

本件は下記の解禁日時でお願いいたします。  
 特許出願案件ですので厳守をお願いいたします。  
 2020年(令和2年)7月13日(月)14:00(日本時間)

2020年(令和2年)7月10日

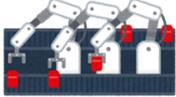
## 意思決定支援システムが示す選択肢の正しさと計算スピードを両立する手法を開発 ～工業製品の品質確認、自動運転、マーケット投資などの戦略の高速計算への道を拓く～

情報・システム研究機構 国立情報学研究所(エヌアイアイ N I I、所長:喜連川 優、東京都千代田区)のアーキテクチャ科学研究系特任研究員 滝坂 透(たきさか・とおる)、同研究系准教授 蓮尾 一郎(はすお・いちろう)らの研究チームは、科学技術振興機構(ジェイエスティー J S T、理事長:濱口 道成、東京都千代田区)の戦略的創造研究推進事業(エラト ERATO)蓮尾メタ数理システムデザインプロジェクト<sup>(\*)1</sup>(ERATO MMSD、研究総括:NIIアーキテクチャ科学研究系准教授・蓮尾 一郎)のもと、ゴールの達成確率を最大化する戦略を精度保証しながら高速に計算する手法(以下、改良 BVI 法)を開発しました。この手法は、工業製品の品質確認や自動運転の制御、マーケット投資戦略の策定など、幅広い分野の意思決定を支援するシステムに応用可能で、システムが示す選択肢の正しさ(精度)を保証しつつ高速に計算して結果を提示できます(図1)。

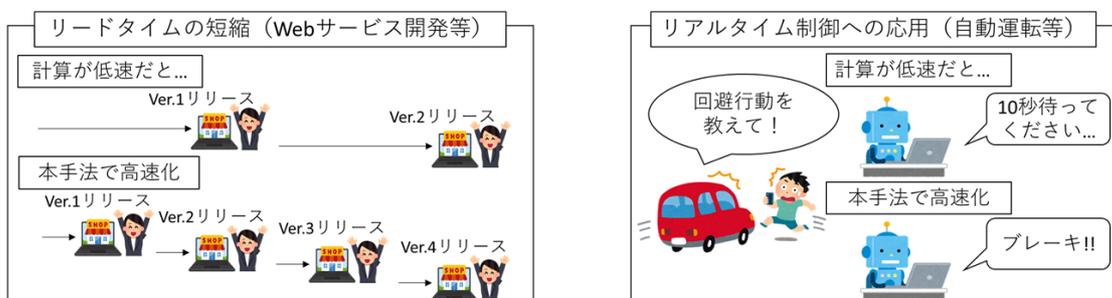
改良 BVI 法では、従来手法が精度保証のために時間をかけて行っていた計算の省略に成功しました。従来の手法<sup>(\*)2</sup>との性能比較実験では、バーチャルマシンへのウェブ・アプリデプロイメント問題<sup>(\*)3</sup>の例題で100倍~数千倍の高速化、マーケット投資戦略の策定問題<sup>(\*)4</sup>の例題で3倍弱~5倍弱の高速化が見られました。これにより、工業製品やウェブ・サービスの開発時のリードタイム短縮や、自動運転や投資戦略判断などの分野におけるリアルタイム制御への応用が期待されます。

本研究成果は、2020年7月19日(日)からオンライン開催される、形式検証<sup>(\*)5</sup>についてのフラッグシップ国際会議 CAV 2020<sup>(\*)6</sup>の本会議で発表されます。

### システムの最適制御問題

課題	工業製品の品質確認 	自動運転 	投資戦略の策定 
	自動生産ラインのロボットにタスクを完了させるには?	事故を起こさず目的地に到達するための運転方法は?	ある金融商品を購入・売却する最適なタイミングは?
本手法の応答例	「このように制御すれば99%以上の確率でタスクを完了できます。一方、どのように制御しても完了確率が99.9%を超えることはありません」	「このように運転すれば99.9%以上の確率で事故なく目的地に到達できます。一方、どのように制御しても無事故での到達確率が99.9%を超えることはありません」	「このタイミングで購入・売却すれば、期待値で100万円以上の利益を上げられます。一方、いつ購入・売却しても期待利益は105万円を越えません」

