

レーザの遅延カオス同期を用いた強化学習と意思決定の性能評価

Evaluation of reinforcement learning and decision making using lag synchronization of laser chaos

埼玉大¹, 情報通信研究機構²

○巳鼻 孝朋¹, 三井 湧方¹, 菅野 円隆¹, 成瀬 誠², 内田 淳史¹

Saitama Univ.¹, NICT²

○Takatomo Mihana¹, Yusuke Mitsui¹, Kazutaka Kanno¹, Makoto Naruse², and Atsushi Uchida¹

E-mails: t.mihana.229@ms.saitama-u.ac.jp, auchida@mail.saitama-u.ac.jp

はじめに: 強化学習[1]とは、未知の環境に対してエージェントが行動を選択することにより、環境からより多くの報酬を得るための選択を行う問題である。意思決定の問題例として、多本腕バンディット問題が挙げられる。この問題は、複数のスロットマシンから意思決定により1つのスロットマシンを選択することを繰り返すことで報酬を最大化する問題である。バンディット問題を解くためには、多くのスロットマシンを選ぶことによりどのスロットマシンが当たりやすいかを探索し、かつ探索から得られた知識を利用して最も当たり確率の高いスロットマシンを多く選択する必要がある。

近年、光を用いた意思決定として半導体レーザカオス波形を用いた実装が報告されている[2]。一方で、半導体レーザカオスには同期現象も知られている。特に相互結合された半導体レーザにおける遅延カオス同期は、先行振動するレーザが自発的にスイッチングすることが知られている[3]。そこで、結合強度により先行振動するレーザを制御することで意思決定に利用可能であると期待できる[4]。しかしながら、遅延カオス同期を用いた意思決定における最適なパラメータ設定方法については未だ評価されていない。

そこで本研究では、相互結合された2つの半導体レーザにおける遅延カオス同期を用いた意思決定を数値計算で行い、パラメータ変化に対する性能評価を行うことを目的とする。

方法: はじめに、意思決定方式について説明する。Fig. 1に示すように相互結合された2つのレーザ (Laser 1, Laser 2) が遅延カオス同期するパラメータ値に設定する。2つのレーザの時間波形からそれぞれのレーザが先行振動する場合の短時間相互相関値を算出する。短時間相互相関値を大小比較することによりどちらのレーザが先行振動しているかを判定できる[4]。そこで Fig. 1のようにスロットマシンとレーザを対応させ、先行振動しているレーザに対応するスロットマシンを選択する。結合強度は選択したスロットマシンの結果に応じて強化学習により段階的に変化させる。これを繰り返すことにより意思決定を行う。本研究では意思決定に用いる短時間相互相関値のサン

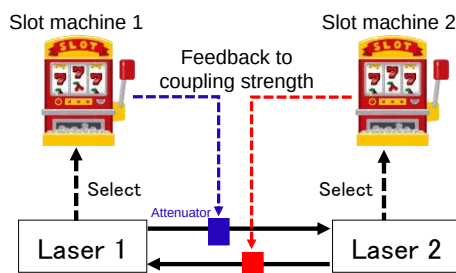


Fig. 1 Schematic diagram of decision making using lag synchronization of chaos in mutually-coupled semiconductor lasers.

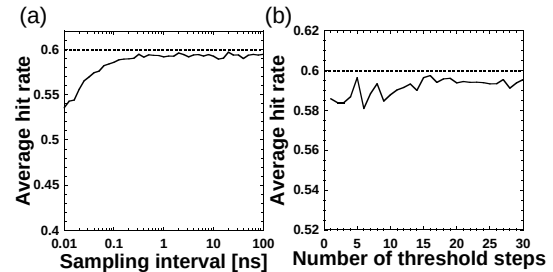


Fig. 2 Average hit rate as functions of (a) sampling interval in decision making and (b) number of threshold step.

プリング間隔と設定する結合強度の段数を変化させた。

意思決定の評価として平均当たり確率 (Average hit rate: AHR)を用いる。平均当たり確率は以下の式で表される。

$$AHR(t) = \sum_{i=1}^n \sum_{j=1}^m H(i, j) \quad (1)$$

ここで $H(i, j)$ は i サイクル j 回目の意思決定においてスロットマシンが当たりならば1を返し、はずれならば0を返す関数とする。平均当たり確率はスロットマシンの中で最も高い当たり確率に近い程良い性能であることを示している。

結果: 相関値のサンプリング間隔の変化に対する平均当たり確率の評価を Fig. 2(a) に示す。本研究では2つのレーザの結合遅延時間を5 nsに設定した。そのため、先行振動するレーザは5 nsごとにスイッチングする。サンプリング間隔が結合遅延時間と同じオーダーである1 nsよりも大きい場合には性能にあまり差が見られない。一方で1 nsよりも小さい場合には平均当たり確率が悪化することが分かった。

また、結合強度の段数を変化させた場合の平均当たり確率を Fig. 2(b) に示す。段数が少ない場合には平均正答率が低下することが分かった。これは探索の回数が少ないため、当たり確率の低いスロットマシンを選択する回数が増えるためだと考えられる。

まとめ: 本研究では、相互結合された2つの半導体レーザにおける遅延カオス同期を用いた意思決定の数値計算を行い、パラメータ変化に対する性能評価を行った。サンプリング間隔は結合遅延時間と同じオーダーで良い性能を示すことが分かった。結合強度の段数は多い場合に正しく意思決定を行えることが分かった。

参考文献

- [1] R. S. Sutton, et al., "Reinforcement learning," The MIT Press (1998).
- [2] M. Naruse, et al., Sci. Rep. 7, 8772 (2017).
- [3] K. Kanno et al., Phys. Rev. E 95, 052212 (2017).
- [4] 巳鼻 ら, 第 79 回応用物理学会秋季学術講演会 (2018).