

# 発火閾値の均一性がもたらす汎化能力の低下と繰り返し行動 —自閉スペクトラム症に対する認知ロボティクス—

Reduced Generalization and Repetitive Behavior Caused by  
Little Heterogeneity of Intrinsic Neuronal Excitability:  
A Neuro-Robotics Study on Autism Spectrum Disorder

1W130044-6 出井 勇人 指導教員 尾形 哲也 教授  
IDEI Hayato Prof. OGATA Tetsuya

概要: 自閉スペクトラム症に対するこれまでの研究により、遺伝的要因及び認知的基盤について、一定の理解や仮説が得られてきている一方で、それらの関係性や症状に至るまでのメカニズムについては、未知な部分が多い。これに対し、神経科学における理論や計算モデルを精神医学に応用し、現象の観察を通じてその病態原理を理解しようとする計算論的精神医学が注目されている。本研究では、再帰型神経回路モデルを用いた予測学習に基づき、ニューロン間の興奮性の均一さが、感覚情報に対する不確実性の減少につながることを示した。また、人型ロボットに神経回路モデルを導入し実験を行った結果、汎化能力の低下、不安定な運動制御、繰り返し行動等が観察された。これらは、認知科学における理論及び、観察される症状と合致し、自閉スペクトラム症の病態がニューロン間の興奮性の均一さと予測誤差最小化に基づく計算処理過程で説明できる可能性を示す。

キーワード: 自閉スペクトラム症, 計算論的精神医学, 認知ロボティクス, ニューロン間の異質性, 不確実性

Keywords: Autism Spectrum Disorder, Computational Psychiatry, Cognitive Robotics, Neuronal Heterogeneity, Uncertainty

## 1 はじめに

自閉スペクトラム症 (Autism Spectrum Disorder : ASD) は、対人やりとりにおける質的欠陥や常同行動に特徴付けられる脳の発達障害である。これまでの研究により、その認知基盤や、生物学的要因について、各分野で一定の理解や仮説が得られてきている。しかし、認知的特徴と遺伝的要因との関係性、及び、実際の症状に至るメカニズムについては、未知な部分が多い。ここで近年注目されているのが、神経科学における理論やモデルを精神医学に応用し、その病態原理をシステムレベルで理解する、計算論的精神医学である [1]。この考え方に基づき、本研究では、ASD の神経学的要因について提案し、神経回路モデルを用いて検証した。また、ネットワークの機能と、身体行動とを区別するために、人型ロボットに神経回路モデルを導入し、生成行動を ASD の行動的特徴と対比するアプローチをとった。

## 2 手法

ASD の病因の一つとして、ニューロンの興奮抑制制御の不全について研究が進められており、興奮作用と抑制作用との間に偏りが生じているとする仮説が提唱されている [2]。しかし、興奮作用の相対的な増加・減少ともに ASD との関連が報告されており、症状をもたらす具体的なメカニズムについては、明らかになっていない。ここで、本研究では、興奮作用の相対的な増加及び減少の共通要素をなすモデルとして、ニューロン間における発火閾値の均一

性、つまり興奮性の多様性の減少を提案する。

本研究で用いた神経回路モデルは、非決定論的な時系列データの予測学習が可能な再帰型神経回路モデル S-CTRNN である [3]。S-CTRNN は、予測対象の現在の状態と、内部に保持するコンテキスト情報から、次の時刻における状態の予測と不確実性の推定を出力し、不確実性に逆重み付けされた予測誤差を最小化する過程で学習を行う。したがって認知理論研究において提唱されている、感覚情報に対する不確実性の推定の減少により ASD の多様な症候を説明する仮説について検証が可能である [4]。

S-CTRNN におけるコンテキスト層のニューロンの内部状態は、以下の式で表される。

$$u_{t,i}^{(s)} = \frac{1}{\tau_i} \left( \sum_{j \in I_1} w_{ij} x_{t,j}^{(s)} + \sum_{j \in I_C} w_{ij} c_{t-1,j}^{(s)} + b_i \right) + \left( 1 - \frac{1}{\tau_i} \right) u_{t-1,i}^{(s)} \quad (2.1)$$

ここで  $b_i$  はバイアスと呼ばれ、個々のニューロンの本来的な発火頻度の上がりやすさを表す。したがって、ニューロン間のバイアスに正規分布を仮定し、分散  $\sigma^2$  により興奮性の多様性を定めた。そして、バイアスの分散の異なるネットワークを複数用意し、それぞれ学習させた。

ロボットに課した行動は、実験者との右手によるボール転がし (Right) と左手によるボール転がし (Left) の 2 パターンである。ネットワークに学習させるデータとして、各行動パターンに対して 3 本ずつ時系列データを用意した。また、各ネットワークが不確実性を正確に推定できているかを判断しやすくするため、あらかじめ一定の不確実性を持つノイズを加えた。

本研究では、実験者がインタラクションの途中で、ポー

ルのダイナミクスを各パターン間で変化させる実験を行った。この時、ロボットは、環境に合わせて適応的に行動を切り替えることを求められる。したがって、S-CTRNNには、大局的な時間表現を持つパラメトリックバイアス(PB)と呼ばれるニューロン群を付加し、学習の過程で各時系列パターンに対する内部表現を獲得させておいた。また、インタラクション中に、不確実性に逆重み付けされた予測誤差(負の尤度)を最小化するように、PBの内部状態を更新させた。

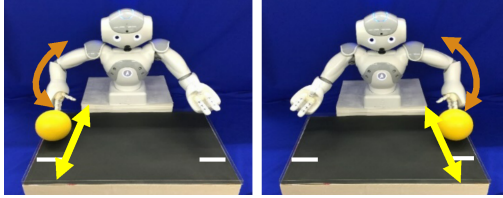


図1 ロボットに課した行動パターン。(左)Right, (右)Left.