### 本手法: 上記の応答を既存手法より高速に計算



＜図1＞今回開発した手法(改良 BVI 法)では、さまざまな意思決定を支援するシステムの最適制御問題において、応答例の精度を保証しながら高速計算を実現できる。

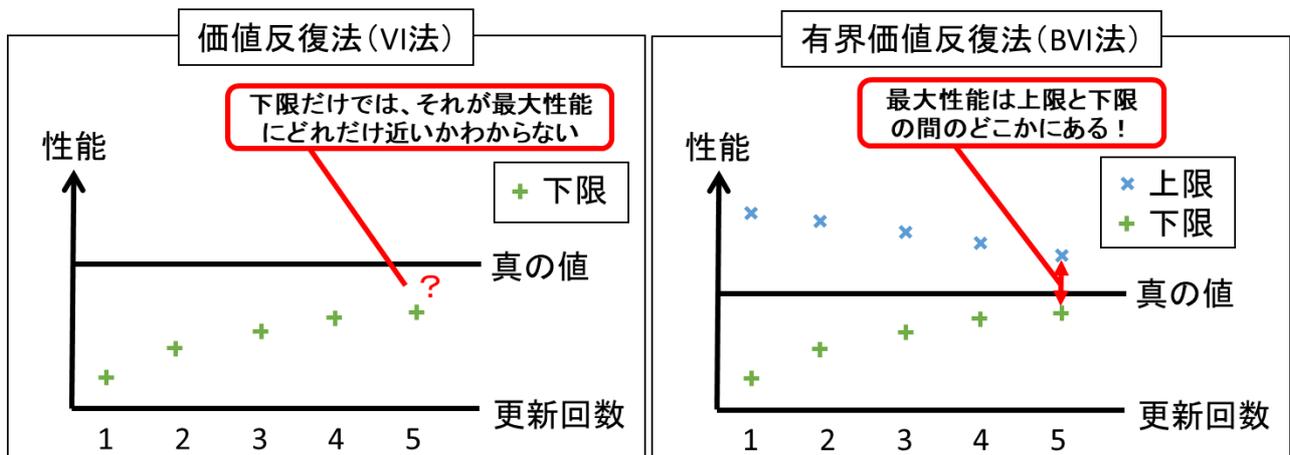
## 【背景】

「自動生産ラインのロボットに、流れてくるさまざまな工業部品をミスなく仕分けさせるにはどう制御すれば良いか?」「さまざまな動き方をする歩行者がいる環境で、事故を起こさず自動車を自動運転するには?」「ある金融商品を購入・売却して利益を上げるのに最適なタイミングは?」など、実世界には意思決定を伴う問題が多く存在します。これらの問題で知りたいのは、最適な戦略は何か、つまり「いつ、どのような選択をすれば目的を最も良く達成できるか?」という点です。このような問題は確率的ゲーム (Stochastic Game) としてモデル化できるため、計算により最適戦略を自動生成する技術が盛んに研究されています。また、囲碁などのゲームや投資判断では、未知の新しい戦略を発見することもあるため、確率的ゲームは近年大きく注目されています。

この確率的ゲームで最適戦略を計算で近似的に求める良く知られた手法の一つに「価値反復法 (Value Iteration : VI 法)」があります。これは、情報収集を繰り返すことで戦略を徐々により良く更新していく手法で、時間をかけるほど最適に近い戦略を見つけやすくなります。

しかし、VI 法には「ある一定の時間をかけて計算された戦略は、最適にどのくらい近いのか?」がわからない、つまり精度保証ができないという欠点があります。精度保証ができないと、例えば「99%の確率で安全に自動運転されることを確かめたかったが、計算された戦略 (= 運転方法) では 90%の安全性しか保証されなかった」といった場合に、もっと時間をかければ安全な運転方法を生成できるのか、それとも仮定した状況のもとで安全な運転がそもそも不可能なのかが判断ができません。

このような問題に対応するため、近年研究が進んでいるのが「有界価値反復法 (Bounded Value Iteration : BVI 法)」です。通常の VI 法では「この戦略に従えば、少なくともこれだけの性能が達成できる」という、性能の下限が保証された戦略を計算できます。対して、BVI 法では性能の下限に加えて、「どのような戦略に従ってもこれ以上の性能は達成できない」という性能の上限を同時に計算します。この上限と下限によって真の最大性能値を「はさむ」ことで、計算された戦略がどの程度最適に近いのかを評価できます (図 2)。

