

栗川知己氏の博士論文 *Spontaneous and Evoked Neural Dynamics Shaped by Embedding Memories* は記憶の埋め込みに関して、力学系から新しい視点を提示し、それにより自発神経活動と刺激による誘起活動間の関係、さらにそれが学習によりいかに形成されるかを示したものである。本論文は5章118ページからなり、第1章は導入説明、第2, 3, 4章では神経ネットワークのモデルと解析が行われ、第5章では全体の考察と今後の展望が述べられている。

神経系がどのように情報を記憶し、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。入力刺激に応じて、ある出力が現われるように神経系は学習により形づけられていく。近年、神経細胞の同時計測技術の発達により、明確な刺激の与えられていない状態でも神経系は自発的に活動しており、これ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、またこれと刺激後の誘起神経活動との連関も見出され、自発神経活動と神経系の情報処理がいかに関連しているかの理解が待たれている。

一方で、記憶の伝統的な仮説としては、多数のアトラクタを持つ力学系を用意し、入力に応じてそれぞれのアトラクタに落ちるよう初期条件が選ばれるという描像がとられてきた。この立場での連想記憶モデルは広く研究されてきた。しかし、この描像では、自発神経活動の解析を行うことは困難である。というのは、この見方では、入力がないと初期条件を設定できなくなり、時間発展を記述できないからである。仮にそれが克服できたとしても、この描像では神経活動の時間発展は入力の有無によらず同一の力学系で記述されているので、自発活動と誘起活動の関連を議論するのは困難である。

この点を考えて、栗川氏の論文では、力学系の分岐としての表象という新しい枠組みが提示される。この枠組みでは入力は系の初期状態ではなく、力学系そのものを変える入力項で与えられる。入力の強さはその項のパラメータとして導入され、それに応じて神経活動の状態は変化し、入力がないときの自発活動状態から、誘起活動状態へと変化（分岐）する。この描像では、多くの入出力関係を記憶するために複数のアトラクタが存在する必要はない。入力毎に系の相空間自体が変化し、適正な状態がつくられればよいからであり、記憶はこうした分岐構造の集合を与えるような力学系として埋め込まれることになる。この新しい枠組みに基づき本論文では、3つのネットワークモデルを導入し、その振る舞いが調べられ、入出力関係の記憶がいかに神経活動の力学系にうめこまれるか、そして自発神経活動と誘起神経活動の関係が詳しく解析されている。

第1章で上記のような本論文の方向が示された後、第2章では入出力関係を埋め込むように、従来の連想記憶モデルを改訂する。具体的には、学習すべき入力とそれに対応する出力パタンのペアを  $(\mathbf{u}^{\mu}, \mathbf{v}^{\mu})$  ( $i$ は神経素子、 $\mu$ は入出力ペアのインデックス) としたと

きに、出力の相関だけで結合行列を決める従来の形ではなく、入出力で決める以下の形

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

が提唱される。理論的考察とシミュレーションからこのモデルが多数の入出力関係を学習できることが示される。具体的には、神経素子数を  $N$  とすると、その記憶容量は  $0.7N$  程度となる。この時、入力がない状態での自発神経活動は一般にはカオスを示すが、これはランダムではなく、記憶されているターゲットたちに選好して近づく挙動を示す。入力を与えると、このカオス状態から、入力に対応した出力を示す誘起活動への選択的变化が生じる。この結果は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、最近の実験結果と対応している。

第2章では多数の学習が可能であるように結合強度をデザインしたものであった。では、入力下で出力すべきターゲットの情報を受けていって、多数の入出力関係を学習できるであろうか。この問いに答えるべく第3章では、シナプス前細胞、後細胞の二体相関で定まるヘブ則タイプの簡単な学習則が導入される。ただし、出力がターゲットとずれていると結合強度を変えるという点が重要である。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、このモデルが逐次に多数の学習ができることがシミュレーションにより示された。形成されたネットワーク構造はもちろん第2章のモデルと同一にはならないが、興味深いことに、結合強度と入出力  $\xi$ 、 $\eta$  との符号の関係では先のモデルに適合した構造が形成されていることが見出された。そして、形成された自発活動は第2章で見出されたと同様に、ターゲットたちに選好性を持ったカオスの振る舞いを示す。

以上の第3章の学習規則では、ターゲットの全情報が必要とされていた。ターゲットと出力の間の差という一つの情報だけを用いた強化学習で、入力関係を学習できるであろうか。第4章では、大域的結合ではなく、入力、中間、出力の3層からなる型神経ネットワークを用いて、層間に2つのシナプスの時間スケールを導入することでこの問いに答えている。まずヘブ則を拡張した学習規則を用いて、入出力学習の逐次記憶が可能であることが示される。この場合、最大の記憶容量を実現するにはフィードバックのシナプス変化がフィードフォワードのよりも十分速いという条件が見出される。次いで、この系においてもターゲットを変遷するような自発神経活動が形成されている。

第5章は得られた結果の脳科学への意義、今後の展望にあてられている。

以上のように、栗川知己氏の学位論文では、神経系の形づくる記憶に関して、従来のアトラクタ描像と異なった、力学系のフローの入力による変容（分岐）に基づく、新しい描像が提示されている。記憶を多重安定系でとらえる従来の描像に替り、入力による力学系の変化の仕方を埋めこんでいくのが記憶である、という新描像である。その結果、自発脳活動と誘起脳活動の連関が議論出来るようになる。この描像を具現化した、3つのモデルが導入され、学習による力学系変化と入力による力学系変化をつなぐ解析がなされている。その結果、誘起活動で現われるパターンへの選好性を持った自発的カオス活動が見出された。

出力を準備した自発活動が入力のない時に既に存在しているという栗川氏の知見は、近年の嗅覚系などの実験との対応でも興味深い。もちろん、この新描像により、汎化、カテゴリー化を含む脳理論をいかに展開するか、本論文で示唆された実験との対応をいかにして明確にして検証していくか、など今後なすべきことは多い。とはいえ、本論文は、今後展開されるべき脳理論の新しい方向への基盤をあたえたものとして大いに注目されるものである。

なお、本論文の2, 3, 4章は金子邦彦との共同研究であるが、いずれも論文の提出者が主体となってモデル化、シミュレーション、理論解析を行ったもので、論文提出者の寄与が大であると判断する。よって本論文は博士（学術）の学位請求論文として合格と認められる。