

# 身体運動のフラクタル次元に基づく行為模倣へのアプローチ

Simulation approach to imitation based on fractal dimensions of bodily movements

鳥居拓馬 \*<sup>1</sup>      日高昇平 \*<sup>1</sup>

Takuma Torii      Shohei Hidaka

\*<sup>1</sup>北陸先端科学技術大学院大学

Japan Advanced Institute of Science and Technology

For learning actions from others, one needs to infer the goals behind observed movements. Recent psychological studies suggested that children can infer the goals of adults, without observing their intended effect of actions. Inspired by the psychological studies, we seek for a mechanistic account for goal inference by observation of unfulfilled actions. We conducted simulations to examine how discrimination of movements by their hidden intentions is possible. In the simulations, two agents control pendulums by their motor controls that greatly differ but produce very similar movements. Our analysis shows the seemingly similar movements can be well discriminated correctly by using, as a feature for classifiers, pointwise dimensions of the movements. Pointwise dimension is a dynamical invariant estimated with less knowledge of the movement-generating system. Our finding thus suggests that dynamical invariants, such as pointwise dimension, can be a foundation for inference on hidden intentions and goals from movements.

## 1. はじめに

人間の子供は一見よく似た運動の背後にある意図の違いを識別できる [Warneken 06]。本稿では、ある目的に向けた運動を行為といい、目的に向けた運動の計画を意図という。[Warneken 06] の心理実験では、子供は、大人の行為の目的を知らされていないにも関わらず、意図通りに運動できていない大人を見ると、その大人が遂行しようとする意図した行為を自らの行為として遂行できる。こうした意図推論は、人間が他者から行為を学習する際に必要な能力と考えられる。

これに対して、行為の社会学習に対する既存の工学的アプローチでは、学習主体は教師の運動の目的に関する知識（目的関数）を所与とすることが多い [Breazeal 02]。学習主体は何らかの目的関数に照らして運動制御（意図）を学習する。教師の目的関数を所与とする場合、制御器の評価は明示的なので、教師の運動は学習を底上げするために使われる [Schaal 97]。目的関数を所与としない場合、制御器の評価自体が学習主体の判断に委ねられる。加えて教師の運動が意図通りでなく目的状態を明示的に含まない場合（[Warneken 06] の心理実験で扱われる）、既存のアプローチでは扱えていない。この場合に向けた第一歩は、観察された複数の運動が同じ制御に従うか、それとも異なる制御に従うかを識別する能力だと考えられる。

そこで本稿では、目的や意図を所与とせずに、いかにして一見よく似た運動の背後にある意図（制御）の違いを識別できるかを問う。本稿では、身体運動系を力学系とみなす立場から、意図（制御）の相違がその意図（制御）の下で為される運動の力学的不変量の相違として観察されるという仮説を検討する。この仮説を検討するため、本稿では [Warneken 06] の心理実験に着想をえたシミュレーションを行う。

## 2. シミュレーション

### 2.1 実験設計

[Warneken 06] の心理実験では、2種類の異なる意図の下で為された2つの類似した大人の運動に対して、子供（18ヶ月児の被験者）が意図の違いを識別できるかが検証された。具体的には、ある目的を達成しようとする大人の運動に対する子供の反応が観察された。子供は大人の行為の目的や意図を聞かされていない。一方の条件では、大人は意図した運動を行い、目的を達成する（成功条件）。他方の条件では、大人は意図した運動を行えず、目的を達成できない（失敗条件）。これら2つの条件は統制されており、目的や意図は異なるが、運動は類似している。そのため、もし子供が失敗条件においてのみ大人を手伝う行動を示せば、子供が意図の違いを識別できたと考えられる。[Warneken 06] らは、子供が意図の違いを識別し、子供が失敗条件においてのみ大人を手伝う行動を示すことを示した。

[Warneken 06] の心理実験は、異なる意図の下で一見よく似た運動が観察されることを利用している。本稿では、運動生成の計画にあたる制御を「意図」と捉える。そして、この心理実験の論理を借用し、一見よく似た運動を生成する異なる制御の違いを識別できるかを数値的に検証する。具体的には、ある目的を異なる制御で達成しようとする2体のエージェントを用意する。一方のエージェントは、その制御の学習時と同じ環境で運動を行い、目的を達成する。他方のエージェントは、その制御の学習時と異なる環境で運動を行い、目的を達成できない。心理実験のうち、前者は意図した運動を行う成功条件に、後者は意図した運動を行えない失敗条件と対応する。以下に述べるように学習時と実演時の2つの環境を工夫することで、制御は異なるが、類似した運動を行うよう工夫した。もしイミテータ（子供に相当する第3のエージェント）が一見類似した運動をその制御の違いと一致するように分類できれば、イミテータが意図の違い（制御の違い）を識別できたとと言える。

