

2024年10月15日

報道機関 各位

国立大学法人東北大学

「好奇心」によって性能を改善した対話型 AI を開発 - 次世代タスク指向対話システムの高度化に貢献 -

【発表のポイント】

- チケット予約など、人から頼まれたタスクを遂行するためのシステムをより賢くする対話型人工知能 (AI) ^(注1) の新しい手法を開発しました。
- AI 自体がシステムの学習時に報酬を受け取り、新しい状況を探求するための「好奇心」を持つことで、より効率的な対話が可能になります。
- 複数の AI 学習者 (エージェント) を競合させ、最も優れたエージェントを選び出すことで、対話の成功率を向上させました。

【概要】

ChatGPT などに代表される、自然言語を使う対話型 AI 技術は、ここ数年で劇的に発展し、チャットボットをはじめとした様々なシステムに活用されるなど、日常のタスクの自動化や効率化が進んでいます。しかし、現状の対話型 AI は、タスク達成に不要な質問をしてしまうなど、必ずしもタスク達成の効率が十分ではありませんでした。

東北大学大学院工学研究科の牛雪澄大学院生と伊藤彰則教授らの研究チームは、映画のチケット予約などのタスクを効率的に行うための新しい対話システムを開発しました。このシステムは、AI がユーザと対話しながら、適切な行動を学習する技術を使っています。

対話システムが適切に行動するためのシステム開発手法に、将棋や囲碁の AI にも使われている強化学習 ^(注2) があります。今回の研究では、強化学習に「好奇心駆動型の探索方法」を導入しました。さらに、強化学習を行う AI エージェント ^(注3) を複数競合させ、もっとも適切に振る舞うことのできたエージェントを選別することで、対話によるタスク達成の成功率を向上させ、また達成までの対話回数を減らすことに成功しました。

本研究の内容は、9月17日に、米国電気電子学会 (IEEE) の学術誌「IEEE Access」に掲載されました。

【詳細な説明】

研究の背景

ChatGPT などに代表される、自然言語を使う対話型 AI 技術は、ここ数年で劇的に発展しました。自然言語の対話によって特定の業務を行うシステム（タスク指向対話システム^(注 4)) は、問題解決を目的としたチャットボットなどに広く利用されています。このようなシステムでは、ユーザの入力に対して適切に応答し、必要に応じて不足する情報を質問するなど、柔軟な対話が求められます。例えば、映画のチケット予約を行うシステムでは、ユーザが「映画のチケットを予約したい」と言った場合、システムは適切な質問を重ねて必要な情報を集め、最終的にチケットを予約するというタスクを完了します。こうしたシステムを構築するためには、システムがどのように対話を進めるかを学習することが必要で、その手法の一つとして強化学習があります。強化学習を用いることで、システムは対話の事例から適切な応答を学習し、より短い対話で確実にタスクを達成できるようになります。

今回の取り組み

本研究では、強化学習に「好奇心駆動型の探索方法」を導入しました。これは、システムが新しい応答を試した際に「好奇心報酬^(注 5)」を与えることで、システムが積極的に新しい対応方法を探索できる仕組みです。また、複数のエージェント候補を用意し、その中から最もバランスの取れた行動をするエージェントを選び出す仕組みを採用しています。これにより、システムが早い段階で誤った行動に固執するのを防ぎ、柔軟な学習が可能になります。

さらに、訓練の後半でパフォーマンスが低下した場合には、エージェントを交換する仕組みも導入しました。これにより、全てのエージェント候補を最大限に活用し、タスクの成功率をさらに高めることができます。

本成果は、ユーザとの対話をより柔軟かつ効率的に行い、複雑なタスクを効果的に達成できる対話システムの実現に寄与します。

今後の展開

本研究の成果により、タスクを達成するタイプのチャットボット（受付、案内、予約など）の性能が向上し、より適切な応答を返しなが、短い対話で作業を終了できるようになります。さまざまな場面でチャットボットによる省力化・自動化が行われていますが、その性能は現在必ずしも十分ではありません。本研究成果により、チャットボットによる作業がより適切に行えるようになります。

