

配布先: 総務省記者クラブ、テレコム記者会、  
文部科学記者会、科学記者会、  
埼玉県政記者クラブ

プレスリリース  
2017年8月22日  
国立研究開発法人情報通信研究機構  
国立大学法人埼玉大学  
慶應義塾大学 SFC 研究所

## AI の重要課題である強化学習をレーザーカオスを用いて超高速に実現 ～周波数の割当てなどで「瞬時の適応」を可能に～

### 【ポイント】

- レーザーから生じる光カオスを用い、AI の基本的な問題の一つの「強化学習」を超高速に実現。
- 光の高速性を生かし、自然界の物理現象を用いて瞬時の“意思決定”を実現。優れた性能を確認。
- 周波数の瞬時の割当てなど、AI や IoT の基盤技術としての貢献に期待。

国立研究開発法人情報通信研究機構 (NICT、理事長: 徳田 英幸) の成瀬 誠主任研究員、国立大学法人埼玉大学 (学長: 山口 宏樹) 大学院理工学研究科 内田 淳史教授、慶應義塾大学大学院政策・メディア研究科 (研究科委員長: 清木 康) 金 成主特任准教授らは、半導体レーザーから生じる光カオス<sup>1</sup> (以下レーザーカオス) を用いて、適応速度 1GHz (ギガヘルツ: 1 秒間に 10 億回) を実現する超高速フォトンクスを応用した強化学習に世界で初めて成功しました。

NICT らは、光の高速性に着目し、半導体レーザーにおいて生じるカオス現象が生み出す乱雑な信号と、独自に開発した強化学習方式を組み合わせることで、「当たり確率の未知な 2 台のスロットマシンから当たり確率の高い台を選ぶ問題」(2 本腕バンディット問題) を、光の極限性能を生かし、高速に、物理的に解決をすることに成功しました。レーザーカオス現象の超高速性により、情報が入力されてから出力されるまでの時間(レイテンシ<sup>2</sup>) が 1ns (ナノ秒: 10 億分の 1 秒) という高速な意思決定が確認され、また、仮想的に生成した高速な擬似乱数<sup>3</sup> (カラーノイズ) に比べても優れた性能を示すことが確認されました。

より高速なコンピューティングのため計算資源を瞬時に調停するアービトレーション<sup>4</sup> や無線通信における周波数の瞬時的な割当てなど、AI や IoT の基盤技術として大きく貢献することが期待されます。なお、本研究成果は、「Scientific Reports」に日本時間 8 月 18 日 (金) 18 時に掲載されました。

### 【背景】

現代の人工知能 (AI) を支える機械学習では、画像認識などに優れた能力を発揮する「深層学習」と呼ばれる技術に並び、「強化学習」と呼ばれる技術が重要です。強化学習とは、未知な環境で試行錯誤をしながら学習を行う方法であり、強化学習の中心となる課題に、多数のスロットマシンが並んだカジノで、儲けを最大にするにはどのようにするとよいか? という問題があります。儲けを最大化するには、「当たり台」を見つけるための試行錯誤＝「試し打ち」をしなければなりません。しかし、過剰な試し打ちは損失になることがあり、当たり台が時々刻々と変わることもあり得ます。逆に、