### 2.2 倒立振り子課題

本稿では、異なる意図（制御）をもつ一見よく似た運動を生成するため、運動制御の単純な課題として倒立振り子課題（振り

連絡先: 北陸先端科学技術大学院大学

石川県能美市旭台 1-1

E-mail: {tak.torii,shhidaka}@jaist.ac.jp

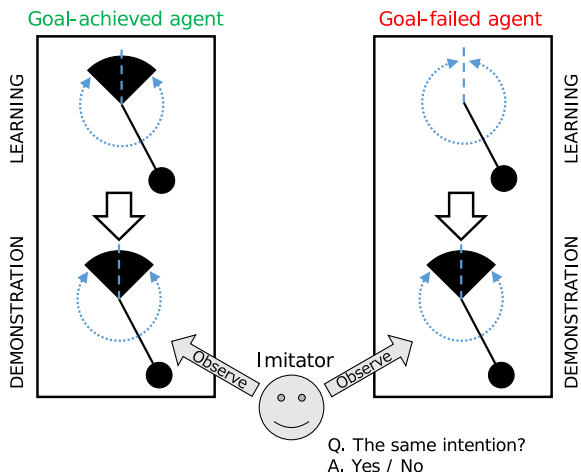


図 1: 実験設計. GA エージェント (左) は制約振り子で制御を学習し, 制約振り子で実演する (成功条件). GF エージェント (右) は標準振り子で制御を学習し, 制約振り子で実演する (失敗条件). イミテータは制御の違いを見抜けるか?

子の振り上げ課題) [Doya 99, Astrom 00] を扱う. 倒立振り子は長さ  $l = 1.0$  の棒の先端に質量  $m = 1.0$  をもち (図 1 右上), その状態は角度と角速度の対  $(\theta, \dot{\theta})$  で定まる. 倒立振り子の運動方程式は

$$ml^2\ddot{\theta} - 9.8ml \sin \theta = -b\dot{\theta} + u \quad (1)$$

ここで, 摩擦  $b = 0.01$ , 制御トルク  $u \in [-5, 5]$  である. 倒立状態は  $\theta = 0$  である.

振り子の状態を記述する量である力学的エネルギー  $E$  は運動エネルギー  $K$  と位置エネルギー  $U$  の和として

$$E(\theta, \dot{\theta}) = \frac{1}{2}ml^2\dot{\theta}^2 + 9.8ml(\cos \theta - 1) \quad (2)$$

である. 倒立停止状態は  $E(0, 0) = 0$  と表せる.

### 2.3 成功条件と失敗条件

倒立振り子課題において, 異なる制御で一見よく似た運動を生成するため, 本稿では制約振り子を導入する. 制約振り子 (図 1 左上) は, 標準的な振り子 (標準振り子) と異なり, 角度  $\theta \in \pm\pi/8$  を可動域としない (この範囲に反発のない壁がある). 成功条件 (goal-achieved; GA) では, GA エージェントは制約振り子を用いて最適な制御を学習し, 制約振り子を用いて実演を行う (図 1 左). したがって, GA エージェントは意図通り (制御通り) の運動を行う. 他方, 失敗条件 (goal-failed; GF) では, GF エージェントは標準振り子を用いて最適な制御を学習し, 制約振り子を用いて実演を行う (図 1 右). したがって, GF エージェントは意図通りでない運動を行う. 学習用の振り子は各条件で異なるが, 実演用の振り子は両条件で一致している. 本研究ではこの実験設計により, 異なる制御から類似した運動を生成可能にした. イミテータ ([Warneken 06] でいう子供) が類似した運動の背後にある制御 (意図) の違いを識別できるかを本稿では検討する (図 1 下).

