

F-019

マルチエージェントによる学習効果の考察 Consideration for Learning Effect by using Multi Agent

隅田 篤十
Atsushi Sumida

原嶋 勝美†
Katsumi Harashima

1. はじめに

人は教育や経験による学習を通して成長していく。そこで本稿ではマルチエージェントを用いて、学習の違いによる成長を比較し、効果的な学習について考察する。

2. 問題設定

本稿では効果的な学習につながる教育と経験を比較する為、 N 匹のハンター H_1, H_2, \dots, H_N と、 M 匹の獲物 P_1, P_2, \dots, P_M 、 L 匹の親 O_1, O_2, \dots, O_L で構成される3種類のエージェントによる狩り問題を対象とする。ハンターの目的はより多くの獲物を獲得する事である。そのため、ハンターは獲物に対する狩りの経験を繰り返し積みかたわら、親からも狩りの教育を受け学習する。本研究では一定の期間ハンターに狩りをさせ獲物獲得数を比較し、学習効果を確認する。

2.1 エージェント

学習するエージェントはハンターのみである。ハンターは①獲物の探索、②追い込み、③捕獲の3つの行動で獲物を捕まえる。効率良い狩りには、獲物を早く見つける広い視野持ち、素早く動き、獲物を倒す強力な力、そして獲物に負けない体力が必要である。そこで、これらを能力と定義した。以下、エージェントの持つ能力及び行動について、ハンターを中心に説明する。

2.1.1 能力

時間 t におけるエージェントの持つ能力を $C_e(i, k, t)$ と表すものとする。ここで、 e はエージェントの種類 (h : ハンター、 p : 獲物)、 i はエージェント e における固体番号、 k は能力の種類を表す。能力は以下の4つがある。

- ・力: $C_e(i, 1, t)$ ・ 獲物の体力を奪い、倒す力
- ・行動力: $C_e(i, 2, t)$ ・ 素早く行動できる速さ
- ・視野: $C_e(i, 3, t)$ ・ フィールドを見る範囲
- ・体力: $C_e(i, 4, t)$ ・ 力に抵抗する屈強さ

ハンターは能力の中の1つを長所とする。長所は他の能力より学習効率が良く成長し易いものとする。

2.1.2 行動

獲物を捕まえるハンターの行動は、以下の通りである。
①探索: ハンターと獲物は相手を発見するまでランダムに移動する。

②追い込み: 図1に示すように、ハンター H_i は視野内に存在する獲物 P_j を追い込む時、複数のハンターが協力する場合を想定して移動方向を3通り設ける。一方、狙

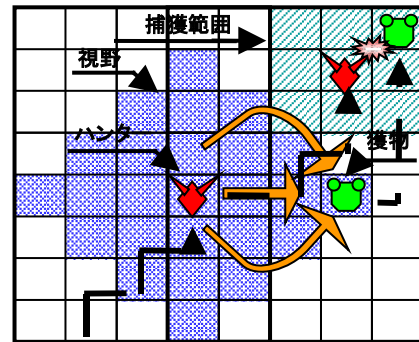


図1 ハンターエージェント

われた獲物 P_j はハンター H_i に対して確率 $ad(j, i, t)$ で追いつく、あるいは $ae(j, i, t)$ で逃走する。式(2)(3)に $ad(j, i, t)$ 、 $ae(j, i, t)$ を示す。

式(2)は獲物 P_j の能力の、ハンター H_i と獲物 P_j の能力の和に対する割合を表している。 $ad(j, i, t)$ が大きい程、獲物は相対的にハンターに抵抗する能力が高いことになる。したがって、 $ad(j, i, t)$ を追返し確率とする。

$$ad(j, i, t) = \frac{\sum_{k=1}^4 C_p(j, k, t)}{\sum_{k=1}^4 \{C_h(i, k, t) + C_p(j, k, t)\}} \quad (2)$$

$$ae(j, i, t) = 1 - ad(j, i, t) \quad (3)$$

③戦い: 図1に示すように、ハンター H_i の捕獲範囲に獲物 P_j が進入した場合、 H_i は P_j に攻撃をする。一方、捕獲範囲に入った P_j は逃げる事ができないため、必ず反撃する。戦いは H_i 、 P_j のいずれかの体力が0になった時点で決着するものとし、体力が0にならなかったエージェントの勝利となる。

戦いにおけるハンター H_i の体力変化を式(4)に、獲物 P_j の体力変化を式(5)に示す。戦いにより0~相手の力の範囲で自身の体力が減少するように設定している。これにより、能力が劣っている者でも勝てる可能性を持たせている。

$$C_h(i, 4, t+1) = C_h(i, 4, t) - rand1 \times \{C_p(j, 1, t)\} \quad (4)$$

$$C_p(j, 4, t+1) = C_p(j, 4, t) - rand1 \times \{C_h(i, 1, t)\} \quad (5)$$

($rand1$: 0~1の乱数)

2.2 学習

学習効果によりハンターは能力値を上昇させる。しかし、学習効果には幅を持たせている。

† 大阪工業大学 電子情報通信工学科

2.2.1. 狩りの経験

狩りの経験には成功と失敗の2つのパターンがある。

○成功：獲物を獲得

狩りが成功したハンター H_i は自身の能力を上昇させる事が出来る。これを上昇値と呼ぶ。ハンター H_i と獲物 P_j との能力比較に応じて得られる上昇値 $im(i, j, t)$ を式(6)に示す。

