

複数種類のシグナルからなる通信手段を持つ マルチエージェント系における強化学習

進士 純之介[†] 山口 文彦[†]

長崎県立大学 情報システム学部 情報セキュリティ学科[†]

1. はじめに

「知能とは何か」を考えると、知能と言語の関係に着目されることがある。知能の低かった原始の人類には言語がなかったことを考えれば、言語の発生過程を知ることによって知能の発生過程についての知見も得られる可能性があるのではないかと考えられる。

本研究では言語の発生として、強化学習を用いてエージェント間の通信にルールを獲得できるかについて実験を行った。

2. 過去の研究

通信手段を持つマルチエージェントの研究としては山本らの研究[2]がある。この研究ではルールのないエージェント間通信が強化学習を行うことで通信ルールを獲得したことを示した。

本研究では複数のハンターからなるマルチエージェント系に強化学習を適用し、ハンターが獲物を捕らえる追跡問題を用いて、各ハンターエージェントに通信手段を持たせたときの効果を検証する。また、その通信に用いられるシグナルの種類数による捕獲までのステップ数の変化を観察する。

3. 学習の適用

本研究では、学習手法として強化学習の一つである Profit Sharing を用いる。この手法は報酬を獲得したときに、それまでに選択したルールの評価を強化関数によって更新するものである。本研究では強化関数として等比減少関数を用いた。

また、ハンターの行動選択方式にはルーレット選択を用いた。

4. 実験環境

本研究では、トーラス状のグリッドワールド上に二体のハンターと一体の獲物を毎回ランダムに配置する。グリッドワールドの大きさは実験ごとに変化させる。二体のハンターはそれぞれ毎ターンごとに上下左右のいずれかに1マス分移動することが出来る。

獲物の捕獲条件はハンターと獲物が同じマスに移動したとき、または二体のハンターが獲物と異なる二辺で隣接したときである。獲物の動きは常にランダムである。ハンターは視覚範囲内にいる他のハンターと獲物を知覚することが出来る。ハンターの視覚範囲の大きさを変えて複数回の実験を行った。

5. ハンター間の通信とメモリ

本研究では、ハンターは上下左右の移動に加えて、移動するごとに必ずシグナルの発信を行う。発信するシグナルの種類数は実験ごとに異なり、1, 2, 4, 8, 16 のいずれかである。なお、シグナル種類数1は実質的にシグナルが無いことと同じである。発信されたシグナルは発信したハンターの視覚範囲内にいるもう一方のハンターにのみ受信される。

シグナルの発信については最初に、どのシグナルをいつ発信するというような情報を一切与えないため、学習の初期段階ではランダムな行動が多くなるが、学習が進むにつれてハンターが通信のルールを獲得していく可能性があると考えられる。

