

令和 5 年 6 月 6 日現在

機関番号：15301

研究種目：国際共同研究加速基金（国際共同研究強化(A））

研究期間：2022～2022

課題番号：21KK0206

研究課題名（和文）未知の協調・環境の動的性質に追従するエージェント群の協調戦略学習と知識進化

研究課題名（英文）Cooperative Strategy Learning and Knowledge Evolution to Adapt to Dynamism in Unknown Cooperation and Environment in Multi-agent System

研究代表者

上野 史（Uwano, Fumito）

岡山大学・自然科学学域・助教

研究者番号：30880687

交付決定額（研究期間全体）：（直接経費） 5,300,000円

渡航期間： 6ヶ月

研究成果の概要（和文）：本研究では、知覚エイリアシングにより従来の知識が利用できず、その手がかりもない未知の状況において、他の領域で学習した知識から特徴的な部分を抽出して組み合わせ、更に未知の状況においてはその特徴を階層的に組み合わせることで、時系列的に遷移する状態の共通点と未知の状況との特徴の共通点の双方を考慮した学習が可能となる手法を提案した。これにより、未知の状況であっても、過去に観測した情報を利用することで適応可能であり、また状態遷移を繰り返しても未知の状況に陥り続けて過去の情報が役立たない時であっても、階層構造を用いて観測情報の時系列的なパターンを学習することで追従可能であることがわかった。

研究成果の学術的意義や社会的意義

本成果は、エージェントの数や環境の変化によって未経験の状況に直面しても、自身の知識を最大限に利用してそれに追従し、目的を達成する上で最適な行動を学習可能であるという点において重要な成果である。それに加えて、学習結果を人間が解釈可能な知識として保存可能な点も、今後のロボット系研究において重要である。なお、本成果は当該分野のトップカンファレンスであるGECCO 2023において発表する予定である。

研究成果の概要（英文）：This study proposed a new method for learning by considering the similarity between sequential transition states and unknown situations in the environment with perceptual aliasing by constructing new knowledge from the parts of the knowledge. That method contributes the learning in an unknown situation such that the robots can adapt to the unknown situation using the past observed data, and adapt to the sequential aliased states using hierarchically structured knowledge for expressing the sequential patterns of the environment.

研究分野：機械学習

キーワード：進化計算 強化学習 未知環境 知識転移 マルチエージェントシステム

様式 F-19-2

1. 研究開始当初の背景

近年、目標達成時に得られる報酬を手掛かりに、人が事前に教示せずに適切な行動を学習する強化学習メカニズムを、複数エージェントの協調行動獲得のために各個体に導入したマルチエージェント強化学習の実用化が始まっている。基課題では、実問題のように学習する状況が多数あり、次の状況が予測困難である問題を想定し、エージェントが学習結果からモジュール化された知識（知識モジュールと呼ぶ）を抽出し、それを組み合わせて生成する新たな知識を用いて未知の協調行動の学習する手法の提案、および未知の環境への適応を目指している。しかしながら、基課題では未知の協調・環境が静的であるため、実問題に向けた動的な未知の協調・環境への追従に関して問題がある。

2. 研究の目的

本国際共同研究は解釈可能な知識を生成する進化的機械学習を利用して、限られた状況から学習した知識を適応的に進化させ、エージェント群がその転移により多様な協調行動を学習することで、協調・環境の未知の性質が動的に変化する際にも追従して協調行動を学習し、解釈可能な知識として保持する手法を提案し、その有効性を検証することを目的とする。

3. 研究の方法

本研究は、ロボットにおけるナビゲーションタスクを取り上げ、またルールをベースとした解釈可能な知識の生成と効率的なナビゲーションを可能とする最新の進化的機械学習手法 FoRsXCS (Frames-of-Reference-Based XCS) [1]を基に、未知の状況における学習方法およびその知識の生成手法について探求し、ロボットのナビゲーションにおいてその性能を検証する。なお、未知の状況は知覚エイリアシングと呼ばれるロボットの知覚の問題により状況の違いが認識できず学習が失敗する問題を拡張し、実際の環境でも状況の違いが区別できないものを未知の状況として定義し、その上での学習方法について探求する。

4. 研究成果

本研究では、階層構造を用いた新たな知識構造を FoRsXCS へ導入し、未知の状況における適切な行動の学習及び解釈可能な知識の生成手法を提案した。具体的には、図 2 に示すような迷路における学習を取り上げた。図は S_{3,V_0} で示す場所をゴールとした迷路問題とその時のロボットの観測を変数で示している。この迷路は中央の通路における観測情報が同様

