはさみ減衰振動をしながら収束しているのが分かる。このことによりJの変化がなくなりダイナミクスは収束する。但し、上にも述べたが今モデルは$\{0,1\}^N$に収束しやすいので一般的にそれと異なる$d=\epsilon$のパタンには収束しにくい。

各ϵ毎に多数のターゲット探索を行い収束したときのネットワークの出力層細胞とターゲットとの距離を平均したものが図3である。ターゲットと出力との平均距離はϵが小さくなるにつれ減少する。この平均距離の減少には連続的減少と断続的減少の2種類有ることが判る。連続的減少は$d=\epsilon$に収束する分布のピークが平均値にあたえる効果である。このピークの平均値への寄与はϵの減少に対して連続的に変化するので平均距離の連続的な減少を起こす[16]。一方断続的減少は$d=i/N_{out}$に収束する分布のピークが平均距離にあたえる効果である。ϵがi/N_{out}を跨いで小さくなると、ターゲットパタンとi成分が異なっているパタンの区別が付くようになり、$d=i/N_{out}$に収束する分布のピークは消滅し、平均距離は非連続的に減少する。そしてターゲットと1成分だけ異なる出力も異なるパタンとみなせる$\epsilon < 1/N_{out}$になると,ターゲットと収束したパタンの距離はほぼ0になっている事が判る。以上の事からから$\epsilon < 1/N_{out}$であればほとんどの場合でターゲットに収束することが判る。またタイムスケール$\tau^{FB,FF}$をかえても同様で適当なϵをとることで、ほとんどの探索でターゲットに収束する。今節以降の結果はとくに断らない限り、ほとんどの場合でターゲットに収束するような$\epsilon=10^{-4}$を用いている。

[15]ある意味ネットワークにとってはこのときもターゲットに収束している。
[16]Nが小さい時のほうが連続的減少をしめすのは、$d=\epsilon$の値に収束する割合が比較的多いからである。

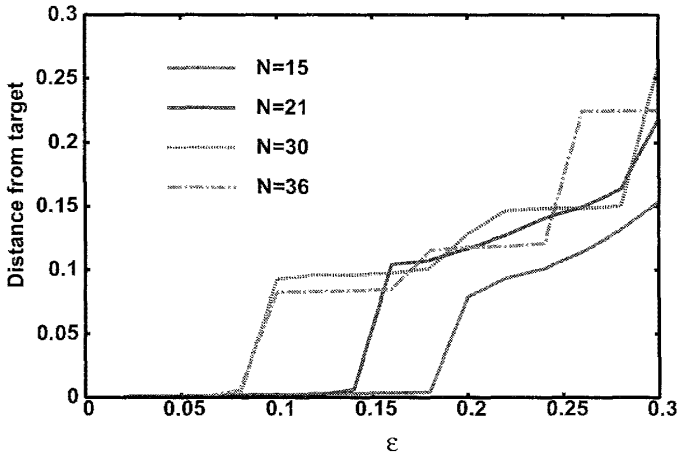


図 3: 探索の ϵ 依存性:
各 ϵ 毎に 1000 回ランダムな初期値、ランダムな学習セットでターゲット探索を行い収束したときのターゲットとの距離の平均値を求めプロットした。$\tau_{FB} = 8, \tau_{FF} = 64$ で固定した。

ϵ が小さくなるに連れて平均距離が小さくなっているのが判る。連続的に減少する場合と非連続に減少する場合がある。連続的に減少するのは、距離 d=ϵ に収束する場合の寄与が効いている。断続的に減少するのは、ϵ の値が i/N_{out} を跨ぐと、$d = i/N_{out}$ に収束する場合が消えることの寄与がきいている。前者は N が大きいときにはほとんど見られないため N が大きいところでは断続的な減少のみが起きている。

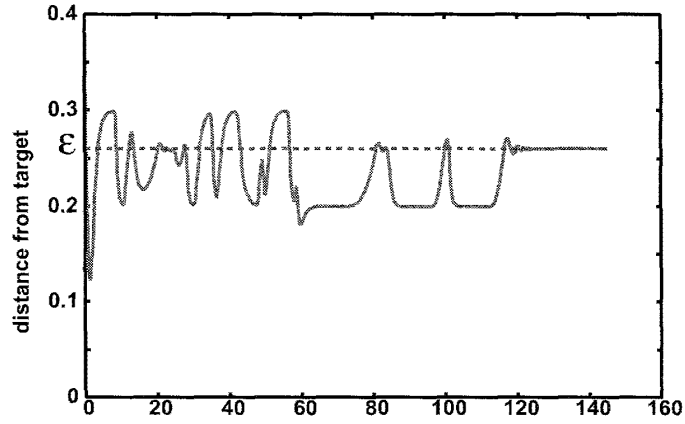


図 4: $d = \epsilon$ に収束する例:
$d = \epsilon$ の状態に収束する場合のターゲットとの距離の時系列をプロットした。$d = \epsilon$ の値をはさみ振動することで実質的にシナプスの固定点をつくり出している。そして振動は減衰し最終的に $d = \epsilon$ に収束していることが分かる。このような収束を行う例は学習前の相空間と初期値、ターゲットの関係に依存して決まるが、一般に N が大きくなるに連れてなくなっていく。$N = 30, \epsilon = 0.26, \tau_{fast} = 8, \tau_{slow} = 64$

3.2 学習

前節では適当なパラメタの元で今モデルが提示されたターゲットにネットワークの出力を収束させられることを述べた。今節ではまず本モデルが 1.2 節の最後で述べた要請を満たした学習を行えることをしめす。

具体例として 2 セットの学習過程を示す。図 5 は 2 セットの学習を行ったときの (a) ターゲットとの距離の時系列、(b) 出力層細胞の時系列をプロットしたものである。図 5(c) はこの学習で形成されたネットワークにおいて入力 $I^{1,2}$ を印加したときの軌道をプロットしたものである。本論文での学習は 2.2 節で定めたように、ターゲットの探索過程とターゲットに収束後の安定過程を繰り返す事で行われる。図 5(a)(b) において、およそ時間 0-200,700-1100 の間が探索過程を、そのほかの時間が安定過程に対応した時系列になっている (ここで時間はシミュレーション時間)。前節で示したように探索過程において出力層の細胞のパタンが次々変遷し、最終的にそれぞれ対応するターゲットに落ちているのが分かる。この学習過程で学習した 2 つの入出力セットが 2.2 節で定めた意味での記憶になっているかどうかを確かめたのが図 5(c) である。本論文では記憶は入力を強度 $\eta = 1.0$ で印加したときの相空間において、対応するターゲットパタンの basin volume が θ^{mem} より大きければよいとしたので、その確認として適当な初期値を多数とったときの軌道を重ねてプロットした。軌道の色の違いは印加した入力のパタンの違いであり、軸はそれぞれターゲットとの距離である。入力 I^1 を印加したときの相空間での軌道 (red line) はすべてターゲットパタン ξ^1 に収束し、入力 I^2 を印加したときの軌道 (green line) はすべてターゲットパタン ξ^2 に収束しているので、2 つの学習した入出力セットは記憶ができていることが分かる。

以上では学習数 2 の場合について具体例で示したが、これをスパース性を満たすパタンの最大数である $N_{pat} = N_{out}$ セットの学習を行ったときに記憶がどの程度できるかを調べた。図 6 は、3 つの時間スケール $(\tau^{cell,FF,FB})$ を $\tau^{cell} = 1, \tau^{FF} = 64$ に固定し $\tau^{FB} = 1, 16, 128$ としたときの記憶数の変化である。点線で示したのは完全学習 (学習数=記憶数) である。時間スケールの違いで記憶数に大きな違いがみえるが、特に適切な時間スケールのとき $(\tau^{FB} = 16)$ を取ったときの学習はほとんどこの完全学習に近い記憶ができていることがわかる。時間スケールを変えたときの解析に関しては次節で詳しく行う。

以上のように本論文の学習モデルが適切な時間スケールにおいて、多数の入出力関係の学習・記憶を行えることが示された。またこのように逐次的に複数のパタンを学習するモデルができた事についての議論は最終章で行う。

3.3 学習過程の時間スケール依存性

本論文のモデルでは時間スケールは細胞のスケール τ^{cell},FF 結合のスケール τ^{FF},FB 結合のスケール τ^{FB} の 3 つの時間スケールが存在する。前節で 3 つの時間スケールの関係によって記憶数が大きく異なることを少し触れた。今節では時間スケールの関係を色々変えたときにどのように学習過程が変化するかについて詳しく述べる。以下では $N = 30, N_{out} = 10, N_{pat} = 10, \epsilon = 0.001$ とした。

記憶できるパタンの数の時間スケール依存性をみるために $\tau^{cel,FF}$ は固定し τ^{FB} を変化させたときの記憶容量を調べたものが図 7 である。ただしここでは記憶容量を各学習段階で記憶数を計り、その最大となる記憶数とした。色の違いは τ^{FF} の違いである。どの τ^{FF} の場合でも記憶数は 1 ピークをもつ分布になっている。すなわち細胞の時間スケールとシナプスの時間スケールに記憶容量の観点から最適な関係が存在し、その関係は $\tau^{cell} \ll \tau^{FB} \ll \tau^{FF}$ と表せる。

このような事が起きる機構について議論する。本論文の学習過程において、入出力関係の記憶は FF 結合がになっていると考えられる。探索過程でターゲットに収束後、ターゲットパタン ξ を安定にするために相空間変化するときに変化するシナプス結合は FF 結合のみである。今モデルではターゲットをスパースなパタンとしているので、各層で活性化した 1 つの細胞を結ぶ FF 結合が選択に強化されることでターゲットパタンに記憶を保持している。実際、図 7 での両端、$\tau^{FB} = 1, 128$ のときは過去に学習した FF 結合が急速に減衰する反面、最大記憶容量をもつ $\tau^{FB} = 16$ では減衰が小さくなっており FF 結合の強さと記憶容

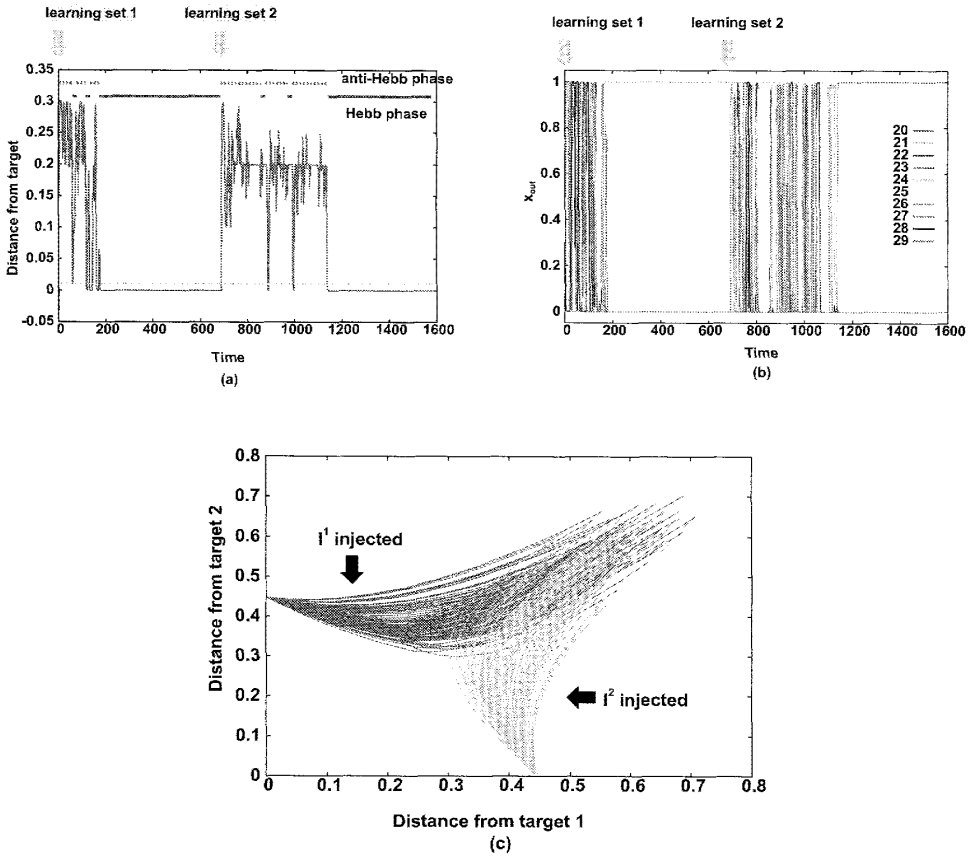


図 5: 2セットの学習過程:
図 (a)(b) は 図 2 と同様にネットワーク・細胞の状態・ターゲットパタンをランダムに選んだ初期条件のもとで、ターゲットの探査をおこないその時系列をプロットしたものである。パラメタは $N = 30, N_{out} = 10, \epsilon = 0.01, \tau^{FB} = 8, \tau^{FF} = 64$
(a) ターゲット Ξ と出力 X^{out} のユークリッド距離の 2 乗 (但し、とりうる最大の距離を 1 に規格化した) の時系列:
green line は $d = \epsilon$ であり、これよりターゲットとの距離が大きいと anti-Hebb phase であり小さいと Hebb phase である。
(b) 出力層細胞 N_{out} 個の時系列の重ね書き:
(c) 学習終了後のネットワークに $I^{1,2}$ を印加したときの軌道:
上図 (a)(b) の学習過程で形成されたネットワークに入力を印加し、細胞の初期値をランダムに 100 個選びそこからの軌道をプロットした。red line は入力 I^1 を印加したときの軌道、green line は入力 I^2 を印加したときの軌道をそれぞれ重ね書きしたものである。横軸、縦軸はそれぞれターゲットパタン ξ^1, ξ^2 との距離をとった。

ターゲット Ξ^1 に収束したのち、ターゲット Ξ^2 を提示することで anti-Hebb phase になり探索を再開している。さらに探索を再開したのちに正しいターゲット Ξ^2 に収束している事が分かる。学習後のネットワークに入力 $I^{1,2}$ を印加したときに全ての軌道が対応するターゲットパタン $\xi^{1,2}$ に収束している。このことから学習後の相空間ではターゲット $\Xi^{1,2}$ 共に記憶できていることが分かる。

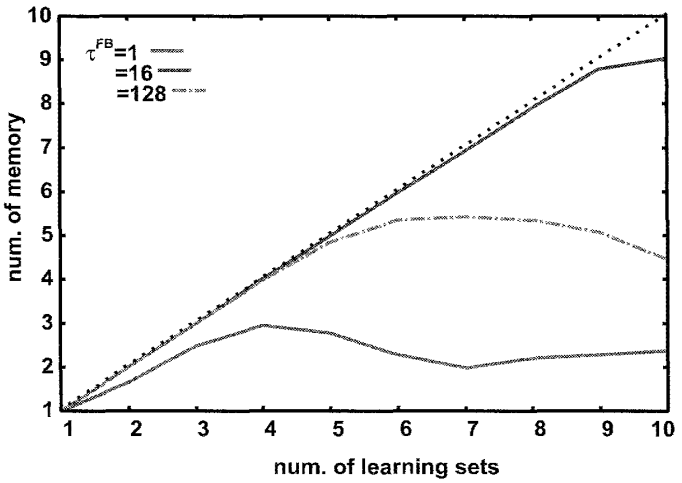


図 6: 記憶数の変化:
$N_{pat} = N_{out} = 10$ セットの学習を行ったときに、どの程度記憶できている のかを多数の学習過程にたいして平均したものをプロットした。色の違いは時間スケールの違い ($\tau^{cell} = 1, \tau^{FF} = 64$ で固定し、$\tau^{FB} = 1, 16, 128$ に関して調べた) で横軸は学習セット数で縦軸は記憶数。図 5 で 2 つ学習セットの記憶をしらべたように、1 セットの学習を終える毎に今まで学習したターゲットについてそれぞれ多数の初期値からの収束先を調べることで記憶できているかどうかを解析した。点線は完全に学習したものを記憶できている場合を示したもので学習セット数=記憶数である。適切な時間スケールの関係をとることで ($\tau^{FB} = 16$) ほとんど全ての学習セットを記憶できていることが分かる。各時間スケール毎の最大の記憶数をその時間スケールの記憶容量とし、この時間スケール依存性を次節で詳しく解析する。

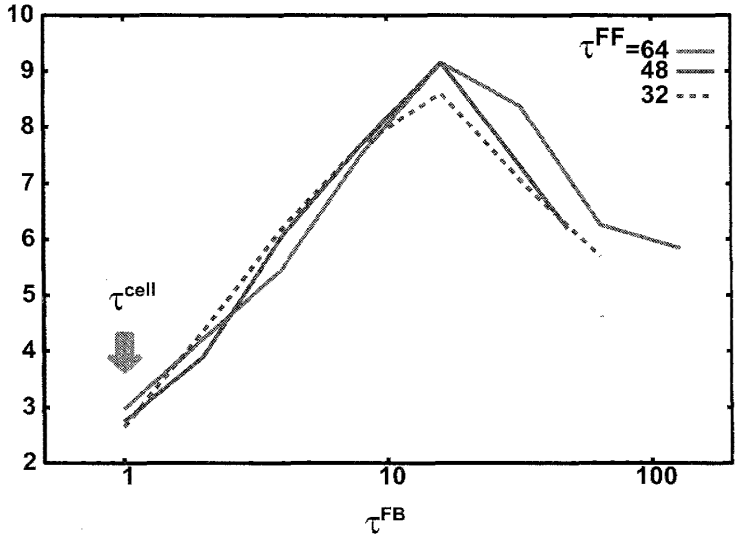


図 7: 記憶容量のタイムスケール依存性:
$\tau^{cell,FF}$ を固定し τ^{FB} を動かしたときの記憶容量の変化を各 τ^{FB} について 100 個の相空間の平均値を求めいくつかの τ^{FF} にたいしてプロットした。横軸は τ^{FB} を対数プロットし、色の違いが τ^{FF} の違いである。ただしここでは記憶容量を各学習段階で記憶数を計り、その最大となる記憶数とした。
どの τ^{FF} の値でも記憶容量が最大値をとる時に $\tau^{cell} \ll \tau^{FB} \ll \tau^{FF}$ という関係が成り立っている。

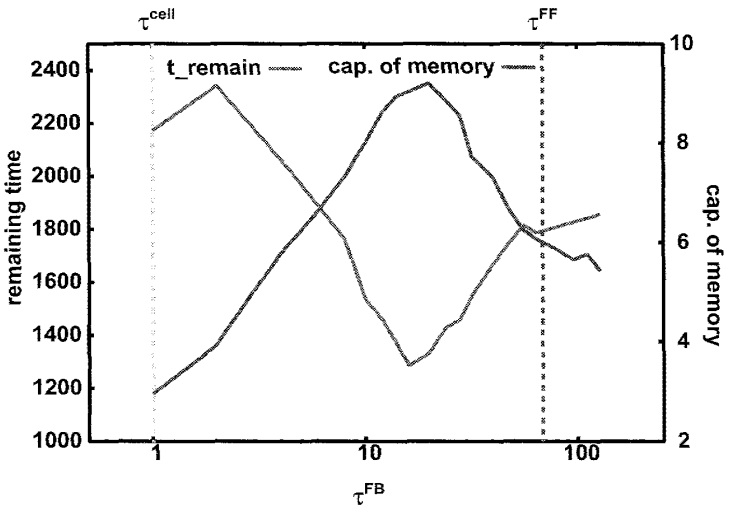


図 8: 過去のターゲットでの滞留時間:
$\tau^{cell}=1, \tau^{FF}=64$ のときの τ^{FB} の変化に対する過去のターゲットの滞留時間の変化 (red line, 左縦軸) と比較のために記憶数の変化 (green line, 右縦軸) をプロットしたもの。過去のターゲットの滞留時間は i 番目の学習セット I^i, Ξ^i を学習中に細胞の状態がターゲットパタン $\xi^j(j<i)$ にどれくらい滞在したかを全セット $(i=2,3,\ldots N_{pat}=10)$ にわたり足し合わせたものである。記憶数は図 7 の $\tau^{FF}=64$ の時と基本的には同じものだが、τ^{FB} に関するサンプル数を大きく増やしている。
記憶数が大きいところほど、過去のターゲットの滞在時間が減少しているのがわかる。

量の関係は明確である。以下ターゲットパタン ξ^a において活性化した細胞 x_a^{hid} と x_a^{out}[17]を結ぶ FF 結合を $J^{FF}_{x_a^{out}x_a^{hid}}$ と表記する。上で述べたようにターゲット ξ^a を学習することで $J^{FF}_{x_a^{out}x_a^{hid}}$ は強化される。その後他のターゲットの学習での探索過程で、この FF 結合 $J^{FF}_{x_a^{out}x_a^{hid}}$ はターゲットパタン ξ^a が活性化し、x_a^{hid} と x_a^{out} が同時に活性化する度に減衰する (anti-Hebb phase での不安定方向への変化)。したがって、学習過程において過去のターゲットパタン ξ^a の滞在時間が長くなると FF 結合 $J_{x_a^{out}x_a^{hid}}$ が減衰し、ターゲットパタン ξ^a の記憶が' 忘却' されると考えられる。この過去のターゲットパタンに滞在した時間と記憶容量の変化について関係をみるために、$\tau^{cell}=1, \tau^{FF}=64$ のときに多数の τ^{FB} に関して過去のターゲットの滞在時間を調べたものが図 8 である。記憶数が大きいところでは滞留時間が小さくなっていることが分かる。

