

研究課題別事後評価結果

1. 研究課題名 「行動系のメタ学習と情動コミュニケーション機構の解明」

2. 研究代表者名及び主たる研究参加者名 (研究機関名・職名は研究参加期間終了時点)

研究代表者 銅谷 賢治 (国際電気通信基礎技術研究所脳情報研究所 室長)

主たる研究参加者

松本 隆 (早稲田大学理工学部 教授)

関野 祐子 (群馬大学大学院医学系研究科 助教授)

澤口 俊之 (北海道大学医学研究科 教授)

山脇 成人 (広島大学大学院医歯薬学総合研究科 教授)

石井 信 (奈良先端科学技術大学院大学情報科学研究科 教授)

3. 研究内容及び成果:

人間や動物は、機械システムとは異なり、環境の変化に適応するきわめて柔軟な学習を実行している。そのためには、いわゆる脳の結合のパラメータであるシナプスを改変する学習の上に、学習の仕方を変化させる学習方式の学習とも言うべき、メタ学習の機構を備えている。脳の「メタ学習」の機構を理論的、実験的手法により解明するとともに、その成果をより人間的なロボットやコンピュータの設計論に結びつけて行くことが本研究の目標である。

これを実現するために、具体的な内容として、

- (1) 学習系のメタパラメータ調節の理論と神経修飾物質系の機構
- (2) 課題や環境に適した内部表現の獲得における大脳皮質の機能
- (3) メタ学習機構としてのコミュニケーションと進化の役割

という3つの主要テーマを掲げ、計算理論 / 生理実験 / 脳活動計測 / ロボット実験の連携という異なる方法を融合し駆使してその解明に当たった。

研究は6個のサブグループに分かれ、それぞれが分散しつつも全体として協調的に展開した。

グループ毎にその成果を述べる。

1) システム統合グループ(銅谷賢治 国際電気通信基礎技術研究所 室長)

a) メタ学習と神経修飾物質系モデル

メタパラメータの最適化を行う階層的強化学習モデルを定式化し、その脳内での実現の可能性を示した。異なる時間スケールでの報酬予測機構に関する脳活動計測実験を行い、強化学習モデルを用いた解析により、短期および長期の報酬予測に関与する脳の部位を固定し、これがセロトニンにより制御されていることを明らかにした。

b) 目的指向的内部表現の獲得

追跡眼球運動の行動心理実験により、ヒトの運動の予測時の脳活動計測を行い、MST 野が視標運動の予測に関与することを示した。

c) ネズミ型ロボット群「サイバーローデント」の学習と進化実験

自己保存と自己複製の要請のもとでの報酬系の設計原理を探るため、ロボット「サイバーローデント」を開発し、行動学習のメタパラメータと報酬関数の進化を実現した。

2) 学習理論グループ(松本隆 早稲田大学理工学部 教授)

機械学習の視点から、強化学習におけるメタ学習のアルゴリズムを主としてベイズ的枠組みから構築し、以下の結果を得た。

a) 強化学習方式を構築し、「逐次周辺尤度」(Sequential Marginal Likelihood) をもとに環境の変化を検出し、その上でメタパラメータの適応的学習が可能である事を示した。

b) 並列モデルにより、価値関数の経験分散最小メタ学習手法を提案し、その有効性を具体例で示した。

c) 強化学習を含む広いクラスの問題に対して Bayes 的枠組を定式化し、その有効性を示した。

3) 適応システムグループ(石井信 奈良先端科学技術大学院大学 教授)

環境の変化、観測の不確かさなどに応じて、適切にメタパラメータの自律制御を行う強化学習の工学的アルゴリズムを開発し、不完全観測環境での制御問題、不安定系の自律制御などに適用した。さらに、そうした強化学習法の脳内実現のモデルを構築し、行動実験やイメージング研究などの手法により検証を行い、情動情報処理における行動系メタパラメータ制御の意義を検討した。さらに、観測に不確かさがある場合に、環境のダイナミクスをベイズ推定しながら、探索と搾取の切替えを行うモデル同定型強化学習アルゴリズムや、環境の変化のベイズ推定に基づき、学習係数を自律的に制御する学習法を開発した。

