

マルチモード半導体レーザのモード競合ダイナミクスを用いた 多腕バンディット問題における意思決定

Decision making for multi-armed bandit problem using mode competition dynamics in multimode semiconductor laser

埼玉大¹, 東京大²

○岩見 龍吾¹, 巳鼻 孝朋¹, 小田 章裕¹, 菅野 円隆¹, 成瀬 誠², 内田 淳史¹
Saitama Univ.¹, The Univ. of Tokyo²

○Ryugo Iwami¹, Takatomo Mihana¹, Akihiro Oda¹,

Kazutaka Kanno¹, Makoto Naruse², and Atsushi Uchida¹

E-mail: riwami.692@ms.saitama-u.ac.jp, auchida@mail.saitama-u.ac.jp

はじめに: 近年、人工知能の強化学習に注目が集まっている。強化学習の問題の一種として、多腕バンディット問題が挙げられる[1]。これは、当たり確率が未知の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んで試行し、それを繰り返した時の総報酬の最大化を目指すという問題である。バンディット問題を解くためのアルゴリズムとして、綱引き理論[2]が提唱されている。このアルゴリズムを用いて、相互結合された半導体レーザの同期状態を制御し、スロットマシンが2台の場合のバンディット問題を解く方法が実装されている[3]。しかしながら、本方式は2台のレーザ間で生じる同期を相互相関値を用いて検出する必要があり、計算処理に時間がかかるという問題がある。

そこで、複数の波長(縦モード)を有するマルチモード半導体レーザに注目する。戻り光を有するマルチモード半導体レーザにおいて、光強度が最も強いモード(支配モード)が一定の発振状態とはならず、時間的に遷移する現象が観測されている[4]。このモード競合現象を複数のスロットマシンの選択に対応させることで、多腕バンディット問題に応用できると考えられる。

そこで本研究では、マルチモード半導体レーザのモード競合ダイナミクスを用いて多腕バンディット問題を解く方法を提案し、数値計算において実装を行うことを目的とする。

方法: マルチモード半導体レーザに対し、外部からレーザ光を注入することで、特定のモードを励起することができる。そこで本研究では、Fig. 1に示すようなモデルを考え、外部光の注入光量を変化させることでモード競合ダイナミクスの制御を行う。

次に、多腕バンディット問題を解くための方法を述べる。Fig. 1のように、モード番号とスロットマシンの番号を対応させる。各モードの時間波形を数値計算し、各モードの時間波形を比べ、光

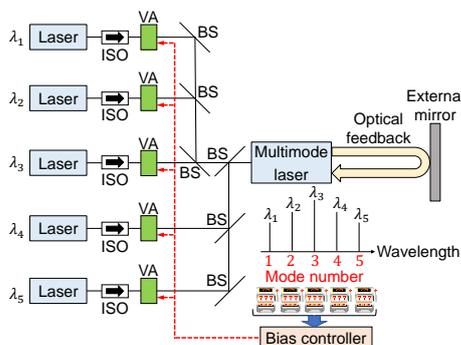


Fig. 1 Schematic of multimode laser subject to optical injection for each mode. λ_i : wavelength of mode i , BS: beam splitter, ISO: isolator, VA: variable attenuator.

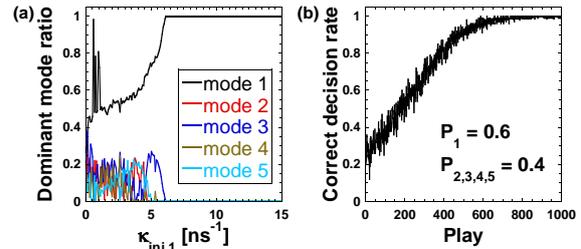


Fig. 2 Dominant mode ratio when injection strength of mode 1 is changed. (b) Correct decision rate as a function of the number of play.

強度が最も強いモードを支配モードと判定し、その支配モードに対応するスロットマシンを試行する。スロットマシンは「当たり」または「はずれ」の試行結果を返すが、試行結果に応じて外部光の注入光量を増減させることで、支配モードとなる確率を制御する。当たりを多く返すスロットマシンに対応するモードが支配モードとなるように、その選択確率を変化させることを繰り返して、多腕バンディット問題を解くことができる。

結果: モード 1 に対して外部光を注入し、その注入光量を変化させた時の支配モード比の変化を Fig. 2(a) に示す。注入光量が増加することで、モード 1 の支配モード比(黒線)が増加していることが分かる。これより、モード 1 の波長に対応する外部光を注入することで、モード 1 が支配モードとなる確率を上昇させることが可能となる。

次に、注入光量を変化させることでモード競合ダイナミクスを制御し、スロットマシンが5台の場合のバンディット問題を解いた。各スロットマシンの当たり確率を $P_1 = 0.6$, および $P_{2,3,4,5} = 0.4$ と設定し、1000回の試行を100サイクル繰り返す。当たり確率が最も高いスロットマシンを選んだ割合である平均正答率の変化を Fig. 2(b) に示す。これより、試行回数が増加することで、最も当たり確率が高いスロットマシンを試行できることが分かった。

まとめ: 本研究では、マルチモード半導体レーザのモード競合ダイナミクスを用いて多腕バンディット問題を解く方式を、数値計算において実装した。

参考文献

- [1] 本多 淳也ら, バンディット問題の理論とアルゴリズム, 講談社 (2016).
- [2] S.-J. Kim, et al., Nonlinear Theory and Its Applications, IEICE, **5**, 198 (2014).
- [3] 巳鼻 孝朋ら, 第 66 回応用物理学会春季学術講演会, 18-039 (2019).
- [4] Y. Liu, et al., International Journal of Bifurcation and Chaos, **8**, 1685 (1998).