

なぞり課題を用いた3者の協調運動における ボトムアップとトップダウン処理に関する検討

Study on Bottom-up and Top-down Processing in Group Behavior of a Triad Using a Coordinated Drawing Task

市川 淳[†], 筒井 和詩[‡], 藤井 慶輔[‡]

Jun Ichikawa, Kazushi Tsutsui, and Keisuke Fujii

[†] 静岡大学, [‡] 名古屋大学

Shizuoka University, Nagoya University

j-ichikawa@inf.shizuoka.ac.jp

概要

本研究では3者で不均一な役割を分担して集団目標の達成を図る実験タスクを用いて、協調運動で重要な他者を助けて状況の改善を図る役割に着目し、その情報処理についてシミュレーションから検討した。結果、ボトムアップとトップダウン処理に対応する深層強化学習とルールベースから当該役割が行動するモデルでは、ルールベースのみやランダムで行動する場合よりもタスクパフォーマンスが高いことが示された。2種類の処理により、役割が十分に機能する可能性がある。**キーワード**: 協調運動 (coordinated group behavior), 役割 (role), 深層強化学習 (deep reinforcement learning), ルールベース (rule-based modeling)

1. 導入

集団スポーツや網漁業に代表されるように、人は他者と取り組むことで単独では難しいタスクを遂行する、単独よりも高いタスクパフォーマンスを達成できる [1]。その協調メカニズムは、研究分野を横断して幅広く議論されている。本研究では、集団目標を共有して3者で不均一な役割を分担する協調運動に着目し、重要な役割の基盤にある情報処理について検討する。

神経科学や運動科学では主に、集団目標を共有しない協調が検討されている。例えば、ダイナミカルシステムアプローチにおいて、他者とのインタラクションを通して感覚からボトムアップに運動を調整する低次元な処理が議論されている [2]。一方で、認知科学では集団目標を共有した協調を対象に、より高次元な処理として知識や思考に焦点をあて役割の分担・交替を通して他者視点を取り入れることで多様な解の発見や問題解決が導かれることが示されている [3, 4]。

上記を踏まえると、協調の基盤にはボトムアップと

トップダウン処理の両側面があると言える。両者を議論することで、理論の統合や発展が期待される [5]。しかし、研究分野の違いもあり、2種類の処理から協調を検討した研究は少ない。近年、非線形時系列解析を用いて運動データから会話の状態を統計的に推定する手法が提案されている [6]。ただし、これは現象の推定に重きが置かれている点に留意する必要がある、人の情報処理モデル自体を説明する上で議論の余地がある。他方で、集団スポーツは、ボトムアップとトップダウン処理に基づく他者とのインタラクションがみられる典型である。集団目標やチームの方針、役割等からトップダウンに対応するだけでなく、視覚等の感覚からボトムアップに柔軟な対処を示すこともある [7]。しかし、集団スポーツの多くが3者以上で構成され、身体運動を介してインタラクションが連続的に展開する複雑で動的な性質があるため [8, 9]、2種類の処理とタスクパフォーマンスの関連が検討しづらい問題がある。

そこで、本研究のアプローチとして、3人1組で不均一な役割分担が求められる協調運動の実験タスクを用いる。なぞり課題 [10] では、各自がリールを操作して糸の張力を変え、3本の糸に接続された1本のペンを動かし長さ 30 cm、幅 2 cm の正三角形の辺をなぞる (図 1)。手元へ引き寄せてペンを動かす「張る」役割、そのペンがスムーズに動くように対応する「緩める」役割、そして2つの役割に介入し、ペンの移動の過度なタイムロスが生じない程度に軌道を修正する「適度に張る」役割への分担が求められる。なぞる辺が替わると役割を反時計回りに交替する (表 1)。繰り返しなぞり課題を行わせた行動実験では、回帰モデルから3つ目の役割が小さな逸脱を維持しつつ、素早くペンを移動させる高いタスクパフォーマンスに関連する可能性が示された [11]。仕様上、「張る」役割と