H_i より P_j の能力が高い程、 H_i の得られる上昇値が高くなり、逆に P_j の能力が低い程、得られる上昇値は低くなる。更に単調な成長にならないように、1~5の乱数をかける。但し、上昇値の最小値は1とする。

$$im(i, j, t) = rand2 \times \frac{\sum_{k=1}^4 C_p(j, k, t)}{\sum_{k=1}^4 C_h(i, k, t)} \quad (6)$$

(rand2: 1~5の乱数)

能力の上昇は以下の手順で行なう。

- 1: 4つの能力の中からランダムに1つ選択する。(但し長所である成長し易い能力は4割、それ以外能力は2割の割合で選択される)
- 2: 選ばれた能力を1上昇させる。
- 3: 上昇値 $im(i, j, t)$ の数だけ1,2を繰り返す。

○失敗：獲物に逃げられる、追返される

狩りが失敗した場合、その原因となった能力(逃げられた場合は行動力・視野、追返された場合は力・体力)を改善する。能力改善は以下の手順で行なう。

- 1: 1~5の数値からランダムに選択する。
- 2: 改善する2つの能力の内、1つをランダムに選ぶ。
- 3: 選ばれた能力を1改善する。
- 4: 1で選んだ数字の分だけ2・3を繰り返す。

2.2.2. 親の教育

親は教育方針に従ってハンターに教育を施す。教育方針は、①力重視、②行動力重視、③視野重視、④体力重視、⑤長所を伸ばす、⑥弱点を補う、⑦平均的に教育の7つである。教育を一回行くと、ハンターの能力は1上昇する。教育が施される能力は、ランダムに選択される。但し、ランダム選択において、①~⑥の教育方針では、重視する能力が選択されやすくする為、選択に重みを設けている。一方、⑦の教育方針では能力成長の幅が全て均等である為、選択に重みを設けていない。

3. 実験

学習効果を評価するために、前章で示したエージェントを用いて以下の条件で実験した。

- ・フィールドは100×100の格子状
- ・ハンターの数 $N=50$, 獲物の数 $M=200$, 親の数 $L=7$
- ・教育の学習割合を変更 (0, 0.2, 0.4, 0.6, 0.8, 1)
- ・学習行動: $T=50000$
- ・学習評価: 学習を伴わない狩りをさせ ($T=50000$)、獲物獲得数を比較した

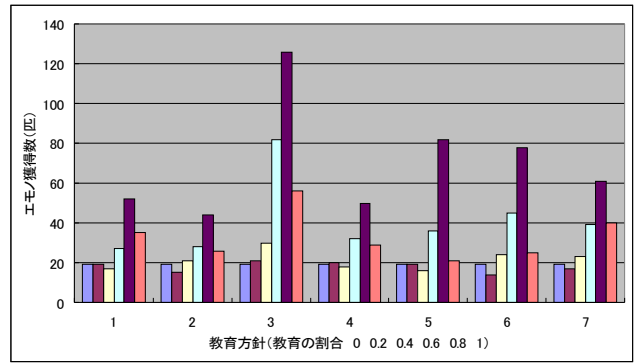


図2 実験結果

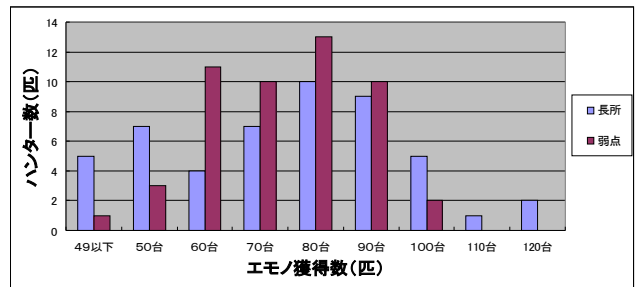


図3 実験結果

図2に、捕らえた獲物の総数をハンターの数で割った獲物獲得数の平均の比較結果を示す。縦軸は獲物獲得数の平均、横軸の数字は各教育方針(①力、②行動力、③視野、④体力、⑤長所、⑥弱点、⑦平均)であり、各教育方針において、親からの教育の割合0, 0.2, 0.4, 0.6, 0.8, 1として示している。

結果より教育方針が視野重視で、各教育方針において教育割合0.8の時に最も高い結果を残していた。この結果から、フィールドを広く見渡し、獲物を見つける視野の広さが獲物獲得に大きく影響することがわかる。したがって、大きく成長するには親の視野に対する教育を十分に受ける必要があると考えられる。

長所重視と弱点を補う事は相反する学習方法であるが、図2に示す様に、最大の獲物獲得はほぼ等しい。そこで各学習方法で獲得した獲物の数に対する、その数の獲物を獲得したハンターの数と比較した。比較結果を図3に示す。

長所重視の場合、成績が悪いハンターは弱点重視の場合より多い。しかし、逆に成績が良いハンターは長所重視の場合の方が多かった。この結果から、親はハンターの長所を的確に把握し、十分に成長させてやる事が必要であるといえる。

4. まとめ

本稿では効果的な学習を考察する上で、エージェントを用い教育と経験のどちらを重視して学習させれば良いかを考察した。

実験により、教育方針は視野重視・長所重視・弱点を補う事を重視の順に高い教育効果が現れていた。更に、十分に教育を行い基礎が出来た上での狩りが最も効果的であった。今後、学習方法の仕組みの改善を計り、集団学習の導入、能力等の見直しを考える。