### 2.4 強化学習による制御関数の構築

倒立振り子課題において, 累積報酬  $\sum_t \cos \theta_t$  の最大化を目的とする運動を生み出す制御関数  $u = g(\theta, \dot{\theta})$  は強化学習 (アクタークリティック法) [Sutton 98] を用いて構築した [Doya 99,

Grondman 12]. 状態空間  $(\theta, \dot{\theta}) \in [-\pi, \pi] \times [-2\pi, 2\pi]$  を  $40 \times 40$  分割のタイルコーディングとし, 正規乱数を用いた方策勾配降下で制御トルク  $u = g(\theta, \dot{\theta})$  を求めた. 学習は 5000 試行 (各試行 10000 ステップ) 以内には収束し, ほぼ最適に近い期待報酬をだす制御関数を構築できた.

## 3. 点次元による制御の識別

本稿では, 制御の相違がその制御で生成される運動の力学的不変量の相違として観察されるという仮説と検討する. この仮説を検討するため, 後続の節では, イミテータの立場から, GA エージェント (成功条件) の運動と GF エージェント (失敗条件) の運動だけから, その運動をその背後にある制御関数の違いとして区別できるかを分析する.

力学的不変量のひとつに点次元 [Cutler 93] がある. 点次元はフラクタル次元の一種で, 点集合 (たとえば軌道) の各点についてその各点の自由度を特徴づける. 点次元の推定方法が提案されている [Hidaka 13].

### 3.1 振り子の運動と次元

倒立振り子課題の成功条件では, GA エージェントの制御が生み出す運動は制約振り子に対して最適化されており, 「なめらかで」自由度の低い (次元の低い) 運動を生成すると予想される. 他方, 失敗条件では, GF エージェントの制御が生み出す運動は制約振り子に対して最適化されておらず, 「ぎこちなく」自由度の高い (次元の高い) 運動を生成すると予想される.

図 2 は GA エージェント (成功条件), GF エージェント (失敗条件) それぞれの制御する制約振り子の軌道と, その軌道から計算された点次元の時系列を示す. 制約振り子の軌道はとても類似しており, 軌道 (運動) を観察すると一見区別がつかない. しかし, 運動の背後にある次元を比較するとより区別が付きやすい. GA エージェントの運動は, 約 2000 ステップで最大振幅に到達した以降, 次元が平均的に低い. 他方, GF エージェントの運動は, 約 2000 ステップ以降, 逆に次元が平均的に高い. この結果は, 制御の相違がその制御で生成される運動の力学的不変量 (次元) の相違として観察されるという仮説を支持する. また, 制御通りの運動は次元が低く, 他方, 制御通りでない運動は次元が高いという予想と一致している.

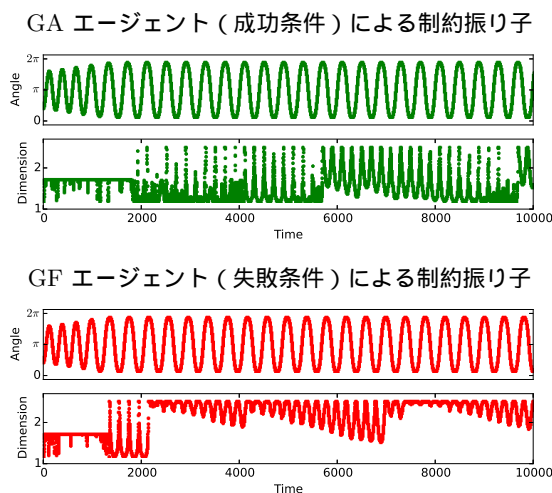


図 2: 制約振り子の角度と点次元の時系列

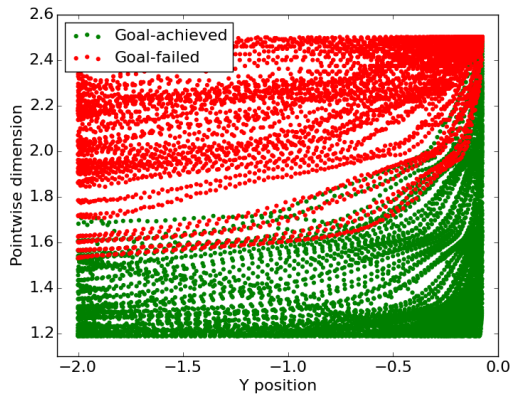


図 3: 点次元と振り子の高さ方向の位置  $\cos \theta - 1$

GA エージェントと GF エージェントの制御の違いは、GF エージェントの制御が可動域制限なしを前提としたものである点で、制約振り子の可動域境界（壁）付近にあると考えられる。図 3 には、振り子の高さ方向の位置（ $\cos \theta - 1$ ）に対して点次元を図示した。可動域境界は  $\cos \theta - 1 = -0.076$  である。図から、どちらのエージェントの制御する運動も可動域境界付近では次元が上昇している。この結果から、点次元は制約振り子の可動域境界（壁）に対する衝突を捉えていると考えられる。

