# 電気光遅延システムのリザーバコンピューティングにおける強化学習を用いたモデル選択の実験的評価

**Experiment on adaptive model selection using reinforcement learning in electro-optic delay-based reservoir computing**

埼玉大[1]、東大情理[2]
○水戸 遼平[1], 菅野 円隆[1], 成瀬 誠[2], 内田 淳史[1]
Saitama Univ[1], Univ. Tokyo[2]
Ryohei Mito[1], Kazutaka Kanno[1], Makoto Naruse[2], and Atsushi Uchida[1]
Emails: r.mito.047@ms.saitama-u.ac.jp, auchida@mail.saitama-u.ac.jp

## はじめに:

近年、機械学習（人工知能）における研究開発が盛んに行われている。特に光を用いた人工知能が注目されている。通常の人工知能とは異なり、光の高速性を活かした情報処理が可能となる。また、機械学習手法の一つとして、リザーバコンピューティングが挙げられる[1]。特徴は、学習対象が出力部の重みのみであり、ネットワーク内部のノード間の重みを学習する必要がないため、少量データで学習が可能である。そのため、物理実装に適した情報処理手法であることが知られている。

リザーバコンピューティングによる情報処理において、学習モデルとは異なるモデルが入力された場合に、情報処理精度が悪化することが知られている。この問題に対処する手法として、強化学習を用いたモデル選択が報告されている[2]。しかしながら、実験による調査は未だ行われていない。

そこで本研究では電気光遅延システムのリザーバコンピューティングを用いて、2 種類の異なるモデルから生成された時系列データを入力信号とし、強化学習を用いたモデル選択を実験的に行うことを目的とする。

**方法**: 本研究では、電気光遅延システム[1]と呼ばれる装置を用いる。これは半導体レーザと電気-光変換を行う光学機器の強度変調器と信号を遅延させるための光ファイバをループ状に接続した構成である。また、レスラーモデルおよびローレンツモデルから生成された時系列が、500 点ごとに切り替わる時系列をソフトウェア上で作成して入力信号に用いる。予測対象としては切り替わる時系列を想定し、二つのモデルについては予め学習を行い、学習済みの重みを二つ用意する。信号を入力したときに用意した重みを用いて予測を行い、二つの予測結果$p_R, p_L$を得る。ここでより高い予測精度を出す重みを強化学習[3]により選択する。

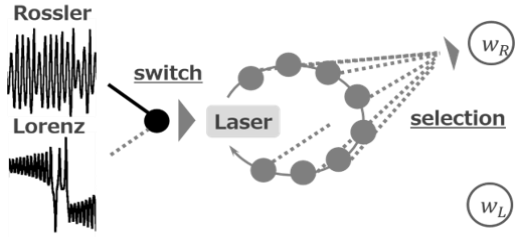


Fig. 1 Scheme of adaptive model selection using time-series prediction.

この手法ではしきい値$T(n)$と$[-1,1]$の一様乱数を使用し、しきい値と乱数の大小関係より、どちらかの予測結果を選択する。次に 1 点先の入力信号がリザーバに入力されたときに、先ほど予測した結果と入力信号を比較することで、誤差$e_R$, $e_L$を以下のように算出する。

$$e_{R,L} = |p_{R,L}(n) - u(n+1)| \quad (1)$$

二つの予測誤差から、しきい値制御変数$TA(n)$は以下のように表される。

$$TA(n+1) = \begin{cases} \alpha TA(n) - 1 & (e_R(n) \le e_L(n)) \\ \alpha TA(n) + 1 & (e_R(n) > e_L(n)) \end{cases} \quad (2)$$

この時、しきい値$T(n)$はしきい値制御変数$TA(n)$から以下のように計算される。

$$T(n+1) = k[TA(n+1)] \quad (3)$$

$T(n)$は一様分布乱数の範囲$[-1,1]$と同じであると仮定する。式(3)の$k$はしきい値段数と呼び、$T(n)$の取りうる値の数が決定される。本研究では$k = 0.1$に設定した。式(2)および式(3)より、しきい値$T(n)$は$e_R(n) \le e_L(n)$の場合に減少し、逆の場合には増加する。すなわち、予測誤差が小さい方が選ばれやすくなるようにしきい値$T(n)$が変化する。しきい値$T(n)$が1または$-1$まで変化したとき、$p_L(n)$または$p_R(n)$のみが選ばれるようになり、適切な重みを用いて予測を行うことができる。

**結果**: Fig. 2 に強化学習によるモデル選択を 100 回行った時の平均正答率（Correct Decsision Rate, CDR）の実験結果を示す。平均正答率は正しいモデル選択を行った割合であり、1 に近い値であるほど正しくモデル選択を行ったことを示す。横軸は時系列のサンプル点数$n$に対応しており、$n = 500, 1000, 1500$においてモデルが切り替わっている。モデルの切り替え後に平均正答率は 0 から 1 まで増加しており、モデル切り替えに追従したモデル選択が正しく行えていることが確認できた。

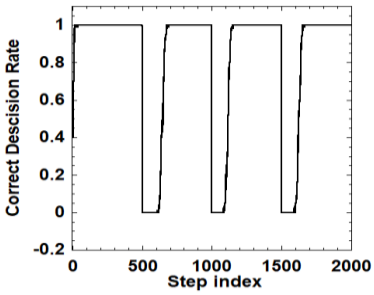


Fig. 2 Result of Correct decision rate.

**まとめ**: 本研究では、電気光遅延システムのリザーバコンピューティングの、2 種類の異なるモデルから生成された時系列データを入力信号とし、強化学習を用いたモデル選択を実験的に行った。学習モデルが変わると情報処理精度が悪化する問題に対し、強化学習によるモデル選択の有用性を示した。

**参考文献**

[1] L. Larger, et al., Opt. Express, 20, 3 (2012).
[2] K. Kanno, et al., arXiv:2004.12575 (2020).
[3] M. Naruse, et al., Sci. Rep.,.7, 8772 (2017).