

機械学習により「動物の行動戦略」を解読

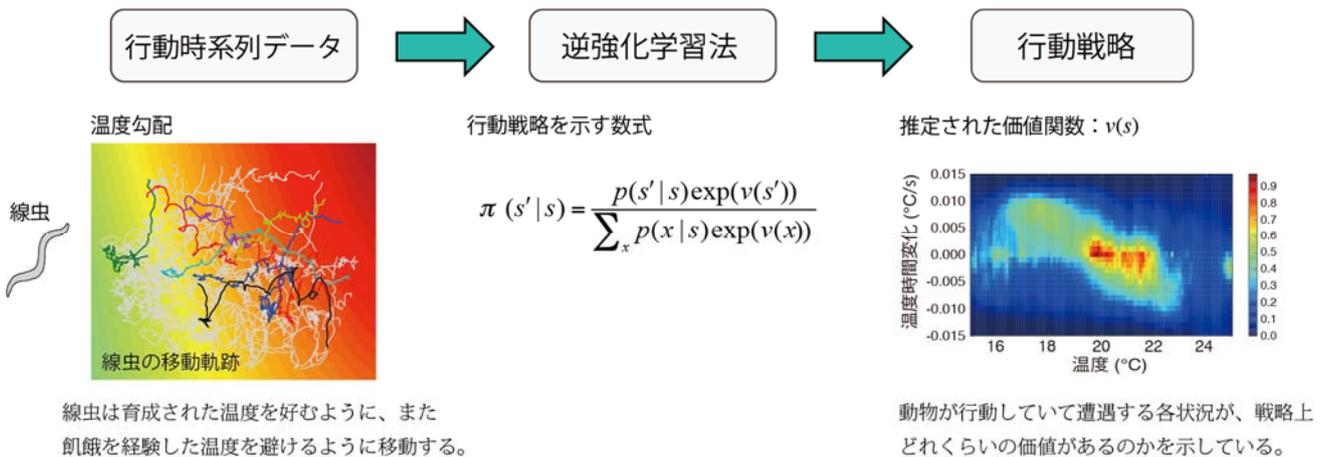
—動物は何を報酬として行動しているのか?—

概要

私たちヒトや動物は、より多くの報酬を得るため、状況に応じた「行動戦略」を持って生きています。報酬には食べ物やお金など実態の伴うものだけでなく、間接的にそれらに結びつくものも含まれているため、自由に行動している動物を単に観察しているだけでは、「動物が何を報酬として意思決定を行い、行動しているのか？」を知ることは困難でした。

そこで京都大学大学院生命科学研究科 本田直樹 准教授、同情報学研究科 山口正一郎 修士課程学生（現：株式会社 Preferred Networks）、石井信 同教授らのグループは、動物の行動データから報酬に基づく行動戦略を明らかにする機械学習法を考案しました。そして、名古屋大学大学院理学研究科の森郁恵教授らと共同で、この手法を線虫の行動へと応用することで、その有効性を示しました。本手法によって、従来の行動が制限された行動実験系から開放され、より自然な状況において自由に振る舞う動物の行動戦略の研究が進むことが期待されます。さらに、本手法は神経活動とその表現形である行動戦略をつなぐ基盤技術を提供するもので、今後、動物の行動戦略を司る神経メカニズムの解明に大きく貢献することが期待されます。

本成果は、2018年5月15日付けで米国の学術誌「PLoS Computational Biology」に掲載されました。



1. 背景

ヒトや動物は生命維持や子孫繁栄のため、外界の状況に応じて、より多くの報酬が期待できる行動を選択していると考えられています。これまでの行動実験を用いた研究では、動物に対して、食べ物などの報酬の伴う課題を課し、動物が取る行動やそれに伴う脳活動が調べられてきました。つまり、課題中において何が動物にとって報酬となるのかは実験者が研究目的に合わせて決めていた訳です。しかしながら、自由に行動している動物の場合、何を報酬として行動しているのかは全く不明でした。また脳内では、報酬はドーパミンと呼ばれる神経伝達物質によって表現されていることから、動物にとって何が報酬となっているのかを明らかにすることは、行動戦略を司る神経メカニズムの理解のためにも重要です。