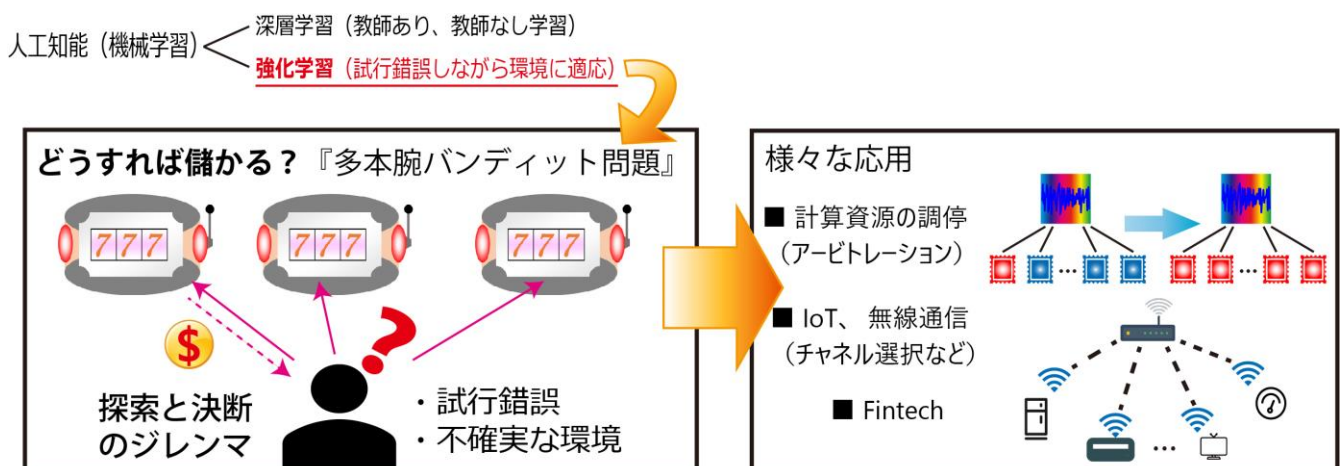


図 1: 研究の背景

早々に試し打ちを打ち切ってしまうと、肝心の当たり台を見逃しているかもしれません。このように、「探索」と「決断」に難しいジレンマが存在しています。この問題は「多本腕バンディット問題<sup>5)</sup>」として知られ、ワイヤレス通信における周波数の割当て、データセンターでの計算資源の割当て、ロボット制御、Web 広告など、非常に重要な応用の基礎になっていることから、熱心な研究が行われています(図 1)。

この問題の解決には、当たり台の探索において、「ランダムに」台を選ぶことが不可欠です。例えば、プレイの序盤では「ランダムに」選び、時間の経過とともに、それまでの勝率を参考に選んでいく方法などが知られています。

これらの従来手法では、ランダムに台を選ぶために、コンピュータ上で生成した「擬似乱数」と呼ばれる数を用いています。しかし、擬似乱数は、コンピュータ上で計算手順(アルゴリズム)に基づいて作られるため、高速に生成することは困難です。また、乱数の「質」にも限界があります。一方で、昨今発展の著しい AI に見られるように、コンピューティングを、より高速に、効率的に行うことは重要であり、ロボティクスやITを使った新たな金融サービスである Fintech (フィンテック)などの新たな応用の出現を受けて、従来の強化学習の限界を克服する高速な原理と技術が強く望まれています。

NICT らは、コンピュータ上のアルゴリズムを用いるのではなく、「自然界の物理現象を直接用いた強化学習」を世界に先駆けて提案し、これまでに、単一光子を用いた方式の実現に成功していました。

### 【今回の成果】

今回の成果は、光の高速性というメリットを生かし、新たに、レーザーを用いた時に生じるカオスを超高速な強化学習に応用できることを世界に先駆けて実証したものです。

本研究では、半導体レーザーを用いた時に生じる光の「カオス現象」を用いています。半導体レーザーから出射する光を、鏡で反射させてレーザーに戻すと、レーザーの動作が不安定化しカオスが生じます。このような不安定化を避けるために、通常、レーザーでは光の逆戻りが起きないように工夫がなされています。しかし、近年では、この現象を、逆に、積極的に生かして、毎秒 1 兆ビットを超える超高速な物理乱数生成技術などに展開されています。本研究では、このレーザーカオスを用いて、2 台のスロットマシン(以後、「スロットマシン 1」「スロットマシン 2」)から、「当たり台」(=当たり確率が高い台)を選ぶ問題の超高速な解決を実現しました。半導体レーザーから生成したレーザーカオス光を高速にサンプリングし、「閾値」との大小判定のみで意思決定を行います。ここで、閾値を「過去の戦歴」に基づいて上げ下げすることが重要であり、以前の成果でも用いた「綱引き原理」と呼ばれている独自の方式を応用しています(図 2 上)。

