

マルチパーティ戦闘ゲームにおける ユーザーからの教示を導入したエージェント学習機構の検討

筒井壮太郎[†] 福田直樹[‡]

[†]静岡大学情報学部 [‡]静岡大学大学院情報学領域

1 はじめに

マルチパーティ戦闘ゲームでは、ユーザーが操作するキャラクター(PC)と味方側のNPCが、協調動作をとることでゲームを優位に進行させることが可能である。味方側のNPCに協調動作を獲得させる方法のひとつに、マルチエージェント学習の手法の適用がある[1]。一方で、マルチエージェント学習の手法を単純に用いて味方側のNPCに協調動作を獲得させる場合、必ずしもPCとの協調動作を獲得できるとは限らないという点が課題である。本稿では、教示を導入したエージェント学習機構について検討する。

2 研究の背景

2.1 マルチパーティ戦闘ゲーム

本稿で述べるマルチパーティ戦闘ゲームとは、キャラクターを操作して敵側のキャラクターを倒すことを目標としたゲームである。マルチパーティ戦闘ゲームにおいてゲームを優位に進行させるためには、PCと味方側のNPCが協力して敵側のNPCを倒すことが必要である。本稿では、マルチパーティ戦闘ゲームを追跡問題の条件を拡張したものとして実装を行った。味方側のキャラクターには上下左右1マスへの移動と攻撃、上下左右4マスへの捕獲行動、敵側のNPCには上下左右1マスへの移動と停止を行動として用意した。ゲームは、PC、敵側のNPC、味方側のNPC、敵側のNPCの順に行動することで1ターンが経過し、20ターンで1ゲームとする。味方側の一方のキャラクターが隣接している敵側のNPCに捕獲行動をとった場合、次の敵側のNPCは行動不能となる。行動不能な敵側の

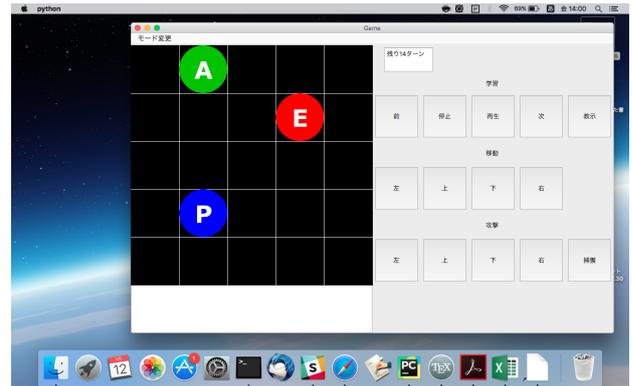


図 1: マルチパーティ戦闘ゲームの外観

NPCにもう一方の味方側のキャラクターが攻撃をした場合ゲームに勝利し、20ターン経過した場合はゲームに敗北する。本稿で実装したマルチパーティ戦闘ゲームの外観を図1に示す。

2.2 マルチエージェント学習

マルチエージェント学習の具体的な手法として、マルチエージェント強化学習があげられる[1]。マルチエージェント強化学習では、エージェントの数や選択可能な行動の種類などによって観測可能な状態の組み合わせが爆発的に増加する。マルチパーティ戦闘ゲームにおいてマルチエージェント強化学習を用いた場合、学習が終了するまでに時間がかかる場合や終了しない場合が考えられ、PCとの協調動作を獲得することが困難となることが予想される。

3 検討手法

本稿では、マルチパーティ戦闘ゲームにおける味方側のNPCにPCとの協調動作の獲得を実現させるために、エージェントの学習にユーザーからの教示の導入を試みる。エージェントの学習過程の任意の状態において、教示によって指定された行動をエージェントに選択させる。教示によって指定した行動より学習によって得

A Preliminary Approach of Multi-agent Learning with Users' Advises on a Multi-Party Hunting Game

Sotaro TSUTSUI[†] and Naoki FUKUTA[‡]

[†]Faculty of Information, Shizuoka University
432-8011, Hamamatsu, Japan

[‡]College of Informatics, Academic Institute Shizuoka University
432-8011, Hamamatsu, Japan