以下時間スケール τ^{FB} が速い側 (τ^{cell} 側) と遅い側 (τ^{FF} 側) で滞留時間が増加していることについて解析する。

1) 速い側について

今モデルでは、現在の細胞の状態がターゲットから遠ければ相空間が anti-Hebb phase で不安定化し近ければ Hebb phase で安定化すること、スパースなパタンが安定になりやすい事の 2 点によってターゲットを探索することができる。このことにより細胞の状態がスパースなパタンを遷移することで最終的にターゲットにたどり着き安定化することができる。ここで細胞の時間スケールはスパースなパタン間を遷移する時間スケールであり、シナプスの時間スケール $\tau^{FB,FF}$ は不安定化・安定化の時間スケールだと考えられる。τ^{FB} が τ^{cell} に近付くということは、遷移する時間スケールと不安定化・安定化する時間スケールがちかづくという事である。この時どのような現象がみられるかを考える。ある時点の相空間構造がスパースなパタン Θ が不安定化し、ターゲットパタン ξ へと向かう軌道が存在する構造になっており、細胞の状態は Θ であるとしよう。遷移する時間スケールと不安定化・安定化するスケールが離れている場合、Θ から ξ に遷移する間での相空間の変化は小さいと考えられるので、細胞の状態は ξ に到達することができる。一方で、時間スケールが近い場合は遷移する間に相空間構造が変化するので、ξ に到達することができない場

[17] 今モデルでの設定から活性化している細胞は基本的に各層ひとつと考える。

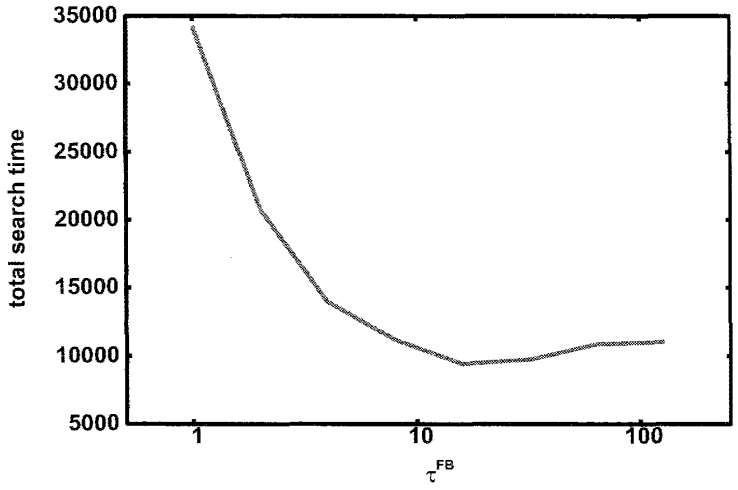


図 9: 探索時間のタイムスケール依存性
図 7 と同様に $\tau^{cell,FF}$ を固定し τ^{FB} を動かしたときの最大容量の変化を各 τ^{FB} について 100 個の相空間の平均値を求めプロットした。横軸は τ^{FB} を対数プロットで $\tau^{FF}=64$ に固定。
τ^{FB} が τ^{cell} に近付くと急速に遷移時間が増加している反面、τ^{FF} に近付いてもあまり遷移時間は増加していない。

合がありうる。以上のように遷移と不安定化・安定化の時間スケールが分離していないことでターゲットの探索がうまくいかず、結果探索時間が長くなることが考えられる。実際に全学習セットを学習するための時間 (主に探索過程のながさに由来する) を調べたものが図 9 であり、上で述べたように探索時間が τ^{cell} 側で大きく増加していることが分かる。このように τ^{FB} が速い側で探索時間そのものがおおきく増加した結果、過去のターゲットパタンに滞在する時間が増加する。これにより記憶容量が減少すると考えられる。

2) 遅い側について

上で述べたように $\tau^{FF,FB}$ は不安定化・安定化の時間スケールだとみなせる。したがって τ^{FB} が $\tau^{FF} \gg \tau^{cell}$ に近付くと遷移の時間スケールと不安定化・安定化の時間スケールが分離するので、探索過程においてパタンの滞留時間がパタン間の遷移時間に比べ長くなる。また過去に学習したターゲットパタンは未学習のパタンに比べ探索過程中に出現する確率が高いと考えられる。これらの理由から τ^{FB} が遅い側では、より過去に学習したターゲットパタンの滞在時間が増加し、記憶数が減少すると考えられる。言い方をかえると過去のパタンに' 固執' し過ぎる事で記憶が減衰してしまうのである。

3.3.1 記憶の壊れ方について

上で述べたように今モデルでは学習したパタンを全て記憶できるわけではない。ここでは学習パタンが記憶できなくなるときにどのように記憶が壊れるかについて述べる。

記憶できないパタンの壊れ方は最大の記憶容量をもつ $\tau^{FB}=16$ をはさみ異なる。ここでターゲットパタン ξ^a で活性化する隠れ層、出力層の細胞を $x_a^{hid,out}$、一番新しい学習ターゲットパタン ξ^{latest} で活性化する細胞を $x_{latest}^{hid,out}$ とする。τ^{FB} が $\tau^{FB}=16$ より大きい時間スケールではターゲットパタン ξ^a が記憶できないときに代わりに活性化する偽記憶パタンは隠れ・出力層ともに 1 細胞が活性化するパタンである。さらに活性化する細胞は隠れ層では決まっていて x_a^{hid} である (出力層では任意の 1 細胞である。)。一方 $\tau^{FB}=16$ より小さい時間スケールの領域では隠れ・出力層ともに 2 細胞が活性化する偽記憶パタンが現れる。このとき活性化する細胞は隠れ・出力層ともにターゲットパタン ξ^a で活性化すべき $x_a^{hid,out}$ と各層細胞 1 つ $x_{else}^{hid,out}$ の組合せである。x_{else}^{hid} は任意の細胞であるが、x_{else}^{out} は $\tau^{FB}=8$ の時は x_{latest}^{out} となっている。($\tau^{FB}=1$ で

は出力層も任意の1細胞が活性化している。)

このように容量のピークとなる時間スケールを挟んで時間スケールが速い側と遅い側で壊れ方が異なる。$\tau^{FB}=8$ の場合については5章でもう少し詳しく解析を行う。

3.4　まとめ

今章では2章のモデルをもちいた学習過程について解析を行った。本論文での学習はターゲットを探索する過程とターゲットを安定化させる過程からなっていた。そのためまず3.1節で適切なパラメタのもとでターゲットの探索が上手くいくことを確認し、ターゲット探索とターゲット安定を繰り返す学習が実際にできること示した。次に3.2節でこの学習過程により複数の入出力関係が記憶できることを示した。本論文の学習ではターゲットの情報が1次元誤差情報Rとしてのみあたえられ、一つの学習セットを完全に学習した後に次を学習するという逐次学習をおこなっているので1章での要請を満たしている。したがって本論文の学習モデルで要請を満たした学習を行えることを示したことになる。このような事が可能になった今モデルの性質にかんしては最後に議論する。最後の節でこの学習モデルにおいて時間スケールが重要であることを述べた。3つの時間スケール $\tau^{cell,FB,FF}$ が適切な関係 ($\tau^{cell}<\tau^{FB}<\tau^{FF}$) を満たすところで最大の記憶数をとり、$\tau^{FB}$ が τ^{cell} に対して速過ぎても、τ^{FF} に対して遅過ぎても記憶が上手くできない。そしてこれらは、速過ぎる側では探索がうまくいかない事から、遅過ぎる側では過去のパタンに'固執'することから起こると考えられる事を述べた。

4　学習による相空間構造の変化

前章までで学習過程においてどのような現象が見られるのかを述べた。学習過程は細胞の活動に依存してシナプスが変化する過程なので、ネットワークの相空間そのものが変化していく。そしてその変化の時間スケール $\tau^{cell,FF,FB}$ の関係によって記憶容量が大きく異なることをみた。今章・次章では学習過程を通して相空間がどのように変化していくか、学習したパタンがどのように埋め込まれ、想起されるのかについて、主に記憶容量が最大にちかい時間スケールの場合について解析を行う。以下2章ともにパラメータは主に $N=30, N_{out}=10, N_{pat}=10, \tau^{FF}=8, \tau^{FB}=64$ とし、このパラメタ以外についての結果にはそれを明記する。また学習前のネットワークとして抑制性結合以外の結合 $J_{ij}^{FF,FB}$ はすべて0であるようなネットワークを用いた学習過程により形成されたネットワークを用いた。

今章・次章の解析は学習過程の各時点(入出力パタンの学習を一つ終えた時点)での相空間を調べることで学習過程による相空間の変化を調べる。したがって今章・次章での解析はシナプス強度は固定し細胞のダイナミクスだけを動かしたものである。まず今章では入力を印加しない状態 ($\eta=0$) での相空間が学習過程でどのように変化していくかについて解析し、次章では今章で解析した入力のない相空間に入力を印加した時にどのように相空間の構造が変化し想起がなされるのかを解析する。

詳しい解析を行う前に今章の構成について説明する。まず始めに学習過程によって入力を印加しない相空間構造がどのように変化していくか概観を述べる。その後4.1節では、多数の学習過程の統計的な結果を元に相空間構造の変化を解析する。そして4.2節では学習過程の各段階について具体例を参照しつつ詳しく解析する。最後にそれらを踏まえ4.3節でまとめを述べる。

学習前のネットワークはFF,FB結合共に0としたので相空間の構造は極めて簡単である。この簡単な相空間構造が学習が進むに連れて以下のように変化していく。まず大きな1つの固定点しかない相空間が学習がすすむと徐々に複数の固定点アトラクタのベイシンに分割され、さらに学習がすすむとリミットサイクルが出現する。またアトラクタに落ちるまでの軌道に着目すると学習がすすむにつれて軌道の長さが長くなり複雑になる。長くなった軌道はターゲットパタンの近傍に近付くよう形成され、その意味でターゲットパタンは軌道に埋め込まれる。学習がすすむと出てくるリミットサイクルもターゲットパタンの近傍をとおる

軌道を形成し、その結果ターゲットパタンを変遷するようなリミットサイクルが現れる。以上のように学習により相空間構造は

・アトラクタは固定点からリミットサイクルになり

・軌道は長くなりターゲットパタンを変遷するよう

に変化する。図 10 はネットワークの出力層細胞の時系列をいくつかの学習段階から選びプロットしたイメージ図である。

このように相空間が変化する機構は以下のように考えられる。3 章でも述べたが, 今モデルでは FF 結合にターゲットパタンの履歴が残っている。すなわち学習したターゲットパタンを ξ^a とすると $x_a^{in} \to x_a^{hid}, x_a^{hid} \to x_a^{out}$ の FF 結合が強化され、$x_a^{in} \to x_{else}^{hid}, x_a^{hid} \to x_{else}^{out}$ の FF 結合は弱くなる ($x_{else}^{hid,out}$ は $x_a^{hid,out}$ 以外の隠れ・出力層細胞)。今章の解析では入力は印加されていないので入力層と隠れ層間の FF 結合は無視する。このとき強化された FF 結合 (それぞれ $J^{FF}_{x_a^{out}x_a^{hid}}$ と表記する。他の結合も同様) の値が他の結合 $J^{FF}_{x_{else}^{out}x_a^{hid}}$ の値よりも大きければ x_a^{hid} が活性化した時に x_a^{out} も活性化されることが期待される。また x_a^{out} が活性化することで他の出力層細胞は抑制される。このように $J^{FF}_{x_a^{out}x_a^{hid}}$ が大きければ x_a^{hid} を活性化させる状態は同時に x_a^{out} を活性化させる。この状態はターゲットパタン ξ^a が活性化した状態に他ならず、ネットワークの状態はターゲットパタン ξ^a に近付く。しかし x_a^{out} からの FB 結合にはターゲットパタン ξ^a の履歴は残っていないので $J^{FB}_{x_a^{hid}x_a^{out}}$ が強化されている訳ではない。従って x_a^{out} が活性化すると次は探索の履歴に応じて強化されている結合[18]の postsynaptic cell x_i^{hid} が活性化する。この細胞が別のターゲット ξ^b の x_b^{hid} であれば、$J_{x_b^{hid}x_b^{out}}$ が強化されているので、今度はターゲット ξ^b が活性化する。このようにネットワークの状態は学習したターゲットパタンに一度近付いたのちに他のパタンに遷移する構造をもつ。学習が進むとこの遷移構造が連結されることで複数の学習したターゲットパタンを遷移する軌道がうまれてくる。あるいはこの構造が複数つながった結果また初めのターゲットパタンとつながることがありえて、このときはターゲットパタンを遷移するリミットサイクルが形成されると考えられる。

以上の機構を力学系的な見方で捉え直してみる。今モデルでは学習はターゲットの探索と収束したターゲットパタンの安定化により行われる。この機構により学習したターゲットパタンは一度安定な構造をしたのちに、その後の別のターゲットを学習するときの探索で不安定化される。ただし最後の議論で詳しく述べるがこの不安定化は全方向で不安定化するわけではなく、一部が不安定化すると考えられる。したがって学習したターゲットパタンは多くの方向からは収束するが一部に不安定多様体が伸びているサドル様の構造をとる。そして不安定多様体がつながることでターゲットパタン付近を遷移する構造が生まれ、また軌道が周期的になることでターゲットパタンを遷移するヘテロクリニック的なサイクルが現れてくると考えられる。このように今モデルの学習を行うことで自然と学習したターゲットパタンを遷移するような構造ができあがることが判る。

以下でこれらの構造の変化を詳しく解析していく。4.1 節では上で述べた相空間の変化を、学習過程に沿って解析し、4.2 節では学習過程をいくつかに分けて、その段階ごとに解析を行う。

4.1 相空間の変化

今章のはじめにどのように相空間変化がおこるか概観した。今節ではこの変化をふたつに分けて解析を行った。4.1.1 節では相空間構造の変化そのものを解析し、続く 4.1.2 節ではその変化と学習したターゲットパタンとの関係について解析した。

[18] この結果たまたま x_a^{hid} が安定になるとターゲットパタン ξ^a が安定固定点になる。

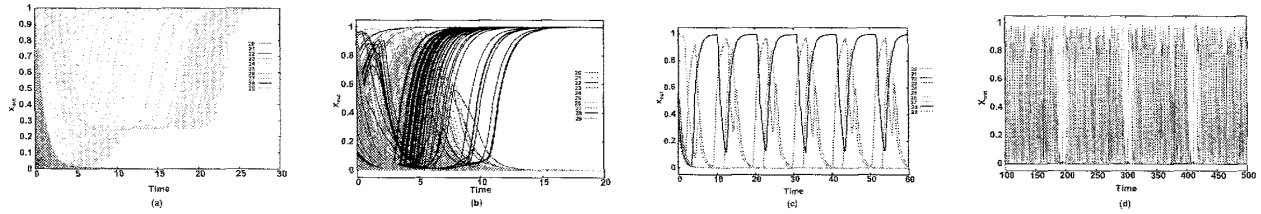


図 10: 学習による相空間の変化:
学習過程が進むに連れて (a) 単純な軌道を描き、1 つのアトラクタに収束する相空間 → (b) より複雑な過渡期を経てから複数のアトラクタに収束する相空間 → (c) 少数のパタンからなるリミットサイクル → (d) 多数のパタンからなるリミットサイクルというように相空間はより"複雑"に変化していく。図にはそれぞれ学習段階 1,5,6,9 の時のネットワークの出力層時系列をプロットした。

4.1.1　アトラクタと軌道の長さの変化

学習していくにしたがって細胞ダイナミクスの相空間がどのように変化していくかを調べるために、アトラクタ数、ベイシンエントロピー、リミットサイクルの存在する割合、アトラクタに収束するまでの軌道の長さを各学習段階の相空間で計ったものが図 11 である。

アトラクタ数は各学習段階のネットワークにおいて細胞の初期点を多数とったときの収束先の数としてはかり、同じ学習段階の多数のネットワークで平均した。したがってここでは不安定なアトラクタや basin volume が極めて小さなアトラクタは考慮されていない。この値を各学習段階の相空間でのアトラクタ数として図 11(a) にプロットした。

ベイシンエントロピー S を、情報エントロピーの確率分布を各アトラクタの basin volume 分布に置き換えて計算した $S = -\sum_i vol_i * log(vol_i)$ として定義する (i はアトラクタの index)。S=0 のときは相空間のどの点から出発しても同じアトラクタに収束する状態で、S が大きくなるとアトラクタ数が増加し相空間が多数に分割されている状態を示す。したがって S が大きいほどランダムに初期点を選んだときにどのアトラクタに収束するかが分からない事を意味する。この S が学習によってどう変化するかを各学習段階の相空間の S を計り図 11(b) にプロットした。

リミットサイクルの存在する割合は各学習段階のネットワークの相空間のうち、リミットサイクルの存在する相空間の割合である。この割合と存在するリミットサイクルの basin volume の大きさをプロットしたものが図 11(c) である。実線がリミットサイクルが存在する割合でこの値が 1 のときはこの学習段階ではどの相空間でも必ずリミットサイクルがあることを意味する。また点線はリミットサイクルの basin volume の全試行での平均値を示しているので、存在するリミットサイクル全てが basin volume = 1 をもつと完全に実線に一致する。

図 11(d) は初期状態から固定点に収束するまでの軌道の長さの平均値をプロットしたものである。軌道の長さの平均値は相空間上に初期点をランダムに多数を選び、それらがアトラクタに収束するまでの長さを初期点について平均することで求めた。ただし、リミットサイクルなど固定点以外のアトラクタに収束するものは軌道の定義が難しいので、固定点に収束する初期値のみについて計算した。学習後半はリミットサイクルが basin volume に占める割合が大きくなるので、固定点に収束するものが小さくなり分散が大きくなっている。以上 4 つの図の横軸は全て学習段階であり、学習がすすむにつれてどのように相空間が変化していくかの指標になっている。この 4 つの指標をもとに学習過程に置ける相空間の変化を議論する。

学習の最初期の相空間には少数の固定点のみが存在し、ベイシンエントロピーも低い。次節で詳しく述べるが大きな basin volume をもつ 1 つの固定点アトラクタが存在し残りは小さな固定点アトラクタ、という構造ができている。またその固定点への収束する軌道の長さをみると最初期の相空間での平均長が一番短かい。したがって相空間上のほとんどの初期点が 1 つの固定点に向かって比較的速やかに収束する相空間になっている事がわかる。

学習が進むにつれ (学習セット数が 3,4)、相空間上の固定点アトラクタの数は急速に増加し、それに伴いSも大きく増加する。単に数が多くなるだけでなく比較的大きな basin volume をもつ固定点が複数個存在する相空間が形成される。また収束するまでの軌道の長さも学習が進むに伴い単調に長くなり、複雑な軌道を形成する。このように初めほとんどの初期点が単一の固定点に向かって速やかに収束するような相空間が、入出力セットを学習することで収束先の固定点アトラクタが多様化しまたそこへ収束する軌道も長く歪むような相空間に変化する。

さらに学習がすすみ学習過程の後半 (学習セット数が 6 を越える当たり) の相空間からは固定点アトラクタのかわりにリミットサイクルの個数が増加し、リミットサイクルの平均 basin volume もそれに伴い増加する。リミットサイクルの割合と平均 basin volume は単調に増加しているので、学習が進めば進むほどリミットサイクルが出やすく、また相空間で大きな volume を占めるようになっているのがわかる。一般にリミットサイクルは固定点アトラクタに比べ 1 つ当たりの basin volume が大きいため個数はそれほど多くなく、固定点アトラクタ・リミットサイクル合わせたアトラクタ全体の数としては減少する。これに伴いSも減少していく。リミットサイクルが存在するネットワークにおいても固定点アトラクタは存在するが、個数・basin volume ともに減少する。したがって多くの初期状態からはリミットサイクルに収束しパタンを変遷する軌道におちる。またリミットサイクルの軌道がどのように変化するかを調べると、固定点アトラクタに収束する軌道と同じく学習が進むに連れてより軌道が長くいり組んだものになっていく。この点に関して 4.2 節で詳しく解析を行う。

以上のように、初め 1 つの固定点とそこに向かう歪みの少ない軌道しか持たない単純な相空間が、学習を経て固定点アトラクタの数が増え歪みをました複雑な軌道をもつ相空間に変化する。さらに学習がすすむとリミットサイクルをもつ相空間が出現し、リミットサイクルの軌道も固定点への収束軌道と同様に長くなりその意味で複雑な相空間になっていく。

4.1.2 相空間の変化と学習したターゲットとの関係

学習が進むにつれてどのように相空間が変化していくか、アトラクタと軌道の長さの点から述べた。ここではこのような相空間の変化と学習したターゲットパタンにどのような関係があるかを解析し、章冒頭でも触れた通りターゲットパタンは軌道に埋め込まれる事を示す。