計測した信号レベルが「閾値より大きいとき」には「スロットマシン 1」を選択すると意思決定し、「小さいとき」には「スロットマシン 2」を選択すると意思決定します。閾値が十分大きいときは、計測される信号レベルは閾値より小さくなる場合が多くなり「スロットマシン 2」を選択するケースがほとんどとなります。ところが、カオスの乱雑さのために、時として、

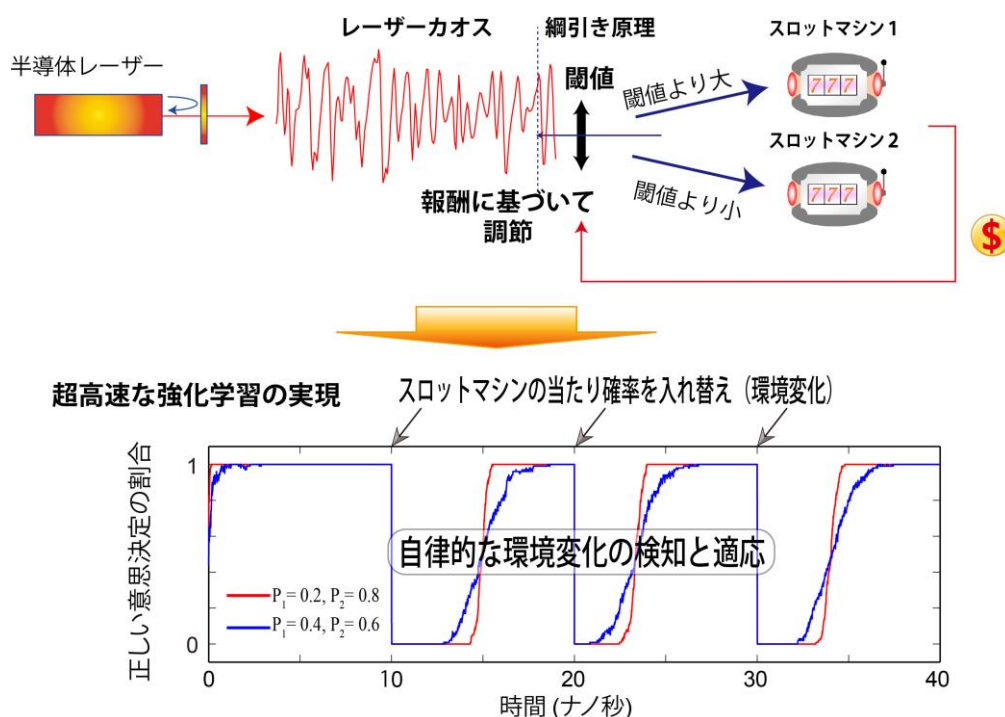


図 2: レーザーカオスを用いたシステムと超高速な強化学習の実現

計測する信号レベルが閾値よりも大きくなることもあり、逆のロットマシン「ロットマシン 1」を選択することも生じま  
す。

このような、カオスから生まれる特長が、「当たり台を速やかに発見する」ことに生かされた結果、適応速度 1 GHz  
を実現する超高速フォトニクスを応用した超高速な強化学習が実証されました。また、仮想的に生成した高速な擬似乱  
数(カラーノイズ)に比べても優れた性能を示すことも確認されました。

## 【今後の展望】

強化学習は、これまで計算機上のアルゴリズムとして実現されてきましたが、高速化には大きな壁がありました。本  
研究によって、レーザーカオスが強化学習という人工知能分野に貢献できると実証されたことで、様々な展開が期待さ  
れます。まず、レーザーカオスを用いた方式は、広帯域性という光のメリットを生かしており、従来技術では不可能な高  
速性が実現できます。また、半導体レーザーなど成熟したフォトニクス技術を用いているため、安定性や集積性に優れ  
たシステムの実現が可能です。応用面では、高速・低レイテンシの実現が強く期待されている、システムのアービトレー  
ション(調停)応用や超高速取引での意思決定としてフィンテックなどへの展開を検討する予定です。

