

マルチエージェント追跡問題の視野制限におけるモジュラー強化学習

日大生産工 ○沖田 大地 日大生産工 山内 ゆかり

1 まえがき

近年、一体の獲物を複数のハンターで捕獲する追跡問題が注目されている。マルチエージェント強化学習における協調課題として渡邊らによって状態をモジュール化することで次元の呪いを回避する階層型モジュラー強化学習で提唱された[1]。

本研究では複数のエージェントの視野を制限し近傍は正確に、遠方は大雑把に把握することで状態数の削減を行いモジュラー強化学習によって獲物を捕獲するまでの高速化を目指す。

2 強化学習による追跡問題

強化学習はエージェントが試行錯誤に行動し目的を達成できたら報酬を与えエージェントを学習させるものである。追跡問題では強化学習によってハンターが獲物を捕獲するまでのステップを削減する。図1に示す8×8のグリッドワールド環境においては、エージェントの認識する状態が絶対座標で構成される場合、8×8の64通りであるが、エージェントの位置を基準として±7の広がりで状態空間を相対座標で構成した場合は、15×15の225通りとなる。

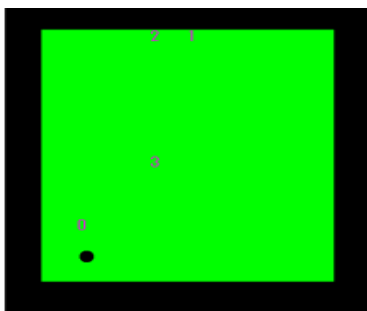


図1, グリッドワールド8×8

渡邊らのハンターエージェントは、相対座標系での状態数の削減として、視野の範囲をエージェントを中心に5×5の25グリッドに制限している (図2)。

モジュラー強化学習とはハンターが複数いる環境下では、状態空間はエージェント数分の次元が増加し、グリッドサイズや仲間エージェ

ント数が増えると爆発的に増大し学習性能が低下してしまう。その状態を圧縮するため、各エージェントが認識する状態を部分空間に分割したモジュラー構造を適用する。

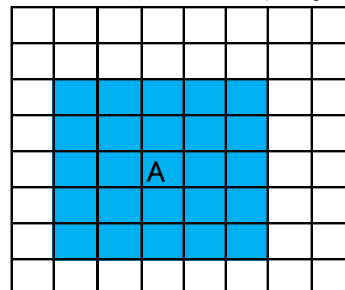


図2, 従来のエージェントから見える視野

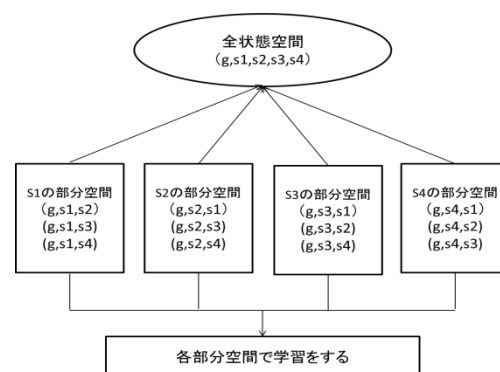


図3, エージェントの状態モジュラーモデル

3 提案手法

本研究ではハンターエージェントの視野を人間に近い視野にし、また獲物を捕獲するまでのステップの高速化を提唱する。具体的にはエージェントの近傍は正確に状態を把握するが、遠方を大まかに把握することで少ない状態数で多くのエリアをカバーできるように視野を設定する。図3に提案手法の視野の範囲を示す。Q学習を行いエージェントが獲物を捕獲するまでのステップを削減できるか検証する。

Modular Reinforcement Learning Algorithm with Vision Restricted for Multi-Agent Pursuit Problem

Daichi OKITA and Yukari YAMAUCHI

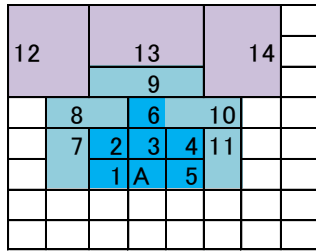


図4, エージェントから見た視野

4. 実験環境

本実験で用いたグリッドワールドは従来研究と同様に図1に示される縦8マス横8マスの64マスを採用し、ハンターエージェント4体、獲物エージェント1体をランダムに初期配置した。