今章で注目しているのは入力 0 の相空間なので学習したターゲットパタンがその相空間上で固定点アトラクタとなっているかは自明ではない。したがってまず学習したターゲットパタンの入力 0 の相空間上での安定性を調べた (図 12)。各学習段階の相空間についてターゲットパタンの basin volume の分布を計り重ね書きしたものが図 12(a) である。ここでの basin volume は学習したターゲットパタンを別々に計ったものではなく、学習したターゲット全ての basin volume の和を計ったものである。したがって vasin volume=0 でのピークは学習したターゲットパタンが固定点アトラクタとして一つも存在していないネットワークを意味する。一方でターゲットパタンを別々に計った、安定なターゲットパタンをもつネットワークの頻度分布図が図 12(b) である。横軸は学習したターゲットパタンの index をとり (小さいほど初期に学習したパタンを意味する)、縦軸にそのターゲットパタンが安定な相空間の割合をとり、各学習段階の相空間について分布をプロットした。この分布ではターゲットパタンがどんなに小さくても安定であればよいとしてカウントしている (図 12(a) において $vol \neq 0$ の分布に対応している)。学習のどの段階においても多くのネットワークでターゲットパタンは小さな basin volume をしめるに過ぎず、そもそも固定点アトラクタですらないネットワークも多く存在する事が図 12(a) から判る。また、学習した順序の依存性を見ると最も新しい学習ターゲットパタンが一番安定である割合が大きく、覚えたパタンが過去であるほど安定である割合は小さくなっていく。しかし一番安定なターゲットパタンでも安定なネットワークの割合は 0.5 を越えない。学習したターゲットパタンがあまり固定点アトラクタとして埋め込まれていない一方で、学習セット数が小さいときは学習したターゲットパタンをほぼ完全に記憶できている. このことは入力がないときの相空間におい

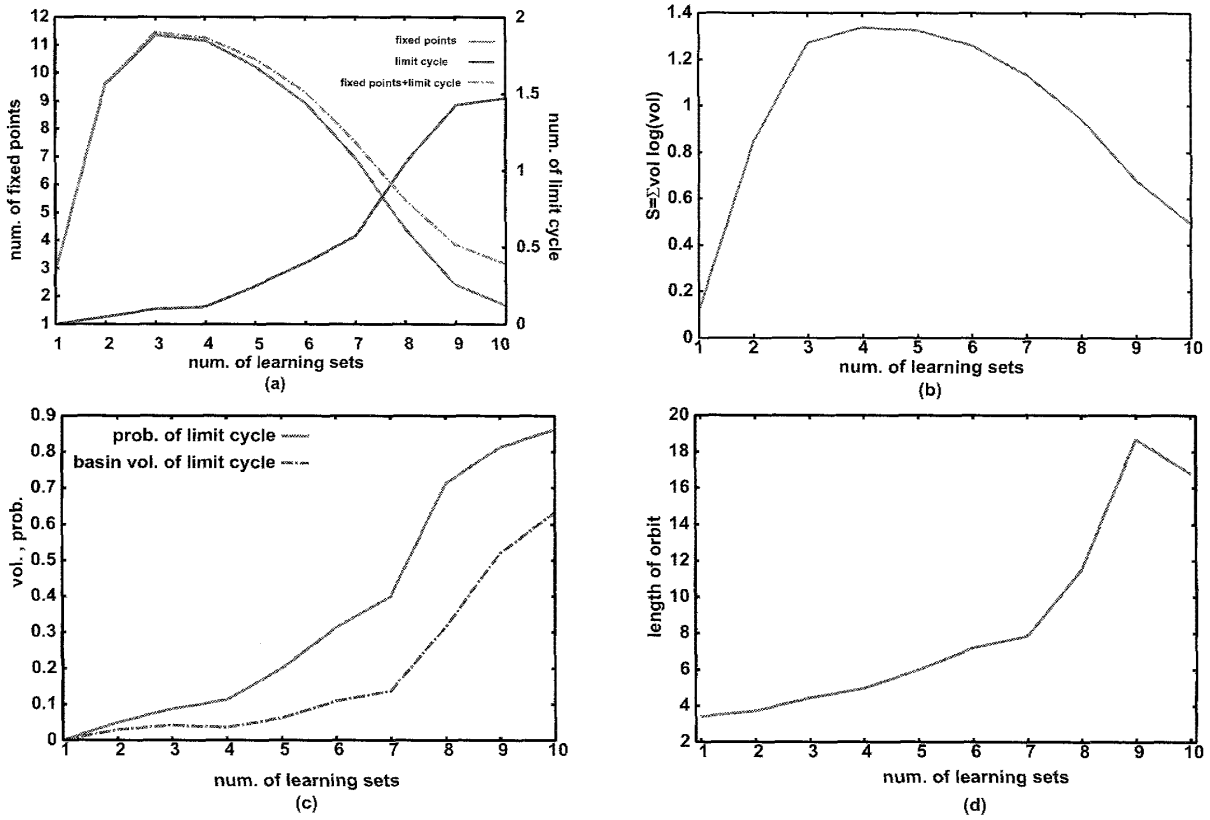


図 11: 学習による相空間の変化:
各学習段階のネットワークについて入力を印加しないときの相空間構造を調べた。各図は下で説明する各指標を 80 本の独立な学習過程について平均したもの。学習試行は細胞の初期値・学習する入出力パタンが異なる。学習前のネットワークは $J^{FF,FB} = 0$ で全て共通である。横軸は調べた相空間の学習段階を示している。
(a) アトラクタ (固定点、リミットサイクル) の個数の学習過程での変化: 多数の細胞の初期状態に対してその収束先をもとめ、分類・カウントした。左の縦軸が固定点とアトラクタ全体の数、右の縦軸がリミットサイクルの数に対応する。
(b) ベイシンエントロピー $S = -\sum_i vol_i * log(vol_i)$ の学習過程での変化: (a) で求めたアトラクタについて basin volume vol_i を求めベイシンエントロピーを計算した。
(c) リミットサイクルが存在する確率 (実線) とリミットサイクルの basin volume (点線): ここでリミットサイクルが存在する確率を 80 試行のうちどれくらいの試行でリミットサイクルが存在するかという割合として計った。またそのときのリミットサイクルの basin volume の 80 試行での平均を点線としてプロットした。存在するリミットサイクルがすべて basin volume = 1 であれば実線と点線は一致する。
(d) 初期状態から固定点に収束するまでの軌道の平均長: 多数の初期値に対して、固定点に収束するまでに相空間中を移動する距離の平均値を求め軌道の平均長とした。

各々指標から, 相空間が学習によって固定点からリミットサイクルへ、簡単な軌道から複雑な軌道へと変化していくことが判る。

てターゲットパタンが固定点アトラクタかどうかという事が、そのターゲットパタンが記憶できているかどうかということと完全な対応関係はないということが判る。

次に収束するまでの過渡期の軌道、リミットサイクルの軌道と各ターゲットパタンとの距離を調べた。ここでは軌道とターゲットパタン ξ^i との距離 d^i を $d^i = \min_{0<t<\infty} |x(t) - \xi^i|$ と定義する[19]。軌道とターゲットパタンとの平均距離が学習によりどのように変化するかを調べたものが図 13 である (平均距離の定義は図 13 の説明を参照)。縦軸は軌道とターゲットパタンとの平均距離であり、横軸はターゲットパタンの index である。図 12 と同様にターゲットパタンの index は小さいほど過去に学習したターゲットパタンであることを示している。各学習段階の相空間を多数用意し、それらの相空間上の軌道とターゲットパタン (未学習のものも含む) の平均距離をプロットした。各ラインは各学習段階の相空間の平均を表している。調べるターゲットパタンは最終的に学習させるセット全てであり、学習初期のネットワークがつくる相空間からみれば未学習のターゲットパタンも含んでいる。例えば学習セット数が 2 の相空間においては ξ^3 以降は未学習のターゲットパタンという事になる。未学習のターゲットパタンの情報は全く与えられていないのでコントロールとみなすことができる。図 14 では図 13 のように平均距離ではなく、軌道とターゲットから距離を多数のネットワークでそれぞれ多数の初期値について計り分布をとった。図 14(a) は学習段階 4 の相空間での各ターゲットパタンとの距離の分布、図 14(b) は各学習段階の相空間でのターゲットパタン ξ^4 との距離の分布である。

図 13、図 14 から

1) ある学習段階の相空間 (図 13 上のあるライン) に着目すると、学習したターゲットパタンと軌道との平均距離は未学習のターゲットパタンに比べて小さくなっている事

2) あるターゲットパタンを固定してみると学習が進めば進むほど平均距離は小さくなっている事

の 2 点が分かる。学習したターゲットパタンと軌道の距離が小さくなっているということは、学習することでターゲットパタンの近傍に軌道が形成される、言い替えると学習によりターゲットパタンは軌道のトランジェントパタンとして埋め込まれる事を意味している。実際ターゲットパタンと軌道の距離の分布図 14 から、学習したターゲットパタンと未学習のターゲットパタンの距離の分布を比較すると距離が小さいところ (d=0.05) に分布のピークができていることが判る[20]。このピークの高さは過去に学習したものほど高くなっているので、古いパタンほど近付く軌道が増えていることを意味する[21]。またこのピークの上昇があるターゲットパタンに着目したとき、学習が進めば進むほど着目したターゲットパタンとの平均距離が短くなっていることの原因にもなっている。このことはターゲットパタンとの平均距離にターゲットパタンの学習順序が反映されていることを意味している[22]

以上のように固定点に収束するまでの軌道やリミットサイクルの軌道はターゲットパタン近傍を通るようになる事、またターゲットパタンとの平均距離を調べるとターゲットパタンの学習順序も反映されていることが示された。

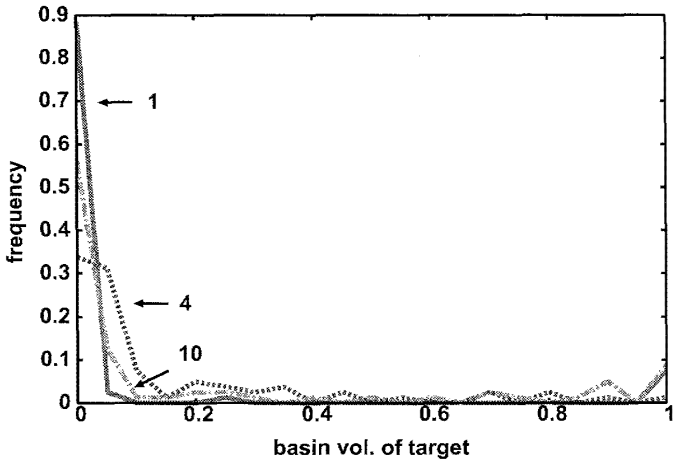
ターゲットパタン近傍を通る軌道の理解のためにに軌道を離散的なパタンの列とみなす解析を後に行う。その解析でもちいる表記をここで定める。今モデルでは activation 関数により 0,1 付近をとりやすくなっているので適当に閾値を設ける事で細胞の状態を活性・非活性の 2 値に離散化でき、相空間を $\{0,1\}^N$ の点に離散化する事ができる。したがって相空間上の軌道も離散化し、$\{0,1\}^N$ の点列とみなすことができる。このように軌道を離散的なパタンの列とみなしたものを遷移シークエンスと呼ぶことにする。またリミットサイクルを表すものとしてこの遷移シークエンスを用いることもあるので、ここで記法を定めておく。リミットサイクルの軌道がパタン a→ パタン b→ パタン a→ パタン b... のときに、このリミットサイクルを (a,b)

[19]この距離は 図 2 3 4 での距離と定義が異なる。これは探索段階でのターゲット Ξ は出力層と同じ次元であるが、ここでのターゲットパタン ξ は隠れ層も含んでいるので次元が隠れ層＋出力層になっている。(2.2 節を参照)

[20]但し、この図からも判るように全ての軌道が学習したターゲットに近付く分けではない。

[21]あくまで近傍を通る軌道が増えているのであり、より近くを通る軌道が現れているわけではない

[22]このように学習がすすむにつれて昔の学習したターゲットパタンへの距離が小さくなることは、過去の学習パタンにたいして何の直接的関与をしない今モデルではかなり非自明な事だが、どのような機構があるのかは判っていない。



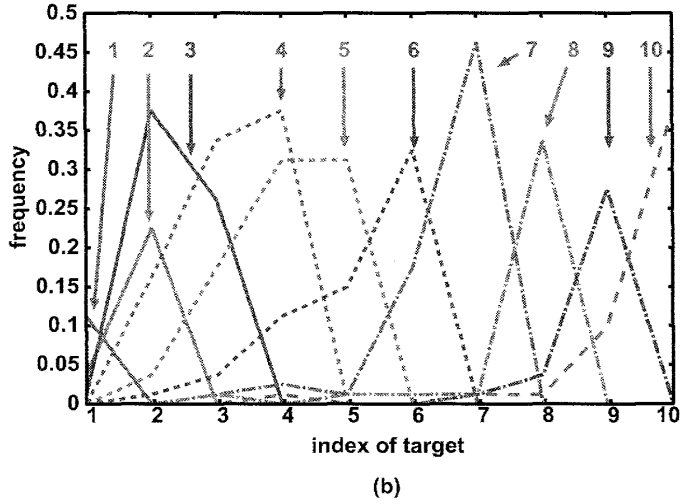


図 12: ターゲットパタンの入力 0 の相空間での安定性:
以下の分布を学習段階ごとに入力 0 の相空間で計った。
(a) 学習段階 1,4,10 の相空間において学習したターゲットパタンの basin volume の分布:
ここでの basin volume は学習したターゲットパタンの basin volume の総和であり、basin volume = 0 のピークは学習したターゲットパタンが全て不安定な構造になっているネットワークの頻度を示す。どの学習段階でも半分程度はターゲットパタンが全て不安定であるようなネットワークになっている。
(b) 各学習段階でのターゲットパタンが安定なネットワークの頻度分布:
各ターゲットパタンについて、そのパタンを安定にもつネットワークの頻度を計った。安定なターゲットパタンは過去に行くほど急激に減衰していく事が判る。特に学習段階が後半になると直前に学習したターゲットパタンしか安定なターゲットパタンが存在しないというネットワークが多い。

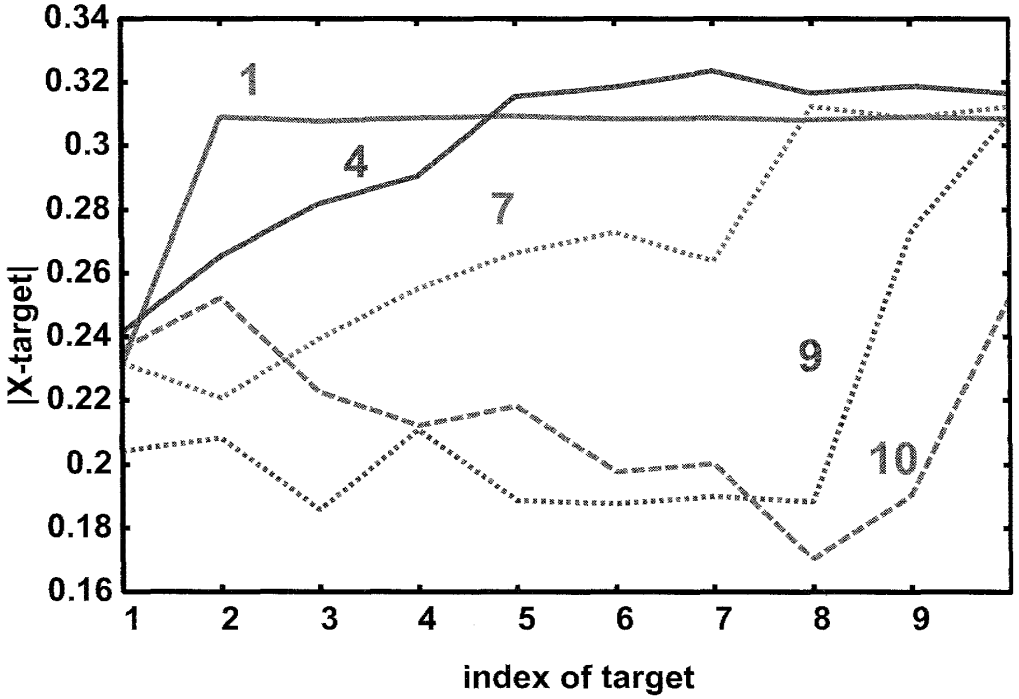


図 13: 学習によるターゲットパタンと軌道との距離の変化:
学習がすすむにつれて、各ターゲットパタンとの平均距離がどのように変化するかを調べた。軌道とターゲットパタン ξ^i との距離 d^i を $d^i = \min_{0<t<\infty}|x(t) - \xi^i|$ と定義する。初期状態 i からアトラクタに収束するまでの軌道 (リミットサイクルの場合はその軌道も含める) と各ターゲットパタン ξ^j との距離を $d_i^j (j = 1, \ldots, Npat)$ とする。d_i^j を多数の初期状態 i に関して平均したものをターゲットパタン ξ^i との平均距離 $\bar{d}^i$ として定義し、相空間の指標とする。さらに $\bar{d}^i$ を各学習段階の多数の相空間で平均したものをここでの平均距離とした。図の数字は学習段階を示す。横軸に何番目に学習したターゲットパタンかを表す index をとり、そのパタンとの平均距離をプロットした。

学習したターゲットパタンは未学習のターゲットパタンに比べ距離が小さくなっている。また学習がすすむにつれてその距離は小さくなっている。このことから、学習によりターゲットパタンのより近くをとおる軌道が形成されている事・学習の順序関係が平均距離に反映されていることが判る。

と表記する。またリミットサイクルの軌道が離散化したときは同じくパタン a→ パタン b→ パタン a→ パタン b... だが実際の連続系での挙動をみるとパタン a→ パタン b で 1 周期ではなくパタン a→ パタン b→ パタン a→ パタン b ではじめて 1 周期のリミットサイクルも存在する。このようなリミットサイクルは $(a,b)^2$ と表記する[23]。もっと複雑になりパタン a→ パタン b→ パタン a→ パタン b→ パタン c→ パタン a→ パタン b→ パタン a→ パタン b でようやく 1 周期のリミットサイクルは $((a,b)^2, c, (a,b)^2)$ と表記する。

4.2 各学習段階での相空間の構造

4.1 節では相空間の変化を学習過程にそって解析を行った。今節では学習過程を学習前・学習初期・学習前期・学習後期に分類し、その分類ごとに詳しく相空間の構造を解析して行った。

[23](a,b) と $(a,b)^2$ の違いは離散的に見るだけではわからないので、相空間上の実際の軌道をみて判断する。

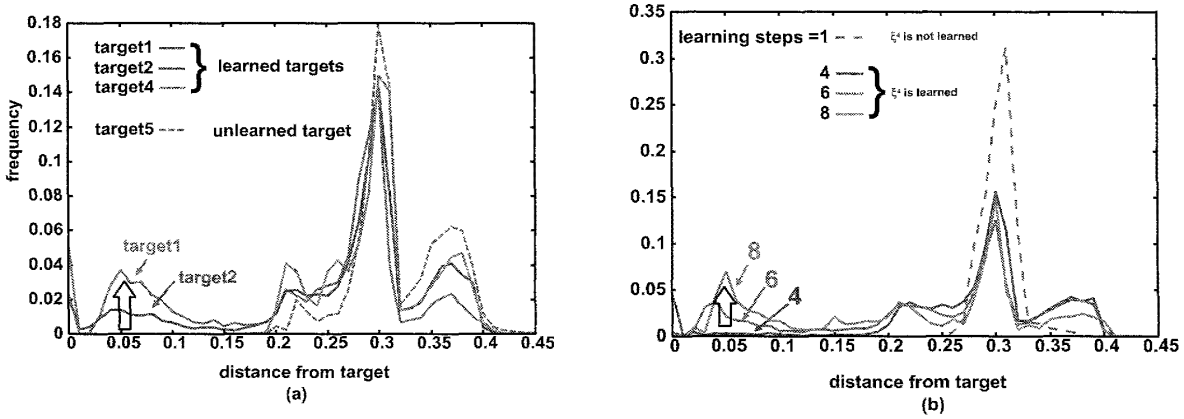


図 14: ターゲットパタンと軌道との距離の分布図:
適当な初期値からスタートする軌道とターゲットパタンとの距離を、ランダムに選んだ初期値に付いて計り分布を取った。
(a) 学習セット数が 4 のときの相空間におけるターゲットパタン $\xi^{1,2,4,5}$ と軌道の距離の分布図。ラインの違いはそれぞれ距離を計ったターゲットパタンの違いを意味する。調べた相空間は学習セット数 4 なので、ターゲットパタン $\xi^{1,2,4}$ は既に学習したターゲットパタンであり (学習したターゲット全てをプロットすると繁雑なので ξ^3 は省略した。)、ターゲットパタン ξ^5 は未学習のターゲットである。学習したパタンと学習済みのパタンを比較すると、d=0.05 付近のピークが学習することで形成されている事が分かる。また過去の学習パタンであるほどこのピークが高くなっていることが分かる。このことが、 図 13 において学習がすすむにつれてターゲットとの平均距離が小さくなる要因になっている。
(b) 学習セット数が 1,4,6,8 の相空間におけるターゲットパタン ξ^4 と軌道の距離の分布図。ラインの違いはそれぞれ分布をとった相空間の学習セット数の違いを意味する。学習段階 (図中の learning steps) が 1 の相空間においてはターゲットパタン ξ^4 は未学習のパタンであり、学習段階 4,6,8 の相空間に置いては ξ^4 は学習済みのパタンとなる。学習段階が進むに連れて、ピークの大きさが大きくなっている。

