

2026 年 8 月 19 日

報道機関 各位

国立大学法人東北大学
ウィタヤシリメティー科学技術大学院大学（VISTEC）

昆虫の「歩き方」を AI が解読しロボットへ —数歩分のデータで六脚ロボットが適応歩行—

【発表のポイント】

- ナナフシ^{（注1）}のわずか数歩分の歩行データから、歩き方を生み出す「報酬構造」^{（注2）}と脚の協調則を、人の設計なしに抽出しました。
- 平地の歩行データのみを教師としながら、凹凸地形や脚を 1 本失った状況でも、生物に近い柔軟な脚の協調^{（注3）}が現れました。
- 昆虫から抽出した報酬構造は、体格も構造も異なる六脚ロボットへ移植でき、歩行学習を約 3 倍高速化しました。
- 実機の六脚ロボットによる実証に成功し、災害現場など未知の環境にも柔軟に対応可能な多脚ロボットの実現につながります。

【概要】

瓦礫の山やでこぼこの地面は、車輪では越えられません。そこで期待されるのが多脚ロボットですが、脚が増えるほど制御は難しく、歩き方のルールを人が設計する必要がありました。一方、昆虫は約 20 万個の神経細胞（ヒトの脳は約 860 億個）で、地形の乱れや脚の欠損にも即座に対応して歩き続けます。

東北大学大学院工学研究科の王 昱晨（ワン ユチェン）日本学術振興会特別研究員（DC2）、大脇大准教授らの研究グループは、タイ王国ウィタヤシリメティー科学技術大学院大学（VISTEC）の Poramate Manoonpong（ポラメート マヌーンポン）教授らとの国際共同研究により、ナナフシのわずか数歩分の歩行データから、歩行を生み出す報酬構造と脚の協調則を AI が抽出する枠組みを構築しました。抽出した報酬構造は体格の異なる六脚ロボットへ移植でき、実機での歩行も実現しました。

本研究成果は、2026 年 8 月 11 日に英国物理学会出版局（IOP Publishing）の科学誌 Bioinspiration & Biomimetics に掲載されました。

【詳細な説明】

研究の背景

瓦礫が積み重なった災害現場や、月・火星のような未知の地表面などで活動できるロボットとして、多数の脚をもつ「多脚ロボット」への期待が高まっています。脚は、車輪では越えられない段差や隙間を踏み越えられるからです。しかしその代償として、多数の脚と関節をどのタイミングでどう動かすかという「協調」の問題を解かなければなりません。

この問題に対し、これまで大きく三つのアプローチが取られてきました。①足先の通る軌道をあらかじめ設計する手法、②生物の神経回路を模した中枢パターン発生器^(注4)などを用いる手法、③強化学習^(注5)によってロボット自身に歩き方を習得させる手法です。しかしいずれも、歩き方のルール、制御パラメータ、あるいは学習の指標となる報酬関数のいずれかを、研究者が経験に基づいて手作業で設計する必要がありました。設計者の想定を超える状況には対応しづらく、機体や環境が変わるたびに作り直しが求められるという課題が残されていました。

一方、自然界の昆虫はこの問題をはるかに小さな神経系で解いています。昆虫の神経細胞はおよそ 20 万個であり、ヒトの脳の約 860 億個とは比べて非常に小規模です。それでも昆虫は、地形の乱れ、荷重の変化、さらには脚の欠損にまで柔軟に対応し、歩行を続けます。しかも昆虫の歩き方は決まったパターンに固定されているのではなく、速度や地形に応じて連続的に移り変わることが知られています。この「小さな神経系が生み出す高い適応性」の裏にある原理を計算モデルとして取り出せれば、ロボット制御に大きな指針が得られるはずです。

ここで発想を逆転させたのが本研究です。報酬関数を人が設計するのではなく、すでに上手に歩いている昆虫の運動データから、「昆虫は何を良しとして歩いているのか」という報酬構造そのものを推定する、逆強化学習^(注6)と呼ばれる考え方を取り入れました。逆強化学習は動物の移動や採餌といった大まかな行動の解析には用いられてきましたが、18 の関節が関わる多脚歩行のように高次元で複雑に絡み合った運動に適用し、さらに実機ロボットの制御まで到達した例はありませんでした。

今回の取り組み

研究グループは、公開データセットに収められたナナフシ (*Medauroidea extradentata*) の平地直進歩行のうち、わずか一区間 (18 関節・約 3~4 歩行周期分) を教師データとして用いました (図 1)。この極めて限られた生物データを出発点に、敵対的逆強化学習^(注7)という枠組みを適用します。この手法では、昆虫の動きと学習中のエージェントの動きを見分けようとする「識別器」と、見分けられないほど昆虫らしく動こうとする「制御方策」^(注8)とを競わせなが