エージェントの行動は上下左右の4方向に1マス移動するか現状のマスに留まるかの5つの行動ができる。Q学習を用い、ランダムに動く獲物を捕獲したら報酬を与え初期配置に戻る。捕獲条件は獲物を上下左右4方向からハンターを囲むことができたなら捕獲成功(図5)、また角を使い2方向(図7)、壁を使い3方向から囲む(図6)ことも捕獲成功とした。斜め四方向囲む(図8)という形は獲物が移動可能であるため未捕獲とする。これを1000エピソード行いステップ数が削減されているか検証する。

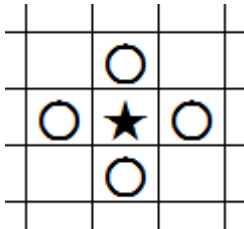


図5, 4方向

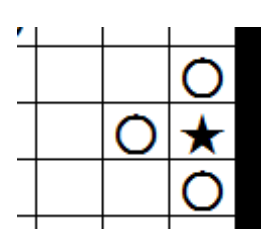


図6, 壁使用時

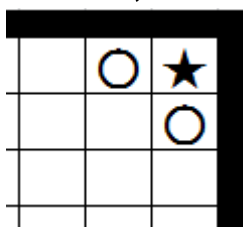


図7, 角使用時

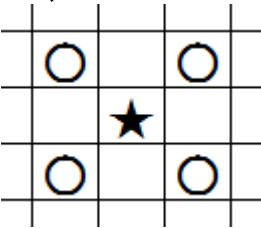


図8, 未捕獲

5 実験結果および検討

絶対座標系と相対座標系で視野の制限を変え実験をおこなった。図9、及び図10にそれぞれの学習の結果を示す。グラフの横軸はエピソード回数、横軸は捕獲までのステップ数である。

図9より、絶対座標系はエピソードが進むにつれて、壁をうまく使いながら獲物を捕獲しておりステップ数を削減できている。

図10より、相対座標系では学習回数の減少傾向は見られないが、絶対座標系より少ないステップ数で捕獲できている。

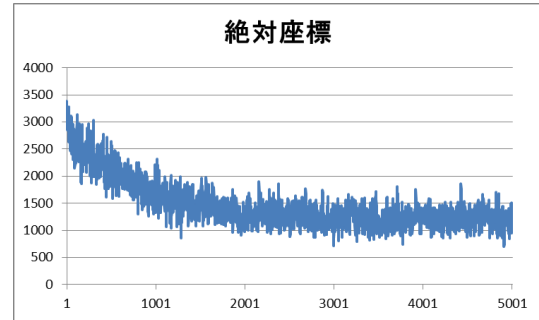


図9, 絶対座標系学習曲線

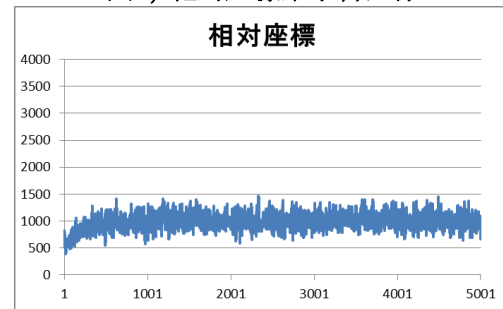


図10, 相対座標系学習曲線

相対座標系はハンターからの獲物位置という考え方から壁を使うという認識が薄いですが、獲物や他のエージェントとの距離に応じて、過去に経験してない場所でも学習の成果を適用できることにより、絶対的な位置に依存しない行動即が得られていると考えられる。

6 まとめ

本研究では複数のエージェントの視野を制限し近傍は正確に、遠方は大雑把に把握することで状態数の削減を行いモジュラー強化学習によって獲物を捕獲するまでの高速化を行った。今後の課題として、視野の違いによって学習回数に変化があるか調べてみる必要があると考える。

「参考文献」

- [1] 渡辺、和田、マルチエージェント追跡問題のための相対座標系に基づく階層型モジュラー強化学習、バイオメディカル・フュージ・システム学会誌 No2, pp. 65-74 (2010)
- [2] 高橋、柿倉、複数台移動のロボットによる協調捕獲に関する基礎研究、日本機械学会論文集(C編)71巻702号(2005)