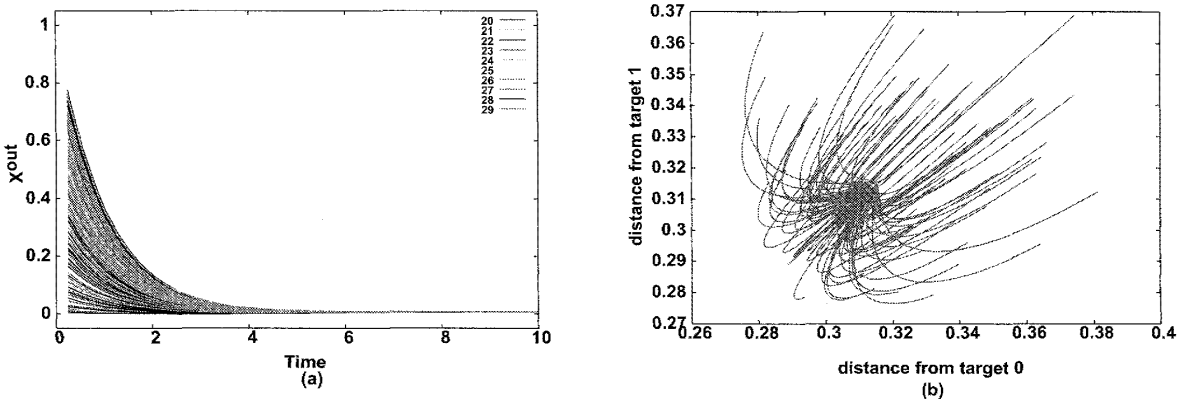


図 15: 学習前のネットワークの相空間構造
学習前のネットワークについて 100 個の初期値をランダムに選び、
(a) 出力細胞 (i=20...29) の時系列
(b) ターゲットパタン $\xi^{1,2}$ との距離をそれぞれ横軸、縦軸に取った軌道
を重ね合わせてプロットした。
一つの固定点に向かって単調に収束する軌道のみの簡単な相空間構造をもっている。

4.2.1 学習前の相空間構造

学習前のネットワークは抑制結合以外は全て 0 なので、各細胞への興奮性の入力も 0 である。したがって抑制性の入力がなければ全ての細胞が自発発火率 r に収束する。つまり $x_i = r \quad (i \in$ 隠れ・出力層) が唯一のアトラクタとして存在する。今回は抑制結合は-1 なのでそれを考慮した値 ($x = 1/(1+exp(-\beta J^{-}(N/3-1)x+\theta))$ の解; 現在のパラメタでは 0.0065366) がアトラクタになっている。図 15(a) は学習前のネットワークの出力層細胞の時系列を重ね書きした図であり、(b) ターゲットパタン ξ^1 とターゲットパタン ξ^2 との距離で張られる 2 次元平面への射影した軌道を重ね書きした図である。ともに多数の細胞の初期値から出発した軌道が速やかに固定点に収束していることを示している。このように学習前の相空間構造は一つの固定点アトラクタとそこへ単調に収束する簡単な軌道からなる事がわかる。

4.2.2 学習過程初期の相空間構造:大きな一つの固定点アトラクタ

1 セットを学習した後の相空間がどのようなアトラクタ、軌道をもつ相空間かを調べる。1 セットを学習後の相空間のアトラクタ数は図 11(a) では固定点アトラクタが平均で 3 個程度である。しかし各相空間の固定点アトラクタの basin volume のランクを横軸に、basin volume を縦軸にとった図 16(a) に示されているように、3 つの内訳は basin volume が 0.9 を越えるような大きな吸引領域を持つ 1 つの固定点と、basin volume が 0.1 以下の小さな固定点からなっている。このように実際はほとんど一つの固定点アトラクタが存在する相空間になっている。おおきな固定点は全ての学習段階 1 の相空間に共通した、2 タイプに分けられる。この固定点の一つのタイプは隠れ層の細胞は活性化しているものがなく、出力層の細胞も大きな値をとらない固定点ともう一つは隠れ層、出力層ともに活性化している固定点である。どちらの場合もターゲットパタン ξ^1 自体はアトラクタにはなっていない。実際これらの固定点 (他の小さな固定点の多くも) は出力層のターゲット細胞 x_1^{out} が活性化する一方で、隠れ層はターゲット x_1^{hid} が活性化するものはほとんどない。これらの軌道とターゲットパタン ξ^1 との関係を見るために, 各ターゲットパタン $\xi^i (i = 1, \ldots, N_{pat})$ との距離の時系列をプロットしたものが図 17 (a) であり、ターゲットパタン $\xi^{1,2}$ の距離への射影が図 17 (b) である。固定点にいたる軌道は図 17 (b) から分かるように一般的に大きく分けて P1 に集まってから固定点に収束する軌道と P2 に集まってから収束する軌道がある。P1 を経由する軌道は一度原点付近を通りその

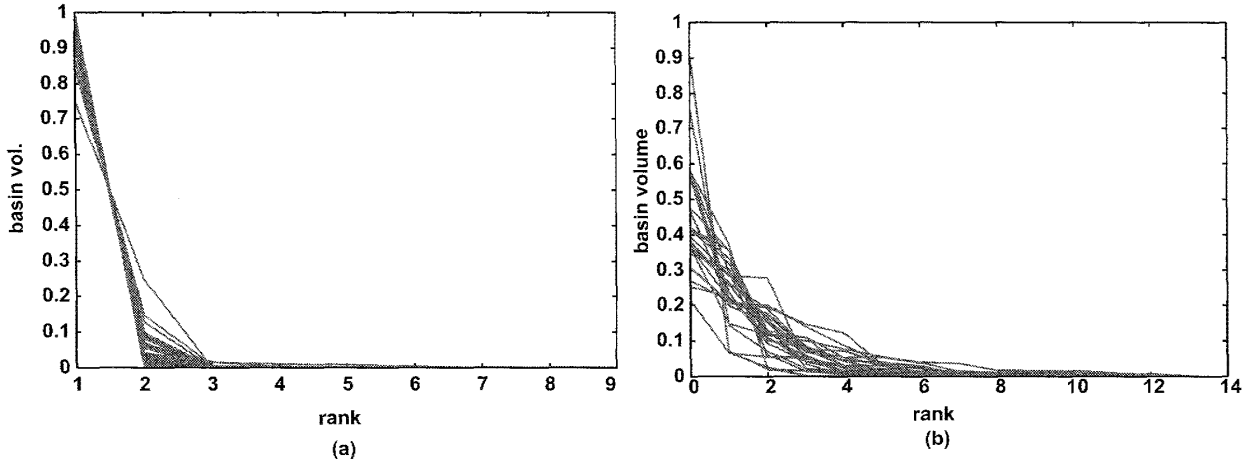


図 16: アトラクタの basin volume 分布:
各相空間のアトラクタを basin volume が大きい順に並べプロットした。一つのラインは一つの相空間を意味する。横軸にランク、縦軸に basin volume を取った。
(a) 学習セット数 1 の相空間を 20 個調べ重ね書き
(b) 学習セット数 4 の相空間を 20 個調べ重ね書き

学習過程の初期では basin volume がおおきい 1 つのアトラクタと他に小さい少数のアトラクタという相空間である。それに対して学習が進んだときの相空間の例として調べた学習セット数が 4 の相空間にはいくつかの大きな basin volume をもつ主要なアトラクタが存在している。また、アトラクタ数自体も増加している。

後固定点に収束する軌道に対応している。原点付近に収束する軌道は学習前の軌道の構造と類似していて、興奮性のシナプス $J^{FF,FB}$ の強度が小さいことと相互抑制の影響が効いている。すなわち興奮性の入力は小さい一方で、抑制結合はパラメタではじめから-1.0 をとるので細胞の状態としてランダムな初期値をとるとその値は減少する傾向にある。P2 を経由する軌道は一度出力層のターゲット細胞が活性化してから収束する軌道に対応し、学習セット数が 1 の段階で相空間には学習したターゲットパタンが軌道の経由点として埋め込まれている事を示している。ただし出力層ターゲット細胞 x_1^{out} への結合は大きいために活性化しやすいが、隠れ層ターゲット細胞 x_1^{hid} への結合はそこまで大きくないために活性化しない。このことによりこの段階では出力層ターゲット細胞のみが活性化し隠れ層ターゲット細胞は活性化しないためターゲットパタンとの距離はあまり小さくならない。このように出力層のターゲット細胞だけが活性化するパタンを経由するときの軌道とターゲットパタンとの距離が分布図 (14) において d=0.2 付近のピークと対応している。

以上のように 1 セットの学習後の相空間は軌道が一度学習したターゲットパタンを通るように歪み、最終的にほとんど一つの固定点に収束する構造をもつ事が判る。

4.2.3　学習過程前期の相空間構造:複数の固定点アトラクタ

学習したセット数が増えるにしたがって固定点アトラクタの数は増加する。basin volume の分布 (図 16(b)) をみてもセット数 1 の場合のように大きな basin volume をもつ固定点が一つだけではなくていくつかの大きな固定点が形成される。学習過程がすすむにつれて大きな一つのアトラクタしかない相空間がいくつかの固定点の basin に分割されていくことがここでも判る。そしてこの分割のされ方は入力-ターゲットの組合せの違い・提示順序で異なるのはもちろんだが、最初のセットを学習させるときの細胞の初期条件によっても異なる。すなわち初めの学習時の細胞の初期条件だけが違う学習過程でも分割のされ方は大きく異なる。これは同じ学習セットが提示されても初期条件が異なると探索をおこなう軌道が異なるので、それによ

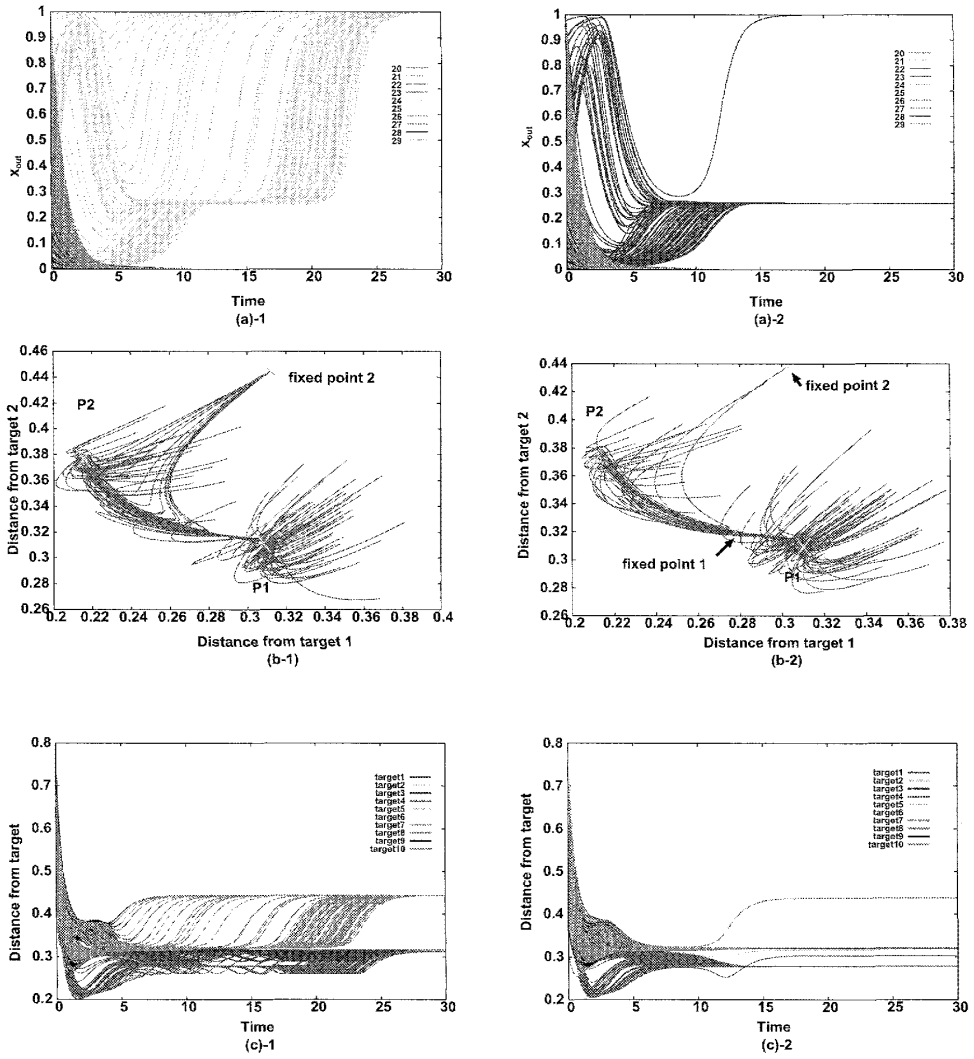


図 17: 1 セット学習後の相空間:
ランダムに初期値を 100 個選び、(a) 出力細胞の時系列、(b) ターゲットパタン $\xi^{1,2}$ との距離をそれぞれ横軸、縦軸に取った軌道、(c) 各ターゲットパタンとの距離の時系列、を重ね合わせてプロットした。
1 セット学習後の相空間には 2 種類ありそれぞれ (a)(b)(c)-1 が隠れ層、出力層ともに活性化している固定点が存在する相空間、(a)(b)(c)-2 が隠れ層が非活性で、出力層も大きな値を取らない固定点が存在する相空間、である。

(a)(b)(c) よりほとんど全ての初期状態からの軌道が 1 つの大きな固定点アトラクタに収束しているのが判る。P2 付近を通る軌道はターゲット ξ^1 に一度近付いてから固定点に収束している。また 固定点 2 に収束する図 1 の相空間でも一度固定点 1 近傍を経由している。

り形成される相空間も異なることに起因していると考えられる。また過渡期の軌道も長くなり、いくつかのターゲットパタンを遷移した後に固定点に収束するような軌道も見られるようになる。複数のターゲットパタンを遷移する場合はターゲットパタン ξ^a での隠れ層活性化細胞 x_a^{hid} → 出力層活性化細胞 x_a^{out} → ターゲットパタン ξ^b の隠れ層活性化細胞 x_b^{hid} → 出力層活性化細胞 x_b^{out} → ... となり学習セットが 1 のときとは異なり、隠れ層・出力層の細胞が同時に活性化するのでターゲットパタンとの距離がより小さくなる。このときの軌道とターゲットパタンとの距離が 0.05 前後であり距離の分布図 14 の d=0.05 付近のピークに対応する。複数のターゲットパタンを遷移するといっても、一つの軌道が過去の学習パタンを全て遷移するのではなく、ある軌道はターゲットパタン ξ^1 をとおり、またある軌道はターゲットパタン $\xi^{2,3}$ を通るという様に軌道ごとに近付くターゲットパタンは異なる。さらに全ての軌道を考慮しても必ずしも全ての学習パタンを網羅するとは限らない[24]。具体例を見る。図 18 では学習セットが 5 の場合の相空間について、(a) 出力層のダイナミクス・(b) 各ターゲットパタンとの距離の変化、(c) 細胞のダイナミクスをターゲットパタン $\xi^{2,4}$ との距離で張られる 2 次元平面に射影したものをそれぞれふたつの固定点 1,2 に分けプロットした。固定点への軌道に注目すると固定点 1 への軌道はターゲットパタン $\xi^{2,3,4}$ に近付いた後に収束するのに対して、固定点 2 への軌道はターゲット ξ^1 にのみに近付いたのちに固定点に収束している。しかしターゲットパタン ξ^5 に近付くものは、すくなくともサンプルした限りでは見付からないので存在しても極めてレアな軌道である。このように複数の固定点アトラクタがあり、そこへの収束軌道は各々異なったターゲットパタンを経て固定点に収束する。またどの軌道も近傍に集まらないターゲットパタンも存在する。このときのターゲットパタンの遷移順序とターゲットパタンの学習順序には明確な関係は認められない。これはフィードバック結合が学習の度に大きく変化し前回までの履歴を残さないことが効いていると考えられる。ただし前にも述べたが平均距離の図にあるように多数の学習過程にたいして平均をとると学習が古いほどターゲットパタンとの距離が小さくなることから、学習順序とターゲットパタンの遷移順序は無関係ではない事は示唆される。

さらに学習が進むと一つの軌道に同じターゲットパタンを複数回経由するような軌道も出てくるが、リミットサイクルでも同じような構造ができるためそこで議論する。

4.2.4　学習過程後期の相空間構造:リミットサイクル

これまで述べたように学習過程の後期では一般的にリミットサイクルをもつようなネットワークの相空間が形成される。今小節では後期に見られるリミットサイクルについてのべる。

リミットサイクルの軌道は章冒頭で見たようにパタンの遷移シークエンスとみなせる。この遷移シークエンスが学習が進むにつれてどのように変化するかについて解析した。遷移シークエンスの長さが学習によりどのように変化するか、学習の各段階毎に長さの分布、平均値について調べたものがそれぞれ図 19(a)(b) である。この図から学習が進むにつれてリミットサイクルの遷移シークエンスは長くなることが分かる。また遷移シークエンスの長さの分布を見てみると学習が進むにつれて値がシークエンスの長さが 4,5 付近のピークが大きくなるとともに長周期のリミットサイクルが出現していることも分かる。このようにリミットサイクルの場合も固定点アトラクタの場合と同じく軌道の長さが徐々に長くなっていく事が分かる。

初期段階で出現する短い遷移シークエンスのリミットサイクルは、1 周期のうちで同じターゲットパタンに複数回接近することはなく、すべて異なるパタンを遷移する。例えば図 20-1 で見られる (ξ^2, ξ^3, ξ^4) の繰り返しのように同じターゲットパタンが 1 周期の中で一度しか出てこない。このような相空間ではあるターゲットパタンの次に現れるターゲットパタンは一意に定まる。学習がすすむにつれて現れる長い遷移シークエンスをもつリミットサイクルの中には、その 1 周期のなかに複数の同じターゲットパタンを含むものがでてくる (固定点の最後でも触れたように固定点アトラクタへの収束軌道でもこのような例は出てくる)。例えば、図 20-2 ではリミットサイクルの遷移シークエンスの中にターゲットパタン $\xi^{4,6}$ が複数回出現していることが判る。このような相空間では周期全体をみればターゲットパタンは決まった順序で出てくるが複

[24] ターゲットパタン近傍そのものを初期値にもつ軌道は勿論取れる。

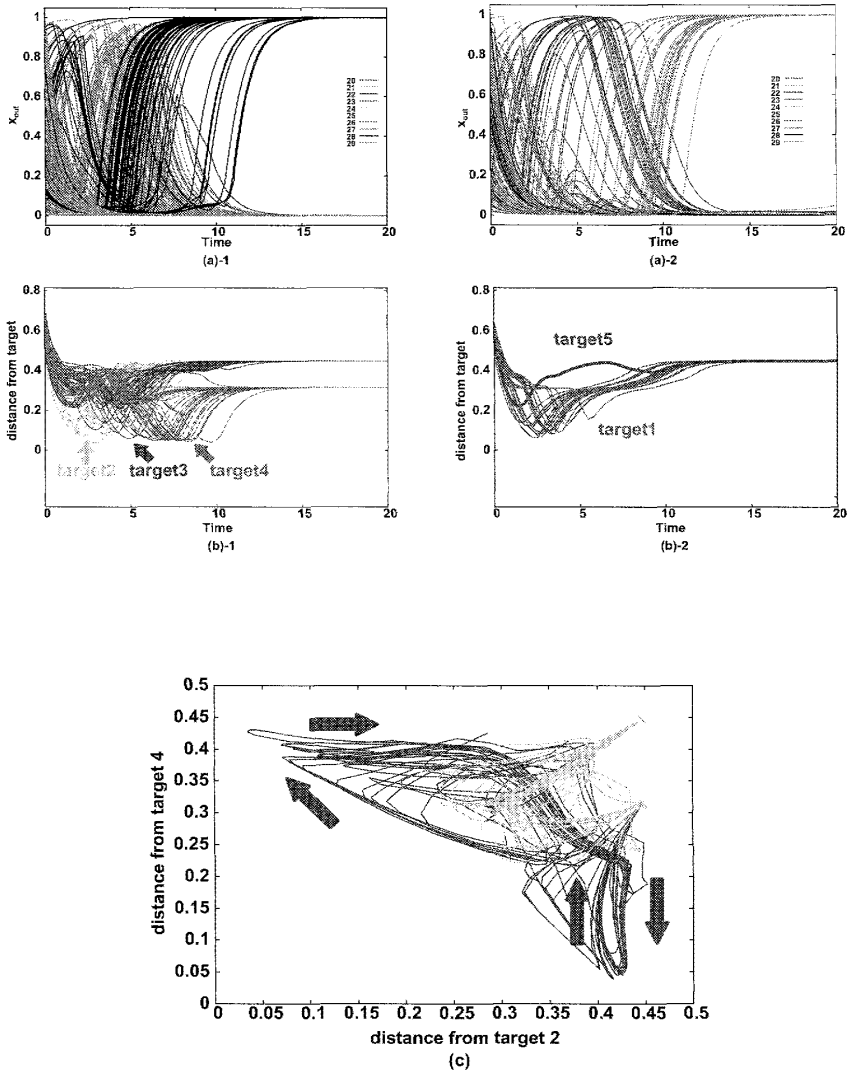


図 18: 学習セット数 5 のときの相空間構造:
ランダムに選んだ 100 個の初期値からの軌道のうち (a)(b)-1:固定点 1 に収束した軌道、(a)(b)-2:固定点 2 に収束した軌道の 2 種類に分けてそれぞれプロットした。
各図はそれぞれ (a) 出力層の時系列、(b) ターゲットパタンとの距離の時系列、(c) ターゲットパタン $\xi^{2,4}$ との距離をそれぞれ横軸、縦軸に取った軌道
を重ね合わせプロットしたものである。
固定点 1 へはターゲットパタン ξ^3 を経てターゲットパタン ξ^4 をとおり収束する経路とターゲットパタン ξ^2 を経てターゲットパタン ξ^4 をとおり収束する経路がある。(b)-1 ではターゲットパタン $\xi^{2,3,4}$ との距離の時系列を重ねてプロットしているが、これは $\xi^3 \to \xi^4$ の軌道群と $\xi^2 \to \xi^4$ の軌道群の二つが重ねてプロットしてある。また、$\xi^{2,3,4}$ 以外のターゲットパタンとの距離も見やすさのためにプロットしていないが、距離 0.2-0.4 の値をとる。固定点 2 へはターゲットパタン ξ^1 を経て収束している。(b)-2 も (b)-1 と同様に着目するターゲットパタン以外の軌道はプロットしていない。(比較のためにターゲットパタン ξ^5 を 1 本だけ書いた。) このようにどの固定点に収束するかで軌道はきれいに分かれる。