## <掲載論文>

掲載誌: *Scientific Reports*

DOI: 10.1038/s41598-017-08585-8

URL: <https://www.nature.com/articles/s41598-017-08585-8>

掲載論文名: Ultrafast photonic reinforcement learning based on laser chaos

著者名: Makoto Naruse<sup>1</sup>, Yuta Terashima<sup>2</sup>, Atsushi Uchida<sup>2</sup> & Song-Ju Kim<sup>3</sup>

1 Strategic Planning Department, National Institute of Information and Communications Technology, 4-2-1  
Nukui-kita, Koganei, Tokyo 184-8795, Japan

2 Department of Information and Computer Sciences, Saitama University, 255 Shimo-Okubo, Sakura-ku,  
Saitama, Saitama 338-8570, Japan

3 Graduate School of Media and Governance, Keio University, 5322 Endo, Fujisawa, Kanagawa 252-0882,  
Japan

本研究の一部は日本学術振興会科学研究費補助金基盤研究(A)『ナノ光学とレーザーカオスを用いた超高集積・  
超高速意思決定の創製(JP7H01277)』(研究代表者: 成瀬 誠(NICT)、研究分担者: 赤羽浩一(NICT)、堀裕和、内  
山和治(山梨大学)、金成主(慶應義塾大学)、西郷甲矢人(長浜バイオ大学)、岡村和弥(名古屋大学))、同挑戦的萌  
芽研究『単一光子を用いた意思決定機能構築の研究(JP15K13387)』(研究代表者: 成瀬 誠(NICT)、研究分担者:  
金成主(慶應義塾大学)、高橋泰城(北海道大学))、同基盤研究(B)『複雑系フォトニクスに基づく超高速物理乱数生  
成の高機能化(JP16H03878)』(研究代表者: 内田淳史(埼玉大学)、研究分担者: 菅野円隆(福岡大学))の支援を  
受けています。

## <参考文献>

プレスリリース「単一光子を用いた意思決定の実証に成功」(2015年8月19日、国立研究開発法人情報通信研究機  
構、国立研究開発法人物質・材料研究機構、フランス国立科学研究センター Institut NEEL/ジョセフ・フーリエ大学)  
<http://www.nict.go.jp/press/2015/08/19-1.html>

MIT Technology Review “First Demonstration of Photonic Intelligence”

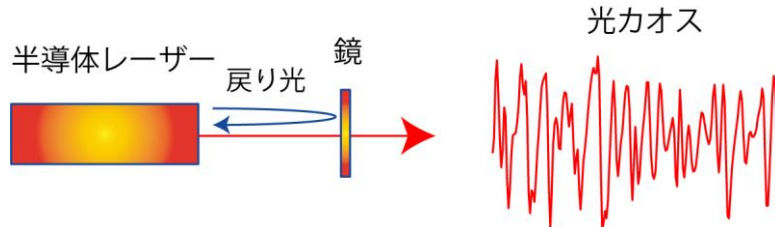
<https://www.technologyreview.com/s/541196/first-demonstration-of-photonic-intelligence/>

## <用語解説>

### \*1 光カオス

カオスとは、初期状態における僅かな違いが、時間経過とともに爆発的に増大する現象のこと。初期値は無限の精度での情報が必要になるため、通常のコンピューターでは再現することは厳密にはできない。

半導体レーザーから出射する光の一部を、鏡で反射させ一定の時間遅延の後にレーザーに戻すと、レーザーの非線形性によって発振がカオスとなる。これは半導体レーザーの戻り光カオスなどと呼ばれている。



