

神経細胞の分散培養系の深層強化学習

澤田晴登 高橋宏知

Dissociated culture, neural network, deep reinforcement learning, animat, modeling

1. 序論

生物知能の理解を深め、新たな制御方法を開発する手法の一つとして、ロボットを身体、神経細胞の分散培養系を脳として、身体から得た情報を脳にフィードバックする閉ループ系を構築し、その人工生物の能力や機能を調べる方法がある[1]。従来の研究では閉ループ系を特定の状態に収束させ、一定の出力を取り出すことには成功している[2]が、閉ループ系を用いて目的志向行動を実現した例はない。本研究では、神経細胞の分散培養系の活動でロボットを制御する実験系を構築し、強化学習を利用し、目的志向行動を実現することを目的とした。

2. 方法

神経細胞の分散培養系とロボットからなる閉ループ系を構成し、強化学習を利用してロボットの目的志向行動を得べく学習を行なった。学習には数10万の試行回数が必要であったため、試行回数を確保するために、神経細胞の分散培養系の活動を事前に計測し、モデル化して用いた。強化学習系として Bonsai (Microsoft) を、ロボットとして倒立振り子 (Microsoft) と二輪ロボットの2つを用いた。制御の目的は、倒立振り子では「既定の試行回数の間、台車・振り子を範囲内に維持すること」、二輪ロボットでは「フィールド内の目標地点に到達すること」とした。強化学習系はロボットの状態を取得し、神経活動モデルに入力する刺激を決定した。神経活動モデルの刺激応答からロボットモーションを生成する方法を 1 threshold (行動空間を2分割), 9 thresholds (行動空間を10分割), Control (行動空間を2分割し, 入力刺激をランダム化) に分け, 各条件で合計エピソード数が規定値になるまで学習した後, 成績を評価した。

3. 結果

Cartpole, Wheel Robot いずれの系においても, 1 threshold 条件の学習完了時の成績は, 1 threshold 条件の学習開始時の成績, Control 条件双方に対して優位に高かった。また, 9 thresholds 条件の学習完了時の成績は, 1 threshold 条件の学習完了時の成績に対して優位に高かった。

4. 考察

分散培養系と倒立振り子/二輪ロボットからなる閉ル

ープ系の制御を強化学習することで制御の成績が向上し, さらにその成績はチャンスレートに比べて有意に高いことが分かった。この結果は, 神経細胞の分散培養系の活動でロボットを制御する閉ループ系に, 強化学習を利用して目的行動を実現することができることを裏付けていると考えられる。さらに成績を向上させるためには, ロボットの行動空間の分割数を増やすことが有効であることも示された。

5. 結論

以上の結果より, 神経細胞の分散培養系の活動でロボットを制御する実験系を構築し, 強化学習を利用し, 目的志向行動を実現した。

参考文献

- [1] Thomas B. Demarse, Daniel A. Wagenaar, Axel W. Blau, et al., "The Neurally Controlled Animat Biological Brains Acting with Simulated Bodies," *Autonomous Robots*, vol.11, (2001), pp.305-310
- [2] Douglas J Bakkum, Zenus C. Chao, Steve M. Potter, "Spatio-temporal electrical stimuli shape behavior of an embodied cortical network in a goal-directed learning task," *Journal of Neural Engineering*, vol.5, no.3, (2008), pp.310-323

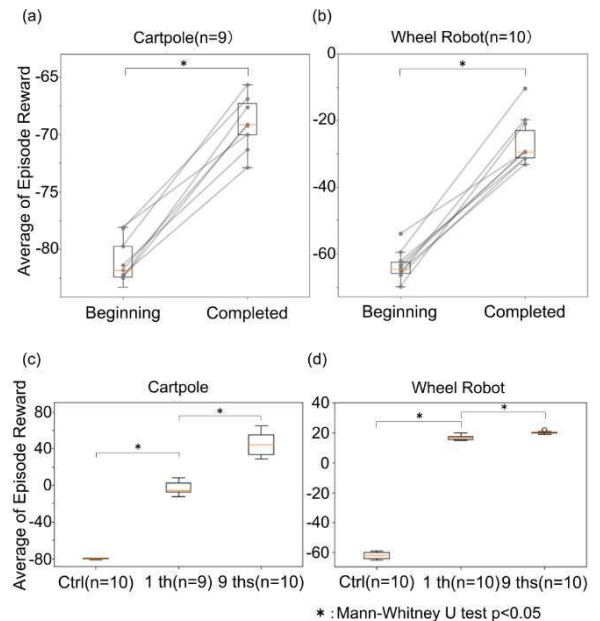


Fig. Results of experiments
Episode reward at beginning and completed of 1 threshold learning of Cartpole(n=9)(a) and Wheel Robot(b) (n=10)
Average of Episode rewards of all conditions of Cartpole(c) and Wheel Robot(d)