

複数の半導体レーザカオス波形比較による意思決定手法の性能調査 Investigation of decision making by comparing multiple chaotic waveforms in semiconductor lasers

埼玉大¹, 東大情理²

○森尻 健正¹, 巳鼻 孝朋¹, 菅野 円隆¹, 成瀬 誠², 内田 淳史¹

Saitama Univ.¹, Univ. Tokyo²

○Kensei Morijiri¹, Takatomo Mihana¹, Kazutaka Kanno¹, Makoto Naruse², and Atsushi Uchida¹

E-mails: k.morijiri.632@ms.saitama-u.ac.jp, auchida@mail.saitama-u.ac.jp

はじめに: 近年, 多くの分野で強化学習が用いられている. 強化学習の問題例の1つとして, 多腕バンディット問題が挙げられる[1]. 多腕バンディット問題とは, ある確率に基づいて報酬が得られる複数台のスロットマシンから 1 台のスロットマシンを選択することを繰り返し, 得られた結果から学習を行うことで報酬の最大化を目的とする問題である. 多腕バンディット問題を解決するためには, 「探索」と「知識利用」の2つの行動をする必要がある. 探索とは, スロットマシンの試行を通じて, 最も報酬確率の高いスロットマシンを調査する行動である. 知識利用とは, 探索で得られた知識を基に, スロットマシンを選択する行動である. 探索と, 知識利用のどちらに偏りすぎても正しい意思決定を行うことはできず, 2つの行動の間にはトレードオフの関係が存在している.

多腕バンディット問題を解決する手法として, 綱引き理論が提唱されている[2,3]. また, 綱引き理論を用いた戻り光を有する半導体レーザカオスによる実装が行われている[4]. この手法では, 半導体レーザカオス波形に対してしきい値を設け, カオス波形としきい値の比較を行い, 選択するスロットマシンを決定する. さらに階層構造を利用することで, スロットマシンの台数を拡張させる手法が報告されている[5]. しかしながら, 従来法による実装では, 最大で 64 台のスロットマシンの多腕バンディット問題までに対応している.

そこで本研究では, 複数の半導体レーザカオス波形比較による意思決定手法を提案し, スロットマシンの更なる台数の拡張を目的とする. 特に, 多腕バンディット問題を解決するソフトウェアのアルゴリズムとして知られている, ϵ -greedy 法, Softmax 法, UCB1-tuned 法, Thompson Sampling と本手法との意思決定性能の比較を行う.

手法: 本研究で用いる意思決定手法を以下に述べる. N 台のスロットマシンに対し, それぞれ異なる半導体レーザカオス波形を用意する. 時刻 t における i 番目のスロットマシンに対応するカオス波形の強度を $S_i(t)$ とする. また, 各スロットマシンに波形バイアス $B_i(t)$ を与える. 各スロットマシンに対応するカオス波形の振幅 $S_i(t) + kB_i(t)$ を比較し, 最大値を示すスロットマシンを選択する. ここで, k はバイアス係数を表すパラメータである. バイアス係数とは, 探索と知識利用のバランスを決定する値であり, k が小さいと探索を重視し, k が大きいと知識利用を重視する. 選択したスロットマシンの結果に応じて, 波形バイアス $B_i(t)$ を綱引き理論に基づいて更新する. 以上の操作を繰り返すことで, 意思決定を達成する.

評価指標: 本研究では, 意思決定性能の評価指標として平均正答率(Correct Decision Rate, CDR)を用いる. CDR とは, 各試行回数に対して報酬確率が最大となるスロットマシンを選択した回数

の割合を示す指標である. 意思決定において m 回の選択を n サイクル行う場合, t 回目の試行における CDR は以下の式で与えられる.

$$CDR(t) = \frac{1}{n} \sum_{i=1}^n C(i, t), \quad t = 1, 2, \dots, m$$

ここで, $C(i, t)$ は i サイクル t 回目の試行において, 最大の報酬確率を持つスロットマシンを選択すれば 1 を返し, そうでなければ 0 を返す関数である. CDR が 1 に近いほど, 正しい意思決定が行われていると言える.

結果: 本研究では, 数値計算により生成した半導体レーザカオス波形を用いて, スロットマシンの台数が $N = 1024$ 台までの多腕バンディット問題における意思決定を行った. スロットマシンの台数が $N = 1024$ における CDR の比較を Fig. 1 に示す. $t = 40000$ における CDR は提案手法が一番高く, その値は 0.990 である. これは意思決定に成功したと言える. しかしながら, その他のアルゴリズムの手法では, 40000 回の試行では CDR の値が低く, 意思決定が上手くいかなかった. このように, 本研究で提案した複数の半導体レーザカオス波形比較による意思決定手法は, 従来の手法と比較して高い性能を示すことを実証した.

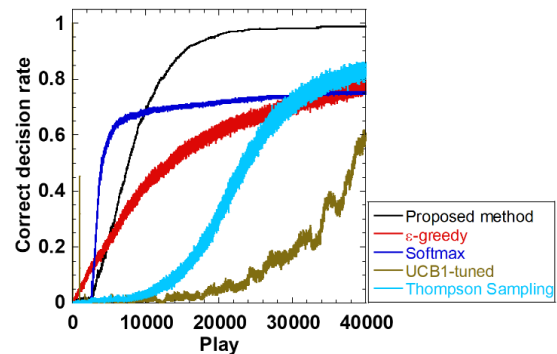


Fig. 1. Comparison of correct decision rate (CDR). Black : Proposed method using chaotic waveforms, red : ϵ -greedy, blue : Softmax, brown : UCB1-tuned, light blue : Thompson sampling.

まとめ: 本研究では, 複数の半導体レーザカオス波形による意思決定手法を提案した. ソフトウェアのアルゴリズムの手法と比較して, 高い性能を実証することに成功した.

参考文献

- [1] H. Robbins., et al, Bull. Am. Math. Soc., **58**, 527 (1952).
- [2] S. J. Kim., et al, New J. Phys., **17**, 083023 (2015).
- [3] S. J. Kim., et al, NOLTA, IEICE, **5**, 198 (2014).
- [4] M. Naruse., et al, Scientific Reports, **7**, 8772 (2017).
- [5] M. Naruse., et al, Scientific Reports, **8**, 10890 (2018).