また、これら強化学習アルゴリズムで必要な脳内機能の局在モデルを提案し、その一部を検証するための機能的磁気共鳴図を用いたイメージング研究を行い、前部前頭前野および内側前頭前野が不観測状態の推定に関わるという知見を得た。

4) 情動記憶グループ(関野祐子 群馬大学大学院医学系研究科 助教授)

情動による記憶制御機構を明らかにする目的で、海馬における新規情報処理の神経回路メカニズムの研究を行った。海馬における記憶形成は、脳内アデノシンにより抑制的に制御され、また上乳頭体からの入力により促進的に制御されるとの仮説の検証実験を行い、海馬 CA2 領域が新規環境に置かれたラットの探索行動で、海馬神経活動を促進していることを明らかにした。

5) 学習修飾グループ(澤口俊之 北海道大学医学研究科 教授)

試行錯誤学習の基礎となると考えられる、環境の変化に応じて行動指針(ルール)の切り替えに関わる前頭前皮質ニューロン機構に焦点を当て、主にサル前頭前皮質における単一ニューロン活動記録法により研究を進めた。この結果、報酬獲得時に見られる神経活動が行動の評価に関わること、前頭前皮質は報酬期待に基づいた行動指令の生成過程に関与すること、また変化する環境に応じて行動指針を変換し保持する過程に関わることを示した。これにより、前頭前皮質は、自身の行動の結果を評価し、その評価に基づいて行動やルールを生成する過程に関与すると考えられる。

6) 精神薬理グループ(山脇成人 広島大学大学院医歯薬学総合研究科 教授)

人間の学習、情動などの高次脳機能のメタ学習の制御に関与する神経修飾物質系の機能、特にセロトニン系が短期的コストと長期的メリットのバランスを決定するという仮説の検証に焦点をあて、実験的な検討を行った。すなわち、健常者およびうつ病患者を対象に将来の報酬予測に基づく意思決定課題遂行中の脳活動を機能的核磁気共鳴画像法(fMRI)を用いて測定し、将来の報酬予測および反応抑制に関する脳機能を検討した。この結果、中枢神経系セロトニンが短期的および長期的期待にかかわる気分に影響を与えていると考えられることを示した。

4. 事後評価結果

4 - 1. 外部発表、特許、研究を通じての新たな知見の取得などの研究成果の状況

脳の学習はきわめて柔軟なもので、その仕組みは何重にも組み合わさっている。すなわち、通常想定されているシナプス可塑性に基づく学習を基礎に、その学習過程を修飾し、学習や探索を加速したり逆に精密化したりする過程が考えられる。この他に、学習の目標に対する長期的または短期的な戦略のどちらに重点を置いて選択を行うかを決めるパラメータもある。これらの学習を修飾するパラメータをメタパラメータと呼び、これをセロトニン、ドーパミン、アセチルコリン、ノルアドレナリンなどの神経修飾物質の機能と結びつけて考察し、ロボットに実装しようという、大変野心的な研究であった。これを目指して、生理学、生化学、精神科学、脳機能測定、さらに理論とロボティクスを統合した研究を行ったもので、数々の重要な知見を得た。これは数多くの報道発表となって表れ、また多数の学術雑誌および国際会議の広い分野での研究発表から明らかである。

4 - 2. 成果の戦略目標・科学技術への貢献

本研究は高い戦略目標を掲げ、そのために異なる分野の最先端の研究者を組織して効果的な研究チームを作り上げた点で、競争的研究資金による共同研究の見本を示したといえる。とくに、脳の学習機能の研究を軸に、精神疾患による機能不全から、ロボットの学習までを統合して研究する点で、科学技術の異分野の交流、融合のあり方として優れたものであった。

しかし、研究は完成したものではなくて、いくつかの分野でそれぞれの拠点となる成果を得たもの

の、それらが繋がっているとはまだ言い難い。今後の成果に大いに期待したい。

4 - 3 . その他の特記事項(受賞歴など)

国際学術誌「Neural Networks」は、本研究を中心に特集号を組んでこれを世に広め、国際的な反響を呼んだ。また、本研究代表者は、沖縄新大学院大学設立の先行研究に選ばれ、そこで国際夏の学校を組織するなど、国際的な活躍も目立っている。脳研究にひとつの新しい流れを作るものと期待したい。