

予測精度の予測に基づいた能動的・受動的な適応行動の生成学習

Learning to generate proactive and reactive behavior
by predicting the precision of prediction errors

村田 真悟*¹
Shingo Murata

山下 祐一*²
Yuichi Yamashita

有江 浩明*³
Hiroaki Arie

尾形 哲也*³
Tetsuya Ogata

谷 淳*⁴
Jun Tani

菅野 重樹*⁵
Shigeki Sugano

*¹早稲田大学理工学術院 創造理工学研究科 総合機械工学専攻

Department of Modern Mechanical Engineering, Graduate School of Creative Science and Engineering, Waseda University

*²国立精神・神経医療研究センター 神経研究所 疾病研究第七部

Department of Functional Brain Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry

*³早稲田大学理工学術院 基幹理工学部 表現工学科

Department of Intermedia Art and Science, School of Fundamental Science and Engineering, Waseda University

*⁴Department of Electrical Engineering, Korea Advanced Institute of Science and Technology

*⁵早稲田大学理工学術院 創造理工学部 総合機械工学科

Department of Modern Mechanical Engineering, School of Creative Science and Engineering, Waseda University

This paper presents a novel hierarchical neural network model that can learn to predict not only the next sensory input but also the precision of prediction errors. The proposed model was implemented in a humanoid robot that was required to learn to generate adaptive behavior depending on the situation. Our experimental results demonstrate that the robot controlled by the trained network can generate both proactive and reactive behavior based on the predicted precision.

1. はじめに

人間は膨大な知覚・行動経験を通じて、生じた現象の解釈や構造化（学習）を行い、外部環境や他者に対する適応行動を実現していると考えられる。例えば、適応行動の例として他者との協調を考えてみると、単に自身の行動を相手に対して受動的に合わせるだけではなく、その相手が何をしようとしているのかを能動的に予測しながら自身の行動を選択することで、円滑なコミュニケーションが可能になるであろう。しかしこの実現のためには、自己と独立した他者のふるまいを全て予測することは不可能であるため、一連の経験から予測が容易である部分と困難である部分を切り分けて、予測学習や行動生成をする必要がある。

これまで我々は、人間の脳の予測メカニズムに着目し Recurrent Neural Network (RNN) をベースとした様々な神経回路モデルを提案してきた [Ito 06, Yamashita 08]。本研究では、適応行動の実現にはこれまでと同様の予測メカニズムに加え、さらにその「予測精度の予測」[Feldman 10] が重要であると考え、それを実現可能な神経回路モデルを提案する。そして、提案モデルを小型ヒューマノイドロボットに実装し、人間によって操作される他者ロボットとの協調行動の生成学習実験を行う。これらの実験を通して、予測精度の予測に基づいた能動的・受動的な適応行動が学習によって実現されることを示す。

2. 学習モデル

本研究では、Yamashita ら [Yamashita 08] によって提案された Multiple Timescale RNN (MTRNN) と Namikawa ら [Namikawa 13, Murata 13] によって提案された Stochastic

Continuous-Time RNN (S-CTRNN) をベースとした新たな神経回路モデルとして、Stochastic MTRNN (S-MTRNN) を提案する。図 1 に本ネットワークの概略図を示す。実験では、ロボットの行動経験によって得られた感覚時系列データを用いて、この S-MTRNN の学習を行う。

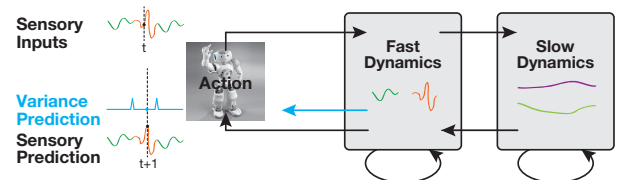


図 1: S-MTRNN の概略図。遅いダイナミクスを表現する上位ネットと速いダイナミクスを表現する下位ネットから構成される。ネットワークは次状態の感覚入力を予測するだけでなく、その予測精度も予測可能である。学習後には、予測された感覚入力のうちロボットの姿勢に関するものを関節角度指令値として用い、ロボットの行動生成を行う。

S-MTRNN は時定数が大きな（活動が遅い）上位ネットと、時定数が小さな（活動が速い）下位ネットによって構成される。このように、上位と下位のネットワークに異なる時定数を設けることで機能的階層性が自己組織化されることが Yamashita らによって示されている [Yamashita 08]。ネットワークの学習は、ロボットの現時刻における感覚入力（視覚や姿勢情報）とネットワークの内部状態から次時刻におけるロボットの感覚入力の平均と分散（予測精度の逆数）を予測するように、勾配法を用いた最尤法によって行われる [Namikawa 13]（学習の詳細は [Murata 13] を参照）。

連絡先: 連絡先: 村田真悟, 早稲田大学理工学術院, 東京都新宿区大久保 3-4-1, murata@sugano.mech.waseda.ac.jp

3. ロボット実験

3.1 タスク

図2に実験タスクの概略図を示す。実験には、人間によって操作される「実験者ロボット」とS-MTRNNによって駆動される「被験者ロボット」が用いられた。これらのロボットはお互いに向き合い、実験者ロボットは保持するオブジェクトをホームポジションから左か右へ等確率で動かすという行動を行った。そのような行動に対して、被験者ロボットは実験者ロボットの動きを予測し、その物体に触りにいくという行動を学習した。