(S_1) であり、ロボットがその未知の状況における状態の違いを見分けられず、追加の変数 (V_1, V_2, V_3) を用いてその状態を見分ける必要がある。しかしながら、FoRsXCS では固有の状態 (S_0, S_2, S_3) を基準として同じ観測状態の違いを見分けていたため、図 2 に示すような連続した場合は基準がなく未知の状況を見分けることができない。そこで、図 1 のように、同様の状態が連続する場合はそれを統合し、その下での位置関係によりバージョンを設定して見分け、それぞれの状態における最適行動を学習するように改良した。これにより、未知の状況のマクロな位置関係とミクロな位置関係を同時に表現可能で、マクロな方針とミクロな方針の同時獲得が可能となっている。これにより、例えば複数ロボットの協調を考えた際に、個々の衝突を回避するためのミクロな方針転換をしつつ、目的地到達などの最終的なマクロな方針を維持するといった高度な協調が可能であり、例え協調行動が未知であってもマクロな方針から協調行動の学習が可能となる点において重要である。

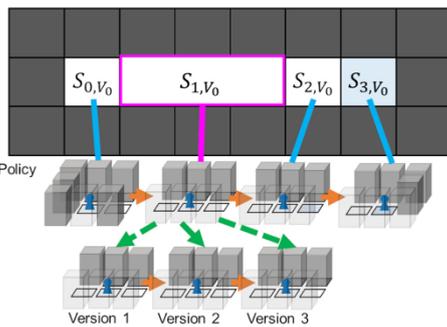


図 1 未知の状況と階層構造

実験では、図 3 に示す 9 種類の環境におけるロボットのナビゲーションタスクでその性能を検証した。図において色付きのマスは知覚エイリアシングを引き起こす場所であり、ロボットはランダムに配置された場所からゴールまでの経路を学習する。ベースラインの手法には FoRsXCS とメモリをベースとした代表的な深層強化学習手法である DRQN (Deep Recurrent Q-Network) を用いて目的地到達までのステップ数を比較した。表 1 は実験結果を示している。Type は知覚エイリアシングの違いを示し、提案手法 (Hi-FoRsXCS) はエイリアシングの種類に関わらず最小のステップ数でゴールへ到達する知識の獲得を可能としていることがわかる。なお、Littman89 において FoRsXCS の数値が最良だが、優位差はなく学習の収束速度や収束値の推移を比較しても提案手法との差はなかった。図 5 は左下を出発地点とした提案手法の獲得経路を示している。中でも緑線と赤線はそれぞれ最小ステップ数及び最小距離の経路を示している。FoRsXCS では獲得が難しい赤線の経路を提案手法は獲得可能であり、提案手法ではより人間の感覚に近く解釈しやすい経路としての知識が獲得できていることがわかる。本成果は機械学習における新しい知識の形とその学習方法を提案するものであり、ロボットの協調における個々の学習の効率化と

	S_{0,V_0}	S_{1,V_1}	S_{1,V_2}	S_{1,V_3}	S_{2,V_0}	S_{3,V_0}	

図 2 迷路問題における未知の状況

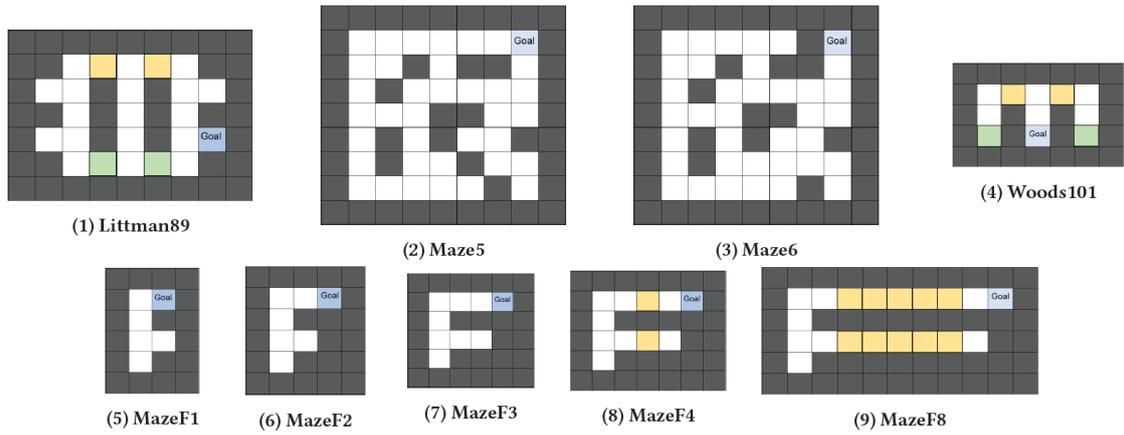


図 3 実験で用いた迷路

その利用範囲の拡大に寄与する。また、そのインパクトは大きく、当該分野のトップカンファレンスである GECCO 2023 へ採択されている [2]。

また、現状の提案手法の限界として、知識獲得のための試行錯誤の行動がランダムな動作に依存していることが挙げられる。例えば、図 4 は下段の未知の状況から行動を開始した場合、ロボットの獲得した経路を示しているが、これら全ての経路を利用して