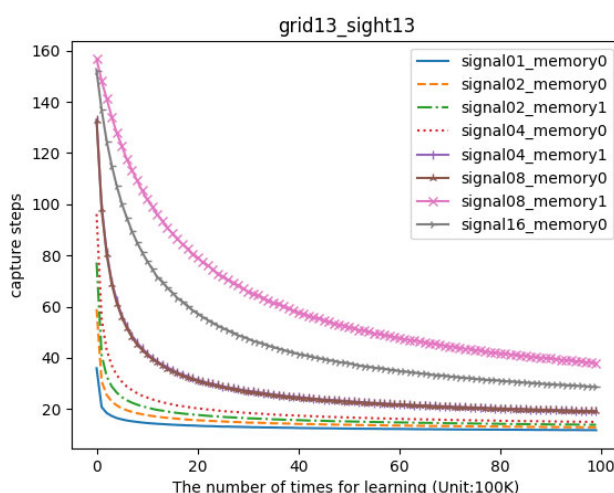
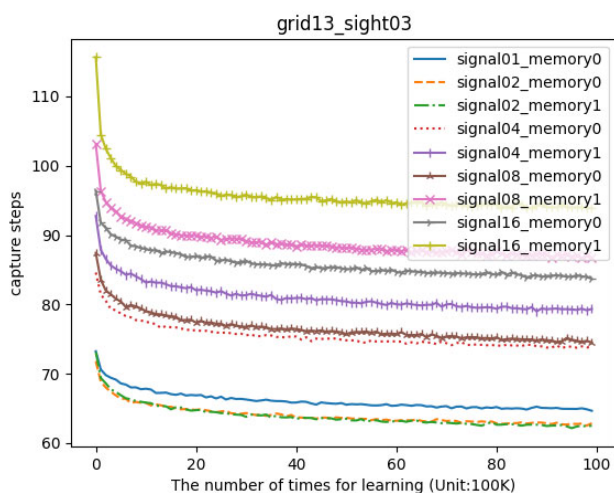
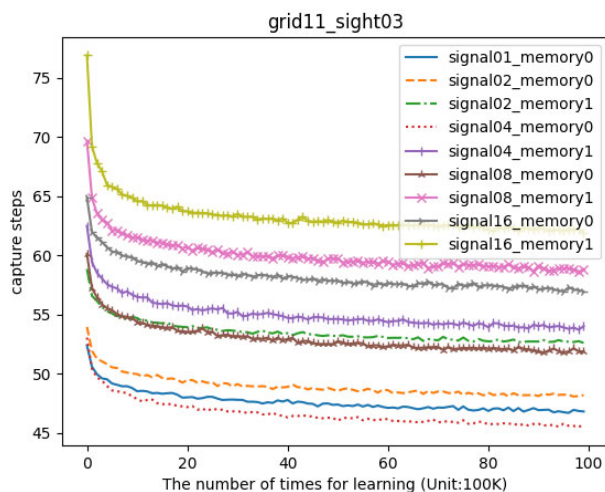
また、それぞれのハンターが、自分が最後に出したシグナルを記憶しておくメモリを持たせた場合の実験も行った。

6. 実験結果

実験結果を以下に示す。縦軸はハンターが獲物を捕まえるまでのステップ数、横軸は評価値を更新した回数を示している。また、図中の memory0 はメモリがないことを、memory1 はメモリがあることをそれぞれ示している。

Reinforcement learning in multi-agent which communicate by multiple signals

[†]Junnosuke SHINSHI, Fumihiko YAMAGUCHI, University of Nagasaki



7. 考察

grid11_sight03 と grid13_sight03 の図において、特定のシグナルの種類数を持つ場合にシグナル数が1の場合と比べてステップ数の改善が確認できた。このことから、ハンター間の通信

に何らかのルールが創発されたのではないかと考えられる。

また、grid13_sight13 の図を見ると、複数種類のシグナル通信を持つときのステップ数はシグナルの種類数が1つの場合のステップ数に収束しているように見えることから、ハンターの視覚範囲が広いほど、通信によるステップ数の改善は発生しなくなると考えられる。

メモリについては、すべての図においてほとんどの場合に、同じシグナルの種類数であればステップ数が増加することが確認された。

8. まとめ

実験の結果から、グリッドワールドの大きさに対してハンターの視覚範囲が小さいときには一部の通信を行うハンターにステップ数の減少が見られた。このことから学習によって捕獲のための通信ルールが創発されたのではないかと考えられる。

一方、シグナルの種類数やメモリの有無によっては通信を行わない場合よりもステップ数が増加していることも確認できる。このことから、シグナルの種類数やメモリの有無によって状態数がグリッドワールドの大きさに対して過剰になり、学習がうまくいかなくなったのではないかと考えられる。

また、メモリの効果については良い結果を確認できなかった。

9. 今後の課題

本研究でシグナルの種類数が結果に影響を与えることが確認できた。今後の課題としてはグリッドワールドの大きさやハンターの視覚範囲の大きさとシグナルの種類数やメモリの有無について適切な組み合わせを分析することが考えられる。

創発された通信のルールについて、どのようなルールがどのような過程で発生したのかを分析することも今後の課題である

参考文献

- [1] 大須賀節雄, 言語と知能, オーム社, 2010
- [2] 山本真也, 山口文彦, 中西正和, 通信手段を持つマルチエージェント系における強化学習, 情報処理学会第57回全国大会, 1998
- [3] 高玉圭樹, マルチエージェント学習一相互作用の謎に迫る一, コロナ社, 2003