ら学習を進めます。この競争の過程で、識別器の内部に歩行の良さを表す報酬構造が浮かび上がってくる仕組みです。

本研究の工夫は、この識別器を二つの部品（ニューラルネットワーク）に分けて設計した点にあります。ひとつは身体の違いに左右されない「歩行としての望ましさ」を表す報酬ネットワーク、もうひとつはその機体固有の事情を吸収する部品（ネットワーク）です。この切り分けによって、前者だけを別の身体へ持ち出すことが可能になりました。

得られた制御方策は、まず平地でナナフシに似た安定した歩行を再現しました。六脚のタイミング関係を解析すると、四脚支持と波状歩容の中間にあたる、中速歩行に適した柔軟な歩容^(注 9)が自然に現れていました。後脚の接地時間が長いという特徴も、推進と安定を後脚が担うという昆虫の実データと一致していました。さらに一本の脚の中の三つの関節^(注 10)の協調を解析すると、体幹に近い関節が安定したリズムを刻み、先端の関節が接地のタイミングを細かく調整する「近位は安定、遠位は柔軟」という制御構造が明瞭に現れました。これは、昆虫が脚先の感覚フィードバックと局所的な修正によって安定を保つという生物学的知見と符合するものです。

次に、この枠組みが教師データの範囲を超えて働くかを検証しました。教師データは健全な個体の平地歩行のみであるにもかかわらず、凹凸のある地形では、脚のタイミング関係が前脚から後脚へ順に波が伝わる形へと自律的に組み替わり、不整地で生物がよく用いる歩き方に近づきました。また右中脚を無効化した脚欠損条件では、残る脚が対角方向に支持の鎖を組み直し、前脚と後脚の接地時間の割合^(注 11)を増やして、失われた支持を補うという代償動作が現れました。いずれの条件でも、脚の中の関節協調の基本構造は保たれたままでした。限られた生物データから抽出された制御原理が、想定外の状況にも耐える骨格をもつことを示す結果です。

最大の成果は、抽出した報酬構造の「乗り換え可能性」です。研究グループは、ナナフシを模したモデルで学習した制御方策を、体の比率・関節配置・モータ特性が大きく異なる六脚ロボット「RedMirror」（全長約 90cm、幅約 54cm、重量約 4.5kg、18 個のサーボモータで駆動）へ直接移植することを試みました。予想どおり、制御方策をそのまま持ち込んでも安定した歩行は成立しませんでした。ところが、報酬ネットワークのみを移植し、それを指標としてロボット側で改めて学習させたところ、協調した方向性のある歩行が速やかに獲得されました。学習に要したステップ数は約 7 万（約 1 時間 8 分）で、従来型の単純な報酬のみを用いた場合の約 20 万ステップ（約 4 時間 7 分）と比べて約 3 倍高速でした。しかも後者は、前進はするものの脚の協調も前後の区別も獲得できませんでした（図 2）。

最後に、シミュレーションで学習した制御方策を実機の RedMirror ロボットへ搭載しました。シミュレーションと実機ではセンサ信号の分布が異なるため、

慣性計測装置^(注 12)の信号の平均と散らばりを合わせる簡便な補正^(注 13)を施したうえで制御方策に入力しました。その結果、実機ロボットは協調した歩行を実現し、生物の歩行データから学びとった制御則が現実の機体上で動作することを実証しました(図 3)。生物の行動データからロボット実機までを一本の流れとしてつないだ枠組みです。

今後の展開

本研究の枠組みは、歩き方のルールや報酬関数を人が設計することなく、生物の運動データから直接、移植可能な制御則を取り出せることを示しました。これは、多脚ロボットの制御開発の考え方を「設計する」から「生物から学びとる」へと転換させる可能性をもっています。抽出された報酬構造は環境や機体を越えて再利用できるため、新しいロボットを開発するたびに制御をゼロから作り直す必要がなくなることが期待されます。また、識別器の内部に解釈可能な報酬ネットワークが得られるため、「昆虫はなぜそのように歩くのか」という生物学の問いに定量的に迫る道具にもなります。

一方、課題も残されています。現在の制御方策は現時刻の観測のみに基づいて動作するため、経験の蓄積や長期的な戦略の切り替えといった、生物が備える記憶の働きを明示的には扱えていません。今後は、記憶機構の導入や、中枢パターン発生器・反射といった生物的な低次制御と組み合わせた階層的な制御へと発展させることで、生物の運動学習により近い枠組みを目指します。将来的には、災害現場の探査、インフラ点検、惑星探査など、事前の想定が難しい環境で活動する多脚ロボットの実用化に貢献することが期待されます。