2. 研究手法・成果

そこで私たちは、動物の行動時系列データから報酬に基づく行動戦略を明らかにする機械学習法（逆強化学習法）を提案しました。この手法は一般的に知られる強化学習の逆問題を解くという意味で、逆強化学習と呼ばれます。強化学習では、どの状況でどれくらい報酬を得られるのかはあらかじめ決められており、試行錯誤によって得られる報酬を最大化する行動戦略を見つけ出すことが目的です。一方で逆強化学習では、動物はすでに最適な行動戦略を獲得しているとして、計測された行動時系列データから未知の報酬を推定することが目的です。

逆強化学習法の応用先として、シンプルなモデル動物である線虫 *C. elegans* の温度走性行動に注目しました。一定の温度で餌を十分に与えて成育した線虫は、その成育温度を記憶し、温度勾配（温度にムラのある空間）下では成育温度を目指して移動し、逆に一定の温度で餌のない飢餓状態で成育した線虫は、温度勾配下で成育温度から逃げて遠ざかることが知られていました。しかし、線虫がどのような戦略にしたがって行動しているのかはこれまで謎でした。線虫を温度勾配においてトラッキングすることで、行動時系列データを取得し、そして逆強化学習法により、線虫にとって何が報酬となっているのかを推定しました。

その結果、餌が十分ある状態で育った線虫は、「絶対温度」および「温度の時間微分」に応じて報酬を感じていることが明らかとなりました。この報酬に基づく戦略は2つの異なるモードから構成されていました。一つは効率的に成育温度に向かうモード、もう一つは同じ温度の等温線に沿って移動するモードを表していました。また飢餓状態で育った線虫は「絶対温度」のみに依存した報酬により、成育温度を避ける戦略を持っていることが明らかとなりました。さらには、推定された報酬を用いて、線虫行動をコンピュータでシミュレーションしたところ、線虫の温度走性行動が再現され、逆強化学習法の妥当性が示されました。

3. 波及効果、今後の予定

本研究は、自由に行動する動物の行動時系列データから行動戦略を解読するという先駆的な挑戦でした。提案した逆強化学習法を用いることで、線虫の温度走性における行動戦略を明らかにすることができました。本手法は行動に伴う状態変化が計測できる場合、他の動物へも適用することが可能です。また、本手法によって、従来の行動が制限された行動実験系から開放され、より自然な状況において自由に振る舞う動物の行動戦略の研究が進むことが期待されます。さらに、本手法は神経活動とその表現形である行動戦略をつなぐ基盤技術を提供するもので、今後、動物の行動戦略を司る神経メカニズムの解明に大きく貢献することが期待されます。

4. 研究プロジェクトについて

本研究は、日本学術振興会（JSPS）科学研究費助成事業「逆強化学習法による「動物の行動戦略を制御する神経基盤」の同定」（若手研究(B) No. 16K16147、代表：本田直樹）および日本医療研究開発機構（AMED）創薬等ライフサイエンス研究支援基盤事業 生命動態システム科学推進事業「多次元定量イメージングに基づく数理モデルを用いた動的生命システムの革新的研究体系の開発・教育拠点」（本田直樹・石井信）の支援のもと行われました。

<論文タイトルと著者>

タイトル：Identification of Animal Behavioral Strategies by Inverse Reinforcement Learning

著者：Shoichiro Yamaguchi, Honda Naoki, Muneki Ikeda, Yuki Tsukada, Shunji Nakano, Ikue Mori, Shin Ishii

掲載誌：PLoS Computational Biology DOI：10.1371/journal.pcbi.1006122

<http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1006122>

<お問い合わせ先>

本田直樹（ほんだ・なおき）

京都大学大学院生命科学研究科 理論生物学分野 准教授

TEL: 075-753-4422

E-mail: honda.naoki.4v[at]kyoto-u.ac.jp