

## 論文の内容の要旨

### 論文題目 Spontaneous and Evoked Neural Dynamics Shaped by Embedding Memories

(記憶の埋め込みによる自発及び誘起神経活動の形成)

氏名 栗川 知己

生物の神経系がどのように情報を処理しているのかを理解することは神経科学の重要な課題の一つである。適当な認知的な課題あるいは感覚刺激を与えその応答を調べることで、その課題に必要な認知機能または感覚情報の処理が脳のどこで処理されているかが多くの研究を通して、これまで明らかにされてきた。このように神経科学は古典的には、何か外部から刺激を与えその応答を測ることで脳の機能の解明を行ってきた。通常、同じ刺激を与えても同じ応答は返ってこないで、同じ刺激を繰り返し与えてその平均を取ることで、刺激に対する応答を測定する。ここでは、試行間の応答の違いは単なるゆらぎとして処理され、一つの試行がもつ情報は平均化により失われてしまう。

しかし、近年の実験技術、特に多点観測による多数の神経細胞の同時計測技術の発達により、一度の試行から多くの情報を取り出すことができるようになってきた。それにともない、今まで単なるゆらぎだと考えられていたものが、実は組織だった時空間構造をもっていることが明らかにされつつある。また、刺激が与えられる前の状態に応じて、その後の応答が変化するということが明らかにされてきた。このように自発神経活動が神経系の情報処理に大きく関与し

ていることが明らかにされてきた（明確な刺激の与えられていない状態での神経系の活動は自発神経活動と呼ばれ、それに対して何らかの刺激が誘起する神経活動は誘起神経活動と呼ばれる）。このことは、自発神経活動の理解が神経系の情報処理の理解にとって重要であることを意味している。

同様に、そもそも情報がどのように表象されているかという点も、神経系の情報処理の理解には重要な要素となる。伝統的な仮説として、外部刺激が神経細胞が構成する力学系のアトラクタとして表象されているという仮説がある。この“表象＝アトラクタ”の標準的なモデルとして、ホップフィールドネットワークなどの連想記憶モデルが研究されてきた。このモデルにおいて、刺激の表象はアトラクタとして埋め込まれる。複数の表象があった場合、対応するアトラクタは同じ相空間に埋め込まれるので、多数のアトラクタが共存する多安定状態になる。そして外部からの刺激はその相空間上の一点を選択し、その点がどのアトラクタの収束領域に入っているかに応じて、刺激とそれに対する応答が定まる。

このモデルの利点の一つとして、ゆらぎに対する安定性がある。入力が多少ゆらぎ、選ばれる初期状態が多少変化しても同じアトラクタの吸引領域に入っていれば同じ応答を返すことができる。また同じ理由でダイナミクスにゆらぎが入っていても、アトラクタへ収束する経路が極めて狭くないかぎり、同じ応答をすることができる。このような現象は間接的ではあるが、海馬の場所細胞や嗅覚系の応答などで実験的に支持されている。

このように、“表象＝アトラクタ”モデルは大変有用な概念である。しかし、幾つかの問題点も指摘されている。記憶の数が増えて、それに伴いアトラクタの数が増加する。同一相空間でアトラクタが増加するので、必然的に1つあたりの吸引領域は減少する。つまり記憶の安定性は減少する。また、本論文の文脈でより重要な課題として、現モデルでは自発神経活動の解析を行うことが難しい点がある。なぜなら、入力は同一の相空間上にある多数のアトラクタから一つを選択する初期条件であり、系のダイナミクス自体は変化しない。入力がない（すべての要素がゼロ）場合でも同様である。

上述のように、自発神経活動は脳における情報処理に重要な役割を果たしていると考えられるにもかかわらず、今のモデルでは自発神経活動を解析することが難しい。そこで、本論文では、“表象＝分岐”という新しい枠組みを提示し、それを実装するモデルを導入し解析する。

我々の枠組みでは入力は系の初期状態ではなく、系のパラメタとして導入される。入力がない状態から、印加した状態へパラメタが変化することで系が変化

(分岐) する。この点が、入力系の初期状態の選択であり系自体は変わらない前述の枠組みと大きく異なる点である。入力に対する応答は、入力により変化した後の相空間上のアトラクタとして定義される。この点はアトラクタとして入力に対する応答が表象されることは前述の枠組みと同様である。