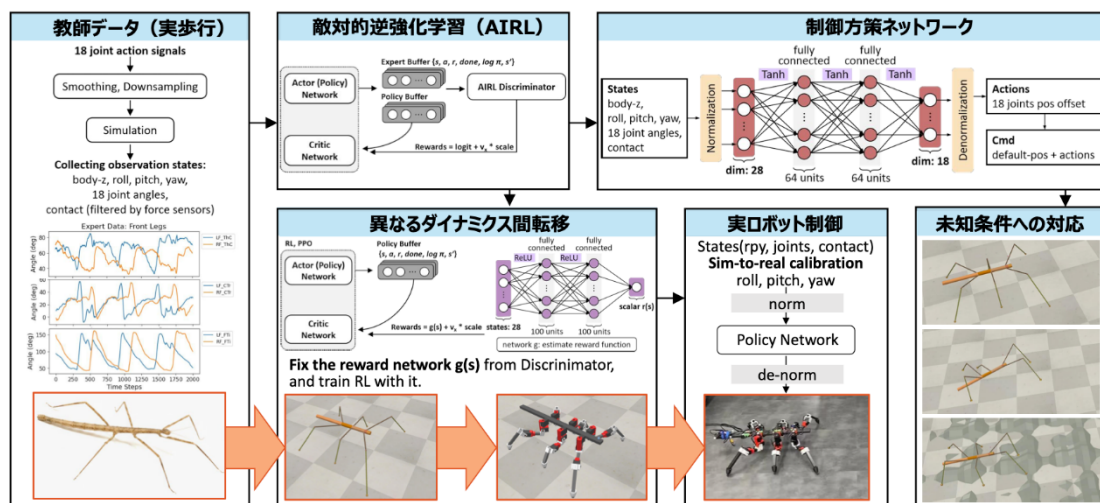


図 1. ナナフシの歩行データから移植可能な歩行制御則を学習する枠組み

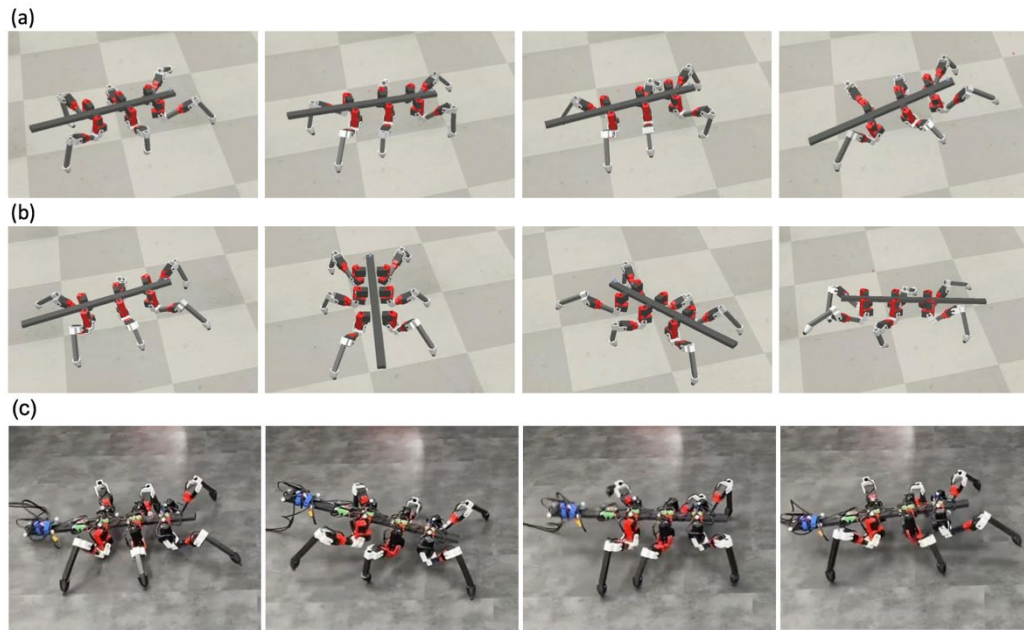


図 2. 報酬構造の移植による歩行学習。報酬構造を移植した場合（a, 上段）は協調した歩行が獲得されたが、単純な報酬のみの場合（b, 中段）は脚の協調を獲得できなかった。（c, 下段）実機六脚ロボット「RedMirror」への搭載と歩行の様子

【関連動画】

From Insect Behavior to Transferable Robot Locomotion

<https://youtu.be/KUCuNRkgSRE>

【データ公開】

本研究を支えるデータおよびプログラムは、以下で公開されています。

<https://github.com/wangyuchen0217/airl-insect-walking>

【謝辞】

本研究は、JSPS 科研費 JP24KJ0341（特別研究員奨励費、研究代表者：王昱晨）、JSPS 科研費 JP23H00481（基盤研究（A）、研究代表者：大脇大）、JP26H02510（基盤研究（A）、研究代表者：大脇大）、JP26K00082（学術変革領域研究（B）計画研究、研究代表者：大脇大）の助成を受けたものです。また、タイ王国高等教育科学研究イノベーション省の Reinventing University プログラムによる支援を一部受けています。