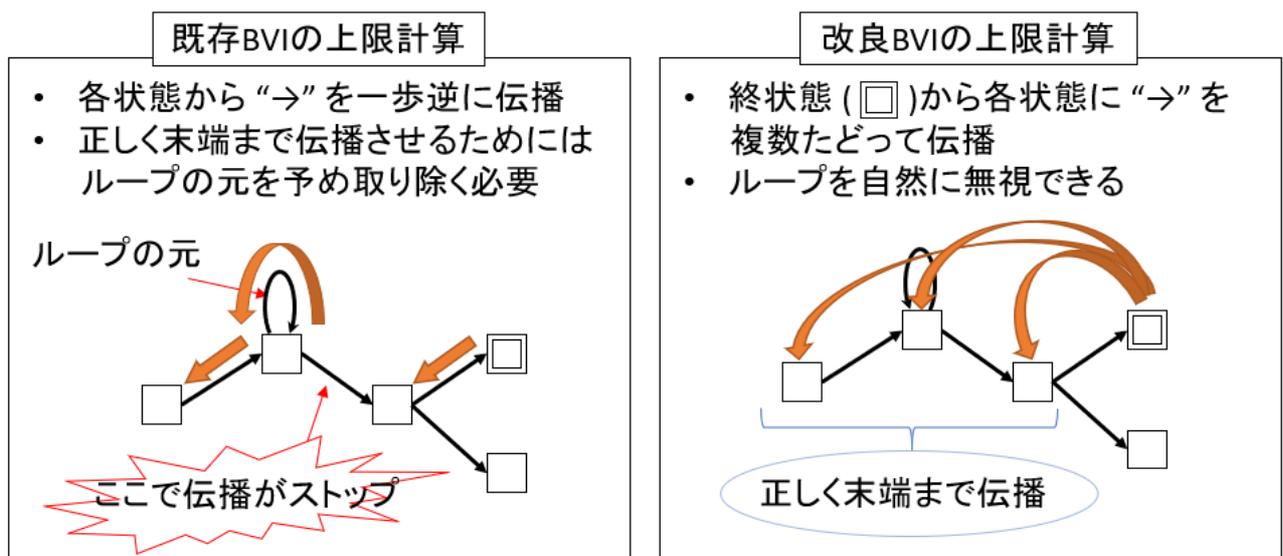


<図 2> 価値反復法 (VI 法) では一方向から真の値に近づくが、有界価値反復法 (BVI 法) では二方向から真の値を「はさむ」形で真の値に近づくため、得られた結果の精度を保証できることになる。

しかし、BVI 法はその計算実行に時間がかかるという課題がありました。これは性能上限の計算が、一般に性能下限の計算よりも難しいためです。単純に下限と同様のやり方で上限の計算を行うと、上限が「はさむ」べき最大値に十分に近づく前に計算がループ（同様計算の繰り返し）に入ってしまう、必要な精度保証が得られないことがあります。従来の BVI 法では、システムモデルに含まれるループの原因を特定し、これを取り除くことによってこの問題を解決してきましたが、この特定にかかる時間が全体の計算時間短縮のボトルネックとなっていました。

### 【研究手法・成果】

本研究では、高速に実行可能な新しい上限の計算手法（改良 BVI 法）を開発しました。従来の BVI 法では①確率的ゲームの各状態で選択可能な各行動をとった後の上限の値を調べ、②これらの上限値を行動前の状態に「伝播」させる、という計算の繰り返しによってシステムモデル全体の上限を計算しています。対して新手法（改良 BVI 法）では、この上限値を調べて伝播させる範囲を各行動（自動運転を例にすると、「加速」、「減速」、「停止」、「左右に曲がる」など）ではなく、終状態<sup>(\*)</sup>に到達するまでの各行動列（「出発地点で加速し、カーブで左に曲がり、その後目的地で停止する」など）に拡大することで、ループの原因をうまく無視しつつモデル全体の上限を更新します（図 3）。



<図 3> 既存の BVI 法では計算した上限値を 1 ステップ前の計算に伝播させる際に、計算がループに入ってしまうと伝播が途中でストップする問題があった。改良 BVI 法では伝播を各行動列に拡大することで、ループの原因をうまく無視してモデル全体の上限値の計算を完了することができる。

既存のいくつかのアルゴリズムと比較して、改良 BVI 法は多くの場合で大幅な高速化を実現します（最悪のケースでも同程度の速度で計算できます）。例えば従来手法<sup>(\*)</sup>との性能比較実験では、バーチャルマシンへのウェブ・アプリデプロイメント問題の例題で 100 倍～数千倍の高速化、マーケット投資戦略の策定問題の例題で 3 倍弱～5 倍弱の高速化が見られました。

BVI 法の応用先として、ウェブ・アプリデプロイメント問題のように、製品やプログラムの品質確認のために使用するケースが想定されています。このような事例では、今回開発した改良 BVI 法による高

速化でリードタイム（計算結果を得るために必要な時間の合計）を短縮できます。また、システムが実際に動いているリアルタイムでの最適戦略生成が必要な自動運転や投資判断など、従来の BVI 法の計算速度では対応できない状況において、改良 BVI 法による高速化は新たなシステム設計の可能性を拓くと期待されます。

### 【今後の展望】

改良 BVI 法は、システムの振る舞いのルールをすべて書き下すことが可能である「ホワイトボックスモデル」を仮定して開発されました。しかし、BVI 法の背景にある VI 法はもともと機械学習において基礎的な技術として用いられてきたもので、その機械学習の大きな強みはシステムの振る舞いのルールがユーザーにはわからない「ブラックボックスモデル」のシステムを解析できる点にあります。今後、ブラックボックスモデルのシステムの解析を含めて、機械学習の諸問題への新たなアプローチとして改良 BVI 法を展開していきます。