制約振り子の制御では、可動域境界に達すると力学的エネルギー（とくに運動エネルギー）が減少し、可動域境界を離れたあとは次の振り上げに向けて運動エネルギーを補充するといったパターンが予想される。このパターンから、制約振り子の運動と次元の関係は、振り子の状態を記述する量である力学的エネルギーを経由して捉えられると考えられる。力学的エネルギー  $E$  は運動エネルギー  $K$  と位置エネルギー  $U$  に分解できる。図 4 には、 $K+U$  と  $K-U$  の平面上に、点次元を図示した。点の色は図 3 と同じくエージェントによる制御の違いを表す。図から、両エージェントの制御から生成される運動は運動エネルギーと位置エネルギーの平面上で周期的な軌道を描いており、可動域境界（壁）に衝突する付近で次元が高まることがわかる。この結果から、点次元が振り子の力学的エネルギーの変化に関して重要な特性を捉えていると考えられる。

### 3.2 運動データから制御の識別

前節までで、制御の違いが運動の次元の違いとして観察できることを定性的に確認した。本節では、イミテータの立場から、GA エージェント（成功条件）の運動と GF エージェント（失敗条件）の運動だけから、その運動をその背後にある制御関数と一致するように分類できるかを定量的に分析する。具体的には、角度、角速度、周波数、力学的エネルギー、点次元という 5 種類の特徴を用意し、これらの特徴の違いによる分類性能（正答率）を調べる。力学的エネルギーは振り子の状態を特徴づける直接的な特徴であり、前節の結果（図 4）からも高い分類性能を与えることが予想される。しかし、イミテータの立場から言えば、力学的エネルギーは振り子の運動方程式を既知としないとえられない。他方で、点次元は振り子の運動（XY 座標の時系列）から計算でき、振り子の知識を前提としない。

イミテータの分類判断には混合正規モデルを用いた、2 条件 × 時系列 15 本分の点をラベル付き訓練データとし、同じく 2 条件 × 時系列 15 本分の点を試験データとし分類させ、正

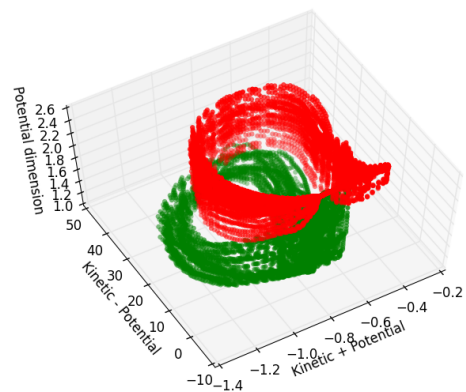
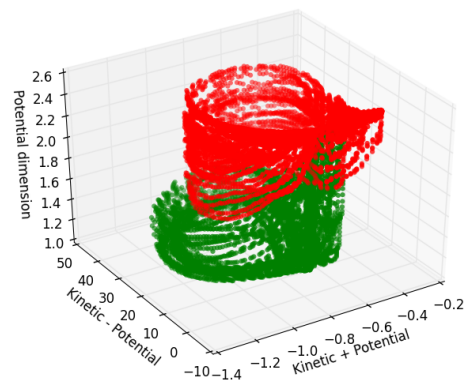


図 4: 点次元（Z 軸）と、力学的エネルギー（ $K+U$ ）、運動エネルギーと位置エネルギーの差（ $K-U$ ）の関係

答率を求めた。訓練・試験データには、可動域境界（壁）に最初に到達する前後（約 2000 ステップ）で、過渡期と定常期に分け、フルデータ（過渡期 + 定常期）、過渡期のみ、定常期のみ、の 3 種類を用意した。

