

## 総括研究報告書

1. 研究開発課題名：強化学習の神経回路機構の解明へ向けて
2. 研究開発代表者：森田 賢治（東京大学 大学院教育学研究科）
3. 相手国研究代表者：Abigail Morrison（ユーリッヒ総合研究機構（ドイツ））
4. 研究開発の成果

（1）皮質 crossed-corticostriatal (CCS)・corticopontine (CPn)細胞モデル

詳細な解剖学および生理学の実験データ（生理学研究所の森島助教・川口教授らによって得られたデータ）に基づく NEURON シミュレータ上の数理モデルのパラメータの推定が近藤将史研究員によって行われ、実験データを再現できるようなパラメータセットが得られた。その上で、得られたパラメータセットを持つモデル（CCS 細胞・CPn 細胞それぞれ二つずつ）を用いて、集団として様々な周波数で振動するようなシナプス入力を与えたシミュレーションが近藤研究員によって行われ、短い時間間隔で連続する二つの活動電位“doublet”の発生の条件などが明らかになった。

（2）大域回路モデル

皮質-基底核-中脳から成る神経回路において基底核直接路・間接路が阻害された場合について大域的モデルのシミュレーションによる検討を進め、直接路が阻害されると正の予測誤差に基づく学習に問題が表れた一方間接路が阻害されると負の予測誤差に基づくことが考えられる反転学習初期に問題が表れたという実験結果などを、両経路が価値情報を時間差を持って表しているという仮説に基づいて説明しうることなどを明らかにした。また、それぞれの経路の阻害と依存症への関わりについて調べ、その関わり方が、行動選択（意思決定）の探索の度合いに応じて大きく変わりうることなどを明らかにした。

（3）ドーパミン濃度変化モデル

ドーパミンが報酬予測誤差を表すことを仮定した大脳皮質-大脳基底核-中脳神経回路の強化学習回路モデルに皮質-線条体シナプスの可塑的变化が時間的に減衰するとの仮定を加えたものによって、報酬希求行動のモデリング・シミュレーションを行い（加藤・森田）、最近複数の実験で報告されたような大脳基底核線条体におけるドーパミン濃度の緩やかな変動（上昇）などの時間パターンが説明されうることを示し、さらに、モデルのシミュレーション・解析に基づいて、かねてから示唆されてきたドーパミンとモチベーションの関係についての回路・シナプスレベルのメカニズムを提案した。

（4）統合的機構の探索

上記の研究をはじめ日本側およびドイツ側でこれまで進めてきた研究、および分野の他の研究者らによる研究も合わせて、強化学習の神経回路機構について、行動選択（意思決定）に焦点を当てつつ総合的な考察を行った。そして、強化学習のアルゴリズム（Q-learning, SARSA）が神経回路によっていかに実装され得るかについての仮説の提案や、最近の実験・理論研究において示唆されているような大脳皮質および大脳基底核線条体の複雑な神経活動パターンがいかなる計算論的意義を持つかについての複数の具体的な可能性の提案などを含めて、日本側・ドイツ側共同の総説論文としてまとめた。