

# 半導体レーザカオスを用いた超高速意思決定における適応性の調査

## Investigation of adaptation of ultrafast decision making using a chaotic semiconductor laser

埼玉大<sup>1</sup>, 情報通信研究機構<sup>2</sup>, Univ. Grenoble Alpes<sup>3</sup>, 物質・材料研究機構<sup>4</sup>

○巳鼻 孝朋<sup>1</sup>, 寺島 悠太<sup>1</sup>, 内田 淳史<sup>1</sup>, 成瀬 誠<sup>2,3</sup>, 金 成主<sup>4</sup>

Saitama Univ.<sup>1</sup>, NICT<sup>2</sup>, Univ. Grenoble Alpes<sup>3</sup>, NIMS<sup>4</sup>

○Takatomo Mihana<sup>1</sup>, Yuta Terashima<sup>1</sup>, Atsushi Uchida<sup>1</sup>, Makoto Naruse<sup>2,3</sup>, and Song-Ju Kim<sup>4</sup>

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

**はじめに:** 人工知能の強化学習は、インターネット広告やゲーム木探索など幅広い分野で利用されている。強化学習の例として、当たり確率が未知である複数のスロットマシンから報酬を最大化するという多本腕バンディット問題が挙げられる。報酬を最大化するには以下の2点が必要となる。1つは最も当たり確率が高いマシンを知るための探索が必要なことであり、もう1つは最も当たり確率が高いスロットマシンをできるだけ多く選択することである。過度な探索を行うと、最も当たり確率が高いマシンの選択回数が減少する。一方で少ない探索で得られた知識を利用すると、最も当たり確率が高いマシンの推定自体が間違いである可能性が高くなる。そのため、探索と知識利用はトレードオフの関係になっている[1]。

近年、多本腕バンディット問題の一解法として、綱引き理論 (Tug-of-war theory) [2]が Kimらにより提唱されている。また、綱引き理論に基づいた意思決定の物理実装が報告されている[3]。しかしながら、綱引き理論で用いるパラメータ値の詳細な調査は未だ行われていない。さらに半導体レーザにて発生するカオス波形を用いることで、意思決定の高速化が期待できる。

そこで本研究では、半導体レーザ実験にて生成された高速カオス波形を用いて、2本腕バンディット問題に対して綱引き理論に基づいた意思決定を行うことを目的とする。特に、しきい値に関する記憶係数の調査を行う。

**方法:** Fig.1に示すように、本研究では半導体レーザカオスの時間波形を取得し、しきい値を変化させることにより意思決定を行う。カオス波形がしきい値よりも上の場合には一方のスロットマシンを選択し、下の場合にはもう一方のスロットマシンを選択する。選択されたスロットマシンが「当たり」を出した場合、そのマシンが選択される領域が広がるようにしきい値を変化させる。一方で「はずれ」を出した場合、そのマシンが選択される領域が狭くなるようにしきい値を変化させる。これを繰り返すことにより強化学習を行う。

しきい値の変化量は、記憶係数  $\alpha$  のべき乗で重み付けされた過去の結果を利用する。また意思決定の評価として、1000回毎に当たり確率の入れ替え ( $P_A = 0.6, P_B = 0.4$ ) を行った場合の、当たり確率の高いマシンを選んだ割合の平均 (Correct Decision Rate : CDR) と、1サイクル (1000回) により得られた報酬の平均の割合

(Average Hit Rate : AHR) を算出する。

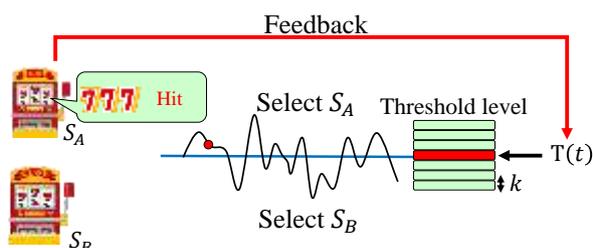


Fig. 1 Schematic diagram of decision making using chaotic laser output.

**結果:** 記憶係数が  $\alpha = 0.990$  と  $\alpha = 0.999$  の場合の CDR を Fig. 2(a) に示す。  $\alpha = 0.990$  の方が当たり確率の入れ替えが起きた後に、CDR が速く 1 に収束することが分かる。一方で  $\alpha = 0.999$  の方が 1 に収束した後の安定性が高い。また  $\alpha$  を連続的に変化させた場合の AHR を Fig. 2(b) に示す。  $\alpha = 0.9838$  の場合に最も AHR が大きくなることが明らかとなった。このように意思決定の記憶係数  $\alpha$  の最適値が存在することが明らかとなった。

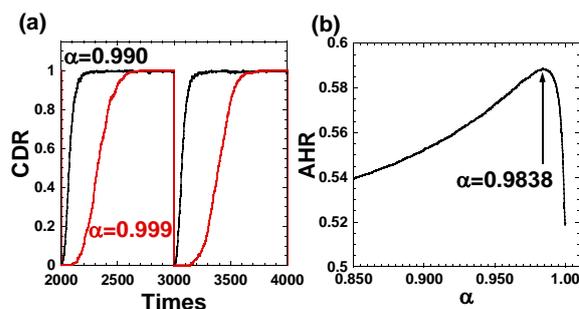


Fig. 2 (a) Temporal evolution of correct decision rate (CDR) for  $\alpha = 0.990$  (black) and  $\alpha = 0.999$  (red). (b) Average hit rate (AHR) as a function of memory parameter  $\alpha$ .

**まとめ:** 本研究では、綱引き理論に基づき半導体レーザカオスを用いた意思決定を行った。記憶係数  $\alpha$  を変化させることにより、意思決定の性能が大きく変化することが分かった。

### 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., New J. Phys. **17**, 083023, (2015).
- [3] 成瀬 誠, 他 第 77 回応用物理学会秋季学術講演会, p. 03-260 (2016).