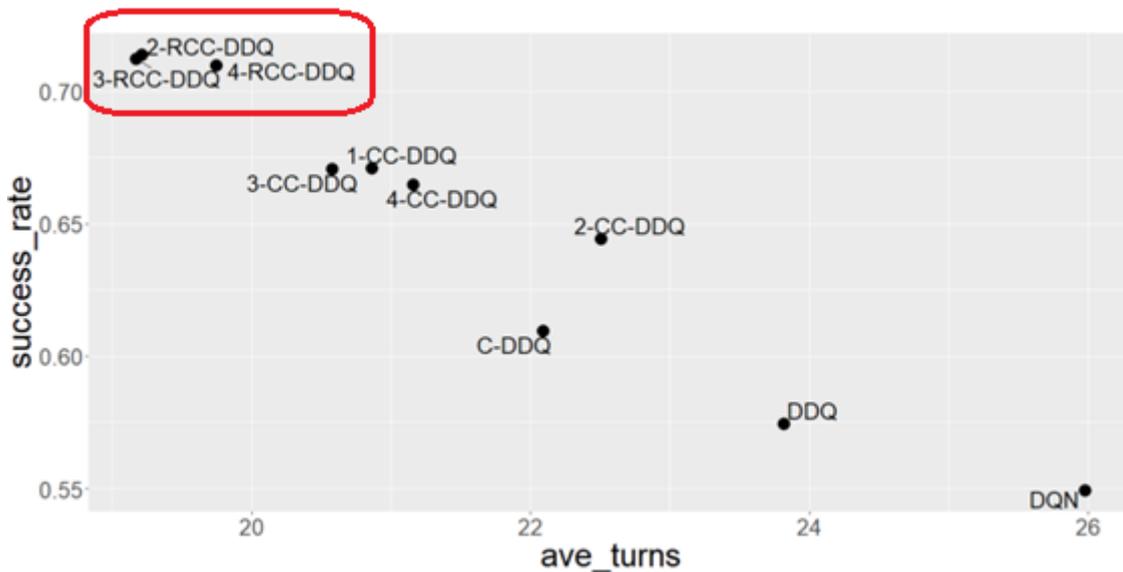


図 1. 様々な対話システムの性能比較（強化学習の一種である Q 学習の派生である DQN (Deep Q-Network) 、DDQ (Deep Dyna-Q) など多数の学習法と比較)。横軸はタスク達成のための平均ターン数（システムとのやり取りの数、小さい方が良い）、縦軸は平均タスク達成率（大きい方が良い）。本研究の提案法（RCC-DDQ：好奇心駆動型の探索方法）は、少ないターン数で高いタスク達成率を得ることができた。

