

障害物あり環境の追跡問題における協調動作の獲得と評価

賀数元春[†] 深海 悟[†]

大阪工業大学 大学院情報科学研究科[†]

1. はじめに

近年、自律移動ロボットの発達に伴い、人物追跡に関する研究が行われている[1]. その中で追跡対象が逃走している場合は追跡者を複数用意し、お互いが協調して追跡・確保することが重要となる. しかし現実世界では、追跡中は障害物により視界や行動が遮られてしまう状況に陥る. そこで障害物の存在する状況を想定した追跡問題について研究を行った.

障害物あり環境での追跡問題については、視界の障害は通信により位置を伝え合うことで解決できることが分かっている[2]. しかしこの通信内容で効果が発揮されるのは大局的な回り込みのみであり、ターゲット周辺での動作についてはさらなる効率化が必要である[2]. そこで本研究では、強化学習を用いて障害物あり環境ならではの協調動作を獲得した. そして得られた動作を従来の追跡エージェントに追加し、有効に機能する地形の特徴を明らかにした.

2. 追跡問題

追跡問題は、ある環境を逃げ回る逃走エージェントを複数の追跡エージェントが確保するシミュレーションである. 本研究では追跡問題を次のように定義する.

- $m \times m$ マスの格子状の環境に、逃走エージェントと複数の追跡エージェントを配置する.
- 各エージェントには視界があり、視界範囲内にいるエージェントしか認識できない. また、視線上に障害物が存在すればその障害物の向こう側のエージェントについては認識できない.
- 各エージェントは上下左右いずれかに1マス移動、もしくはその場に停止の5つの行動を行える. 行動は全てのエージェントが同時に行う.
- 各エージェントが同時に行動する時間単位をステップとする.
- 追跡エージェントに通信が許可されている場合は、通信範囲内に存在する全ての追跡エージェントに任意のメッセージを送信することができる.

- 通信は移動と同時に行うことができる.
- 逃走エージェントの四方全てが追跡エージェントや壁、障害物に囲まれるとそのターゲットは確保される.
- 初期状態から確保されるまでの1回をエピソードとする.

3. 協調動作の獲得

協調動作を見つけるために、強化学習を使用した追跡エージェントを作成した. 用いる手法は学習時間と協調動作の獲得に有効的な Profit Sharing である[3]. 逃走エージェントとの距離が2マス以内のときを学習範囲とし、環境は自身から2マス以内の状態(更地・障害物・追跡エージェント・逃走エージェント)と位置を正確に記述する. この学習エージェントを4体用意し、実験で使用する地形・環境にて各100万エピソードずつ学習を行った. その後学習済みエージェントから確保に有効と思われる協調動作を2つ見つけ出した. 一つ目の動作は「抑え込み」で逃走エージェントの行動を制限させる動作である. 逃走エージェントと隣接している際、自身から見て逃走エージェント方向の両斜め前2マスのうち、片方だけに味方が障害物が存在すれば静止を行う. 二つ目は「囲い込み」で最後の詰めを行う動作である. 逃走エージェントから2マス離れている際、逃走エージェントの2マス以内の状況を精査し、逃走エージェントの隣接4マスのうち自分のみが辿り着ける地点に移動する.

4. エージェントの動作

実験では学習エージェントではなく一定のアルゴリズムで動作するエージェントを用いた. 今回の追跡エージェントの基本動作は、逃走エージェントまでの道のりを、障害物を考慮した最短ルートで追跡を行う. 追跡対象の位置が不明な場合はランダムに移動または静止する. また、逃走エージェントを直接視認している場合は、通信を用いてその位置情報を発信する. そして学習結果より見つけ出した2つの協調動作を実装した.

5. 実験

今回の実験は抑え込みと囲い込みの効果を確かめ

Acquisition and Evaluation of Cooperative Motion in the Pursuit-Problem with Obstacle

[†] Motoharu Kakazu [†] Satoru Fukami

[†] Graduate School of Information Science and Technology, Osaka Institute of Technology.

