

STDP の初期強度依存性問題 - 強化メタ学習則による解決

Reinforcement meta-learning rule solves the problem in the initial-strength dependence of STDP

荒木 尚二郎 (PY)[†], 酒井 裕[‡]

Shojiro Araki(PY), and Yutaka Sakai

[†] 玉川大学大学院工学研究科, [‡] 玉川大学脳科学研究所

arksh4ee@engs.tamagawa.ac.jp

Abstract— It is known that spike-timing-dependent synaptic plasticity (STDP) depends on the initial strength of the synapse, and that the dependence is asymmetric for potentiation and depression. It is pointed out that this fact implies a problem: the destination of a synapse should be restricted in a small region determined by the initial-strength dependence, and little reflect the input-output statistics. In order to solve the problem, we propose a meta-learning rule depending on reinforcement signals. We applied the meta-learning for STDP learning rule that possesses asymmetric initial-strength dependence, and demonstrated that a single model neuron can learn the selectivity reflecting input correlation.

Keywords— STDP, Dopamine, Meta-plasticity

1 背景

シナプスの送信側 (pre) と受信側 (post) のニューロンから発生するスパイクの相対的なタイミングに依存して、シナプス強度が変化する現象が知られている (STDP; Spike-Timing-Dependent Plasticity)[1]. この相対的なタイミング依存性のみ注目してモデル化した STDP 学習則は、入力タイミング相関を反映した選択性の獲得や post 側の発火率調節など、従来の平均発火頻度に依存した Hebb 則より優れた機能をもつことがわかっている [2, 3]. しかし、STDP はスパイクタイミングだけでなく、そのシナプスの初期強度にも依存することが知られている [1]. その依存性は増強と減弱で非対称になっており、ある強度付近以外では、増強と減弱のバランスが取れていない. このような初期強度依存性を STDP 学習則を適用すると、あらゆるシナプスが増強と減弱のバランスが取れるシナプス強度付近に落ち着いてしまい、入力の統計性の学習や発火率調節の機能を失うことがわかっている [4]. 定常状態でのシナプス強度は、pre 側、post 側の活動パターンにほとんど依存せず、初期強度依存性によって決まってしまう [5]. この事実をそのまま受け入れると、シナプス変化則にもとづいた経験

に応じた学習という Hebb が提唱したパラダイムが崩壊してしまう、という問題が生じる. この問題を初期強度依存性問題と呼ぶ.

ここでは、強化信号に応じて学習則のシナプス強度依存性が変化するようなメタ学習則を導入し、この問題の解決を図る. 脳の中では様々な伝達物質が状況に応じて放出されており、その中でもドーパミンは、強化信号に関わると考えられている. ここでは、ドーパミンが直接シナプス強度変化に関わるのではなく、変化則の調整に関与している、という想定のもとでモデルを構築する.

2 シナプス学習則

STDP 学習則の相対スパイクタイミング依存性には増強・減弱に同じ時定数をとった指数関数を用い、スパイクを越えても効果が残る All-to-All 方式を採用する [2]. シナプス増強は初期強度に依存せず、シナプス減弱は初期強度に比例する初期強度依存性を採用する [4]. シナプスの強度 w_{jk} は、pre 側 (k), post 側 (j) でスパイクが発生したときに、次のような変化分 Δw_{jk} だけ、それぞれ変化する.

$$\Delta w_{jk} = \begin{cases} +A^+ P_k u_{jk} & (\text{at } j\text{-spike}) \\ -A^- P_j w_{jk} & (\text{at } k\text{-spike}) \end{cases}$$

ここで、 A^\pm は学習率を表す正の定数であり、 P_k は相対タイミングを表す変数で、次の方程式に従う.

$$\tau \frac{dP_k}{dt} = \sum_i \delta(t - t_i^k) - P_k$$

ここで、 $\{t_i^k\}$ は、 k 番目のニューロンのスパイク時刻系列を表す. 変数 u_{jk} は強化信号が与えられたときのみ変化する変数であり、そのときのシナプス強度 w_{jk} に依存して設定される.

$$u_{jk} = f(w_{jk}) \quad (\text{at reinforcement})$$

変数 u の変化は学習則の変化とみなすことができ、これを強化メタ学習と呼ぶことにする.

3 シミュレーション

ある統計性をもって確率的に発火する (相関ポアソン過程) 1000 個の興奮性ニューロン集団 (pre) と、ランダ

ムに発火する (独立ポアソン過程)200 個の抑制性ニューロン集団から入力を受けてスパイクを発生する 1 個のモデルニューロン (post) を考える . したがって , post ニューロンにあたる添え字 j は省略する . モデルニューロン (post) は , コンダクタンススペースの入力を受ける Leaky integrate-and-fire モデルに従い , モデルパラメータは [2] と同じに設定した . 興奮性シナプス k のピークコンダクタンスは $g_0 w_k$ と表し , $w_k = 1$ のとき , [2] の最大ピークコンダクタンスとなるように , $g_0 = 0.015$ と設定した . ただし , ここでは , [2] のように $w \leq 1$ となるように制限するような操作は行っていない . その他のパラメータは $\tau = 20\text{ms}$, $A^+ = 0.01$, $A^- = 1.05A^+$ と設定し , 強化メタ学習を定める関数は区分線形関数 $f(w) = w (w \leq 1)$, $f(w) = 1 (w > 1)$ と置いた .

