

多本腕バンディット問題における最良選択に対する複雑ダイナミクスの効果

Effects of complex dynamics for the best arm identification in multi-armed bandit problems

金沢大¹, 埼玉大², 東京大³ ○河内 聡志¹, 新山 友暁¹, 内田 淳史², 成瀬 誠³, 砂田 哲¹

Kanazawa Univ.¹, Saitama Univ.², The Univ. of Tokyo³ ○Satoshi Kochi¹, Tomoaki Niiyama¹, Atsushi Uchida², Makoto Naruse³, Satoshi Sunada¹

E-mail: kochisatoshi@stu.kanazawa-u.ac.jp

1. 緒言

近年,機械学習や人工知能の発展に伴い必要となる計算量は爆発的に増大している.しかし,現在のコンピューティング原理では,半導体微細化の限界などの要因により,更なる情報処理量の増大は容易ではない.そのため,自然現象を計算資源とみなす計算原理(自然計算)の研究が活発に行われている.自然計算の応用の一つに WEB 広告割当などに深く関係している意思決定問題がある.そして,意思決定の基本問題は多本腕バンディット問題 (MAB 問題) へと定式化することができる.MAB 問題とはギャンブラーが最適なスロットマシン(SM)を探索する問題であり,最終目標は最大の報酬確率を持つ SM を見つけ出し報酬を最大化することである.MAB 問題を解く上で意思決定の“素早さ”と“正確さ”のトレードオフの解消が必要となる.近年,光を用いた意思決定に関する研究が行われており,特にレーザーカオス波形を用いた意思決定では従来方法に比べ素早くかつ正確な意思決定が可能であることが示されている[1,2,3].

しかし,レーザーカオス波形のどのような性質が意思決定性能に大きな影響を及ぼしているのかは明らかにされていない.そこで本研究では意思決定に大きな影響を及ぼす要因を明らかにすることを目的に,様々な波形の統計的性質と意思決定性能の関係について調査を行った.本予稿では波形の統計的性質の中,特に影響及ぼした確率密度関数について述べる.

2. 意思決定方法

本研究で用いた意思決定方法は文献[1,2]に基づく方法であり,波形信号とある基準値との大小関係に従い SM を選択していく.例えば,時刻 t での波形信号 $s(t)$ と基準値 $x(t)$ を用いて,2台のスロットマシン SM_A と SM_B から報酬確率が大きい SM を選ぶ問題を考える. $s(t) \geq x(t)$ の場合 SM_A を,それ以外の場合 SM_B を選択するように対応づける.そして,選択した SM をプレイし報酬の有無により x を制御する.具体的には報酬を得た場合は次のプレイ時もその SM が選ばれやすくなるように,逆に報酬を得られなかった場合はもう一方の SM が選ばれやすくなるように x を増減させる.2 択以上の選択肢 (2 台以上の SM) がある場合, Fig.1 のような階層構造をとり 2 択選択を繰り返すことで複数の SM からプレイする SM を選択することが可能となる.この意思決定方法では,用いる波形信号によってその選択精度が大きく影響することが指摘されている.例えば, $s(t)$ が高速かつ複雑に変動するレーザーカオス信号を用いると,単純な一様擬似乱数列よりも少ない試行回数で精度よく正しい SM 選択が可能となることが指摘されている[1,2].

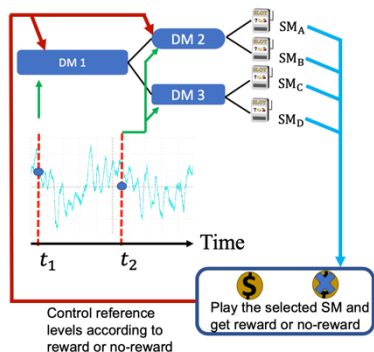


Fig1. Schematic of hierarchical decision-making method.

3. 意思決定シミュレーションとその結果

上述の意思決定方法において,波形 $s(t)$ の影響を詳しく調べるため,本研究では文献[2]において用いたレーザーカオス波形のほかに,レーザーカオス波形をランダムにシャッフルし生成されるサロゲート波形,ロジスティック写像等から生成したカオス波形,更に正規分布に従う擬似乱数列を用意して,その意思決定性能を評価した. Fig.2 は 8 本腕問題にそれぞれの波形を用いた場合の正答率(CDR)の変化である.レーザーカオス波形や乱数列はプレイ回数の増加に伴い CDR が向上しほぼ 1 へ収束する.しかし,カオス写像の収束値は低く,レーザーカオス波形などに比べ意思決定性能が劣るということが分かる.さらに,レーザーカオス波形とサロゲート波形の意思決定性能に大きな差が生じていないことも分かる.これより,意思決定は波形の確率密度関数の形状に大きな影響を受けており,波形による意思決定性能差が生じていると考えられる.

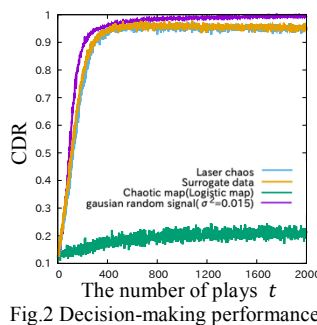


Fig.2 Decision-making performance

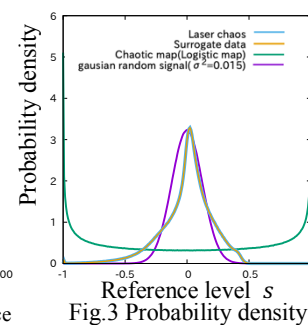


Fig.3 Probability density

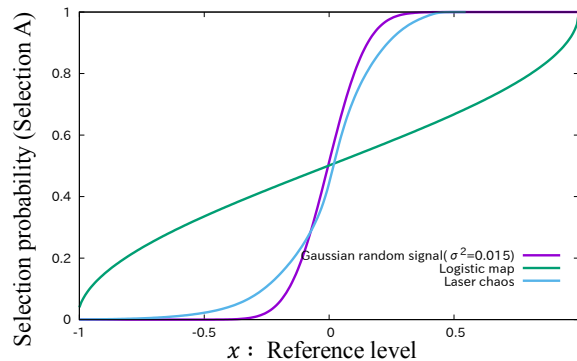


Fig.4 Selection probability(Selection A)

Fig.3 は各波形の確率密度関数である.レーザーカオス波形の確率密度と正規分布は似た形状を持ち,カオス写像の確率密度はそれら 2 つと全く異なる形状を持つ.確率密度関数の形状は意思決定戦略に大きな影響を及ぼす. Fig.4 に,2 択問題における x とある選択肢 A を選ぶ確率との関係を示す.正規分布やレーザーカオスのように $s=0$ で確率密度が鋭いピークを持つ場合,選択確率は x の変化に対してセンシティブになる.つまり,意思決定の初期段階における x の変化に対して,各腕の選択確率は大きく変化をする.逆に,カオス写像のように x の変化に対して選択確率の変化が緩やかな場合,学習を全体を通して探索が行われる.このような戦略の違いにより意思決定性能の差が生じている.各波形の確率密度関数と意思決定戦略や意思決定性能との関係の詳細は当日の発表で議論する.

- [1] M. Naruse et al., Sci. Rep. 7, 8772 (2017).
- [2] M. Naruse et al., Sci. Rep. 8, 10890 (2018)
- [3] R. Homma et al. Sci. Rep. 9, 9429 (2019).