数回出てくるターゲットパタン ξ^4 のようなもの着目するとターゲットパタン ξ^4 の次現在のパタンだけでは決まらずもっと過去をみる必要がある。

複数のリミットサイクルが存在する相空間もみられる。図 21 にこのように 2 つのリミットサイクルが存在する相空間をプロットした。図 21(a) に出力層の時系列、(b) に各ターゲットパタンからの距離の時系列、(c) に適当な 2 つのターゲットパタンとの距離を軸にとった軌道をプロットした。特に図 21(a)(b) はリミットサイクル毎に 1,2 に別々にプロットした。この相空間ではターゲットパタン ξ^2 が二つのリミットサイクルに共通するパタンとなっている。このように複数のリミットサイクルが同じターゲットパタンを遷移シークエンスに含む相空間も存在する。

4.3　まとめ

今章では前章で述べた最適な時間スケール付近の $\tau^{FB} = 8$ の学習で形成されたネットワークを用いて、入力を印加しない場合の相空間の構造が学習によりどのように変化するかについて述べた。その結果学習により、固定点アトラクタからリミットサイクルへ、短い軌道から長い複雑な軌道をもつ相空間に変化することが示された。またこのときの軌道には学習したターゲットパタンが埋め込まれていることも示された。このように本モデルにおいて多数のセットの学習 (学習後期) を行うことで学習したターゲットパタンを次々と変遷するような構造が、入力のない相空間において形成されていることが分かった。

5　入力による分岐

4 章は入力がない状態でのネットワークの相空間構造が学習によってどのように変化していくか、ターゲットがどう埋め込まれるかについて述べた。今章では、入力を印加することで相空間構造がどのように変化するかを述べる。時間スケールに関しては前章と同じである。

今モデルでは記憶できているターゲットパタンを入力を強度 $\eta = 1.0$ で印加したときに、ターゲットパタンが固定点アトラクタになっていて basin volume が 0.5 を越えるものとして定義した。このように本論文では、入出力関係の記憶しているネットワークは入力を印加することでターゲットパタンが大域的な安定固定点として分岐する相空間をもつネットワークであると考える。複数の記憶があるネットワークは印加する入力パタンに応じてそれぞれ大域的な安定固定点として対応するターゲットパタンが分岐するような相空間をもつ (図 22)。

このように入力の違いに応じて複数の方向に分岐が起こり各々が別の安定固定点が出現する構造はどのようなものか 5.1 節で解析した。この節では、入力強度をあげることで対応するターゲットパタンを経由する軌道が増加し、不安定だったターゲットパタンはサドルノード分岐を起こし安定固定点に分岐することを述べる。

続く 5.2 節ではターゲットパタンを全て記憶しているネットワークと、ひとつでも記憶できていないターゲットパタンがあるネットワークを分けて解析する。そして入力を印加したときに前者はターゲットパタンだけが安定固定点として存在する相空間に分岐するのに対して、後者は多数の偽ターゲットパタンが安定固定点として存在する相空間に分岐することを述べる。

5.1　ターゲットパタンの記憶:安定固定点への分岐

学習したターゲットパタンが全て記憶されている相空間の分岐に関して、入力 0 の相空間にリミットサイクルが存在する場合と存在しない場合についてわけて解析を行った。

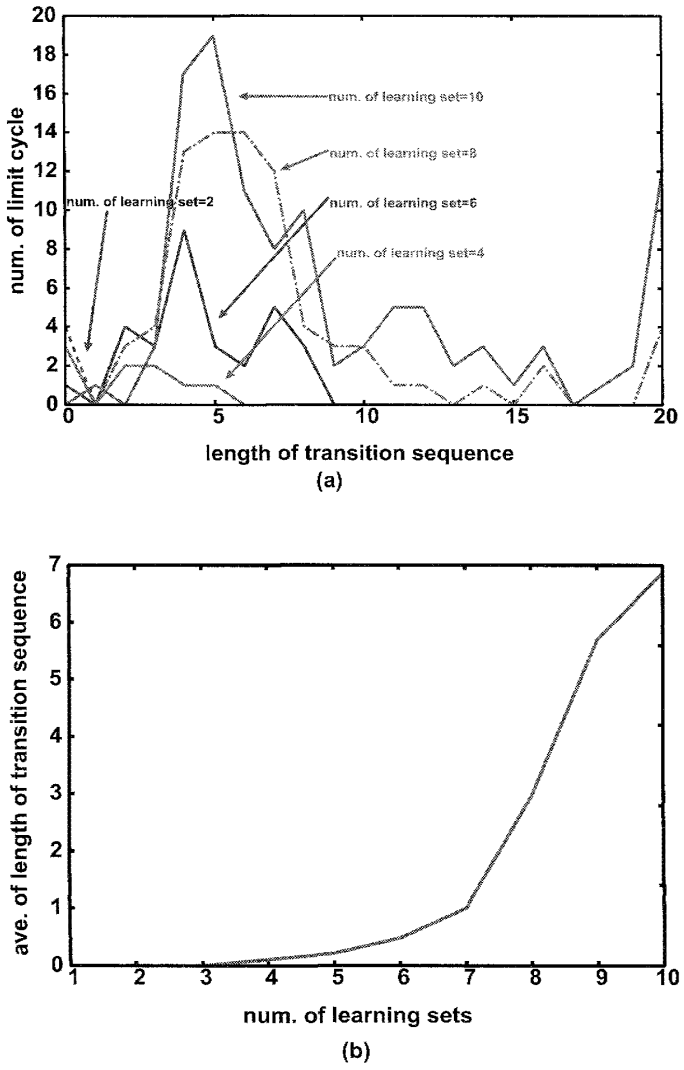


図 19: リミットサイクルの遷移シークエンスの学習による変化:
細胞の状態を閾値 0.5 で離散化することでリミットサイクルの軌道をパタンの遷移シークエンスと見なす事ができる。(a) 各学習段階毎にリミットサイクルの遷移シークエンスの長さ (=1 周期に含まれるパタンの数) の分布をプロットしたもの。長さが 0 のリミットサイクルが存在するのは、全細胞が離散化の閾値 0.5 以下で振動するリミットサイクルが存在することによる。また長さが 20 のリミットサイクルの分布の値は長さが 20 以上のリミットサイクルの和をとった値をプロットした。(b) 各学習段階毎に遷移シークエンスの長さの平均値をとりプロットしたもの。
平均値の挙動をみると学習セット数が 8 を越えるところから急激に長さが長くなっている。分布図をみると同じく 8 を越えるところから長い遷移シークエンスをもつリミットサイクルが増加している、一方で分布のピーク自体は長さ 5 のままである。

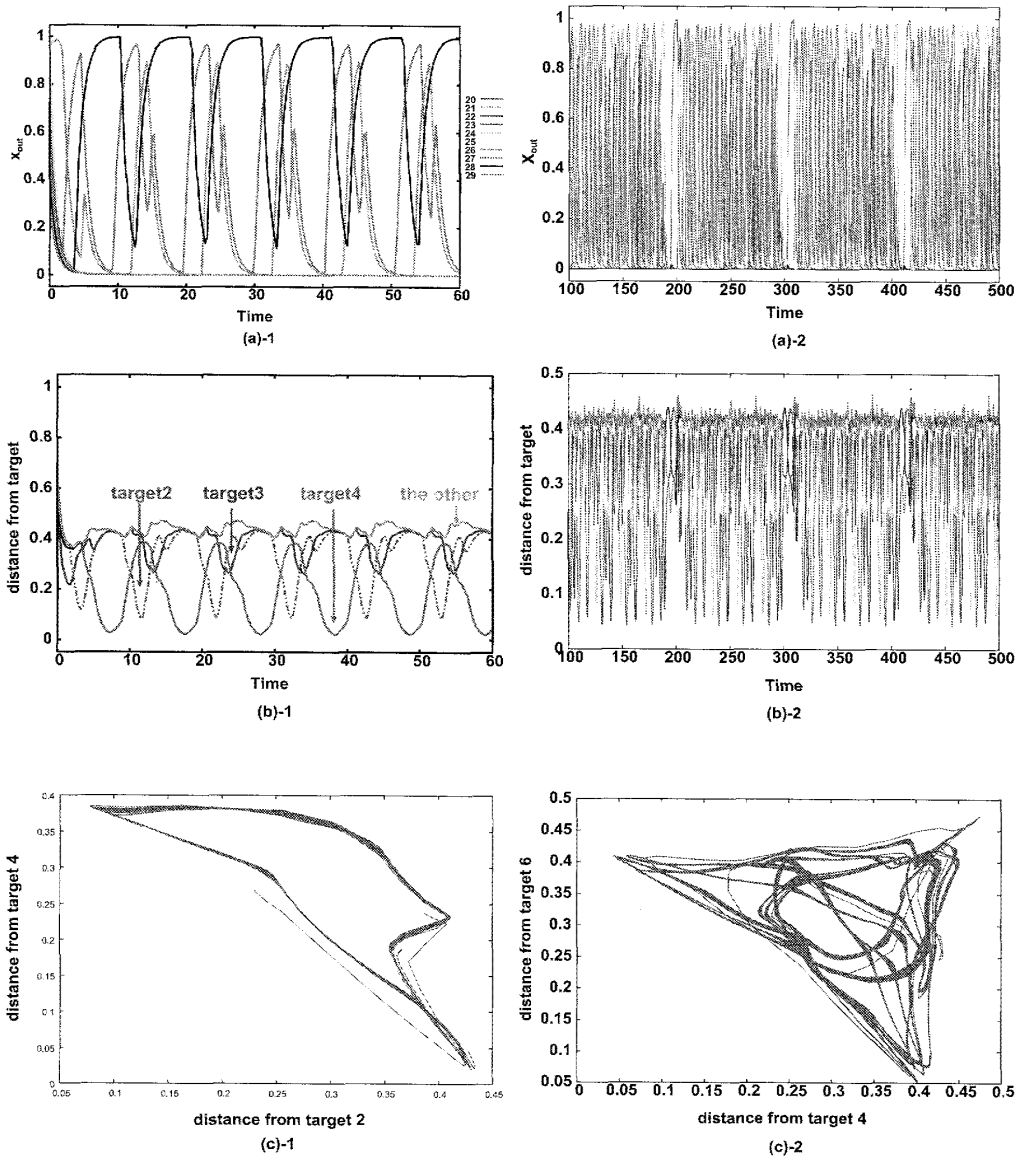


図 20: リミットサイクルの挙動:
(a) 出力層細胞の時系列:
1 初期状態からの全出力細胞の時系列を重ね書きした。
(b) 各ターゲットパタンとの距離:
リミットサイクルの軌道を各ターゲットパタンとの距離でプロットした。
(c) リミットサイクルの 2 つのターゲットパタンの距離で張られる平面への射影:
(c)-1 はターゲットパタン $\xi^{2,4}$ (c)-2 はターゲットパタン $\xi^{4,6}$ との距離をそれぞれプロットした。
(a)(b)(c) 共に*-1 は周期 3 のリミットサイクル、*-2 は周期 20 以上のリミットサイクルについてプロットした。

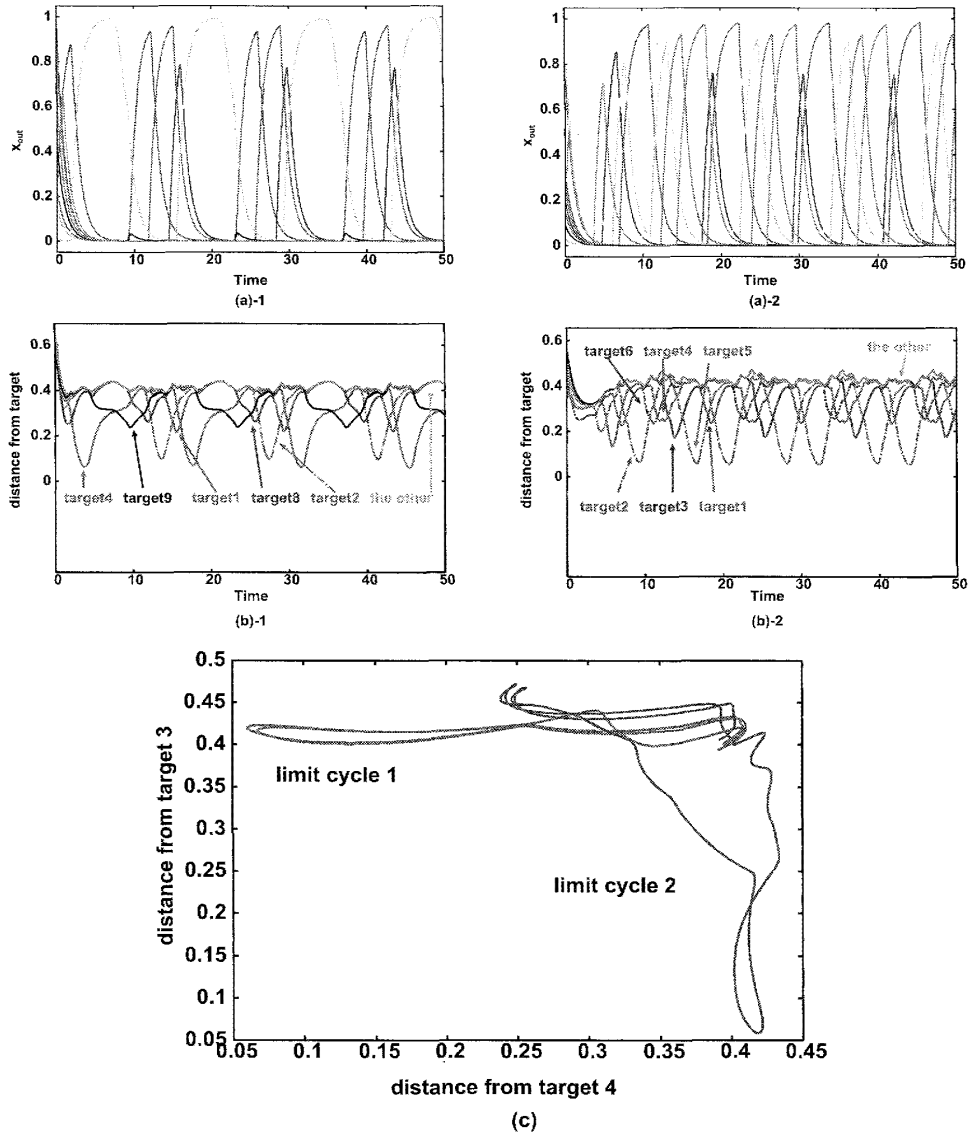


図 21: 共存するリミットサイクルの挙動:
(a) 出力層細胞の時系列:
1 初期状態からの全出力細胞の時系列を重ね書きした。
(b) 各ターゲットパタンとの距離の時系列:
リミットサイクルの軌道を各ターゲットパタンとの距離でプロットした。
(c) リミットサイクルの 2 つのターゲットパタンの距離で張られる平面への射影:
ターゲットパタン $\xi^{4,3}$ で張られる 2 次元平面に軌道を射影し、リミットサイクル 1,2 を重ね書きした。
2 つのリミットサイクルが共存している。このリミットサイクルは共にターゲットパタン ξ^2 を経由する軌道をもっていることが分かる。

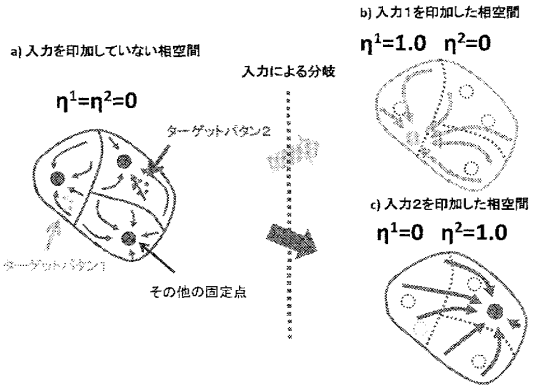


図 22: 分岐によるターゲットの想起のイメージ図:
今モデルの入出力関係の学習では入力 0 の相空間では軌道の経由点で不安定だったターゲット $\xi^{1,2,3}$ が対応する入力 $I^{1,2,3}$ をいれることで各入力方向に分岐し最終的にそのターゲットが安定な構造になるという機構がある。

5.1.1　入力 0 の相空間においてリミットサイクルが存在しない場合

4 章で述べたように複数のターゲットパタンを学習した相空間上には一般に複数の固定点がありそこへの収束軌道は学習したターゲットパタン近傍を通る。またそのうちの少数のターゲットパタンは固定点アトラクタとして存在することもある。しかしアトラクタは全て固定点アトラクタである。今節では学習したパタンは全て記憶できている場合を扱うので、入力強度 $\eta^i(i = 1, \ldots, N_{pat})$ の入力を印加することで、それぞれ対応するターゲットパタン ξ^i が大きな basin volume をもつ固定点アトラクタに分岐する。今小節では以上のような相空間を考える。

このときの分岐構造を具体例をもちいて解析する。具体例として $N_{pat} = 3$ の場合のネットワークの相空間を一つ選び調べる。$N_{pat} = 3$ なので学習したターゲットパタンは 3 つある。このネットワークの入力を印加しない相空間上では 3 つのうち 2 つのターゲットパタン $\xi^{1,2}$ は不安定であり、多くの軌道が一度ターゲットパタン近傍を通った後別の固定点に収束する。残りのターゲットパタン ξ^3 は初めから安定固定点アトラクタとして存在している。このような相空間構造が入力を印加することでどのように変化するかを調べた。ここでは大域的な変化を調べるために、多数の初期状態をとりそれらがどのような軌道を通りどこに収束するのかをいくつかの η で調べた。

まず入力 $\eta^{1,2}$ による変化をみる。$\eta^{1,2} = 0, \eta^1 = 0.015, 0.06, 0.2 \quad \eta^2 = 0.005, 0.02, 0.2$ の入力を印加した相空間でそれぞれ 100 個の初期状態をとり、そこからの軌道をターゲットパタン $\xi^{1,2}$ との距離で張られる 2 次元空間に射影し、100 本の軌道を重ね書きしたものが図 23 である。$\eta = 0$ の相空間においてターゲットパタン ξ^1 に近付いたのちに収束する軌道 (図 23(a) 中の P1) とターゲットパタン ξ^2 に近付いた後に収束する軌道 (図 23(a) 中の P2) が有ることが判る。この相空間に入力 I^1 を印加し強度 η^1 を大きくしていくとき、ターゲットパタン ξ^1 に近付く軌道をもつ初期状態が増加し ($\eta^1 = 0.015$)、最終的にほとんど全ての初期状態からの軌道がターゲットパタン ξ^1 を経由するように変化する。ただしこの入力強度ではまだターゲットパタン ξ^1 は安定ではない。さらに η^1 を大きくするとターゲット 1 が安定固定点に分岐し、ほとんどの初期状態からの軌道がそこに収束するようになる。同じようにターゲットパタン $\xi 2$ に対応する入力 I^2 の強度 η^2 を大きくしたときにも同様である。ターゲットパタン ξ^2 に近付く軌道の初期状態が増加し、ターゲットパタン ξ^2 が安定化しそこに収束するようになる。ただしターゲットパタンが安定化する強度はさまざまであり、ターゲットパタン ξ^1 のようにほとんど全ての軌道が経由するようになってから安定化する場合もあれば、ターゲットパタン ξ^2 のようにまだ他のターゲットパタンを経由する軌道が存在するうちに安定化し、その後その安定化したターゲットパタンに収束する軌道が増える場合もある。