### 3 結果

バイアスの分散が異なるネットワークをそれぞれ学習させた結果、バイアスの分散が、一定以上小さい( $\sigma^2 \leq 1$ )ネットワークにおいて、overfitting, 及び感覚情報に対する不確実性の推定の減少が観察された(図2参照)。また、これらのネットワークにおいて、PBにおける各時系列データの内部表現が一部の値に密集する特徴が現れた。これは、ネットワークの情報処理が、低次の感覚情報に頼るように構造化されたことにより生じたと考えられる。

次に、各ネットワークをロボットに導入し、汎化性をテストした。各行動パターンについて10回ずつボールの転がし合いを行い、ロボットが転がしたボールが白い線(ロボット側から20cm)を越えたら成功とカウントした(図1を参照)。ボールが白い線に届かない、あるいは左右にそれた場合は失敗とした。図3に成功率を示す。バイアスの分散が小さいネットワークでは、成功率が著しく下がり、汎化性が低下していることが分かる。

最後に、ロボットに、PBの内部状態の実時間での更新による適応を行わせた。実験の結果、 $\sigma^2 \geq 10$ の時、ロボットはボール位置の変化に対して適切な内部状態の切り替えを行った。一方で、overfittingしたネットワークにおいては、内部状態を切り替えることができない特徴が生じた。また、感覚情報に対する不確実性の推定の減少により、予測誤差が過度に肥大化し、内部状態が不適切な値に切り替わり固定してしまうこともあった。これらはいずれも、学習していない行動の繰り返しとして観察された。

以上より、汎化性の低下や同一性の保持、繰り返し行動といった、多彩なASDの症状が、興奮性の均一さと予測誤差最小化に基づく情報処理過程により現れることが示された。また、学習におけるoverfittingと、適応過程における重み付き予測誤差の肥大化によって、これらの症状が現れたことは、Cruysらの理論を支持する結果である[4]。

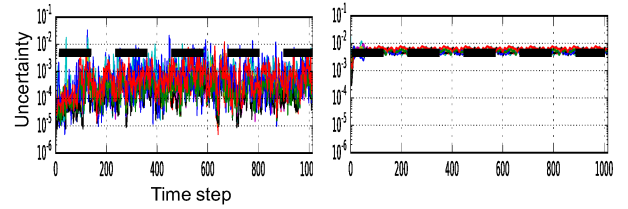


図2 感覚情報に対する不確実性の推定。(左) $\sigma^2 = 0.1$ , (右) $\sigma^2 = 100$ . 太点線は学習データの不確実性。

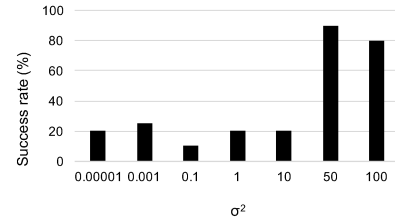


図3 各ネットワークにおける学習行動の成功率。

### 4 まとめ

本研究では、ASDの神経学的要因について提案するとともに、再帰型神経回路モデルと人型ロボットを用いて検証を行った。実験により、ニューロン間の興奮性の均一さと予測誤差最小化による情報処理過程により、ASDの理論と合致する認知的特徴、及び、多彩な症状が現れることが示された。これらの結果は、計算論的精神医学が、ASDに対する多様な分野の知見の関係を明らかにし、その病態を原理的に理解する上で有効であることを示唆する。今後は、実験で得られた結果について分析を重ねるとともに、統合失調症などの他の神経・精神疾患の病態についても研究を進めていきたい。

### 参考文献

- [1] K. J. Friston, K. E. Stephan, R. Montague, and R. J. Dolan, "Computational psychiatry: The brain as a phantastic organ," *The Lancet Psychiatry*, vol. 1, no. 2, pp. 148158, 2014.
- [2] R. Gao and P. Penzes, "Common mechanisms of excitatory and inhibitory imbalance in schizophrenia and autism spectrum disorders.," *Curr. Mol. Med.*, vol. 15, no. 2, pp. 14667, 2015.
- [3] S. Murata, J. Namikawa, H. Arie, S. Sugano, and J. Tani, "Learning to reproduce fluctuating time series by inferring their time-dependent stochastic properties: Application in Robot learning via tutoring," *IEEE Trans. Auton. Ment. Dev.*, vol. 5, no. 4, pp. 298310, 2013.
- [4] S. Van de Cruys et al., "Precise minds in uncertain worlds: predictive coding in autism.," *Psychol. Rev.*, vol. 121, no. 4, pp. 64975, 2014.