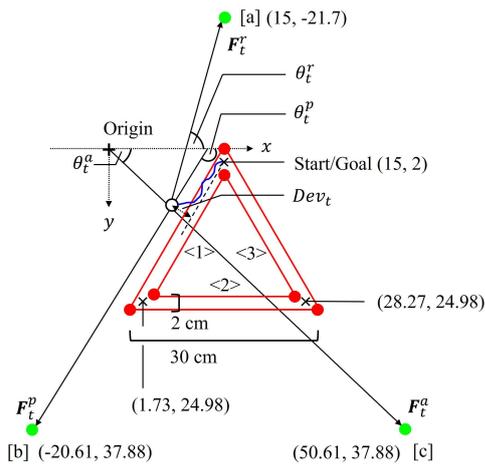


図1 なぞり課題の模式図。括弧は2次元座標である。

表1 操作者の役割分担

辺	操作者		
	[a]	[b]	[c]
<1>	緩める	張る	適度に張る
<2>	適度に張る	緩める	張る
<3>	張る	適度に張る	緩める

「緩める」役割だけではペンの外側へ逸脱するため(図1)、「適度に張る」役割が適宜介入し、助けるかたちで状況の改善を図るレジリエントな操作を行っていると考えられる。レジリエンスはなぞり課題に限らず、集団スポーツや情報システムの運営、軍隊などでその重要性が議論されている一方で[1, 12]、基盤にある情報処理の詳細は十分に議論されていない。

以上を踏まえ、更なるアプローチとして「適度に張る」役割に着目して構成論的な理解に向けてシミュレーションを導入する。状況に応じた柔軟な対処が求められること[11]、協調運動に2種類の情報処理が機能すること[7]を考慮して当該役割にボトムアップとトップダウン処理が働くと仮定する。本研究では、なぞり課題のシミュレーションから各処理を強化学習とルールベースでモデル化し、タスクパフォーマンスとの関連を検討することを目的とする。

2. 方法

2.1 3つの役割におけるトップダウン処理

なぞり課題では糸を張りペンを動かすため、各役割を運動方程式で定式化できる。シミュレーションに向けて3つのエージェントを用意し、現在の時間ステップにおける張力の合成 F_t (N) を式(1)で表す。

$$F_t = F_t^p + F_t^r + F_t^a - f_t' \quad (1)$$

F_t^p , F_t^r , F_t^a はそれぞれ、糸を「張る」役割、「緩める」役割、「適度に張る」役割の張力を表し、 x 成分と y 成分で構成される。3つの役割の張力を合わせた後、2次元に分解された摩擦力を示す f_t' を引く¹。

以降は、マルチエージェントシミュレーションにおける Keep It Simple Stupid (KISS) の原則[13]に従い、可能な限り単純な構成で役割をモデル化する。本研究では、「張る」役割と「緩める」役割は全てルールベースで記述する。これは、「適度に張る」役割に着目して2種類の情報処理とタスクパフォーマンスの関連を検討する点が目的であること、さらには仕様上「適度に張る」役割と異なり状況に応じた対処が必須ではないことに基づいている。

「張る」役割は、行動実験[11]で記録された張力の代表値を使用する。参加グループを通した平均ピーク値の一定張力でペンを引き ($|F_t^p| = 0.57$ N), x 成分と y 成分に分解される ($F_t^p = (|F_t^p| \cos \theta_t^p, |F_t^p| \sin \theta_t^p)$).

また、「緩める」役割では糸を張る必要はないが操作自体は行っているため、わずかな張力が発生するノイズとみなす。ただし、ペンが辺から逸脱して操作者にとって問題がある場合、張力は発生させない。「緩める」役割の張力を式(2)として表す。

$$|F_t^r| = \begin{cases} 0 & \text{if } |Dev_t| > Th^r, \\ noi_t^r & \text{otherwise,} \end{cases} \quad (2)$$

Dev_t (cm) は、現在の時間ステップにおけるペンの逸脱量(ペンとなぞる辺の幅の中央線との距離)を示し(図1)、外側に逸れると正の値、内側に逸れると負の値をとる。 $|Dev_t|$ は逸脱量の絶対値である。 Th^r は、操作を切り替える閾値で辺の幅の半分を示す1 cmである。 noi_t^r はノイズレベルの張力を表す²。 $|F_t^r|$ は x 成分と y 成分に分解される ($F_t^r = (|F_t^r| \cos \theta_t^r, |F_t^r| \sin \theta_t^r)$).