次に入力 η^3 による変化をみる。ターゲットパタン ξ^3 は入力 0 の相空間ですでに固定点アトラクタであり安定であるが、basin volume は極めて小さい。η を大きくしていくとターゲットパタン ξ^3 に収束する軌道はほぼ単調に増加していく。そして最終的にほぼすべての相空間からターゲットパタン ξ^3 に収束するように相空間は変化する。

このようにして入力 I^1 を印加するとターゲットパタン ξ^1 を I^2 を印加するとターゲットパタン ξ^2 を、という様にこのネットワークは学習した複数のターゲットパタンを記憶することができる。この相空間の変化をもう少し定量的にみるために軌道とターゲットパタンとの距離が η によってどのように変化するかを調べた。入力 $I^{1,2}$ を印加し、強度 $\eta^{1,2}$ をそれぞれ大きくしたときに各ターゲットパタン $\xi^i(i=1,\ldots,Npat)$ と軌道の平均距離がどのように変化するかを取ったものが図 24 で、それぞれ (a) は入力 I^1 を、(b) は入力 I^2 を印加したときの変化を示したものである。各々の入力に対応したターゲットとの距離がほぼ単調に減少する一方で他のターゲットは平均距離が増加している。そして最終的に入力に対応したターゲットとの距離が 0 になっていることが判る。

このように入力強度 η^i が大きくなるにつれて対応するターゲットパタン ξ^i と軌道の平均距離は連続的にほぼ単調に小さくなっていく一方、ほかのターゲットパタンとの距離は長くなる。最終的には対応するターゲットパタンに収束するようになるとの距離は 0 になる。

ここまでは軌道がどのように変化するかという大域的な変化を解析したが、以下ではターゲットパタン近傍の相空間を解析しはじめ不安定であったターゲットパタンがどのような分岐をして安定固定点になるのかを述べる。ターゲットパタンの近傍の相空間の変化を調べるために、初期状態を近傍にもつ軌道の振る舞いが η の変化によってどう変わるかを調べた。ターゲットパタン ξ^1 の近傍 $|x-\xi^1|<0.01$ にランダムに初期状態をとった軌道が $\eta^1=0, 0.063., 0.065, 0.066$ と大きくしていったときにどう変化するかを重ね書きしたものが図 25 である。$\eta^1=0$ の時、軌道はターゲットパタン ξ^1 の近傍から速やかに離れ別の固定点に収束する。そこから $\eta^1=0.063, 0.065$ と強度をあげるに従い近傍に滞在する時間が長くなり、$\eta^1=0.066$ において近傍から抜けられなくなる、すなわちターゲットパタンが安定固定点になる。以上のように元々固定点が存在しない領域で徐々に滞在時間が伸び、最終的に安定固定点が出現する挙動からターゲットパタンはサドルノード分岐を経て安定固定点に分岐していると考えれる。今の場合はターゲットパタン ξ^1 についてみたがターゲットパタン ξ^2 についても同じ分岐を起こす。このことは他のネットワークでも同じであり、一般に入力 0 の相空間においてアトラクタではないターゲットパタンはサドルノード分岐をとおして安定な固定点へと分岐する。

以上のように各々の入力の強度 η^i をあげることでそれぞれの対応したターゲットパタン ξ^i が大域的に安定な固定点に分岐することで、複数の入出力関係が記憶されている事が示された。

5.1.2 入力 0 の相空間がリミットサイクルを含むとき

今小節では入力を印加しない相空間上でリミットサイクルが存在しているネットワークの分岐構造について解析した。リミットサイクルを含む相空間の分岐は極めて複雑であり少しのパラメタの変化で系の振る舞いが大きく変わるため、固定点の場合のように入力強度 η^i を大きくしていったときにどのように相空間が変化するかを一般的に記述するのは難しい。そこでいくつか具体例をあげた後にそれらのおおまかな共通点を述べる。

入力 0 の相空間のリミットサイクルの軌道をターゲットパタンの遷移シークエンスとみなせる事は前章で述べた通りである。そこでリミットサイクルの遷移シークエンスに入っているターゲットパタンと入っていないターゲットパタンの分岐を分けて解析する。

ここではまず遷移シークエンスに入っているターゲットパタンの分岐について調べた。学習段階が 7 でリミットサイクルがある相空間を具体例にしてどのような分岐が起こるかをのべる。入力を印加していな

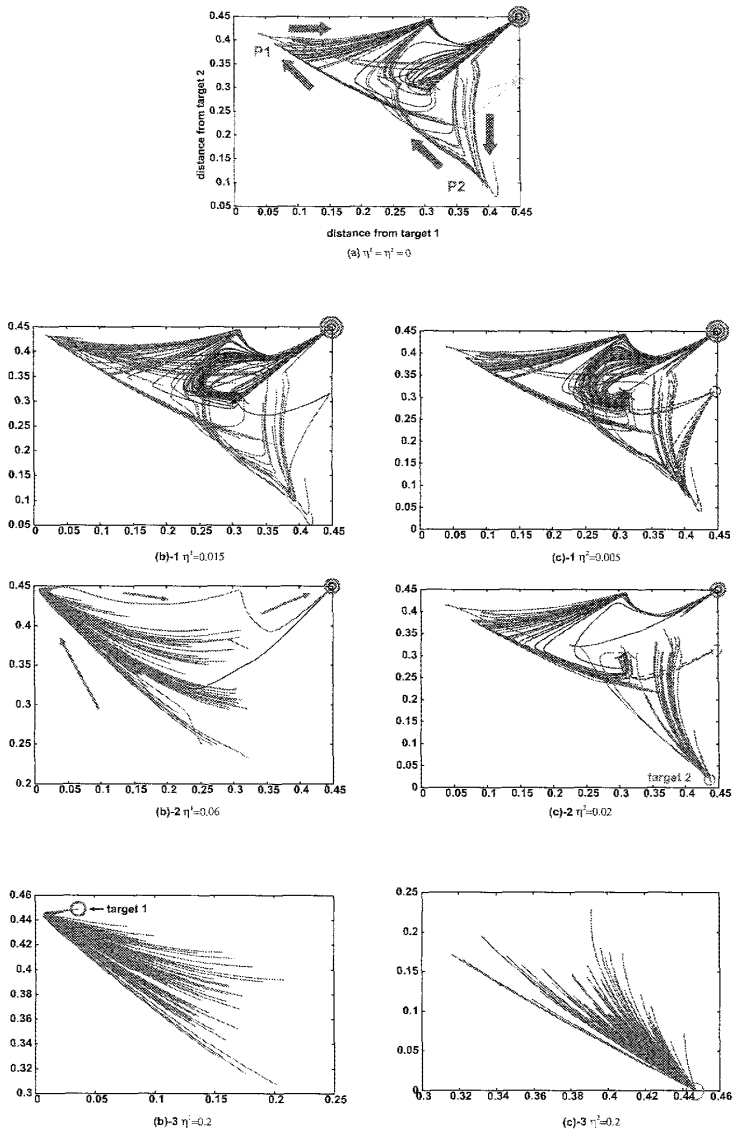


図 23: 入力強度 $\eta^{1,2}$ による相空間の変化:
ランダムに選んだ 100 個の初期状態から固定点アトラクタに収束するまでの軌道の重ね書きした。ただし軌道の極初期 ($t < 2$) の部分は重ねると見にくいので除いてある。横軸はターゲットパタン ξ^1 との距離、縦軸はターゲットパタン ξ^2 との距離をとった。軌道の色の違いは収束先のアトラクタの違いを表している。
(a) 入力のない相空間
(b) 入力 I^1 の強度をあげたときの相空間変化
(c) 入力 I^2 の強度をあげたときの相空間変化

下の図ほど大きい入力強度を印加した相空間になっている。入力強度が小さいときはターゲットの近傍を通らなかった軌道も強度をあげるにつれて対応するターゲットパタンの近傍を通るようになる。さらに強度をあげるとターゲットパタンが分岐し安定固定点になる。

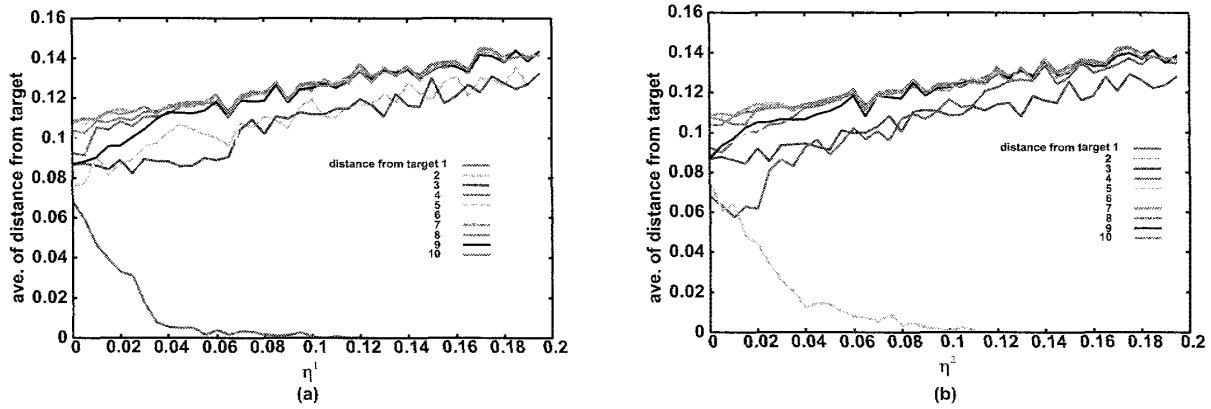


図 24: 軌道とターゲットパタンとの距離の η 依存性:
初期状態 i からアトラクタに収束するまでの軌道が一番ターゲットパタン ξ^j に近付いた距離を軌道とターゲットパタン ξ^j の距離 d_i^j とする。これを多数の初期状態について平均した量 $\bar{d}^j = \sum_i d_i^j (j = 1, \ldots, 10)$ を各入力強度 η について計算しプロットした。
(a) 入力 I^1 の強度 η^1 を印加したときの平均距離の変化
(b) 入力 I^2 の強度 η^2 を印加したときの平均距離の変化

入力強度 $\eta^{1,2}$ を強くすると対応するターゲットパタンとの距離 $d^{1,2}$ が選択的に減少し、他のターゲットパタンとの距離は拡がる。その結果、入力に対応する距離は 0 になる。これにより初め他のターゲットパタンに近付いていた軌道が強度 $\eta^{1,2}$ をあげることで対応するターゲットパタン $\xi^{1,2}$ に近付く軌道に変わり、最終的に全ての軌道がターゲットパタン ξ に近付いていることが分かる。

いネットワークの相空間上のリミットサイクルは $(\xi^2, \xi^3, \xi^6)^2$ の遷移シークエンスから成っている。

まずターゲットパタン ξ^6 について述べる。リミットサイクルの軌道上での出力層のターゲット細胞 x_6^{out} の極大値が η^6 を大きくしたときにどのように変化するかでリミットサイクルの分岐を調べた。各 η^6 において十分長く過渡期をとった後ターゲット細胞の極大値を重ねてプロットしたものが図 26(a) である。ただし固定点に収束した場合は極大値がないのでプロットされない。さらにリミットサイクルの軌道がどのように変化しているかを詳しく調べるために、図 26(b)(c)(d)(e) にそれぞれ $\eta^6 = 0, 0.02.0.026, 0.027$ のときのターゲットパタンとの距離の時系列をプロットした。図 26(a) でラインが二本なのはリミットサイクルが $(\xi^2, \xi^3, \xi^6)^2$ で表記されるように 3 つのターゲットを 2 巡して初めて初期状態に戻るので極大値が二つあるため。$\eta^6 = 0$ でのリミットサイクルはターゲット $\xi^{2,3,6}$ がほぼ対称的な距離、近傍の滞留時間を持っているが、入力強度を $\eta^6 = 0.02, 0.026$ とあげるに従い対称性は崩れターゲットパタン ξ^6 近傍の滞留時間が徐々に長くなる。このパラメタ領域ではターゲット ξ^6 は不安定で、実際近傍の初期状態をとっても近傍を離れリミットサイクルに収束する。さらに入力強度をあげ $\eta^6 = 0.027$ にするとターゲットが安定化し、リミットサイクルが固定点に分岐する。これはターゲットパタンが不安定なパタンからサドルノード分岐を起こし、安定固定点になったためリミットサイクルが崩壊したためにおこるものである。以上のようにリミットサイクルの軌道がターゲットに近付き、さらに η を大きくすることで最終的にターゲットがサドルノード分岐を起こし安定固定点になるという構造をもつ。この分岐の仕方は前小節で調べた固定点しかない相空間のターゲットパタンの分岐と同じであり、不安定なターゲットパタンがサドルノード分岐を起こして安定固定点になるのは、固定点・リミットサイクルに共通な構造だと考えられる。ここの分岐ではターゲットパタンが安定固定点に分岐するまで遷移シークエンスはおなじである。

次に遷移シークエンスの別のターゲットパタンである ξ^3 について述べる。対応する入力 I^3 を印加した時にどのような分岐を起こすのかを調べた。ターゲットパタン ξ^6 と同様にリミットサイクルがどのように

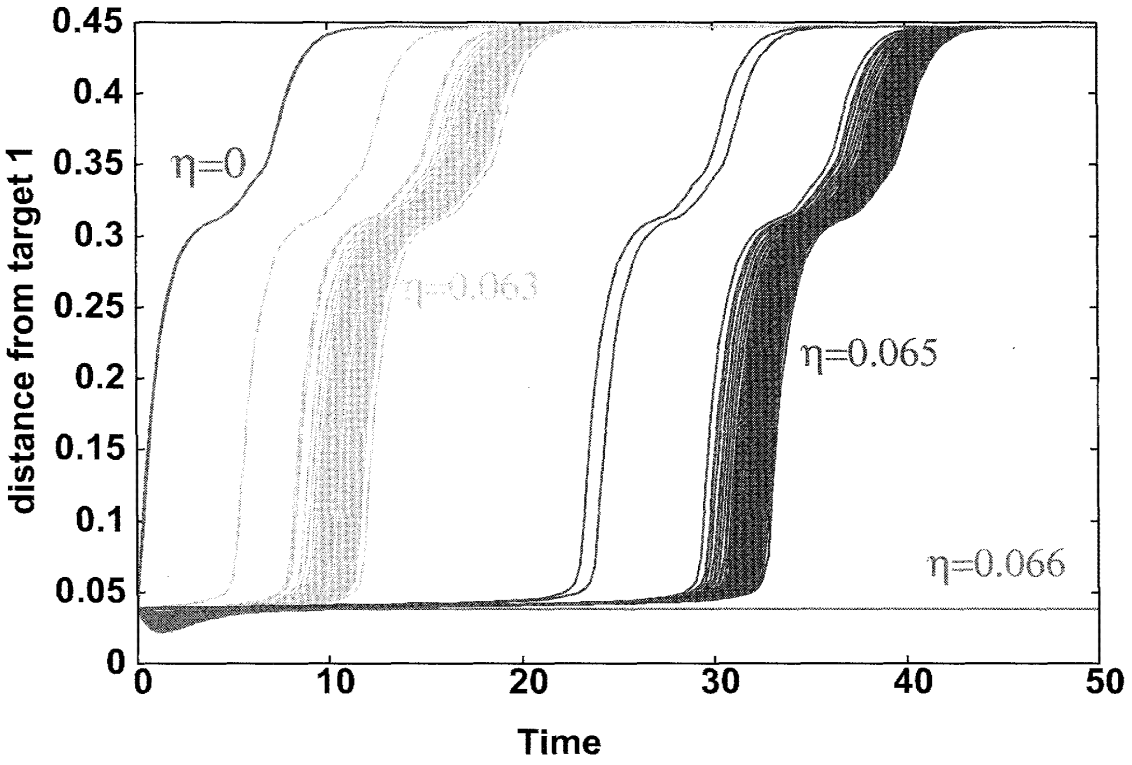


図 25: ターゲットパタン ξ^1 の分岐:
ターゲットパタン ξ^1 近傍の流れが入力強度 η を強くしたときにどう変化をプロットした。入力強度 $\eta^1 = 0, 0.063, 0.065, 0.066$ についてターゲットパタン ξ^1 の近傍に初期状態をもつ軌道を 100 個重ね書きした。ターゲットパタン ξ^1 との距離が 0 ではないのはターゲットパタンの位置自体が入力強度によって多少値が変化するが、ターゲットパタンとの距離というときのターゲットパタンは強度 $\eta = 1.0$ のときの値を使っているため。

入力強度が小さいときはターゲットパタンは固定点ではなく、速やかに近傍から離れ別のアトラクタに向かう。入力強度をあげていくと近傍での滞留時間が長くなり、最後は近傍から抜けられなくなり ($\eta^1 = 0.066$) ターゲットパタン ξ^1 は安定固定点になる。以上のようにパラメタをあげるにつれて固定点ではないところから安定固定点が出現し、そのパラメタ近傍では滞留時間が長くなることからサドルノード分岐だと考えられる。

変化していくかを調べるために、ターゲットパタン ξ^3 の出力層のターゲット細胞 x_3^{out} のリミットサイクル上の極大値が強度 η^3 に依存してどのように変化するかをプロットしたものが図 27(a)(b) である ((b) は (a) の領域 [0.0105:0.01064] の拡大図)。また同じくリミットサイクルの軌道の変化を詳しく調べるために、$\eta^3 = 0.01, 0.012$ の時のリミットサイクルの振る舞いをターゲットパタン $\xi^{2,3,6}$ との距離時系列としてそれぞれ図 27(c)(d) にプロットした。$\eta^3 = 0$ から $\eta^3 = 0.0105$ までは入力 0 の相空間でのリミットサイクル $(\xi^2, \xi^3, \xi^6)^2$ が連続的に変化し、ターゲットパタン ξ^3 の極大値もターゲット ξ^3 近傍に長く留まるように連続的に変化する。これは上のターゲットパタン ξ^3 のときの変化と同様である。$\eta^2 = 0.0106$ 付近でのカオティックな遷移シークエンスをとるパラメタ領域を経て、リミットサイクルの軌道が不連続的に変化しリミットサイクル $(\xi^3, \xi^2, \xi^3, \xi^6)$ に分岐する。このようにリミットサイクルの遷移シークエンスが変わるときにカオティックな領域を経る現象は一般的に見られる。リミットサイクル $(\xi^3, \xi^2, \xi^3, \xi^6)$ のパラメタ領域では印加している入力 I^3 に対応するターゲットパタン ξ^3 とリミットサイクルの距離は η^3 が小さいときに比べて大きくなっている。一方で遷移シークエンス中にターゲット ξ^3 が出現する頻度は多くなっている。さらに強度をあげると $\eta^3 = 0.0162$ あたりでリミットサイクル (ξ^2, ξ^3, ξ^6) が分岐する。このパラメタではリミットサイクル $(\xi^3, \xi^2, \xi^3, \xi^6)$ はまだ安定であり、2 つのリミットサイクルが存在している。このうち新しく分岐したリミットサイクル (ξ^2, ξ^3, ξ^6) ではターゲットパタン ξ^3 に大きく近付き、また滞留時間も長くなっている。$\eta^3 = 0.0166$ でターゲットパタン ξ^3 が安定固定点にサドルノード分岐を起こしリミットサイクル (ξ^2, ξ^3, ξ^6) の軌道は安定固定点に収束する。この強度でもリミットサイクル $(\xi^3, \xi^2, \xi^3, \xi^6)$ はまだ安定であり、さらに強度をあげた $\eta^3 = 0.017$ で $(\xi^3, \xi^2, \xi^3, \xi^6)$ はターゲットパタン ξ^3 に収束する。したがってこのパラメタ領域ではリミットサイクル $(\xi^3, \xi^2, \xi^3, \xi^6)$ とターゲットパタン ξ^3 は共存している。このようにリミットサイクルが壊れ、ターゲットパタンに分岐するときは必ずしもサドルノードで分岐するわけではない。しかしこの場合でもターゲットパタン ξ^3 自体はサドルノード分岐を起こし安定固定点に分岐している。リミットサイクル $(\xi^3, \xi^2, \xi^3, \xi^6)$ とターゲットパタン ξ^3 が共存するパラメタ領域があるのは安定化したターゲットパタンの basin が小さくターゲットパタンからやや離れた軌道をとっているリミットサイクル $(\xi^3, \xi^2, \xi^3, \xi^6)$ を壊さなかったためと考えられる。そして入力強度をあげることでターゲットパタンの basin が大きくなり、リミットサイクル $(\xi^3, \xi^2, \xi^3, \xi^6)$ の軌道もやがて basin の境界と交わりターゲットパタンに収束するようになることで壊れると考えられる。また、一度ターゲットに収束したのち入力強度 $\eta^3 = 0.018$ から 0.02 にかけてリミットサイクルが現れるがこれについては最後に述べる。