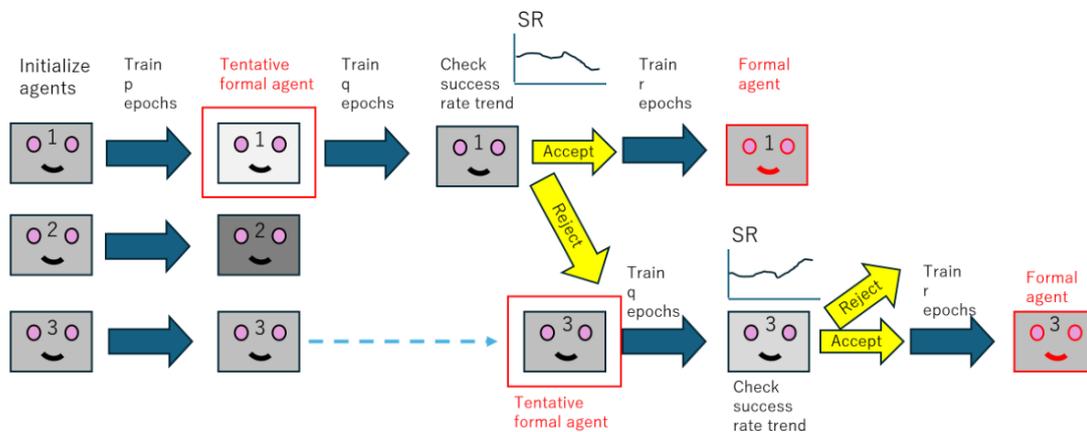


図 2. 複数のエージェント候補から、より良いエージェントを選択する仕組み。まず複数のエージェント（この例では 3 つ）を独立に強化学習で開発し、成績の良いエージェントを一旦選ぶ。その後さらに学習を進め、性能が上がるようならそのままそのエージェントを採用するが、性能が上がらないようなら 2 番目のエージェントを新たに選択する。

【用語説明】

- 注1. 対話型人工知能（AI）：人間同士は、言葉で対話をしながら仕事を進めたり、あるいは楽しんだりします。それと同様に、言葉によって人間と様々なやり取りをするための AI が対話型 AI です。電話応答のオペレーターの代替やヘルスケアといった実用的な分野から、エンターテインメント分野まで、様々な応用があります。
- 注2. 強化学習：強化学習（Reinforcement Learning）は、人工知能（AI）の一分野で、エージェント（学習者）が環境と相互作用しながら最適な行動を学習する方法です。エージェントは行動を選択し、その結果として報酬（または罰）を受け取ります。このプロセスを繰り返すことで、エージェントはどの行動が最も良い結果をもたらすかを学習します。強化学習は、ゲームの AI、自動運転車、ロボット制御、金融取引など、さまざまな分野で応用されています。エージェントが環境との相互作用を通じて、複雑なタスクを効率的に解決できるようにします。
- 注3. エージェント：強化学習におけるエージェントは、環境と相互作用しながら最適な行動を学習する主体です。エージェントは行動を選択し、その結果として得られる報酬を基に学習を進めます。例えば、迷路を解くロボットでは、ロボットがエージェントであり、環境（迷路）の中を動き回りながら、最適な脱出方法を試行錯誤によって学習します。同様に、本研究でのエージェントは、ユーザと対話をしながら、ユーザの要求を最も効率的に満たす対話方法を試行錯誤によって学習します。
- 注4. タスク指向対話システム：タスク指向対話システムは、特定のタスクを完了することを目的に設計された対話システムです。例えば、映画のチケットやレストランの予約、天気情報を提供するシステムなどが挙げられます。このシステムは、ユーザとの対話を通じて必要な情報を収集し、タスクを達成することを目指します。
- 注5. 報酬：強化学習における報酬は、エージェントが取った行動の結果として受け取るフィードバックのことです。この報酬は、エージェントがどの行動が良いか、どの行動が悪いかを学習するための重要な手がかりとなります。報酬を通じてエージェントは最適な行動を学び、効率的にタスクを達成することができます。

【論文情報】

タイトル : A replaceable curiosity-driven candidate agent exploration approach for task-oriented dialog policy learning

著者 : Xuecheng Niu, Akinori Ito, Takashi Nose

*責任著者 : 東北大学大学院工学研究科 大学院生 牛 雪澄

掲載誌 : IEEE Access

DOI : 10.1109/ACCESS.2024.3462719

URL : <https://ieeexplore.ieee.org/document/10681405>

【問い合わせ先】

(研究に関すること)

東北大学大学院工学研究科通信工学専攻

教授 伊藤 彰則

TEL: 022-795-7084

Email: aito.spcom@tohoku.ac.jp

(報道に関すること)

東北大学大学院工学研究科情報広報室

担当 沼澤 みどり

TEL: 022-795-5898

Email: eng-pr@grp.tohoku.ac.jp