

「脳を創る」

平成 11 年度採択研究代表者

銅谷 賢治

(㈱国際電気通信基礎技術研究所 主任研究員)

「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」

1. 研究実施の概要

本研究の目的は、脳の柔軟で効率よい学習を可能にしている高次の制御機構、つまり「メタ学習」のメカニズムを明らかにすることである。そのため、学習の計算理論、神経修飾物質系の生理実験とモデル化、人間の行動実験と脳活動計測、ロボットによる行動学習実験など、多面的なアプローチを採用する。これにより、脳の学習を支える情動やコミュニケーションの役割を解明し、より自律的な学習ロボットや、より人間になじみやすい情報通信の基礎技術を提供することをめざす。

近年、神経回路網と統計的推定の理論の統合により、学習プログラムの音声認識や視覚処理、情報検索やロボット制御への応用が進んでいる。しかし多くの場合、学習プログラムの動作を調節する「メタパラメタ」の設定は、開発者の経験と直感を頼りに設定されており、これが学習プログラムの一般的な普及のネックになっている。

一方近年、脳の学習の回路や物質レベルでのメカニズムの解明は急速に進んでいるが、脳の持つ柔軟な学習機能は、この分散並列的な回路での学習を高次のレベルで調節する「メタ学習」の機構の存在を示唆している。このようなメタ学習の機構の解明は、真に自律的な学習システムの構築のためにも、また脳の高度な適応機構の理解に向けても、非常に重要な課題である。

そこで当研究チームでは、

- i) 学習系のメタパラメタをいかに調整するか
- ii) 適応的な内部表現やモジュール構造をいかに獲得するか
- iii) コミュニケーションや進化により学習がいかに促進されるか

という3つのテーマについて、多面的な研究を展開している。

学習系のメタパラメタ調節に関して、神経修飾物質系の機能に着目し、

- a) ドーパミンは予測される報酬の増減を表す。
- b) セロトニン報酬予測の時間スケールを制御する。
- c) ノルアドレナリンは探索のランダムさを制御する。
- d) アセチルコリンは記憶の更新の速度をを制御する。

という作業仮説のもとに、仮説の検証に向けた実験と、これらの物質と環境、行動の相互作用を予測する理論モデルの構築を進めている。

内部表現とモジュール構造の獲得に関して、大脳皮質の分散表現による隠れ状態の推定モデルや、小脳の予測モデルの競合と協調によるモジュール学習制御方式を定式化し、計算機シミュレーションや、人間の眼球運動計測実験により、それらの妥当性の検討を行っている。

学習を支えるコミュニケーションと進化の役割を探るためのプラットフォームとして、エネルギーの自己補給とプログラムの伝搬機能を備えたネズミ型のロボット集団を開発した。これにより、自己保存と種族保存という生物と同じ拘束条件のもとで、必要なメタ学習の機構に迫ることが目標である。

今後、神経修飾物質系の機能と、環境や行動との相互作用の理論的解明により、これまで「コンピュータにはわからない」とされていた、人間の感情や個性、コミュニケーションにおける感性的、社会的要素のより深い理解と、様々な行動、情動障害の治療や、教育や人間関係のあり方への指針を与えることをめざす。

さらに、脳のメタ学習機構の解明は、未知の環境に適応する真に自律的な学習ロボットの開発や、人間の感情を、単に表情や声色などの識別に留まらず、その自律行動系の制御機構としての本質に迫る形で理解するコンピュータの設計への道を開く。

2. 研究実施内容

学習のメタパラメタ調節、内部表現とモジュール構造の獲得、コミュニケーションと進化の役割の解明に向けて、今年度は以下のように研究の展開を進めた。

1) 計算理論

ベイズ学習理論は、学習系のパラメタ設定とメタパラメタの選択を統一的に行うことを可能にする。その具体的な計算手法として注目されている、マルコフ連鎖モンテカルロ法(MCMC)や、変分ベイズ法(VB)を、非線形の時系列予測、データのクラスタリングや低次元モデル化、神経活動データの相互情報量解析、fMRI 計測データからのモジュール抽出、MEG データからの信号源推定などに適用し、その性能と実用性を確認した。