以上の 2 つの分岐はともにリミットサイクル $(\xi^2, \xi^3, \xi^6)^2$ の遷移シークエンスに属しているターゲットパタンの分岐であるが、分岐の仕方は大きく異なっている。ターゲットパタン ξ^6 は安定固定点に分岐するまで遷移シークエンスは同じでリミットサイクルの軌道は比較的変化しない。一方でターゲットパタン ξ^3 が安定固定点に分岐するまでは遷移シークエンスが変化するので、リミットサイクルの軌道は大きく変化する。残りの ξ^2 も同様にところどころで遷移シークエンスが変化し、カオティックな遷移領域を経て最終的にターゲットパタンに収束する分岐を起こす。

ターゲットパタンが入力 0 の相空間のリミットサイクル上の遷移シークエンスにはない時はどのような分岐が起きるのかを述べる。上と同じネットワークを用いて、ターゲットパタン ξ^4 の入力強度 η^4 への依存性を調べた。極大値の η^4 依存性を調べたものが図 28(a) である。強度が小さいときにはリミットサイクルの遷移シークエンスには入っていないので点がプロットされていない。$\eta^4 = 0.006$ あたりからカオティックな領域をへて図 28(b) のようにターゲットパタン ξ^4 が遷移シークエンス入ったリミットサイクルに分岐する。以上のようにリミットサイクルの遷移シークエンスにターゲットパタンが入っていない場合は、まず遷移シークエンスにターゲットパタンが入ったリミットサイクルに分岐する。このときも遷移シークエンスが変わるときと同様にカオティックな変遷をするパラメタ領域を通る事が多い。また一度遷移シークエンスにターゲットパタンを含むリミットサイクルになると後の分岐はターゲットパタンを含んだリミットサイクルの分岐と違いはないと考えられる。この分岐の場合は最終的にはサドルノード分岐をおこしてリミットサイクルはターゲットパタンに分岐している。

以上のようにリミットサイクルの分岐は複雑で統一した理解は難しい。しかしその中でもいくつかの共通の特徴として

1) リミットサイクルの軌道が分岐し異なる遷移シークエンスになるときにはカオティックな遷移シークエンスのパラメタ領域が存在することが多い

2) 入力 I^a を印加し分岐が起きるときにリミットサイクルの軌道と対応するターゲットパタン ξ^a との関係をみると、徐々にターゲット ξ^a に近付くあるいは遷移シークエンスの中でターゲット ξ^a の頻度が増加するという傾向がある

3) 初めにターゲットパタン ξ^a がリミットサイクルの遷移シークエンスにない場合は、1. と同様にカオティックな遷移シークエンスのパラメタ領域を経て ξ^a が遷移シークエンスの中に入ったリミットサイクルに分岐する

4) ただしターゲットパタンの安定性だけみるとはじめ不安定なターゲットパタンがサドルノード分岐を起こして安定固定点になるという構造は一般的である

といった点が見られる。

5.2　記憶できていないターゲットがあるネットワーク

今節では一つでも記憶できていないターゲットがあるネットワーク、すなわち入力パタン I^i を印加したにもかかわらずに対応するターゲットパタン ξ^i が basin volume が $\theta^{mem} = 0.5$ 以上の安定固定点に分岐しないネットワークの解析を行う。

学習したすべてのターゲットパタンが記憶されているネットワークと一つでも記憶できていないパタンがあるネットワークの性質が、記憶できているかどうかの閾値 (本論文では 0.5) にかかわらず大きく異なることを示す。すべてのターゲットパタンが記憶されているネットワークは、入力 $I^i(i = 1, \ldots, N_{pat})$ を印加するとそれに対応するターゲットパタンが全て大きな basin をもつ安定固定点に分岐するネットワークである。一方で、ターゲットパタン ξ^a は記憶されているが、別の入力 I^b に対してはターゲット ξ^b は想起されないというネットワークが、一つでも記憶でも記憶できていないパタンがあるネットワークである。これはある方向 (今の場合は η^b) の分岐構造が壊れてしまい最終的にターゲットパタンが大きな basin をもつ安定固定点に分岐できなくなったネットワークとみなせる。

上の事を示すため、入力を印加したときの相空間を以下の 3 つに分けて解析を行う。全パタンを記憶できているネットワークに入力を印加したときの相空間、一つでも記憶できていないパタンがあるネットワークに入力を印加したときの相空間のうちターゲットパタンを記憶できている相空間と記憶できていない相空間の 3 つである。

この 3 つの相空間構造の違いを調べるために、ベイシンエントロピー $S = -\sum_i vol_i * log(vol_i)$ を入力を $\eta = 1.0$ で印加した相空間で調べ、分布を比較したものが図 29 である。3 つの相空間は上で述べた順にそれぞれ red line,green line,blue line でプロットした。3 つのうち前者 2 つは定義上はターゲットは記憶できている相空間であるが、記憶できないターゲットがあるネットワークから分岐した相空間の分布 (green line) と全てのターゲットが記憶できているネットワークから分岐した相空間の分布 (red line) は大きく異なる事が判る。前者は S が広く分布しているのでターゲット以外のアトラクタが多く存在している一方で (上でも述べたが) 後者はターゲットしか存在しない。むしろ数多くのアトラクタが存在しているという点では、記憶できないターゲットパタンがあるネットワークから分岐する 2 つの相空間の方が構造は似ている事が分かる。このように、全パタン記憶できているネットワークから分岐した相空間と一つでも記憶できないパタンがあるネットワークでは性質大きく異なる。前者は入力を印加し分岐した相空間上には 1 つの大きなアトラクタであるターゲットパタンしか存在しない。一方で後者は分岐した相空間上ではたくさんのアトラクタ (偽ターゲットパタンと呼ぶ) が存在し、ターゲットパタンはアトラクタであったとしても小さな basin volume しか持たない事が分かる。この性質は記憶できているかどうかの閾値 θ^{mem} には関係のない性質である。

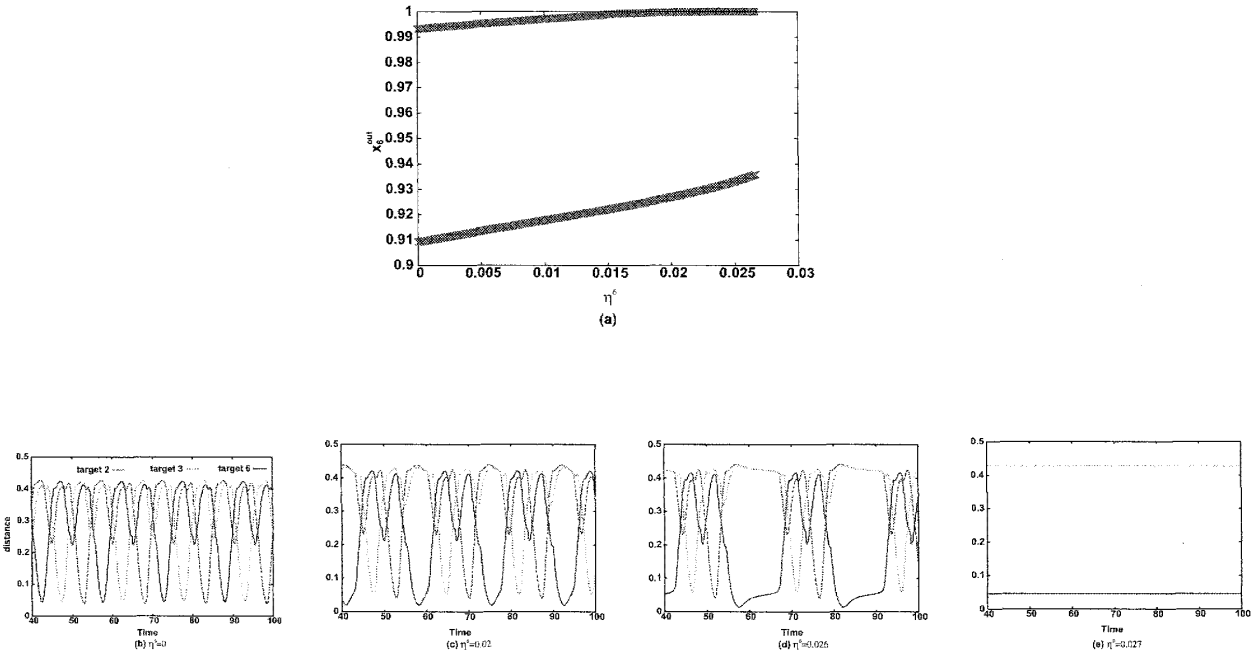


図 26: リミットサイクル $(\xi^2, \xi^3, \xi^6)^2$ の η^6 方向への分岐:
(a) リミットサイクル上のターゲット細胞 x_6^{out} の極大値をプロットしたもの:
十分長い過渡期を経た後の極大値を各 η^6 に関して重ね書きした。以下極大値をプロットした図は同じようにプロットして作成した。二本のラインがプロットされているのはリミットサイクル $(\xi^2, \xi^3, \xi^6)^2$ は 1 周期で二度ターゲットパタン ξ^6 に近付くので、1 周期に二つ極大値を持つことによる。$\eta^6 = 0.027$ 付近で安定固定点に分岐しているためラインが消えている。
(b)$\eta^6 = 0$ (c)$\eta^6 = 0.02$ (d)$\eta^6 = 0.026$ (e)$\eta^6 = 0.027$ のときのターゲットと距離の時系列:
プロットしているのはリミットサイクルの遷移シークエンスにあるターゲットパタンだけである。他のターゲットに大きく近付くことはないので他のターゲットとの距離はプロットしなかった。(b)(c)(d)(e) と図 27(c)(d)(e) は同じ色のラインはおなじターゲットとの距離の時系列を表す。

印加した入力 I^6 に対応するターゲットパタン ξ^6 に着目すると、η^6 が大きくなるに連れて滞留時間が長くなり最終的に安定固定点に分岐している様子が分かる。この挙動は固定点しかない相空間でのターゲットパタンの分岐と同じ挙動であり、サドルノード分岐をしていると考えられる。

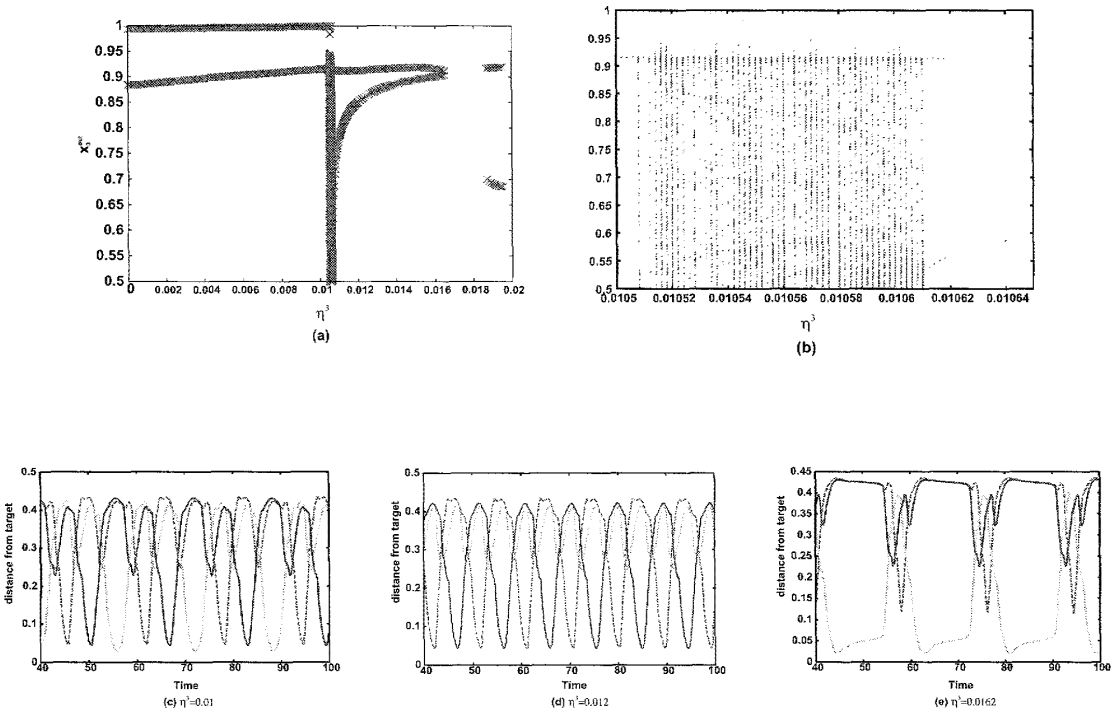


図 27: リミットサイクルの η^3 方向への分岐:
(a) リミットサイクル上のターゲット細胞 x_3^{out} の極大値をプロットしたもの: (b)[0.0105:0.01064] の拡大図: 図 26 と同様に出力層ターゲット細胞 x_3^{out} のリミットサイクル上での極大値が η^3 に依存してどのように変化するかをプロットした。
(c) $\eta = 0.01$ (d) $\eta = 0.012$ の時のターゲットパタン $\xi^{2,3,6}$ との距離の時系列:
リミットサイクルの遷移シークエンスが切り替わるときにカオティックな遷移シークエンスになるパラメタ領域を経ていることが分かる。また $\eta^3 = 0.016$ 付近でリミットサイクルが壊れて安定固定点であるターゲットパタン ξ^3 に収束するようになるので、(a) でプロットがとぎれている。このときの詳しい挙動に関しては本文で述べているが、わずかにターゲットパタン ξ^3 とリミットサイクル $(\xi^3, \xi^2, \xi^3, \xi^6)$ が共存しているパラメータ領域がある。したがってリミットサイクル $(\xi^3, \xi^2, \xi^3, \xi^6)$ が安定固定点に分岐するのは、単にターゲットパタンが安定固定点に分岐したからではない。一度ターゲットパタンに収束した後、$\eta^3 = 0.018$ 付近でリミットサイクルが再び出現している。この現象挙動に関しては 5.5 節参照。
(e)$\eta^3 = 0.0162$ で分岐したリミットサイクル:
上で述べたリミットサイクル $(\xi^3, \xi^2, \xi^3, \xi^6)$ とターゲットパタン ξ^3 が共存しているパラメタ領域より少し小さい $\eta^3 = 0.0162$ でターゲットパタン ξ^3 により近い点を通るリミットサイクル (ξ^2, ξ^3, ξ^6) が分岐する。(e) では、このリミットサイクルの時系列をプロットした。このリミットサイクルの軌道はターゲットパタン ξ^3 付近での滞留が長くなっているのがわかる。最終的にターゲットパタンがサドルノード分岐を起こし安定固定点に分岐する。(a) にはこのリミットサイクルの軌道の極致はプロットしていない。

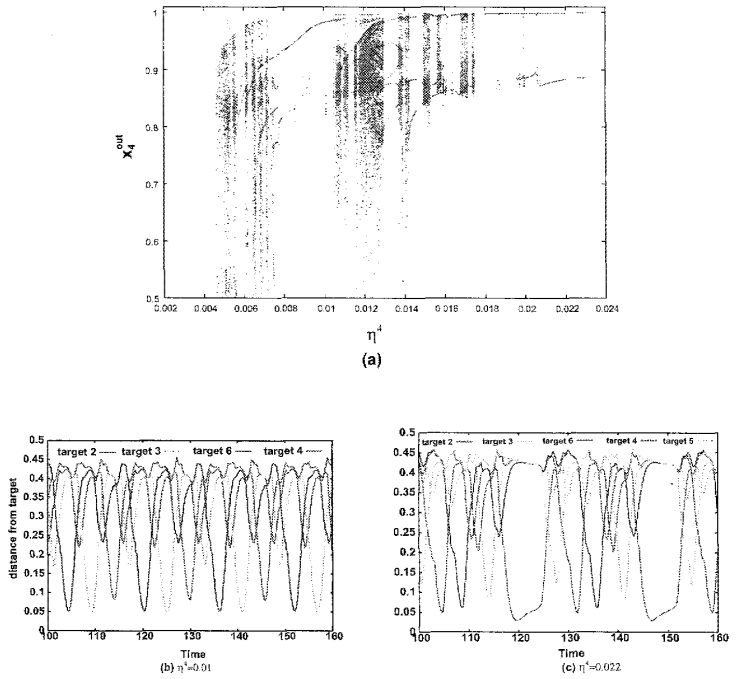


図 28: リミットサイクルの η^4 方向への分岐:
(a) リミットサイクル上のターゲット細胞 x_4^{out} の極大値をプロットしたもの:
図 26 と同様にターゲットパタン ξ^4 の出力層ターゲット細胞 x_4^{out} のリミットサイクル上の極大値が η^4 に依存してどのように変化するかをプロットした。
(b)$\eta^4 = 0.01$ (c)$\eta^4 = 0.022$ の各ターゲットパタンとの距離の時系列:

ターゲットパタン ξ^4 は入力がないときにはリミットサイクルの遷移シークエンスには入っていない (図 26(b))。このようなリミットサイクルが対応する入力強度 η^4 をあげることで遷移シークエンスにターゲットパタン ξ^4 が入ったリミットサイクルに分岐する (図 (b))。一度遷移シークエンスに入った後のリミットサイクルの分岐構造は、元から遷移シークエンスに入っているターゲットパタン (この例の場合ターゲットパタン $\xi^{2,3,6}$) の分岐と同じと考えられる。この場合はリミットサイクルの軌道がターゲットパタン ξ^4 に近付くと同時に滞留時間が長くなり、最終的にサドルノード分岐を起こしターゲットパタン ξ^4 に分岐する (図 (e)))。

次にこの偽ターゲットパタンはどのような固定点なのかを調べる。今モデルでは、各ターゲットとして出力層の細胞のうち1細胞のみが活性化したスパースなパタンを選んでおり、さらに中間・出力層には層内に相互抑制結合があるために中間層の細胞の活性化パタンも1細胞が活性化するスパースなパタンになっているので、ほとんどのターゲットパタンは隠れ・出力層ともに1細胞だけが活性化するパタンである。しかし偽ターゲットパタンの多くは隠れ・出力層ともに2細胞が活性化するパタンが安定固定点になっている。この偽ターゲットパタン ξ^{false} の2細胞活性化パタンを詳しく見ると 出力層の活性化している細胞については $x^{out}_{false} = x^{out}_{ture} + x^{out}_{latest}$ 隠れ層の活性化細胞については $x^{out}_{false} = x^{hid}_{ture} + x^{hid}_{else}$ と分解できる。ここで x_{ture} 等は ture は正しいターゲットパタン、 latest は最後に学習したターゲットパタン、else はどれか1細胞が活性化したパタンの index としたときの記法にしたがった。すなわち出力層の活性化している2細胞は正しいターゲットパタンで活性化するべき細胞と最後に学習したターゲットパタンで活性化するべき細胞の2細胞からなり、隠れ層の活性化している2細胞は正しいターゲットパタンで活性化すべき細胞と他の細胞からなっている。隠れ層のターゲット細胞 x^{hid}_{ture} 以外の細胞 x^{hid}_{else} は、学習時の探索過程がどのように行われるかに依存して決まり一つのネットワークに複数存在する。その結果図29で見たように多数の偽ターゲットが形成される。

以上のように、全ての入力に対してターゲットパタンのみが固定点となる相空間に分岐をするネットワークと、全ての入力に対して多数の偽ターゲットパタンが固定点となる相空間に分岐をするネットワークの2種類に分けることができる。後者はどのような入力を印加しても偽ターゲットパタンが存在する一方で、前者はどのような入力をいれてもターゲットパタンのみの相空間が分岐することが判る。さらにこのような偽ターゲットパタンは上でも述べたように $x^{out}_{false} = x^{out}_{ture} + x^{out}_{latest}, x^{hid}_{false} = x^{hid}_{ture} + x^{hid}_{else}$ というように分解できる。そして、偽ターゲットパタンも正しいターゲットパタン+最後に学習したターゲットパタンに分解できるので、記憶ができていない相空間においてもターゲットパタンの情報が完全に失われているわけではない事もわかる。

5.3　まとめ

本論文では入力に対する系の応答を、入力をパラメタとみなしそれに対する相空間構造の変化として捉えた。今章では学習されたネットワークの相空間構造が入力に対してどのように変化するかを調べることで、今モデルにおける記憶がどのようになされているかを解析した。そして入力を印加しないときに固定点アトラクタしかない相空間においては、

1) ターゲットパタンと軌道との距離はほぼ単調に減少し最終的に0になること

2) 不安定なターゲットパタンがサドルノード分岐をへて安定な固定点に分岐すること

を示した。一方でリミットサイクルが存在する相空間の分岐は複雑であるがターゲットパタンがサドルノード分岐で安定化する構造は共通している。さらに

1) リミットサイクルが分岐し異なる遷移シークエンスになるときにはカオティックな遷移シークエンスを経る

2) 入力 I^a を印加し分岐が起きるときにリミットサイクルの軌道と対応するターゲットパタン ξ^a との関係をみると、徐々にターゲット ξ^a に近付くあるいは遷移シークエンスの中でターゲット ξ^a の頻度が増加する

3) 初めにターゲットパタン ξ^a がリミットサイクルの遷移シークエンスにない場合は、1. と同様にカオティックな遷移シークエンスのパラメタ領域を経て ξ^a が遷移シークエンスの中に入ったリミットサイクルに分岐する