### \*2 レイテンシ

情報が入力されてから出力を得るまでに必要な時間。「瞬時」の処理とはレイテンシが短い処理のことを指す。

### \*3 擬似乱数

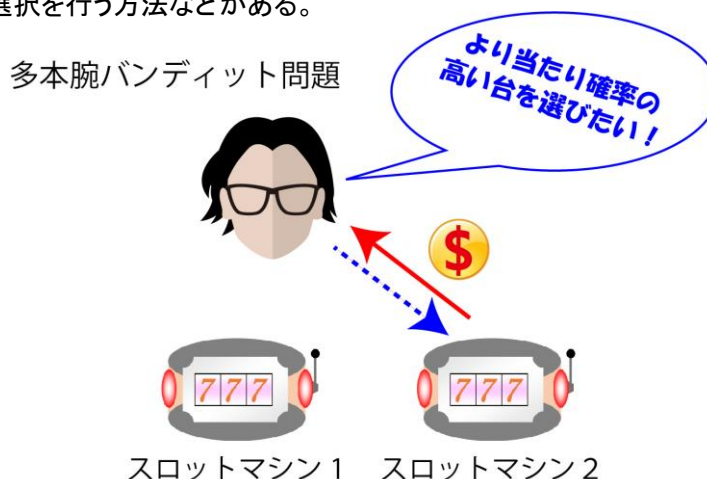
コンピューターにおいて乱雑な信号を生成するには、適当な計算手順(アルゴリズム)を用いる必要がある。このことから、規則性や再現性が全く存在しない乱数を生成することは不可能であり、生成される乱雑な信号は、「擬似的な乱数、すなわち「擬似乱数」と呼ばれる。コンピューターでは無限の桁数を扱うことができないため、得られる乱数の乱雑さには限界がある。また、擬似乱数の生成速度は、コンピューターの計算速度がボトルネックとなるため、高速性に限界がある。

### \*4 アービトレーション

有限の資源に対する複数の要求に対して、それぞれの要求を満たすように調停すること。CPU や GPU などの複数のプロセッサ(コア)を有するシステムでの解探索、並列ロボット学習、IoT など、個々のデータ生成レートは小さなシステムでも、全体としては膨大となり得るデータや計算を、有限な資源で処理する必要があるシステムでは、資源の調停を瞬時に行うことが要求される。

### \*5 多本腕バンディット問題

多本腕バンディット問題の解決では、「当たり台」を知るための「試し打ち」(exploration)と、有利と考える台を集中的に引いて稼ぎを増やすこと(exploitation)のジレンマ(exploration-exploitation dilemma)を解くことが鍵になる。 $\epsilon$ -greedy 法や softmax 法のように、「序盤」ではランダムに台を選び、時間の経過とともにランダムな選択を低減させ、戦歴に基づいた選択を行う方法などがある。



### \*6 擬似周期信号 (補足資料参照)

正弦波のように周期的な挙動を示すが、時に、正弦波からずれた挙動を伴う信号。

### \*7 自己相関関数 (補足資料参照)

ある信号系列に対して、その信号系列を一定時刻ずらした系列と元の信号系列がどの程度同じかを 0 から 1 の間の数値として表現したもの。1 であれば「全く同じ」であることを意味する。自己相関関数が「負の値」を示すとは、元の信号波形と時刻をずらした信号波形が「逆転」していることを意味する。

## ＜本実験の詳細説明＞

本実験では、レーザーカオスを毎秒最大 1000 億個でサンプリングし、データ処理はオフラインで実装しています。システムの全体構造の概略を図 3a、光信号の時系列の例を図 3b(i)に示します。(なお、図 3b(ii)は同一の半導体レーザーの設定を微調整し、光信号が擬似周期信号<sup>6</sup>となったときを示しています。)

代表的な結果を以下に示します。

2 台のslotマシンの当たり確率を  $\{0.2, 0.8\}$ ,  $\{0.4, 0.6\}$  と設定したとき、当たり確率の大きな台を選択した割合 (=正しい意思決定の割合)を図 3c に示します。10 ナノ秒(1 億分の 1 秒、10 ns)ごとに、slotマシンの当たり確率を入れ替えています。このため、入れ替えた直後では意思決定の性能が一端劣化していますが、その後、自律的に回復しています。これは、「不確実な環境における適応」と呼ばれ、強化学習で求められる重要な性質の一つです。