報酬の予測に基づく探索的な行動学習の枠組みである強化学習において、報酬予測の時間スケールを決める「割引率」と、行動探索のランダムさを決める「逆温度」は、学習系の効率と性格を決める重要なメタパラメタである。これらを、環境の特性と学習の経過に応じて適応的に制御するアルゴリズムを定式化し、非線形制御課題や迷路探索課題に適用し、その性能を確認した。

2) 生理実験

サルの前頭前野における神経修飾物質の機能に関して、報酬依存的な行動切り替えにおけるドーパミンの役割、注意におけるノルアドレナリンの役割を明らかにする実験を行った。さらに、注意状態に応じた神経発火パターンの変化を明らかにするための実験を進めている。

海馬における記憶の形成が、上乳頭体核からの入力によりアデノシン受容体を介して制御されているという仮説のもとに、ラットの上乳頭体核の破壊実験を行った。その結果、情動行動に変化が見られることが確認された。さらに、アデノシン受容体過剰発現や上乳頭体破壊が、学習行動に及ぼす影響を調べる実験を進めている。

3) 人間の行動実験と脳活動計測

長期と短期の報酬予測に関わる脳部位を明らかにするため、カード選択やマルコフ状態遷移などの行動課題を開発し、fMRI による脳活動計測実験を行った。その結果、長期の報酬予測には、右外側前頭前野、右頭頂葉に加え、前帯状回皮質、大脳基底核、小脳左半球が関与することが明らかになった。また、一般的な脳賦活課題の fMRI 計測の結果から、うつ病患者では、前帯状回皮質の活動が対照群被験者に対して有意に低いことを確認した。

前帯状回皮質は、脳幹のセロトニン中枢と強い相互結合を持つことが知られており、これらの結果は、前帯状回皮質を中心とした回路が、セロトニンの制御のもとに長期の報酬予測に基づく行動学習に関与し、その機能低下がうつ病の要因であることを示唆している。今後さらに、食餌制限によるセロトニン機能低下が、報酬予測課題の成績と脳活動に与える影響を調べる予定である。

また、予測的な状態表現の獲得に関して、カルマンフィルターを拡張したモデルが大脳皮質の回路に実現される可能性を検討し、その検証に向けた眼球運動計測実験を開始した。

4) ロボット実験

これまで、ロボットの強化学習の研究は、任意に設定された「報酬」の最大化を目標に行われてきた。しかし脳の強化学習のメカニズムを考える場合、報酬系は何らかの形で生物の基本的要件である自己保存と種族保存につながる形で決められているはずである。そこで、生物と同じ拘束条件のもとでの学習とメタ学習の役割を探るため、電池の捕捉による自己充電と、赤外線通信による個体間のプログラム転送機能を持つロボット集団“Cyber Rodents”を開発した。

今後、このプラットフォームを用いた実験とシミュレーションにより、行動学習とメタ学習のアルゴリズム、目的指向的な内部表現の獲得、コミュニケーションと進化の役割についての研究を展開する予定である。

これらの研究活動に加え、この分野の研究全般の活性化と当研究チームの対外アピールのため、2001年4月に、京都府精華町のけいはんなプラザにて、“CREST Workshop on Metalearning and Neuromodulation”を開催した。ドーパミン、セロトニン、ノルアドレナリン、アセチルコリン等に関する国内外の先端的な研究者を招き、海外からの一般参加を含め84名の参加を得た。さらにこのワークショップの参加者を核にした Neural Networks 誌の特集号“Computational Models of Neuromodulation”を、チーム研究代表者がゲストエディターとして企画し、約20本の論文を集め、2002年7月に出版予定である。

3. 研究実施体制

(1) システム統合グループ

- ① 銅谷賢治（国際電気通信基礎技術研究所 主任研究員）
- ② 研究項目
 - ・ メタ学習アルゴリズム
 - ・ 目的指向的情報表現の獲得
 - ・ 神経修飾物質系モデル

- ・報酬予測の脳内機構
- ・ロボット集団の学習実験

(2) 学習理論グループ

- ① 松本 隆（早稲田大学理工学部電気電子工学科 大学院理工学研究科 教授）
- ② 研究項目
 - ・メタ学習の理論と応用
 - ・オンラインベイズ学習と数値実験
 - ・強化学習の基本原則と応用