という傾向があることを述べた。

またターゲットパタンを全て記憶しているネットワークとひとつでも記憶できていないターゲットパタンがあるネットワークを分けて解析を行い、分岐した相空間の構造が大きく異なる事を示した。

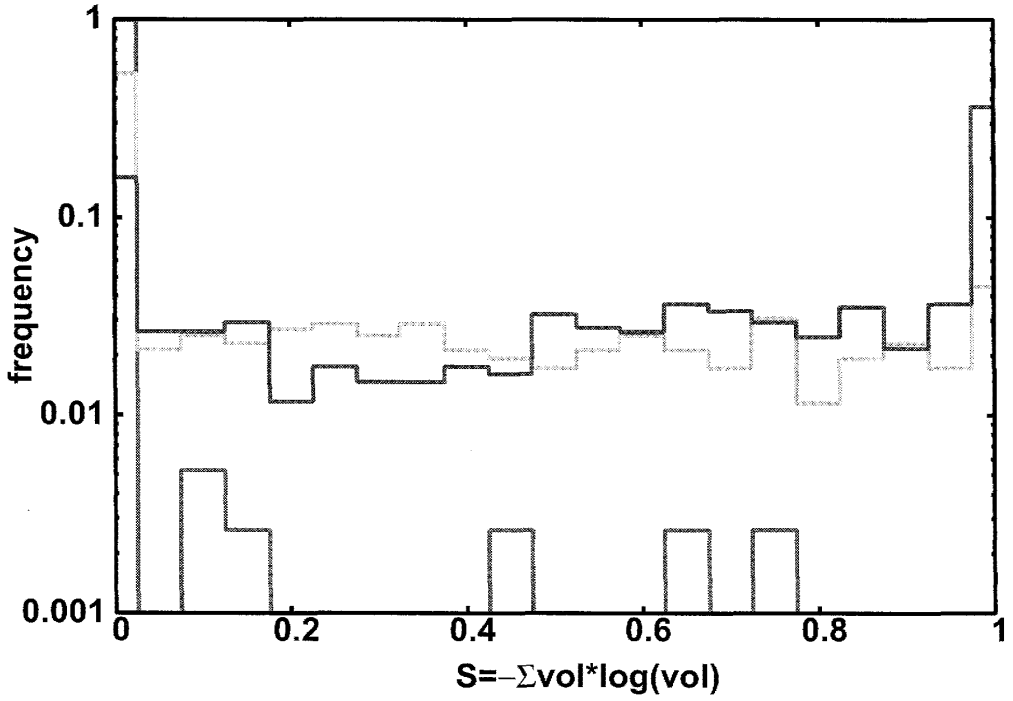


図 29: 入力を印加した相空間のベイシンエントロピー $S = -\sum_i vol_i * log(vol_i)$ の分布図:
多数の学習したネットワークに対して入力 $\eta^i (i = 1, \ldots, N_{pat})$ を印加したときの相空間構造について以下の 3 種類に分けてそれぞれのベイシンエントロピーの分布をとった。全ターゲットが記憶できているときの相空間の S 分布 (red line)、一つでも記憶ができていないターゲットが有るネットワークから分岐した相空間の中で記憶できている相空間の S 分布 (green line)、ターゲットが記憶できていない相空間の S 分布 (blue line)。一つでも記憶できないものがあるネットワークから分岐した相空間はターゲットが記憶されていてもほかの固定点が多数できている。一方で全て記憶できているネットワークから分岐した相空間はほとんどが S=0、すなわちターゲットの basin volume が 1 であり、他のアトラクタはほとんど存在しない。

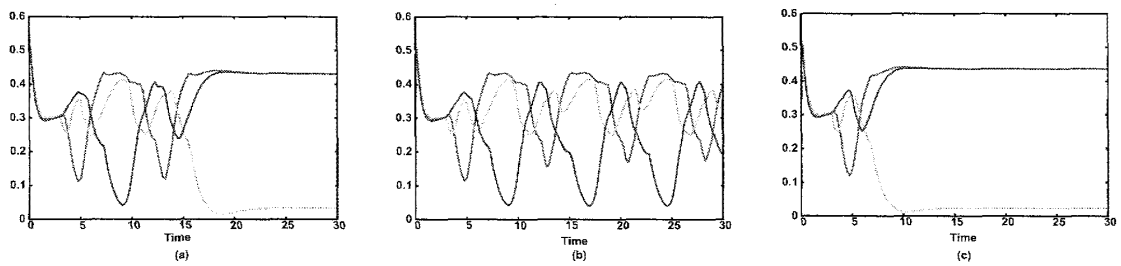


図 30: 一度ターゲットパタンに収束した後にリミットサイクルとなり再びターゲットパタンに収束する軌道例

5.1.2 節で解析したリミットサイクル $(\xi^2, \xi^3, \xi^6)^2$ に入力 η^3 を印加したときの時系列をプロットした。3 つの図はそれぞれ (a)$\eta^3 = 0.018$ (b)$\eta^3 = 0.019$ (c)$\eta^3 = 0.02$ の時のターゲットパタン $\xi^{2,3,6}$ との距離の時系列。これらは全て同じ初期値からスタートした時系列である。

入力強度を $0.018 \rightarrow 0.019 \rightarrow 0.02$ とあげるにつれて、一度ターゲットパタンに収束していた軌道がリミットサイクルになりまたターゲットパタンに収束している。このときの (a)(c) の過渡期を見ると過渡期の長さが短くなっている事が分かる。

5.4 Appendix

5.5 一度ターゲットに収束したのちに、パラメタをあげるとターゲットに収束しなくなる例

5.1.2 節のターゲットパタン ξ^3 の分岐のところで一度ターゲットパタンに収束した後にまたリミットサイクルになることを少し触れた。今節ではこの様な場合について述べる。図 5.5(a)(b)(c) は、それぞれ入力強度 $\eta^3 = 0.018, 0.19, 0.20$ のときのターゲットパタン $\xi^{2,3,6}$ からの距離の時系列をプロットしたものである。(a) ではターゲットパタンに収束していた軌道が、入力強度をあげるにつれて一度リミットサイクルになり (b)、さらにあげるとまたターゲットパタンに収束している様子が分かる。(a)(c) の軌道の過渡期の部分に着目してみると、(a) では (b) のリミットサイクルの軌道を 1 周期通った後に収束しているが、(c) では素早くターゲットパタンに収束している。(b) の入力強度の相空間においてもターゲットパタン自体は安定な構造を取っており、安定固定点のターゲットパタンとリミットサイクルが共存している。このようにある入力強度で長い過渡期の後にターゲットパタンに収束する軌道が、入力強度をあげると一度ターゲットパタンに収束しなくなりリミットサイクルをとり、さらに入力強度をあげるとより短い過渡期でターゲットパタンに収束するという構造は比較的多く見られるが詳しい機構はまだ未解析である。

6 まとめと議論

この章ではこれまでの解析で得られた結果をまとめ、この結果に必要な機構を考察する。またこれらの現象がどの程度一般的なのかあるいはこれらの意義、何を確かめる必要があるのかを議論する。

6.1 まとめ

まず本論文で得られた結果を以下でまとめる。

1. 1 章で要請した学習モデルを構成することができた (3.2 節)。

   モデルを構成するにあたっての 1 章で行った要請、すなわち細胞のダイナミクスと相空間のダイナミクスを分離しない 1 次元誤差情報による逐次学習は今モデルで可能であることを 3 章では示した。

ここで今モデルのどの性質がこの結果をもたらしたのか議論する。

ニューラルネットワークを用いた学習は、細胞の状態が相空間を形成しつつ相空間の規定を受け変化することで最終的にターゲットが安定な相空間を形成する。このような学習方法において学習済みのターゲットパタンの記憶と新しいターゲットの学習は一般的には相反するものである。なぜなら学習したターゲットパタンを記憶することは相空間を維持することであり、新しくターゲットパタンを学習することは相空間を作りなおすことであるからだ。これを避けるために例えば最急降下法は覚えるパタンに合わせて少しずつ相空間を変化させるという方法をとる。

一方本論文のように逐次学習を行うにもかかわらず今モデルで記憶が可能になっているのは、

1.presynaptic cell が非活性な状態では結合が変化しないようにしたこと。

2. あらかじめスパースなパタンが安定になるように相互抑制項をいれている点と入力パタン・ターゲットもそれに対応させスパースなパタンとしたこと。

3.FB 結合が Hebb phase で変化しないようにしたこと。

の三点が重要であると考えられる。

1. は presynaptic cell が活性化しているところだけにシナプス変化をとどめる。2. は presynaptic cell が活性化している細胞を限定する。これらはシナプスの変化を限定的なものにする効果をもち、探索における相空間の作り替えを最低限に抑える意味がある。またこのように限定することで状態を不安定化するために相空間が変化する際も全次元に渡って変化するのではなく限定された方向のみに変化することになる。探索が終わり収束後の Hebb phase において適切な状態を安定にする際にも同様に相空間の変化を抑制する効果がある。3. は Hebb phase での学習の効果を限定することで前回までの記憶が埋め込まれている相空間を壊し過ぎないようにする効果がある。もし FB 結合も同じように Hebb phase で強化してしまうと現在学習しているターゲットパタンが強くなり過ぎてしまう。すなわち以前の学習した相空間構造を壊してしまい記憶ができなくなる。実際 FB 結合も同じように Hebb phase で変化させるモデルでは複数セットは学習できないことを確認している。また FB 結合と FF 結合の振る舞いが異なることが今モデルでランダムネットワークではなく層型のネットワークを用いている理由でもある。このように FF 結合のみが Hebb phase で強化されるためにターゲットパタンの履歴は FF 結合が担っていると考えられる。

以上のようにここにあげたモデル上の設定が、探索と学習の段階において相空間の変化が大きくなり過ぎないように調整するために複数の記憶が可能になっていると考えられる。実際、上の設定をはずしたモデルに置いても適切な状態を探索すること自体は可能である。しかしこの場合複数の記憶ができないことを確かめている。このように今モデルでは学習パタンをスパースなターゲットにするなど性能を制限することで、記憶と学習という相反する性質を上手く両立させているということができる。

2. 学習に最適な時間スケールの関係が見出された (3.3 節) 。

今モデルにおいて記憶容量が $\tau^{cell,FB,FF}$ の関係に依存し、記憶容量の最大値が $\tau^{cell} \ll \tau^{FB} \ll \tau^{FF}$ のときに得られる事を示した。また各 τ^{FB} の学習において記憶できなかったときに現れるパタンには特徴があり、

最大記憶容量をもつ $\tau^{FB} = 16$ よりも τ^{FB} が速い領域では隠れ・出力層ともに 2 細胞が活性化するパタン

$\tau^{FB} > 16$ の領域では隠れ・出力層ともに 1 細胞だけが活性化するパタン

であることも示した。

このことは細胞の時間スケールとシナプスの時間スケール (相空間の時間スケール) の関係の違いにより、学習で形成される相空間構造は大きく異なることを意味している。これらの時間スケールは学

習においてどのような機能のスケールとみなすことができるのだろうか。本論文のモデルにおいて $\tau^{cell,FB,FF}$ の3つの時間スケールのうち、前者2つが相空間上を探索する時間スケール、後者は学習したパタンを記憶する時間スケールに相当すると考えられる。また前者2つのうち τ^{cell} はパタンからパタンに遷移する時間スケールであり、τ^{FB} は遷移がおこる時間間隔の時間スケールとみなせる。したがって上の関係は遷移の時間スケール < 遷移間隔の時間スケール < ターゲットパタンの保持のスケールという関係があるときに記憶容量が最大化するということを示唆している。

3. 学習されたターゲットパタンは入力のない相空間において、固定点への過渡期の軌道・リミットサイクルの軌道に埋め込まれることを示した (4章)。

   学習段階があがるにつれて相空間は複雑になり、固定点からリミットサイクルへ短い軌道から長い軌道へと相空間は変化すること、その軌道にターゲットパタンが経由するパタンとして埋め込まれていくことを示した。

   そしてこのようになる機構として、探索時に過去の学習パタンを不安定化することでサドル様の構造を形成しこれがつながることにより学習したパタンを遷移するような軌道が形成されるのではないかという機構を述べた。

   このことは学習したターゲットパタンが軌道上に埋め込まれる現象は今モデルに特有のことではなく、相空間上を動きつつ相空間を形成し学習を行う機構をもっていれば一般的に起きる現象である可能性を示唆している。

4. 記憶されたターゲットパタンは入力強度をパラメタとするサドルノード分岐を起こし安定固定点に分岐すること、全ターゲットパタンを記憶できているネットワークと一つでも記憶できていないターゲットパタンがあるネットワークの性質が異なることを示した (5章)。

   本論文では入出力関係の学習を入力 I^a を印加することでターゲットパタン ξ^a が大きな basin volume をもつ安定固定点に分岐するネットワークを形成することとして定義した。このときのターゲットパタンは固定点のみの相空間からの分岐でもリミットサイクルから存在する相空間からの分岐でもサドルノード分岐を経て安定固定点に分岐することを示した。また全ターゲットパタンを記憶できているネットワークから分岐した相空間上ではターゲットパタンのみが安定固定点として存在する一方で、一つでも記憶できないターゲットパタンをもつネットワークから分岐した相空間上では多数の偽ターゲットパタンを含む事を示した。

以上のように本論文において、始めに要請として 1.1 次元の情報で学習を行う 2. 逐次的な学習を行うという要請をみたすため、相空間を (再) 形成しながらターゲットを探索するモデルを構成し、このモデルにおいて学習が上手く行われる時間スケールが存在することを示した。本論文では入出力関係の記憶を、"入力によって、対応したターゲットパタンが安定になるように分岐する"事とみなした。そして学習が上手く時間スケールにおいて、ターゲットパタンは入力のない相空間ではそれらを遍歴するような構造として埋め込まれ、入力印加により対応するターゲットパタンがサドルノード分岐を起こして安定化することで記憶がなされていることを示した。

6.2　議論

このような結果を踏まえてモチベーションで述べた入力がどのように埋め込まれるのかという点について本研究がどのような意義を持つのか議論する。

1章で触れたように教師なし学習は入力を有用に (分類、構造化をするように) ネットワークの相空間に埋め込むにはどうすればよいかという枠組みだと考えることができる。例えば自己組織化マップ [8] [13] は入力の特徴に応じてどのようにネットワークに埋め込むかというモデルとみなせるし、Hopfield モデル [7]

は記憶をヘブ則を用いて相空間に埋め込むという記憶のモデルとみなすことができる。これらのモデルも含め一般に教師なし学習のモデルは入力の構造を固定点アトラクタとして埋め込むモデルだといえる。つまり学習によりネットワークが形成されたのちは入力は初期状態であり、初期状態がどのベイシンにいたかによって最終的な細胞の状態が決まるモデルである。この見方は一度収束した後は変化しないという意味で、記憶を静的な構造として捉えている。

それにたいして津田 [3][6] は固定点アトラクタという静的な構造ではなく、もっと動的な構造とする見方を提示している。つまり記憶は固定点アトラクタのような全次元に渡って安定な構造ではなく、そこから不安定な多様体がでて他の記憶とつながってあり、これにより記憶を' 遍歴' するようなカオス的遍歴という見方である。また Rabinovich [2] は脳が行う情報処理の安定性と柔軟性をどのように両立させるかという立場から記憶がヘテロクリニックな軌道でつながっている、サドルのネットワークのような見方を提示している。

これらの見方ではある記憶されたパタンが活性化した時に、そこから抜けでる軌道があるために同じパタンに留まらず次々にパタンを遷移するという動的な構造を記憶はとっている。このような動的な構造として記憶が埋め込まれるという見方は魅力的だが、入力をどう埋め込むかという学習機構自体の力学系的側面はあまり言及されていない。今モデルは学習のモデルの提示し、入力が 0 の相空間では (十分学習がすすんだ段階では) 学習したターゲットはリミットサイクル上を遷移する構造が自然と現れることを見た。そして細胞の状態がネットワークの相空間に制御されながら相空間を形成していくという今モデルの学習機構により上の様な相空間構造が形成されることを示唆した。また入力をいれることで相空間が分岐しターゲットが安定化した結果、入力に対応したターゲットを出力するという関係を系は学習しているという見方も提示した。

このように入力のない状態では学習されたパタンはリミットサイクルに埋め込まれ、入力をいれると相空間構造が変わるとターゲットが安定化するという描像は実験とどのような対応が付けられるのだろうか。最近脳の自発発火がどのような意味をもつのかという点にかんして興味が持たれており、例えば入力のない閉眼時の状況で orientation map の発火パタンがスイッチングしている現象 [5] や EEG でのパタン遷移 [11] などとして報告されている。自発発火の状態は今モデルでの入力のない相空間に対応すると考えられるので、上で紹介したように自発発火の実験と関連づける事により今モデルは新しい脳の見方を提示できる可能性をもっていると考えている。

6.3 課題

今モデルでの未解析な点、問題な点をのべる。

今モデルの結果として学習の時間スケールの依存性を示した。この機構として速い時間スケールでは遷移時間が伸びることで、遅い側では滞留時間が伸びることを指摘した。このうち特に遅い側の原因である滞留時間の増加と時間スケールの関係について理解ができていない点が多い。τ^{FB} は遷移間隔の時間スケールであるので、τ^{FB} が大きくなるところで滞留時間が増加するのは自然である。一方で遷移時間そのものはあまり増加していない。単に遷移間隔が大きくなるだけであれば遷移時間もその分大きくなるのが自然だと思われる。したがって遷移の仕方そのものが変化していると考えられるがこれがどのような変化かよく分かっていない。

また学習したターゲットパタンは軌道に埋め込まれる事を示しその機構として探索過程でターゲットパタンが不安定化することを述べたが、本論文では実際に探索過程の遷移とその後の相空間構造をひとつの学習過程として関連づけて理解する段階まで達してない。あくまで各学習段階での統計的な挙動を解析した結果として上述の結論を得たものである。したがって 1 学習過程に関して今後重点的に解析を行う必要がある。またこれらの解析を異なる時間スケールに関して行うことで上述の問題についてもなんらかの示唆が得られると考えている。

謝辞

研究面における刺激的なアドバイスと、本論文執筆にあたり何度もチェックをいれてくださった金子さんに、研究環境一般を維持してくださった研究室のみなさんに、特に自分以外の super user とそれをサポートしてくれた中島さんと立川さんに感謝します。

最近駄目さを痛感している自分がまがりなりにも楽しくやってこれたのは友人のおかげであり、ここに感謝します。

最後に一年余分な時間をとらせて心配をかけた家族に最大限の感謝と謝罪を述べたいと思います。

参考文献

[1] Per Bak and Dante R. Chialvo, Adaptive learning by extremal dynamics and negative feedback,PRE vol. 63 031912 (2001)

[2] Mikhail I. Rabinovich, Ramon Huerata, Pablo Varona, Valentin S. Afraimovich, Transient Cognitive Dynamics, Metastability, and Decision Making, Plos comput. biol. 4(5) e 1000072 (2008)

[3] 津田一郎、複雑系脳理論、サイエンス社 (2002)

[4] Rumelhart,D.E, and J.L,McClelland, Parallel Distributed Processing, MIT Press (1985)

[5] Ta Kenet ,et.al,Spontaneously emerging cortical representations of visual attributes, Nature 425 954-956 (2003)

[6] 金子邦彦、津田一郎、複雑系のカオス的シナリオ (1996)

[7] Hopfield.J.J,Neural networks and physical systems with emergent collective computational abilities,Proc. Natl. Acad. Sci. USA vol.79 2554-2558 (1982)

[8] T.Kohonen,Self-Organizing Maps,Springer (2000)

[9] George Lakoff, 認知意味論、紀伊之國屋書店 (1993)

[10] R.S. Sutton and A.G. Barto,Reinforcement learning,MIT Press,Cambridge,MA,USA (1998)

[11] Junji Ito, et al,Dynamics of spontaneous taransitions between global brain states, Human Brain Mapping 28 9 904-913 (2007)

[12] 銅谷賢治, 計算神経科学への招待, サイエンス社 (2007)

[13] Willshaw ,D.J, and C. von der Malsburg,How patterned neural connections can be set up by selforganization, Proceedings of the Royal society of London series B, vol. 194 431-445 (1976)

[14] Simon Haykin,Neural Networks ,Prenntice Hall International,Inc (1999)

[15] T.M.Jay,Dopamine;a potential substrate for synaptic plasticity and memory mechanism, Progrss in Neurobiology 69 375-390 (2003)

[16] Eve Marder, Vatsala Thirumalai, Cellular, Synaptic and network effects of neuromodulation, Neural Networks 15 479-493 (2002)

[17] A. G. Barto,R. S. Sutton, and Peter S. Brouwer, Associatice Search Network, Biool. Cybern. 40, 201-211 (1981)

[18] X. Xie and H. Sebastian Seung, Learning in neural networks by reinforcement of irregular spiking, PRE vol. 69, 041909 (2004)

[19] Peter Dayan and L.F.Abbott,Theoretical neuroscience : computational and mathematical modeling of neural systems, The MIT Press, (2001)

[20] Schultz, W., Dayan, P. and Montague, P. R., A neuralsubstrate of prediction and reward. Nature 275 1593-1599 (1997)