

マルチエージェントシステムにおける協調度を用いた行動選択法

情報科学科 河村 将範

指導教員：小林 邦和

1 はじめに

人は、相手に合わせて接し方を変化させながら生活をしている。例えば、友人とのコミュニケーションの取り方と、先生とのコミュニケーションの取り方は異なっている。同様に、それらは家族とのコミュニケーションの取り方とは異なっている。このような接し方の変化をロボットができたとしたら、人の話相手となってくれるなど、社会に役立ってくれることが想定される。そこで、人が相手に対して接し方を使い分けるといった分人という概念が平野により提唱されている [1]。私はこの分人という概念をマルチエージェントシステムに導入することで、接し方の変化を実現しようと考えた。本研究ではこの分人モデルの一部として、協調行動の実装に観点を置いた。その方法として、分人が構成されると同時に、協調度という尺度を導入する。この協調度をマルチエージェントシステムに導入することにより、エージェントがとりうる行動にどのような変化が現れるかを評価する。また、提案法を用いることにより、エージェントの協調行動が促進または抑制されることを計算機シミュレーションにより、確認する。

2 提案法

提案法では、相手が存在する方向を特定し、その方向を a_0 とする。この a_0 は、自エージェントが行う事ができる行動から選択する。そして、 a_0 の方向へ行動する時の確率に、協調度を反映させることで、協調行動の促進と抑制を行う。

目的地への経路の学習は、代表的な強化学習法として、Q 学習 [2] を用いて行う。Q 学習とは、現在の状態 s における行動 a によって報酬 r を与え、Q 値を更新し、学習を行うものであり、式 (1) によって与えられる。

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \cdot \max_{a' \in A} Q(s', a') - Q(s, a)] \quad (1)$$

ここで、 A は行動全体の集合、 α は学習率と呼ばれる学習速度を調節する定数、 γ は割引率と呼ばれる将来の報酬が現在において、どれだけの価値があるかを決定づける役割を持つ定数である。

行動選択確率 $\pi(s, a)$ は softmax 選択法を適用し、式 (2) によって与えられる。

$$\pi(s, a) = \frac{\exp(Q(s, a) * g(c))}{\sum_{b \in A} \exp(Q(s, b) * g(c))} \quad (2)$$

協調度を $g(c)$ という関数を用いることで行動選択確率に反映させる。提案法では、 $g(c)$ を式 (3) のように定義する。

$$g(c) = \begin{cases} 2c & (a = a_0) \\ 1 & (\text{otherwise}) \end{cases} \quad (3)$$

$g(c)$ に関して、協調度が 0.5 の時は $g(0.5) = 1$ となり、通常の行動選択となる。一方、協調度が 0.5 より大きい時は $g(c) > 1$ となり協調が促進され、0.5 より小さい時は $g(c) < 1$ となり協調が抑制される。そのため $g(c)$ は c の 2 倍の値と設定する。

3 計算機シミュレーション

10×10 の格子状フィールドを用意し、2 体のエージェントと 1 つの目的地を配置する。ステップ数の上限は 100 とし、それを 10000 回繰り返して学習を行う。性能評価は、ステップ毎のエージェント間の距離によって行う。

協調度は、相手と協調しやすい時は 0.8、協調しにくい時は 0.2 の値で固定する。(1) 両エージェントの相手への協調度が 0.8 の場合の結果を実線、(2) 両エージェントの相手への協調度が 0.2 の場合の結果を点線として、図に示す。(1) の時、最初の 4 ス

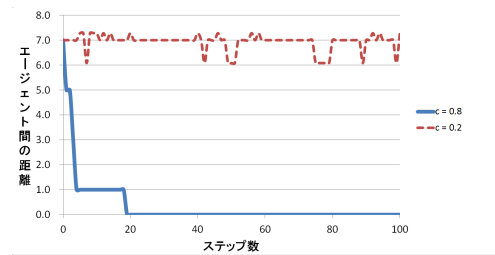


図1 エージェント間の距離の変化

テップで両者の距離が縮まり、その後は短い距離を保ちながら、19 ステップで目的地へ辿り着いていることが見て取れる。従って、お互いが協調していると言えるだろう。(2) の時、いずれのステップにおいても両者の距離が離れていて、最終的にゴールに辿り着いていないことが見て取れる。従って、お互いが協調していないと考えられる。また、目的地へ辿り着くよりも協調しないことを優先している事が分かる。

一方、自分から相手への協調度が高く、相手から自分への協調度が低い時には、自分からの一方的な協調が行われてはいたが、お互いが目的地に辿り着いた結果となった。しかし、両エージェントの協調度が高い時に比べ、ステップ数が増えてしまった。従って、協調度は両エージェントが目的地に辿り着くまでのステップ数に影響を与えられられる。

4 まとめ

本研究では、協調度を用いた行動選択法を提案した。協調度に応じて、協調行動が促進または抑制されることが分かった。

本研究では、協調度を固定した計算機シミュレーションを行っているが、将来的には協調度の更新をすることを視野に入れている。

参考文献

- [1] 平野啓一郎: 「私とは何か「個人」から「分人」へ」、講談社新書 (2012)
- [2] C. J. C. H. Watkins and P. Dayan: "Q-learning", Machine Learning, Vol.8, No.3-4, pp.279-292 (1992)
- [3] Masanori Kawamura, Kunikazu Kobayashi: "An Action Selection Method Using Degree of Cooperation in a Multi-agent System" Proc.of ICAROB2014, pp.118-121 (2014)