そして、本研究で焦点を当てる「適度に張る」役割の張力を式(3)に示す。

$$|F_t^a| = \begin{cases} 0 & \text{if } Dev_t \leq Th^a, \\ A_t^a & \text{otherwise,} \end{cases} \quad (3)$$

Th^a は、操作を切り替える閾値で0.4 cmとする。 Dev_t が閾値以内、つまり内側にペンが逸脱する、あるいは外側への逸脱が一定範囲内に収まっている時は張力を発生させない。一方で、外側に逸脱して閾値を超えると、操作者にとって問題があるため張力を発生させる。

¹摩擦力は動摩擦係数とペンの質量、重力加速度の積で、それぞれ $\mu' = 0.1$, $m = 0.085$ kg, $g = 9.8$ m/s² である。

²各時間ステップで標準正規分布の乱数を発生させ、行動実験の全試行を通した最小ピーク値のグループ間の変動(標準偏差: 0.03 N)との積を求め、ノイズ量を1/10に設定した値である。

本研究では、ペンを引く程度 A_t^a を決めるにあたり、ボトムアップ処理を導入する（詳細は次節を参照）。その後、 $|F_t^a|$ は x 成分と y 成分に分解される ($F_t^a = (|F_t^a| \cos \theta_t^a, |F_t^a| \sin \theta_t^a)$)。

2.2 「適度に張る」役割のボトムアップ処理

深層学習から Q 関数を近似する Deep Q-Network (DQN)[14] を用いる。入力にはペンの逸脱量、なぞる辺のスタートとペンの相対位置、ゴールとペンの相対位置、及び「適度に張る」役割のリールとペンの相対位置に関する 4 種類の連続値とし、出力は 2 種類の離散値とする³。2 種類とはそれぞれ、行動実験 [11] における平均ピーク値 (0.59 N) とノイズレベルの張力である⁴。エージェントは状況に応じて視覚情報から一定張力（平均ピーク値）でペンを引き、軌道修正する対処を強化学習する。

なお、本研究の DQN は、Double Deep Q-Network (DDQN)[15]、Prioritized Experience Reply [16]、Dueling DQN [17] から構成される。

2.3 シミュレーションの学習及び評価

各役割の張力から F_t が求まると、式 (4) よりペンの加速度 a_t が算出される。

$$a_t = F_t/m. \quad (4)$$

a_t を二階積分して次のペンの位置を求める。時間が経ち、ゴールとの距離が閾値 Th^g 以内であれば、なぞり終えたとみなし役割を反時計回りに交替させて次の辺をなぞる。辺 <1> と辺 <2> をなぞる場合は $Th^g = 1$ cm、辺 <3> をなぞる時は $Th^g = 2$ cm とする。

報酬は「適度に張る」役割のエージェントに対して、 Dev_t が 0.4 cm よりも大きくかつ、ノイズレベルの張力を行動選択した時に与えられる。ペスが外側に逸脱しつつあるにも関わらず、軌道修正する行動を選択しないため、逸脱量に応じた負の報酬を $r_{t+1} = -Dev_t/50$ として受け取る。他方で、ゴールに到達すると辺 <1> で +1、辺 <2> は +2、辺 <3> で +3 の正の報酬が与えられる。

学習は 10^4 エピソード（試行）で、計 2000 ステップごとに DDQN のメインネットワークとターゲット

³実装は 8 次元ベクトルにするために、0 が格納されたダミー変数を用意した。

⁴後者は、標準正規分布の乱数と平均ピーク値のグループ間の変動（標準偏差：0.14 N）との積から、ノイズ量を 1/10 に設定した値である。

ネットワークの同期を行った。1 ステップは 0.05 秒である。ネットワークは全結合の 2 層、状態価値関数、アドバンテージ関数で構成され、ユニット数は 64, 64, 32, 32 とした。バッチサイズは 32、学習率は 10^{-6} 、リプレイバッファサイズは 10^4 、割引率 γ は 0.9、探索は ϵ クリーディ法で、探索率 ϵ は計 50,000 ステップかけて 1.0 から 0.1 まで線形減少させた。学習パラメータは、協調運動に関連する研究 [18] を参考に本研究に合わせて改変した。学習は、先行研究 [15] のように複数の乱数 seed を使用し、1 種類の乱数 seed で 10^4 エピソードの学習を行わせ、15 種類で実施した。行動実験 [11] と同様にどのエピソードでも辺 <1> からなぞり始め、1 辺をなぞるために必要なステップ数を 100 とした。なぞれなかった時点でそのエピソードは終了し、次のエピソードに進めた。

