

## フォトニック強化学習における時間割引の解析

### Analysis of time discount in photonic reinforcement learning

白鳥 帆香<sup>1</sup>, ○漆原 昂<sup>2</sup>, ニコラ ショヴェ<sup>1,2</sup>, 砂田 哲<sup>3</sup>, 菅野 円隆<sup>4</sup>, 内田 淳史<sup>4</sup>, 堀崎 遼一<sup>1,2</sup>, 成瀬 誠<sup>1,2</sup>

1. 東大工, 2. 東大情報理工, 3. 金沢大, 4. 埼玉大

H. Shiratori<sup>1</sup>, ○T. Urushibara<sup>2</sup>, N. Chauvet<sup>1,2</sup>, S. Sunada<sup>3</sup>, K. Kanno<sup>4</sup>, A. Uchida<sup>4</sup>, R. Horisaki<sup>1,2</sup>, M. Naruse<sup>1,2</sup>

1. Faculty of Eng., Univ. Tokyo, 2. Grad. School of Info. Sci. & Technol., Univ. Tokyo, 3. Kanazawa Univ., 4. Saitama Univ.

E-mail: t-urushibara@g.ecc.u-tokyo.ac.jp

我々は光を用いた知的機能創生を目指した研究を行っている。1) では、強化学習の基礎問題であるバンディット問題を、レーザーカオスから得た乱数列である光カオス時系列を用いて解決した。2) では、バンディット問題とは異なり、複数の状態間を遷移する一般の強化学習を扱い、図1に示される手法を提案した。以下ではこの手法をフォトニック強化学習と呼ぶ。実験ではCart-Pole Balancingのシミュレーションを取り上げ<sup>3)</sup>、従来法のQ学習よりも速い学習が可能であることを示した。Cart-Poleの失敗時には図1中のThresholds updateでは指数型の割引を用いた。しかし指数型である必然性はないため、本研究では図1のフォトニック強化学習における時間割引について探索する。

まずパラメータチューニング前の双曲割引と比較したところ、双曲割引の方が学習が遅くなった。ここで、指数割引の場合は10ステップ以前の割引が0にみなせるほど小さいことを踏まえ、同一の双曲割引で10ステップ以前の割引を0にカットオフしたものを試すと指数割引の学習速度と同程度になった。ここで、同じく10ステップ以前の割引が0にならないしは0とみなせるよう調整したステップ関数型、シグモイド型、双曲型の割引をそれぞれ用いて指数型と比べたところ、図2のようにどれも同程度となった。図の見方としては横軸に沿って学習が進行し、縦軸はその各学習の進行度合いにおいてCart-Poleを立て続けることができた時間に対応する。図2の差が小さいことから、フォトニック強化学習の学習速度においては関数型よりも割引が有効なステップ数の方が重要だと考えられる。

図3は線形割引で割引時に遡るステップ数を変えたものである。遡るステップ数は図中図のように6ステップから16ステップまで2ステップ刻みで変化させた。図3から、Cart-Poleの場合は10ステップ程度まで遡ってThresholds updateを実施することが最適だと分かる。

今後の方針としては、Cart-Pole以外の問題設定でも有限ステップの遡りによるThreshold updateが最適であることを確かめ、それを踏まえてThreshold updateの方策を単純化することでフォトニック強化学習の汎用性を高めることを考えている。

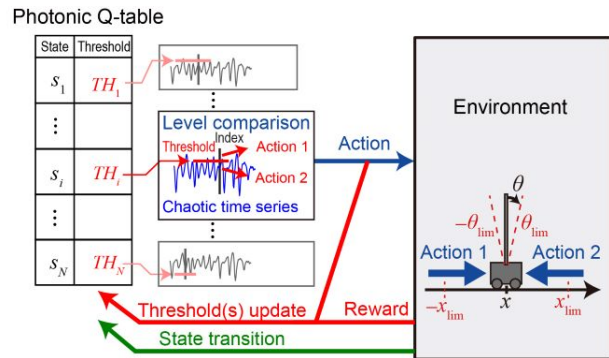


Fig. 1: フォトニック強化学習.

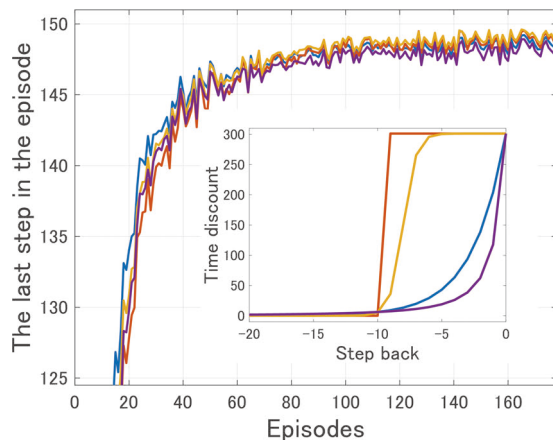


Fig. 2: 時間割引の関数型の比較.

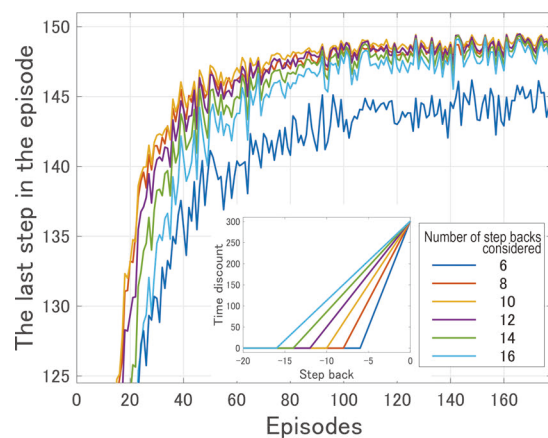


Fig. 3: 有効ステップ数ごとの比較 (線形割引) .

謝辞: 本研究の一部は JST CREST(JPMJCR17N2)、JSPS 科研費 (JP20H00233) の支援を受けた。

参考文献: 1) M. Naruse, et al. Sci. Rep. 7, 8772 (2017).

2) 漆原 昂, 他, 第 81 回応用物理学会秋季学術講演会, 講演予稿集, 19-031 (2020).

3) A.G. Barto, et al. IEEE Transactions on Systems, Man, and Cybernetics, SMC-13, 834 (1983)