1000 個の pre ニューロンから 100 個ずつ 5 グループ (G1, G2, ..., G5) を選ぶ . グループ内では相関があるが , グループ間では独立であるようなポアソン過程を発生する . 発生率 $\lambda(1 - c)$ で独立に発火するポアソン過程に加えて , 発生率 $100\lambda c/C$ でグループ内 100 個のうちランダムに選ばれた C 個が同期発火する事象が発生する . グループに属さないニューロンは , 発火率 λ で独立にポアソン過程を発生する . 各 pre ニューロンのスパイク系列はすべて , 発火率 λ のポアソン過程となり , グループ間にのみ一様な相関が存在する . ここでは , $c = 0.1$, $C = 20$, $\lambda = 20\text{Hz}$ に設定した .

初期条件として全ての変数を $u_k = w_k = 0.5$ に設定し , 1000 秒間シミュレーションした後のシナプス強度の値を図にプロットした . 図 1 は , 強化信号を与えない場合の結果であり , 従来の STDP 学習則 [4] と同等である . 全てのシナプス強度が入力相関のパターンにほとんど依存せず , $w \sim 0.5$ 付近に集まっていることがわかる . 一方 , 図 2 は , 強化信号を 1Hz でランダムに発生した場合の結果である . 相関グループ間に競合が起こり , 勝ち組 (G3,G4) と負け組 (G1,G2,G5) に分かれていることがわかる . さらに相関グループに属さない独立なニューロンからのシナプスはすべて弱くなっていることがわかる . 提案した強化メタ学習により , 入力の統計性に依存したシナプスパターンが獲得できていることがわかる .

4 まとめ

強化信号によるメタ学習を導入することにより , STDP の初期強度依存性問題が解決することがわかった . この強化メタ学習は , 報酬に結びつく結果の原因となった神経活動より報酬が遅れて与えられる状況で , どのようにその神経活動を強化するか , という遅延報酬問題の解決 [6] にも有効であることが予想される . 今回は強化信号をランダムに発生させたが , 本大会では強化信号を遅れて発生させて , 遅延報酬問題を解決できるかど

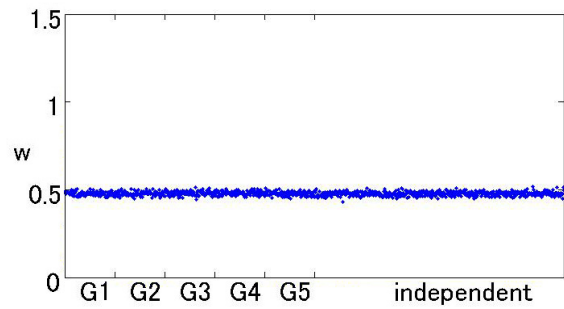


図 1: 強化信号なし

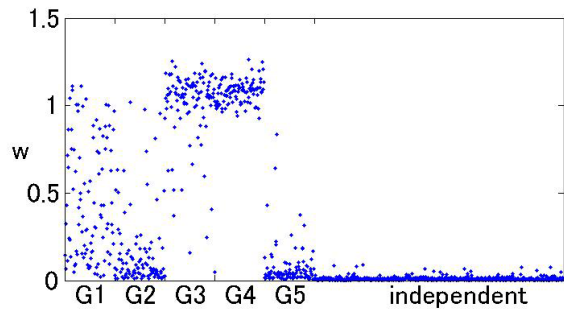


図 2: 強化信号あり

うか検討した結果も報告する .

参考文献

- [1] G.Q. Bi and M.M. Poo (1998) “Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type.” *J. Neuroscience*, **18**, 10464–10472.
- [2] S. Song, K. D. Miller, and L. F. Abbott. (2000) “Competitive Hebbian learning through spike-timing-dependent synaptic plasticity.” *Nature Neurosci.*, **3**, 919–926.
- [3] S. Song and L. F. Abbott (2001) “Cortical development and remapping through spike timing-dependent plasticity.” *Neuron*, **32**, 339–350.
- [4] M. C. W. van Rossum, G. Q. Bi, and G. G. Turrigiano (2000) “Stable hebbian learning from spike timing-dependent plasticity.” *J. Neuroscience*, **20**, 8812–8821.
- [5] 酒井裕 (2005) “STDP 学習則が導く現象 神経系数理における新展開.” *数理科学*, **43**, 46–52, 2005.
- [6] E. M. Izhigevich (2007) “Solving the Distal Reward Problem through Linkage of STDP and Dopamine Signaling.”