[†]ia14057@s.inf.shizuoka.ac.jp

[‡]fukuta@inf.shizuoka.ac.jp

た行動の方が最適解に近くなる場合のことを考慮して、本稿では SA 法における温度パラメータの概念を導入する [2]。SA 法における温度パラメータ $T(0 \leq T < 1)$ は $T_{n+1} = \alpha T_n$ と定義され、ステップ n が経過する度に冷却率 $\alpha(0 < \alpha < 1)$ によって減少する。本稿では、教示を行った状態 S において T の確率でエージェントを教示に従わせ、 $1 - T$ の確率で学習の方策によってエージェントに行動選択させる。状態 S においてエージェントに行動選択させるたびに T を更新する。

4 実験および結果

本稿で提案したエージェント学習機構を用いて、味方側の NPC にユーザーが操作するキャラクターとの協調動作を獲得することが可能であるかを確認するために実験を行った。比較の対象として、マルチエージェント Q 学習 [1] と Felipe らの手法 [3] を用意した。

4.1 実験設定

実験の環境として、2 章で述べたマルチパーティ戦闘ゲームを用い、味方側の NPC と PC にエージェントを導入し、行動を学習させた。学習には Q 学習を用い、学習率 $\alpha = 0.2$ 割引率 $\gamma = 0.9$ とし、方策には ϵ -Greedy 法を用い、 $\epsilon = 0.3$ とした。Felipe らの手法 [3] においては、スケラブル変数を $v_{ask} = 0.75, v_{give} = 1.5$ とした。報酬として、敵を倒したエージェントに 10 点、敵を倒す直前に捕獲行動をしたエージェントに 5 点、ターンが経過するごとに -1 点を与えた。検討手法においては、冷却率 $\alpha = 0.95$ とし、学習エピソードの開始 10 ゲーム分教示を行った。学習が 100 エピソード終了するごとに PC に別途学習をさせたエージェントを導入した上でゲームをプレイさせ、ゲームクリアするまでの味方側のキャラクターの総行動回数を計測した。学習は全部で 10000 エピソード行った。

4.2 結果

実験の結果を図 2 のグラフに示す。行動回数の値が小さいほど味方側の NPC が敵側の NPC を倒すための最適な行動に近い行動を選択していることを表している。図 2 のグラフから、検討手法を用いて学習させた味方側の NPC が従来の手法を用いて学習させた場合に比べて行動回数がほぼ全てのエピソードにおいて小さいことがわかる。このことから、本稿で行った実験においては検討手法を用いて学習させたエージェントに、マルチエージェント学習の手法を単純に用いた場合よりも効果的に協調動作を獲得させることが可能であったと考えられる。

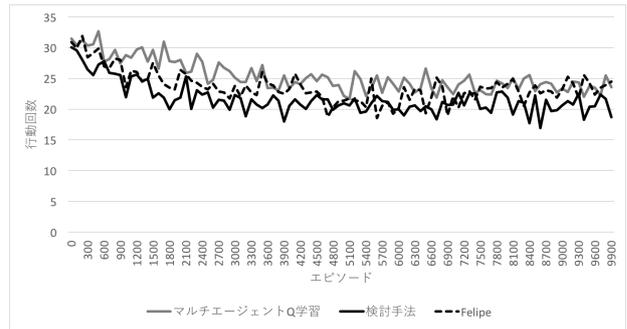


図 2: 実験の結果

5 まとめと今後の課題

本稿では、マルチパーティ戦闘ゲームにおける味方側の NPC に PC との協調動作を獲得させるために、マルチエージェント学習の手法に教示を導入したエージェント学習機構を検討した。検討手法を用いた味方側の NPC が、PC との協調動作を獲得させることが可能かを確認するために、マルチパーティ戦闘ゲームを実装し、その効果を解析するための実験に用いた。本稿で扱った条件下においては、教示を導入し、学習エージェントを教示に従わせることによって、マルチエージェント学習の手法を単純に用いた場合よりも効果的に協調動作を獲得させることが可能であることを確認した。

今後の課題としては、本稿の実験に用いたマルチパーティ戦闘ゲームの敵側の NPC の行動を増やした場合やキャラクターの数を増やした場合におけるエージェントの学習の検討があげられる。

参考文献

- [1] Steven D. Whitehead. A complexity analysis of cooperative mechanisms in reinforcement learning. In *Proceedings of the 9th National Conference on Artificial Intelligence, Anaheim, CA, USA, July 14-19, 1991, Volume 2.*, pages 607–613, 1991.
- [2] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.
- [3] Felipe Leno da Silva and Anna Helena Reali Costa Ruben Glatt. Simultaneously Learning and Advising in Multiagent Reinforcement Learning. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1100–1108, 2017.