(3) 適応システムグループ

- ① 石井 信（奈良先端科学技術大学院大学大学院情報科学研究科 教授）
- ② 研究項目
 - ・行動系のメタ学習アルゴリズム
 - ・注意レベルと単一ニューロンの応答の変化
 - ・メタ学習アルゴリズム

(4) 情動記憶グループ

- ① 関野祐子（群馬大学医学部行動分析学部門 講師）
- ② 研究項目
 - ・ラット行動実験
 - ・オペラント学習実験システムの完成と学習課題の確立
 - ・脳スライス標本を用いた実験
 - ・前脳アデノシン A1 受容体過剰発現マウスの作製

(5) 学習修飾グループ

- ① 澤口俊之（北海道大学医学研究科 教授）
- ② 研究項目
 - ・前頭連合野の報酬依存的な行動切り替えにおけるドーパミンの役割
 - ・前頭連合野における行動切り替えのニューロン過程

(6) 精神薬理グループ

- ① 山脇成人（広島大学医学部神経精神医学講座 教授）
- ② 研究項目
 - ・将来の報酬予測および意思決定に関する脳領域の同定
 - ・うつ病患者の認知課題遂行中の脳活動

4. 研究成果の発表

(1) 論文発表

- Doya, K., Kimura, H., Kawato, M. (2001). Neural mechanisms of learning and control. IEEE Control Systems Magazine, 21(4), 42-54.

- Doya, K., Kimura, H., Miyamura, A. (2001). Motor control: Neural models and system theory. *International Journal of Applied Mathematics and Computer Science*, 11(1), 77–104.
- Doya, K., Samejima, K., Katagiri, K., Kawato, M. (in press). Multiple model-based reinforcement learning. *Neural Computation*, 14(6).
- Fukumoto, T., Morinobu, S., Okamoto, Y., Kagaya, A., Yamawaki, S. (2001). Chronic lithium treatment increases the expression of brain-derived neurotrophic factor in the rat brain. *Psychopharmacology*, 158, 100–106.
- Ikeda, M., Sagara, M., Sekino, Y., Shirao, T., Honda, K., Yoshioka, T., Allen, C.N., Inoue, S. (2001). The sulphhydryl reagent, N-ethylmaleimide, disrupts sleep and blocks A1 adenosine receptor-mediated inhibition of intracellular calcium signaling in the in vitro ventromedial preoptic nucleus. *Neuroscience*, 106(4), 733–743.
- Katagiri, H., Kagaya, A., Nakae, S., Morinobu, S., Yamawaki, S. (2001). Modulation of serotonin-2A receptor function in rats after repeated treatment with dexamethasone and L-type calcium channel antagonist nimodipine. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 25, 1269–1281.
- Kobayashi S., Shirao, T., Sasaki, T. (2001). Drebrin expression is increased in spinal motoneurons after axotomy. *Neuroscience Letters*, 311, 165–168.
- vMatsumoto, T., Nakajima, Y., Sato, M., Sugi, J., Hamagishi, H. (2001). Reconstructions and predictions of nonlinear dynamical systems: A hierarchical Bayesian approach. *IEEE Transactions on Signal Processing*, 49(9).
- Morimoto, J., Doya, K. (2001). Robust reinforcement learning. V.Tresp et al.(eds.) *Advances in Neural Information Processing Systems 13*. MIT Press, 1061–1067.
- Morimoto, J., Doya, K. (2001). Acquisition of stand-up behavior by a real robot using hierarchical reinforcement learning. *Robotics and Autonomous Systems*, 36, 37–51.
- Nakahara, H., Doya, K., Hikosaka, O. (2001). Parallel cortico-basal ganglia mechanisms for acquisition and execution of visuo-motor sequences – A computational approach. *Journal of Cognitive Neuroscience*, 13(5), 626–647.
- Sawaguchi, T. (2001). Unmasking of silent “task-related” neuronal activity in the monkey prefrontal cortex by a GABA(A) antagonist. *Neuroscience Research*. 39(1), 123–31.
- Sawaguchi, T. (2001). The effects of dopamine and its antagonists on directional delay-period activity of prefrontal neurons in monkeys during an oculomotor delayed-response task. *Neuroscience Research*. 41(2), 115–28.
- Sawaguchi, T., Iba, M. (2001). Prefrontal cortical representation of visuospatial working memory in monkeys examined by local inactivation with muscimol. *Journal of Neurophysiology*. 86(4), 2041–53.
- Tamura, T., Morinobu, S., Okamoto, Y., Kagaya, A., Yamawaki, S. (2002). The effects of

antidepressant drug treatments on activator protein-1 binding activity in the rat brain. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 26, 375-381.

