

綱引き理論に基づく半導体レーザネットワークを用いた意思決定 Decision making using semiconductor laser network based on tug-of-war theory

埼玉大¹、東京大²

○巳鼻 孝朋¹、藤井 聖広¹、菅野 円隆¹、成瀬 誠²、内田 淳史¹

Saitama Univ.¹, Univ. of Tokyo²

○Takatomo Mihana¹, Kiyohiro Fujii¹, 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: 近年、強化学習の意思決定において光を用いた実装が多く報告されている[1-3]。これらの報告では、多腕バンディット問題における意思決定が行われている。多腕バンディット問題は、当たり確率が未知である複数のスロットマシンから一台を選択し、選択したスロットマシンから得られる報酬により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。

光を用いた実装例では、半導体レーザカオスを用いた 2 台のスロットマシンのバンディット問題に対する実装が報告されている[1]。この実装では GHz オーダの高速な意思決定が可能である。さらに、この方式をトーナメント方式に拡張することで、任意の台数のスロットマシンのバンディット問題に対応できることが報告されている[2]。

一方で半導体レーザの遅延カオス同期に着目した実装も行われている[3]。この実装では 2 台のスロットマシンのバンディット問題に対して、相互結合された半導体レーザの遅延カオス同期を用いる。相互結合における遅延カオス同期では一方のレーザ（ラグード）が他方のレーザ（リーダー）よりも遅延して同期するリーダー・ラグード関係が生じる[4]。リーダー・ラグード関係は結合遅延時間ごとに自発的にスイッチングし、結合強度変化に対してリーダー確率を制御できることが知られている[5]。この関係を用いることでリーダーとなるレーザに対応するスロットマシンを選択し、スロットマシンの結果に応じて結合強度を変化させることで意思決定が可能になる。しかしながら、3 台以上のスロットマシンに対する遅延カオス同期を用いた意思決定は行われていない。

そこで本研究では、半導体レーザネットワークを用いることにより遅延カオス同期を達成し、3 台のスロットマシンのバンディット問題における意思決定を数値計算により行うことを目的とする。

方法と結果: 遅延カオス同期を利用した意思決定では、スロットマシンとレーザの同期状態（リーダー・ラグード）を対応させる必要がある[3]。例えば、3 つのスロットマシンにおける意思決定では 3 つの同期状態が生じる必要がある。そのため、図 1(a)に示すようなレーザネットワークを構築する。このネットワークでは一方向結合によりリングネットワークを構成することで全てのレーザで異なる同期状態を有することができる。また、結合強度を変化させることによりリーダー確率の制御が可能である。例えば、図 1(a)において結合強度 κ_1 を小さくしていくと、レーザ 3 から 1 への結合が切断される。従って、レーザ 1 からレーザ 2 を経由し

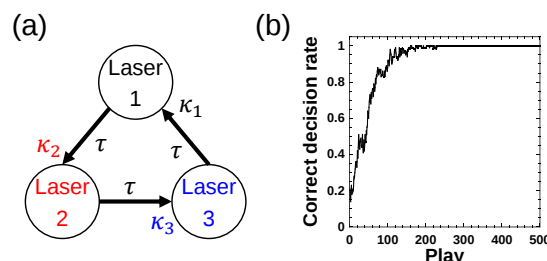


Fig. 1 (a) Schematic diagram for three laser network. (b) Correct decision rate for three-armed bandit problem.

てレーザ 3 への一方向結合となり、レーザ 1 がリーダーになることが分かる。

意思決定方法を以下に示す。まず、レーザとスロットマシンを対応させ、リーダーとなるレーザに対応するスロットマシンを選択する。リーダーとなるレーザの判定には、各々のレーザに対してそのレーザがラグードとなる短時間相互相関値を利用する。短時間相互相関値を大小比較し、最小値となったレーザがリーダーとなる。相関値が最小となるとき相関値に対応する結合が切断されたとみなされ、線形な結合となるためリーダーが判定できる。次に、選択したスロットマシンの報酬によりレーザネットワークの結合強度を変化させる。結合強度の変化量は多腕バンディット問題に対するアルゴリズムである綱引き理論[6]により決定する。レーザにおけるリーダー確率とスロットマシンの選択確率が対応しているため、結合強度変化によりスロットマシンの選択確率が変化する。このように、スロットマシンの選択と結合強度の変化を繰り返すことにより意思決定を達成する。

本手法による意思決定を 3 台のスロットマシンのバンディット問題に対する平均正答率により評価した。平均正答率が 1 に近づくと意思決定が達成できたと言える。評価の結果を図 1(b)に示す。3 台のスロットマシンにおいては、250 回の選択で平均正答率が 1 に到達した。従って、本手法による意思決定を達成できたと言える。

まとめ: 本研究では、半導体レーザネットワークにおける遅延カオス同期を用いることにより 3 台のスロットマシンのバンディット問題における意思決定を数値計算により達成した。

参考文献

- [1] M. Naruse, et al., Sci. Rep. **7**, 8772 (2017).
- [2] M. Naruse, et al., Sci. Rep. **8**, 10890 (2018).
- [3] T. Mihana, et al., Opt. Exp. **27**, 26989 (2019).
- [4] T. Heil, et al., Phys. Rev. Lett. **86**, 795 (2001).
- [5] K. Kanno, et al., Phys. Rev. E **95**, 52212 (2017).
- [6] S. -J. Kim, et al., NOLTA, IEICE, **5**, 198 (2014).