【用語説明】

- 注1. ナナフシ：木の枝に擬態した細長い体をもつ昆虫。本研究では学名 *Medauroidea extradentata* の歩行データを用いた。柔軟で高度に協調した六脚歩行を示すため、歩行研究のモデル生物として広く用いられている。
- 注2. 報酬構造（報酬関数）：学習において「何を良しとするか」を数値で表したもの。強化学習では通常これを人が設計するが、本研究では昆虫の歩行データから逆向きに推定した。
- 注3. 脚の協調（脚間協調・脚内協調）：複数の脚どうしのタイミング関係を脚間協調、一本の脚に含まれる複数の関節どうしの連係を脚内協調と呼ぶ。両者の組み合わせが、昆虫の適応的な歩行を生み出している。
- 注4. 中枢パターン発生器（Central Pattern Generator：CPG）：外部からの周期的な指令なしに、リズムカルな運動信号を自律的に生成する神経回路。歩行・遊泳・呼吸などのリズム運動の基盤とされ、ロボット制御にも広く応用されている。
- 注5. 強化学習：試行錯誤を通じて、報酬が最大となる行動の仕方を学習する機械学習の枠組み。本研究では近位方策最適化（Proximal Policy Optimization: PPO）と呼ばれる代表的な手法を用いている。
- 注6. 逆強化学習（Inverse Reinforcement Learning：IRL）：観察された行動データから、その行動の背後にある報酬構造を逆向きに推定する手法。「何を目指して動いているのか」をデータから読み解く枠組み。
- 注7. 敵対的逆強化学習（Adversarial Inverse Reinforcement Learning：AIRL）：お手本の動きと学習中の動きを見分けようとする「識別器」と、見分けられないよう振る舞おうとする制御方策を競わせることで、お手本の背後にある報酬構造を推定する手法。本研究では識別器を、身体に依存しない部分と依存する部分に分けることで、報酬構造の移植を可能にした。
- 注8. 制御方策（ポリシー）：現在の状態を入力として、次を取るべき動作を出力する規則。本研究では、体の姿勢・18 関節の角度・各脚の接地情報を入力とし、18 関節への指令を出力するニューラルネットワークとして表現されている。
- 注9. 歩容（ガイト）：歩行時における各脚の接地・遊脚の時間的パターン。三脚支持、四脚支持、波状歩容などがあり、昆虫では速度や地形に応じて連続的に変化する。
- 注10. 脚の三つの関節：昆虫の脚を構成する主要な関節。体幹側から順に、胸-基節関節が脚の前後の振り出し、基節-転節関節が上下の持ち上げ、腿節-脛節関節が脚の伸び縮みをそれぞれ担う。
- 注11. 接地時間の割合（デューティファクタ）：一歩行周期のうち、脚が地面に接している時間の割合。値が大きいほど支持に費やす時間が長く、荷重の分担を評価する指標として用いられる。

注12. 慣性計測装置 (IMU) : 加速度と角速度を計測し、機体の傾きや向きを推定する小型センサ。スマートフォンやドローンにも搭載されている。

注13. シミュレーションから実機への補正 (Sim-to-Real) : 計算機上のシミュレーションで学習した制御則を、現実のロボットへ移す技術。両者のセンサ特性や物理特性の差が性能低下を招くため、その差を埋める補正が重要となる。

【論文情報】

タイトル : From insect behavior to transferable robot locomotion: inferring embodied locomotor principles from limited data via adversarial inverse reinforcement learning

(和訳 : 昆虫の行動から移植可能なロボット歩行へ — 敵対的逆強化学習による限られたデータからの身体性歩行原理の推定 —)

著者 : Yuchen Wang, Chuthong Thirawat, Mitsuhiro Hayashibe, Poramate Manoonpong, Dai Owaki*

*責任著者 : 東北大学大学院工学研究科ロボティクス専攻 准教授 大脇大

掲載誌 : Bioinspiration & Biomimetics (IOP Publishing)

DOI : 10.1088/1748-3190/ae901f

URL : <https://iopscience.iop.org/article/10.1088/1748-3190/ae901f>

※ 実験動物に関する倫理的配慮について

本研究で用いた昆虫の歩行データは、すでに公開されているオープンデータセットを利用したものであり、本研究のために新たな動物実験は行っていません。

【問い合わせ先】

(研究に関すること)

東北大学大学院工学研究科

准教授 大脇大 (おおわき だい)

TEL: 022-795-4064

Email: owaki@tohoku.ac.jp

(報道に関すること)

東北大学工学研究科・工学部 情報広報室

担当 沼澤 みどり (ぬまざわ みどり)

TEL: 022-795-5898

Email: eng-pr@grp.tohoku.ac.jp