表 1 実験結果 (括弧内は最小値, 太字は最良結果)

	Type	DRQN	FoRsXCS	Hi-FoRsXCS
Littman89 [6]	I	13.58 ± 13.12 (4.13)	3.68 ± 0.34	3.81 ± 0.36
Maze5 [4]	*	8.28 ± 3.53 (4.2)	4.87 ± 0.28	4.81 ± 0.31
Maze6 [4]	*	9.27 ± 5.54 (5.37)	5.39 ± 0.36	5.31 ± 0.32
Woods101 [7]	III	32.20 ± 18.90 (3.37)	2.92 ± 0.17	2.90 ± 0.15
MazeF1 [11]	*	5.67 ± 5.57 (1.6)	1.81 ± 0.14	1.78 ± 0.14
MazeF2 [11]	*	21.91 ± 23.0 (2.2)	2.48 ± 0.17	2.47 ± 0.15
MazeF3 [11]	*	24.67 ± 20.86 (3.5)	3.32 ± 0.24	3.32 ± 0.22
MazeF4 [11]	II	26.78 ± 19.74 (4.67)	4.50 ± 0.37	4.50 ± 0.35
MazeF8	III	35.23 ± 12.39 (18.23)	-	12.8 ± 2.87

も全ての可能性を網羅した知識とはならず、例えば下段の回廊の右から 3 番目を行動開始地点とした場合、合致する経路はない。そのため、今後の課題として、より網羅性の高い知識生成手法の提案が挙げられる。そして、今後の展望として、複数ロボットにおける協調の種類における対応が挙げられる。本成果では、各ロボットが独立に学習したものを組み合わせて協調を実現する想定を置いているが、より高度な協調行動を実現するためには、個々の独立した学習のみならず全体での学習を必要とするものもある。しかしながら、未知の状況では効率的な学習が難しく、環境次第では協調行動の学習が不可能なものも考えられる。そこで、[3]に示す手法のように、与えられた経路から最適な経路を導き出す機構を取り入れ、効率性を高めつつ経路の最適性を向上させることを今後は目指す必要がある。

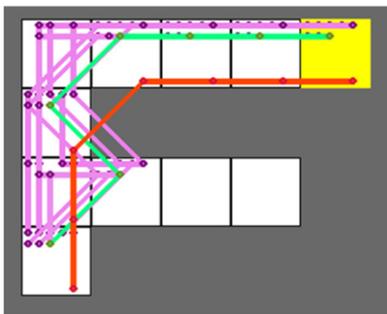


図 5 獲得経路

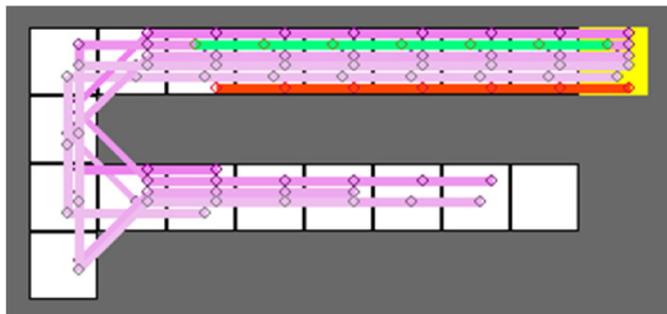


図 4 提案手法の限界

参考文献

- [1] Siddique, A., Browne, W. N., and Grimshaw, G. M., “Frames-of-Reference-Based Learning: Overcoming Perceptual Aliasing in Multistep Decision-Making Tasks,” *IEEE Transactions on Evolutionary Computation*, vol. 26, no. 1, pp. 174–187, Feb. 2022.
- [2] Uwano, F. and Browne, W. N., “Hierarchical Frames-of-References in Learning Classifier Systems,” in *Proc. of Genetic and Evolutionary Computation Conference 2023 (GECCO 2023)*, Lisbon, Portugal, Jul. 2023. (to appear)
- [3] Brown, D. S., Goo, W., Nagarajan, P., and Niekum, S., “Extrapolating Beyond Suboptimal Demonstrations via Inverse Reinforcement Learning from Observations,” in *Proc. of 36th International Conference on Machine Learning (ICML 2019)*, pp. 783–792, CA, USA, Jun. 2019.

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 Fumito Uwano
2. 発表標題 Hierarchical Frames-of-References in Learning Classifier Systems
3. 学会等名 Genetic and Evolutionary Computation Conference 2023 (国際学会)
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
主たる渡航先の主たる海外共同研究者	ブラウン ウィル (Browne Will)	クイーンズランド工科大学・工学部・教授	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関		
オーストラリア	クイーンズランド工科大学		