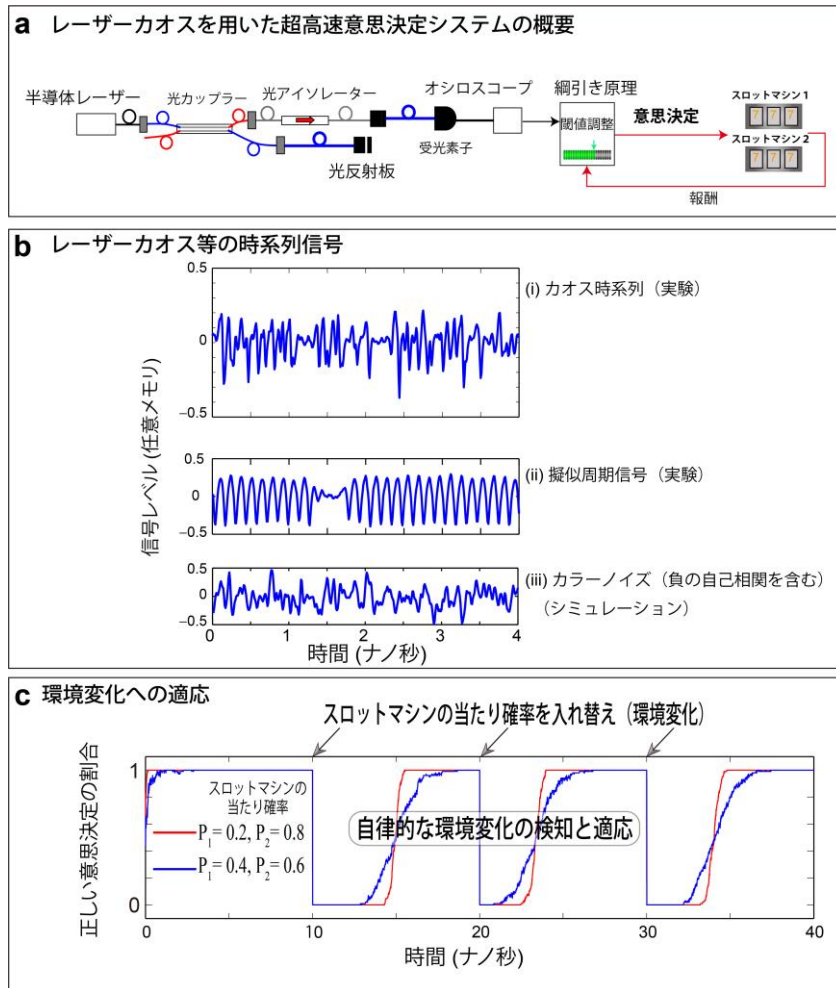


図 3: a レーザーカオスを用いた超高速意思決定システムの概要、  
 b レーザーカオスなどの時系列信号、c 環境変化への適応

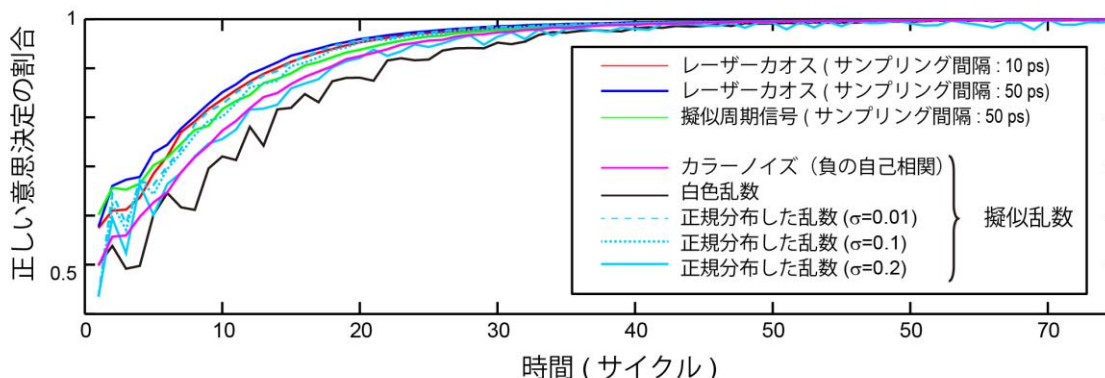
図 4a は事前の知識が全くない状態からの強化学習の実現例を示しています。レーザーカオスの信号を 50 ピコ秒 (1 ピコ秒は 1 兆分の 1 秒、50 ps) 間隔 (毎秒 200 億個) でサンプリングしたときに、最も適応性に優れた性能が得られています。「正しい意思決定」がより速く上昇しています。およそ 20 回の試行で正解率 9 割以上が得られており、事前知識ゼロの状態から約 1 ナノ秒 (10 億分の 1 秒、1 ns) という非常に短い時間 (レイテンシ) で強化学習が実現しています。