るために、抑え込みあり・なし、囲い込みあり・なしを組み合わせる4パターンで検証を行った。また、実験フィールドも以下の図1~図4のMAP1~4を用意した。そしてそれぞれのフィールドで各パターンにつき50,000エピソードの実験を行った。

その他、実験環境は以下の通りである。

- 追跡エージェントは4体
- 視界距離は5マス
- 通信範囲は10マス

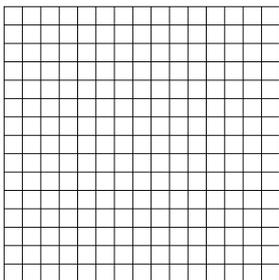


図1 MAP1

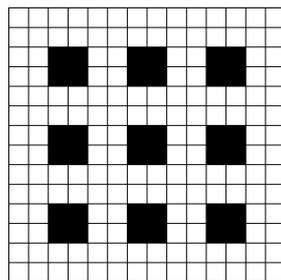


図2 MAP2

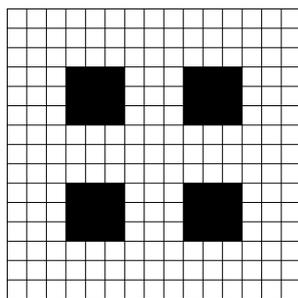


図3 MAP3

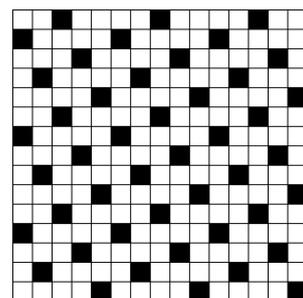


図4 MAP4

6. 結果と考察

実験結果を以下の表1に示す。値は1エピソードにかかったステップ数の平均値である。また、「囲い」は「囲い込み」、「抑え」は「抑え込み」を表している。

表1 MAP1~4の実験結果

MAP1	囲いあり	囲いなし	MAP2	囲いあり	囲いなし
抑えあり	50.3	73.7	抑えあり	54.4	65.5
抑えなし	53.5	77.3	抑えなし	68.6	76.0

MAP3	囲いあり	囲いなし	MAP4	囲いあり	囲いなし
抑えあり	138.4	204.3	抑えあり	46.9	48.9
抑えなし	172.8	241.5	抑えなし	49.2	50.9

表を見ると、どの地形においても抑え込み・囲い込みにより時間が短縮できていることが確認できる。しかしMAP4では時間短縮効果は小さいなど、地形により効果に違いがあることもわかる。

抑え込みの効果が比較的高いのは障害物のあるMAP2とMAP3で、これは障害物が無いMAP1よりも時間の短縮率が高い。このことから抑え込みは障

害物あり環境により適した動作であるといえる。

地形により効果量の違いがある理由としては、障害物の配置が関係していると思われる。効果の高いMAP2, MAP3は障害物により道ができていたような地形である。このような地形で障害物を利用した抑え込みが発動した場合、逃走方向はほぼ1方向に限定されることになる。ここで対面側にも追跡エージェントが構えていれば、確保率が上がるものと見られる。対してMAP1では空間が開けているため、抑え込みが発動しても他にも逃げ道が多い。MAP4は空間が開けているとはいえないが、細かく散らばった障害物により逃走方向を絞らせずに逃げるができる。よってMAP1, MAP4では抑え込み効果が低かったと思われる。追跡エージェントを増やす等でより多方向から迫ることができればMAP4でも抑え込みが有効に働くかもしれない。

また、囲い込みについてはMAP1とMAP3において時間短縮割合が高くなっている。その要因には回り道をする必要の少ない地形が挙げられる。

囲い込みは逃走エージェントからの距離が2マスのときに発動するが、囲いきるためにはあらかじめ逃走エージェントの全方位に追跡エージェントや障害物を用意しておかなければならない。直線的に迫れるMAP1や、大局的な逃走方向が限定され道幅も広いMAP3ではあらかじめ回り込んでおくことが容易になるため、多方向からの囲い込みを行う機会が増えるものと思われる。しかしMAP2やMAP4ならば逃走方向を絞らせないように逃げられるため、囲い込める状況に持ち込むまでに時間がかかってしまう。改善のためには、抑え込みと同様に、大局的な回り込みを行うことが重要だと考えられる。

7. おわりに

今回の実験により、障害物あり環境においても強化学習を用いて有効な協調動作を獲得できることが確認できた。今後は確保時間を短くするためにさらに協調動作を見つけることや、別の研究に用いられている協調動作を今回のものと組み合わせるとどのような結果になるか、などについて研究を進めたい。

参考文献

- [1] 戸上千裕, 長尾智晴: 自律移動ロボットによる特定人物追跡に関する研究, 情報処理学会第76回全国大会, 5T-6(2014).
- [2] 賀数元春, 深海悟: 障害物で視界が限定されたマルチエージェント追跡問題に関する研究, 情報処理学会第78回全国大会, 5P-02(2016).
- [3] 櫻井 祐輔: マルチエージェント強化学習による協調性獲得の検証 — 追跡問題を例として —, 高知大学大学院理学研究科修士論文(2008).