さらに、学習済みモデルから評価を行った。学習時とは異なる 15 種類の乱数 seed を使用し、1 種類の乱数 seed に対して 1 つの学習済みモデルを利用し 10 エピソード分のシミュレーションを実施した。 ϵ は 0 とする。そして、3 辺全てをなぞることができたエピソードを使い、タスクパフォーマンスとして各辺でペンの平均逸脱量を算出した。また、比較対象としてルールベースのみとランダムを用意した。前者は、強化学習のボトムアップ処理をルールベースに置き換えた。ペンの逸脱量が外側に 0.4 cm よりも大きく 0.8 cm 以内の場合はノイズレベルの張力、0.8 cm を超えると平均ピーク値の張力を発生させた。後者は、2 種類の行動選択をランダムに置き換えた。

3. 結果と考察

ボトムアップ処理による行動の選択が 2 種類のため、約 70 エピソードの早い段階で報酬が飛躍的に向上する推移が確認された。学習前後におけるペンの移動軌跡をみると、学習後はどの辺であってもペンを引いて適切に軌道修正する特徴がみられた。各条件のタスクパフォーマンスを図 2 に示す。横軸は辺、縦軸はペンの逸脱量 (cm) で、箱ひげ図は乱数 seed 間の変動である。ペンの逸脱量について、手法要因（ランダム、強化学習とルールベース、ルールベース）と辺要因 (<1>, <2>, <3>) による 2 要因の分散分析を実施した。注目すべき点として、全ての辺で強化学習とルールベースの水準で逸脱量が他の水準に比べて有意に小さいことが確認された ($ps < .05$)。

本研究では、集団目標を共有して不均一な役割を分担する 3 人 1 組のなぞり課題から、協調運動で重要な他者を助けて状況の改善を図る役割をモデル化した。

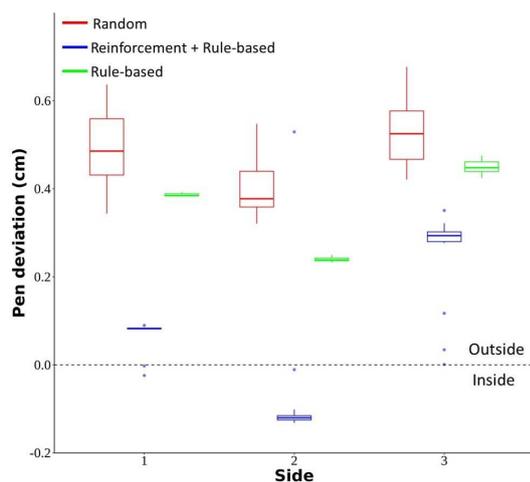


図 2 評価としてのタスクパフォーマンス。

結果、ボトムアップとトップダウン処理に対応する深層強化学習とルールベースから「適度に張る」役割が行動選択することで、より高いタスクパフォーマンスが達成されることが示された。2種類の情報処理が基盤にあることで、当該役割が十分に機能する可能性がある。ペンを軌道修正する適切なタイミングをボトムアップ処理から発見したと考えられ、複雑で動的な協調運動のメカニズムを理解する上で基礎的な知見を提供した。

ただし、KISSの原則[13]に従い、単純な構成で各役割をモデル化したため、行動実験[11]のペンの移動軌跡を再現するモデルとしては検討事項がある。さらに、深層強化学習から動物の協調的狩りを検討した研究[18]では、協調運動に高次の情報処理は必要ないことについて議論している。一方で、集団スポーツの協調運動では2種類の処理が基盤にあることが主張されている[7]。より緻密な動きが要求されるタスクにはトップダウン処理も求められると考えられ、どちらの処理を優先させるかに関係するかもしれない。