また現時点では、改良 BVI 法は「単一のゴールへの到達確率を最大化する」という問題しか扱えませんが、一方、BVI 法は本来「道路からなるべく脱線することなく自動車で目的地に到達したい、途中で給油できるとなおよい」といった、ゴールまでの振る舞いの良し悪しを含めた最適戦略の近似が可能であり、今後はこのような一般的な問題へ改良 BVI 法の適用範囲を広げていきます。

滝坂 透 特任研究員からのコメント：

「囲碁などのゲームにおいて、プレイヤーが取り得る戦略には膨大な数があり、すべてを比較して最適な戦略を見つけるというのは現実的ではありません。価値反復法（VI 法）は、こういった巨大な問題を「ある状況である行動をしたら、状況はどう変わるか？」という 1 回の行動の問題に分解する手法といえます。新手法（改良 BVI 法）のアイデアは「1 回の行動ではなく行動列を評価する」という、一見すると VI 法の単純化のアイデアに逆行するものですが、適切なシステムモデルの解釈手法により高速な計算実行が可能となりました。機械学習をはじめとしたさまざまな分野へと本手法を展開していきたいと考えています。」

### 【研究プロジェクトについて】

本研究は科学技術振興機構 戦略的創造研究推進事業 ERATO 蓮尾メタ数理システムデザインプロジェクト（JPMJER1603）の一環で行われました。

### 【論文タイトルと著者】

タイトル：Widest Paths and Global Propagation in Bounded Value Iteration for Stochastic Games

著者：Kittiphon Phalakarn, Toru Takisaka, Thomas Haas, and Ichiro Hasuo

発表会議：International Conference on Computer-Aided Verification (CAV 2020)

発表日：2020年7月21日（火）からの本会議で口頭発表予定（オンライン開催）

〈メディアの皆様からのお問い合わせ先〉

**大学共同利用機関法人 情報・システム研究機構 国立情報学研究所**  
総務部企画課 広報チーム  
TEL : 03-4212-2164 E-mail : [media@nii.ac.jp](mailto:media@nii.ac.jp)

**国立研究開発法人 科学技術振興機構 (JST)**  
広報課  
TEL : 03-5214-8404 E-mail : [jstkoho@jst.go.jp](mailto:jstkoho@jst.go.jp)

〈JST の事業に関すること〉

**国立研究開発法人 科学技術振興機構 (JST)**  
研究プロジェクト推進部 内田信裕  
TEL : 03-3512-3528 E-mail : [eratowww@jst.go.jp](mailto:eratowww@jst.go.jp)

- 
- (\*1) ERATO 蓮尾メタ数理システムデザインプロジェクト : 国立研究開発法人 科学技術振興機構 (JST) の「戦略的創造研究推進事業 ERATO」に採択されている研究プロジェクトで、Society 5.0 の大きな柱となる CPS の品質保証手法の学術的研究を推進している。特に、CPS の典型例の一つとして注目される自動運転システムを重点応用対象として、その信頼性保証を支えるモデリング手法・形式検証手法・テスト手法、さらにこれらを含む実用的な V&V 技術の研究開発に取り組んでいる。このような大きなチャレンジでは、ソフトウェア・制御・AI といった多様な学術分野の協働が必要となるため、学術分野融合の基礎となる数理的 (メタ) 理論も重視して研究を推進する。略称は ERATO MMSD。プロジェクト詳細は <https://www.jst.go.jp/erato/hasuo/ja/> 参照。
- (\*2) 比較対象の従来手法は以下を参照 :  
Kelmendi, E., Kramer, J., Kretinsky, J., Weininger, M.: Value iteration for simple stochastic games: stopping criterion and learning algorithm. In: International Conference on Computer Aided Verification. pp. 623-642. Springer (2018)
- (\*3) バーチャルマシンへのウェブ・アプリデプロイメント問題 : サーバダウンなどの可能性が考慮された状況で、最大確率でウェブ・アプリのデプロイ (反映) を完了させる戦略を求める問題。この例でのデプロイとは、ウェブ・アプリをバーチャルマシン上に展開して外部からアクセス・使用できるようにすることを意味する。
- (\*4) 投資戦略の策定問題 : 所有する金融商品の期待売却益を最大化する戦略を求める問題。
- (\*5) 形式検証 : ハードウェアやソフトウェアの挙動を数学的な表現に書き下して分析することにより、これらが仕様通りに動作する (またはしない) ことを証明する手法の総称。
- (\*6) CAV 2020 : International Conference on Computer-Aided Verification。CORE と呼ばれる計算機科学系の国際会議ランキングにて最高ランク (A\*、トップ 4%)。
- (\*7) システムモデルに含まれる、タスクの完了を示す状態。