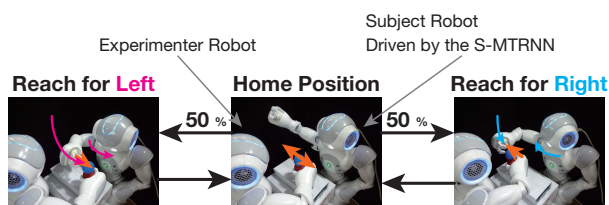


図2: 実験タスクの概略図。

3.2 結果

学習後のネットワークを用い、被験者ロボットが行動生成をした際に取得した時系列データを図3に示す。図から、スパイク上の分散予測を確認することができる。この分散が立ち上がっている部分は、実験者ロボットがホームポジション（分岐点）にいる状態と対応している。すなわち、ネットワークは物体が左か右のどちらへ動かされるか、という予測が不可能であるため、高い分散（低い予測精度）を予測しているといえる。しかし、一度実験者ロボットが物体を左か右へ動かした後は、その後の動きが予測可能になるため、低い分散（高い予測精度）を予測しているといえる。これらの予測精度の予測に応じて、ネットワークは自律的に予測学習の強弱を調整し、ロボットは感覚入力に対して反応する受動的な行動生成と、自身の予測に基づく能動的な行動生成を実現していると考えられる。

4. まとめ

本研究では、適応行動の実現には予測精度の予測が重要であると考え、それが可能な神経回路モデルを提案した。提案モデルを実ロボットに実装し、二台のロボットを用いて適応行動の生成学習実験を行った結果、予測精度が高いと予測される場合には、強い学習が行われることでロボットは能動的に行動を生成し、予測精度が低いと予測される場合には、弱い学習が行われることでロボットは感覚入力に対して受動的に反応し行動を生成することが示された。

謝辞

本研究成果の一部は、文部科学省科研費若手研究 (B) (No. 23700279), 文部科学省科研費基盤研究 (C) (No. 25330301), 文部科学省科研費基盤研究 (S) (No. 25220005), 文部科学省科研費新学術領域研究「構成論的発達科学」(No. 24119003), 文部科学省「卓越した大学院拠点形成支援補助金」, JST さきがけ領域「情報環境と人」の助成を受けたものです。本研究成果の一部は、早稲田大学理工研プロジェクト研究「自然と共生する知能情報機械系に関する基盤研究」の一環として行われたものです。ここに謝意を表します。

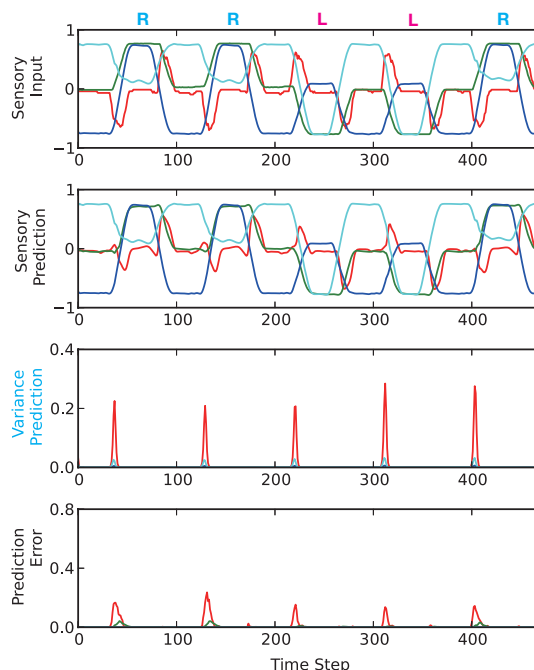


図3: 行動生成実験によって得られた時系列データ。上段から、実際の感覚入力（実験者ロボットの行動ラベルが振られている）、予測された感覚入力、予測された分散、予測誤差を示す。図中の赤、緑、青、シアン軌道はそれぞれ、ロボットに搭載されたカメラの画像における物体の中心位置（横方向）、首（ヨー）、肩（ピッチ）、肘（ヨー）の関節角度を示す。

参考文献

- [Ito 06] Ito, M., Noda, K., Hoshino, Y., and Tani, J.: Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model: *Neural Networks*, Vol. 19, No. 3, pp. 323-337 (2006)
- [Yamashita 08] Yamashita, Y. and Tani, J.: Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment: *PLoS Comput. Biol.*, Vol. 4, No. 11, e1000220 (2008)
- [Namikawa 13] Namikawa, J., Nishimoto, R., Arie, H., and Tani, J.: Synthetic approach to understanding meta-level cognition of predictability in generating cooperative behavior: *Advances in Cognitive Neurodynamics (III)*, Springer, pp.615-621 (2013)
- [Murata 13] Murata, S., Namikawa, J., Arie, H., Sugano, S., and Tani, J.: Learning to reproduce fluctuating time series by inferring their time-dependent stochastic properties: application in robot learning via tutoring: *IEEE Trans. on Autonomous Mental Development*, Vol.5, No.4, pp.298-310 (2013)
- [Feldman 10] Feldman, H. and Friston, K. J.: Attention, uncertainty, and free-energy: *Front. Hum. Neurosci.*, Vol. 3, No. 96, pp.1-14 (2010)