- Yamashita, H., Fujikawa, T., Yanai, I., Morinobu, S., Yamawaki, S. (2001). Clinical features and treatment response of patients with major depression and silent cerebral infarction. *Neuropsychobiology*, 176-182.
- Yamashita, H., Fujikawa, T., Yanai, I., Yamawaki, S. (2002). Cognitive dysfunction in recovered depressive patients with silent cerebral infarction. *Neuropsychobiology*, 45, 12-18.
- Yamawaki, S., Kagaya, A., Okamoto, Y., Takebayashi, M., Saeki, T. (2001). Effects of antidepressants and lithium on intracellular calcium signaling. *Contemporary Neuropsychiatry*, 257-262.
- Yoshimoto, J., Ishii, S., Sato, M. (2001). Application of reinforcement learning based on on-line EM algorithm to balancing of acrobot. *Systems and Computers in Japan*, 32(5), 12-20.
- 伊藤真, 三宅章吾, 沢田康次 (2001). ラットの空間探索課題における海馬-大脳基底核モデル. *電子情報通信学会論文誌*, J84-D-II(9), 2107-2120.
- 射場美智代, 澤口俊之 (2001). 前頭前皮質の選択的注意過程. *脳の科学*, 23(6), 469-478.
- 川人光男, 銅谷賢治, 春野雅彦 (2001). ヒト知性の計算神経科学, 第5回その2:モザイクの拡張とコミュニケーション. *科学*, 71(6), 839-843.
- 川人光男, 銅谷賢治, 春野雅彦 (2001). ヒト知性の計算神経科学, 第6回:ヒト知性に計算理論はどこまで迫られるか. *科学*, 71(18), 1617-1696.
- 鮫島和行, 銅谷賢治 (2001). 強化学習と大脳基底核. *バイオメカニズム学会誌*, 25(4), 167-171.
- 鮫島和行, 銅谷賢治, 川人光男 (2001). 強化学習 MOSAIC : 予測性によるシンボル化と見まね学習. *日本ロボット学会誌*, 19(5), 551-556.
- 鮫島和行, 銅谷賢治, 川人光男 (2001). モジュール競合による運動パターンのシンボル化と見まね学習. *電子情報通信学会論文誌*, J85-D-II(1), 90-100.
- 鮫島和行, 片桐憲一, 銅谷賢治, 川人光男 (2001). 複数の予測モデルを用いた強化学習による非線形制御. *電子情報通信学会論文誌*, J84-D-II (9), 2092-2106.
- 澤口俊之 (2001). 前頭葉と認知. *脳と精神の医学*, 12(1), 23-30.
- 澤口俊之 (2001). 前頭前野の自我の動的オペレーティングシステム. *数理科学*, 39(8), 66-75.
- 澤口俊之 (2001). 「心の謎」の核心へー認知脳科学の成果と将来. *DBCダイヤモンドビジネスコンサルティング*, 140, 4-9.
- 関野祐子, 白尾智明 (2001). 海馬内興奮伝播のゲート機構. *神経研究の進歩*, 45(2), 283.
- 銅谷賢治 (2001). 系列運動の脳内表現と学習のアルゴリズム. 乾敏郎, 安西祐一郎編, 認

知科学の新展開第3巻「運動と言語」, 岩波書店. 29-52.

- 銅谷賢治 (2002). 強化学習とメタ学習の脳内機構—大脳基底核と神経修飾物質系—. 日本神経回路学会誌, 9(1), 36-40.
 - 銅谷賢治, 森本淳, 鮫島和行 (2001). 強化学習と最適制御. システム／情報／制御, 45(4), 30-40.
 - 森本淳, 銅谷賢治 (2001). 階層型強化学習を用いた3リンク2関節ロボットによる起立運動の獲得. 日本ロボット学会誌, 19(5), 574-579.
- (2) 特許出願
国内4件、外国なし