さらに、本実験に用いたレーザーカオスの波形を特徴づける量である「自己相関関数<sup>7</sup>」の値は、強化学習の性能が最も高くなる時間ずれ (サンプリング間隔 (50 ピコ秒)) のときに、負の最大値を示しています (図 4b 青)。このことは、「負の自己相関」と「強化学習の性能」の関係を示唆しています。ところが、負の自己相関の最大値がより大きな擬似周期信号 (図 4b 緑) では、強化学習の性能はレーザーカオスよりも劣っています (図 4a)。また、自己相関に「負の相関」を含むように、コンピューター上で生成させた乱雑な信号 (図 4b ピンク、「カラーノイズ」と呼ばれています) と比べて

も、レーザーカオスの方が優れた性能を実現しています(図 4a)。

これらのことから、レーザーカオスでは、従来の擬似乱数では生成が不可能な超高速な動作を実現できるのはもちろんのこと、レーザーカオスの質的な特徴が、優れた強化学習性能をもたらしていることが分かります。

**a** 事前の知識が全くない状態からの強化学習の実現例  
(レーザーカオス、擬似乱数等の比較)



**b** 実験に用いた信号の自己相関関数

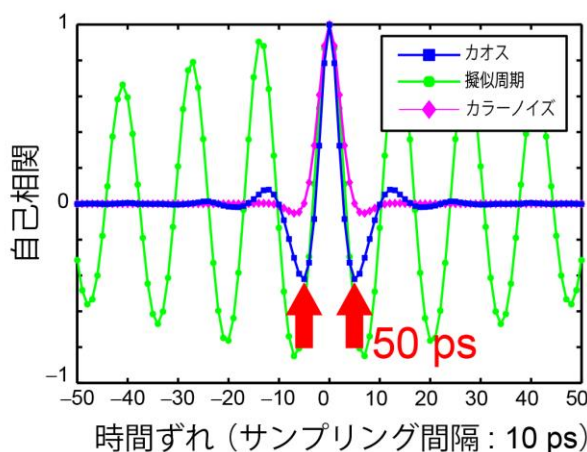


図 4: a 事前の知識が全くない状態からの強化学習の実現例、  
b 実験に用いた信号の自己相関関数

< 本件に関する問い合わせ先 >

国立研究開発法人情報通信研究機構  
ネットワークシステム研究所  
(現: 経営企画部 企画戦略室)  
成瀬 誠  
Tel: 042-327-6794  
E-mail: naruse@nict.go.jp

国立大学法人埼玉大学  
大学院 理工学研究科 数理電子情報部門 教授  
内田 淳史  
Tel: 048-858-3490  
E-mail: auchida@mail.saitama-u.ac.jp

慶應義塾大学  
大学院 政策・メディア研究科 特任准教授  
金 成主  
E-mail: songju@sfc.keio.ac.jp

< 広報 >

国立研究開発法人情報通信研究機構  
広報部 報道室  
廣田 幸子  
Tel: 042-327-6923 Fax: 042-327-7587  
E-mail: publicity@nict.go.jp

国立大学法人 埼玉大学  
広報渉外室 広報係  
二川目 一、武藤 裕  
Tel: 048-858-3932 Fax: 048-858-9057  
E-Mail: koho@gr.saitama-u.ac.jp

慶應義塾大学  
湘南藤沢事務室 学術研究支援担当  
河越 英代  
Tel: 0466-49-3436 Fax: 0466-49-3594  
E-mail: kri-pr@sfc.keio.ac.jp