謝辞

本研究は、JSPS 科研費 21K18033 の助成を受けた。

文献

- [1] Fujii, K., Yokoyama, K., Koyama, T., Rikukawa, A., Yamada, H., & Yamamoto, Y. (2016). Resilient help to switch and overlap hierarchical subsystems in a small human group. *Scientific Report*, Vol. 6, 23911. <https://doi.org/10.1038/srep23911>
- [2] Kelso, S. J. A. (2021). Unifying large- and small-scale theories of coordination. *Entropy*, Vol. 23, No. 5, 537. <https://doi.org/10.3390/e23050537>
- [3] 林 勇吾・三輪 和久・森田 純哉 (2007). 異なる視点に基づく協同問題解決に関する実験的検討. *認知科学*, Vol. 14, No. 4, pp. 604–619.
- [4] Shirouzu, H., Miyake, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. *Cognitive Science*, Vol. 26, No. 4, pp. 469–501.
- [5] Gorman, J. C. (2014). Team coordination and dynamics: Two central issues. *Current Directions in Psychological Science*, Vol. 23, No. 5, pp. 355–360.
- [6] Dale, R. (2015). An integrative research strategy for exploring synergies in natural language performance. *Ecological Psychology*, Vol. 27, No. 3, pp. 190–201.
- [7] Steiner, S., Macquet, A. C., & Seiler, R. (2017). An integrative perspective on interpersonal coordination in interactive team sports. *Frontiers in Psychology*, Vol. 8, 1440. <https://doi.org/10.3389/fpsyg.2017.01440>
- [8] Yokoyama, K., & Yamamoto, Y. (2011). Three people can synchronize as coupled oscillators during sports activities. *PLoS Computational Biology*, Vol. 7, No. 10, e1002181. <https://doi.org/10.1371/journal.pcbi.1002181>
- [9] Braun, D. A., Ortega, P. A., & Wolpert, D. M. (2009). Nash equilibria in multi-agent motor interactions. *PLoS Computational Biology*, Vol. 5, No. 8, e1000468. <https://doi.org/10.1371/journal.pcbi.1000468>
- [10] 丸野 俊一 (1991). 社会的相互交渉による手続き的知識の改善と“自己-他者”視点の分化・獲得. *発達心理学研究*, Vol. 1, No. 2, pp. 116–127.
- [11] Ichikawa, J., & Fujii, K. (2022). Analysis of group behavior based on sharing heterogeneous roles in a triad using a coordinated drawing task. *Frontiers in Psychology*, Vol. 13, 890205. <https://doi.org/10.3389/fpsyg.2022.890205>
- [12] Bowers, C., Kreutzer, C., Cannon-Bowers, J., & Lamb, J. (2017). Team resilience as a second-order emergent state: A theoretical model and research directions. *Frontiers in Psychology*, Vol. 8, 1360. <https://doi.org/10.3389/fpsyg.2017.01360>
- [13] Axelrod, R. (1997). *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. New Jersey: Princeton University Press.
- [14] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu A. A., Veness, J., Bellemare, M., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, vol. 518, No. 7540, pp. 529–533.
- [15] van Hasselt, H., Guez, A., & Silver, D. (2016). Deep reinforcement learning with double q-learning. *Proceedings of the 30th AAAI conference on artificial intelligence*, <https://doi.org/10.1609/aaai.v30i1.10295>
- [16] Schaul, T., Quan, J., Antonoglou, I., & Silver, D. (2015). Prioritized experience replay (2016). *arXiv preprint*, <https://doi.org/10.48550/arXiv.1511.05952>
- [17] Wang, Z., Schaul, T., Hessel, M., Hasselt, H., Lanctot, M., & Freitas, N. (2016). Dueling network architectures for deep reinforcement learning. *Proceedings of the 33rd International Conference on Machine Learning*, pp. 1995–2003.
- [18] Tsutsui, K., Tanaka, R., Takeda, K., & Fujii, K. (2022). Collaborative hunting in artificial agents with deep reinforcement learning (2022). *bioRxiv*, <https://doi.org/10.1101/2022.10.10.511517>