本論文で用いられる枠組みでは、複数のターゲットを記憶する場合でも一つの相空間上に複数のアトラクタが存在する必要はない。なぜなら、入力毎に系の相空間自体が変化し、変化後の相空間構造は入力によって異なり得るからである。したがって、上述の記憶数が増えた場合に記憶の安定性が必然的に減少するという問題は回避できる。実際本論文で導入されたモデルでは、入力により変化した相空間においては少数のアトラクタしか生じない。

また、この新しい枠組みは実験と相反するものではない。例えば嗅覚系では入力が印加、この系ではにおい物質を与えることにより、応答する神経細胞の発火パターンがにおい物質の種類に応じて異なるパターンに変化することが報告されている。また、におい物質を排除するとその発火パターンは元の発火パターンに戻ることも報告されている。

この枠組みに基づき本論文では、以下の3つのネットワークモデルを導入した。そして自発神経活動と誘起神経活動の関係、特に自発神経活動において誘起神経活動に似たパターンが出現するという、実験で特徴的に見られる現象に何らかの機能的な意味があるのかという点に着目して、解析を行った。

1. ホップフィールドネットワーク(HNN)を改変したネットワークモデル。具体的には HNN からネットワークの結合強度を

$$J_{ij} = \sum_{\mu} \xi_i^{\mu} \xi_i^{\mu} \rightarrow \sum_{\mu} (\xi_i^{\mu} - \eta_i^{\mu})(\xi_i^{\mu} + \eta_i^{\mu})$$

のように改変した。但しここで $(\eta_i^{\mu}, \xi_i^{\mu})$ は学習すべき入力とそれに対応する出力パターンのペアを表す。詳細は本論文で述べられるが、新しい結合形より入力印加時のダイナミクスが HNN と同じ形の表現されることになる。これにより、入力印加時にターゲットが安定な固定点となる。

本研究ではこのモデルが多数のペアを学習することができ、その記憶容量は  $0.7N$  (ここで  $N$  は要素数) 程度であることを、数値シミュレーションを用いて示した。また学習数が極少数の場合を除き、自発神経活動はカオスとなり、ターゲットに他の入力パターンやランダムパターンに比べより近づく挙動を示すことを明らかにした。

2. ランダムネットワークに学習を導入したモデル。上のモデルは新たな枠組みの元、多数の学習が可能であることを示したが、結合強度は予め上述の結果が得られるようにデザインされていた。したがってこの結合強度がどの程度特殊であるのかを調べるために、本モデルでは、ランダムネットワークから学習を行うモデルを導入した。採用された学習則はよく用いられるヘブ則のようにシナプス前細胞、後細胞の二体相関で定まる簡単な学習則であり、新しい枠組みのための特殊な学習則ではない。

この結果、学習パラメタと学習時の入力強度がある程度の条件をみたせば、このモデルでも多数の学習ができることを示した。さらにこの場合、ネットワーク構造は先のモデルに類似した特徴を持つ構造が形成されていることを明らかにした。

3. 強化学習を用いた複数のシナプスの時間スケールをもつレイヤー型のネットワークモデル。このモデルではレイヤー間結合強度が学習により変化することで、入力と出力を学習する。この結果、最大の記憶容量を実現する時間スケールの関係が存在し、その関係を満たすところで、ターゲットを変遷するような自発神経活動が形成されることを明らかにした。

以上のように本論文では近年注目を集めている自発神経活動の機能の解析を行うために、伝統的な“表象＝アトラクタ”の枠組みに変わる“表象＝分岐”という枠組みを提案し、この枠組みに基づいた3つの異なるモデルを導入した。そして、それらについて自発神経活動と誘起神経活動の関係に着目した解析を行った。結果として、3つのモデルで記憶が形成される領域では自発神経活動はターゲットを変遷する構造が形成されることを示した。ターゲットは入力パターンを印加した時の応答と考えられるので、以上の結果は新しい枠組みのもとで、記憶を行うために形成される構造として、自発神経活動が誘起神経活動と似たパターンを遷移するが形成される可能性を示唆していると考えられる。