図 5 は各特徴（角度、角速度、周波数、力学的エネルギー、点次元）を用いた場合の分類の正答率を示す。図から、角度、角速度、周波数では、フルデータ、過渡期のみ、定常期のみによらず、ほぼチャンスレベルである。力学的エネルギーでは、定常期のみの場合には約 71% の正答率である。この結果は、力学的エネルギーのように、振り子の知識を前提とした量を用いれば、運動の背後にある制御の違いを識別できることを示唆する。一方で、点次元では、力学的エネルギーほどではないが、定常期のみの場合には約 64% の正答率である。この結果は、点次元という特徴に着目すれば、振り子の知識を前提としなくとも、運動の背後にある制御の違いを識別できることを示唆する。

## 4. 議論

本稿では、目的や意図に関する事前知識を所与とせず、いかにして一見よく似た運動の背後にある意図（制御）の違いを識別できるかを扱った。本稿では、制御の相違がその制御の下で生成される運動の力学的不変量、点次元の相違として観察されるという仮説を検討した。[Warneken 06] の心理実験に着想

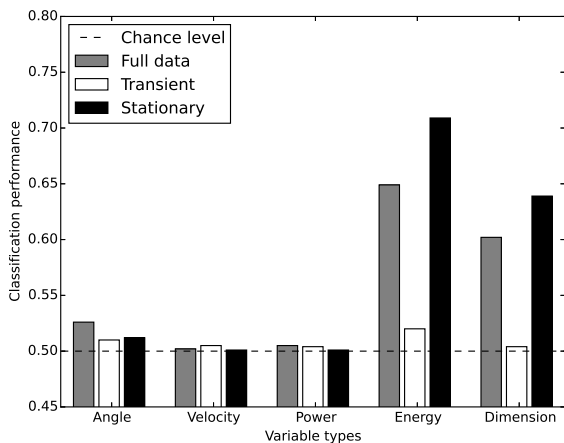


図 5: 運動データから制御の分類性能

をえた実験設計でシミュレーションを行い，点次元という力学系の不変量を特徴とした場合，振り子の知識を前提としなくとも，高い正答率で制御の違いを識別できることが示された．点次元を特徴とした場合の正答率は，振り子の知識を前提とする力学的エネルギーを特徴とした場合に匹敵する．この結果は，力学的不変量である次元に着目することで，意図推定がほとんど事前知識なしに身体運動の観察から可能であることを示唆する．本稿の範囲は意図（制御）の識別に留まるが，本稿の成果は真の模倣 [Breazeal 02] に向けた第一歩といえる．今後は，意図（制御）の推定と合わせて，イミテータによる運動の生成を検討していく．

## 謝辞

本研究は科学研究費補助金 若手研究 A 16H05860 の助成を受けて行われた．

## 参考文献

- [Astrom 00] Astrom, K. J. and Furuta, K.: Swinging up a pendulum by energy control, *Automatica*, Vol. 36, No. 2, pp. 287–295 (2000)
- [Breazeal 02] Breazeal, C. and Scassellati, B.: Robots that imitate humans, *TRENDS in Cognitive Sciences*, Vol. 6, No. 11, pp. 481–487 (2002)
- [Cutler 93] Cutler, C. D.: *A review of the theory and estimation of fractal dimension*, Vol. 1, pp. 1–107, World Scientific (1993)
- [Doya 99] Doya, K.: Reinforcement learning in continuous time and space, *Neural Computation*, Vol. 12, pp. 243–269 (1999)
- [Grondman 12] Grondman, I., Vaandrager, M., Busoniu, L., Babuska, R., and Schuitema, E.: Efficient model learning methods for actor-critic control, *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 42, No. 3 (2012)
- [Hidaka 13] Hidaka, S. and Kashyap, N.: On the estimation of pointwise dimension, *ArXiv:1312.2298* (2013)

[Schaal 97] Schaal, S.: Learning from demonstration, in Mozer, M., Jordan, M., and Petsche, T. eds., *Advances in Neural Information Processing Systems 9*, pp. 1040–1046, Cambridge: MIT Press (1997)

[Sutton 98] Sutton, R. S. and Barto, A. G.: *Reinforcement Learning: An Introduction*, MIT Press (1998)

[Warneken 06] Warneken, F. and Tomasello, M.: Altruistic helping in human infants and young chimpanzees, *Science*, Vol. 311, pp. 1301–1303 (2006)