

実環境競合解消のための通信量を抑えた機械学習手法 —自律移動ロボット群の衝突回避問題—

砂山享祐[†]

三上貞芳[‡]

鈴木恵二[‡]

大沢英一[‡]

1 はじめに

自律エージェントは現代社会において欠かせない技術であり、さまざまな場所で利用されている。しかし、複数のエージェントが有機的に相互作用し、協調するマルチエージェントシステムは未だ解決課題が多いのが現状である。分散ロボットシステムもそのひとつであり、不完全知覚問題をもつシステムである。また、自律ロボットが同一環境内に複数台存在すると、それぞれの移動タスクの中で他ロボットと衝突する可能性が生じてしまう。そこで、本研究では、自律移動ロボット群における衝突回避システムを提案する。衝突回避問題はそれぞれのロボットの移動目標が異なるため、実環境における競合解消問題である。

2 提案手法

本研究では図 2.1 に示す Friendly Robotics 社の Robomower RL500 を改造したものを想定している。RL500 には差分 GPS が搭載されており、誤差 2m 程度の位置同定が可能である。この GPS を用いて位置情報を取得し、搭載された無線 LAN による通信で衝突回避を実現することを目指す。また、実環境を扱う以上、必要となる情報に誤差やノイズが含まれることは避けられないため、一般に普及するようなシステムを開発するためには、用いる情報に誤差が含まれていても衝突回避が可能なシステムを開発することが望まれる。

2.1 通信

分散ロボットシステムにとって衝突回避アルゴリズムは、移動タスクの中のひとつの機能にすぎない。従って、可能な限り少ない情報通信量で衝突回避を実現することが望ましい。そこで、本研究では他ロボットの速度と大きさという情報のみで衝突回避を実現することとする。実環境では必ずしも環境内の全てのロボットについて情報を知ることが出来るとは限らないため、獲得できた情報のみを利用して行動の意思決定を行う。

また、通信範囲とは別にロボット固有の視野を持たせ、視野内に存在するロボットの情報を利用することとする。本研究では、ロボットは独立に行動するため、回避行動を始めるタイミングなどに差異が見られるべきである。これを促進するために情報の使用の可否を決定する視野を導入した。図 2 に 3 台の場合の情報通信と意思決定までの例を示す。各ロボットは獲得できた情報から意思決定を行うため、あるロボットとの通信が失敗したとしても衝突に関して致命的な状況に陥ることは無い。



Fig.1 RL500

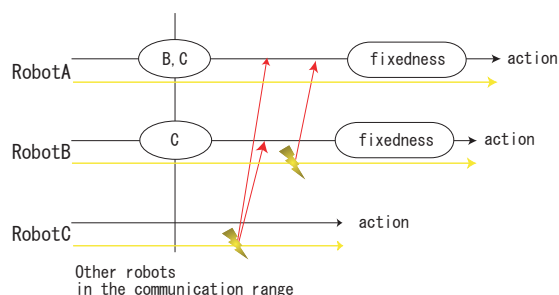


Fig.2 The flow of each robot's decision making

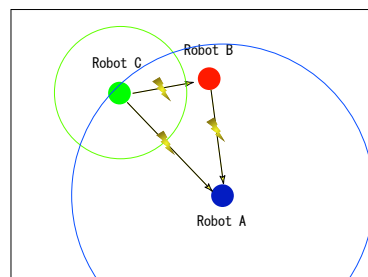


Fig.3 Robot's vision and each available information

また、図 3 に、それぞれのロボットの視野と利用できる他ロボットの情報獲得の例を示す。

2.2 衝突位置の面積

必要最小限の情報から衝突回避を実現することを考える。本研究では、他ロボットの速度と大きさから衝突が起こるかどうかを予測し、衝突する可能性がある場合は回避行動をとることとする。従って、ロボット間で行う通信では、現在の速度とロボットのサイズという二つの情報を交換する。

ロボットは獲得した情報から、衝突位置の面積を算出する。図 4(a) のようにロボット A とロボット B が直進するとき、衝突位置は円形で表現される。ロボット A が加速、ロボット B が減速というように行動変更を行うと、衝突位置の面積は (b) のように変化する。このように、各ロボットが行動を変更することで衝突位置の面積を減少させていき、0 になったときに衝突は回避される。従って、ロボット

[†] 公立はこだて未来大学大学院

[‡] 公立はこだて未来大学

は衝突位置の面積が小さくなる方向に行動を変更することになる。

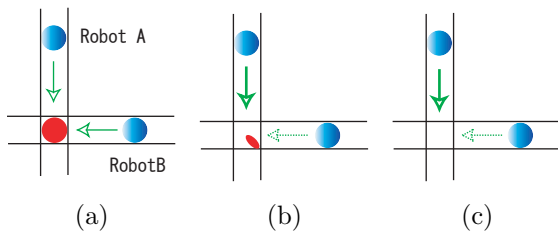


Fig.4 The transition of the collision point area

2.3 強化学習

ロボットの行動決定は基本的に衝突位置の面積に従うが、具体的には強化学習を用いて行動系列を獲得する。強化学習は、環境を完全にモデル化できないマルチエージェント問題に対してある程度の有効性を見込むことができる。[4]

衝突位置の面積、衝突位置の個数、そして最近傍の衝突位置までの距離を状態とする。行動は、加減速、左右の方向転換、そして現状速度の維持とする。ここで、方向転換はポテンシャル法を用いて実現している[1]。報酬は、式(1)のように行動前後の衝突位置の面積の比により計算される。

$$R(t+1) = 1 - \frac{A_{next}}{A_{current}} \quad (1)$$

3 実験

図5に示すように4台のロボットが1点で交差する場合についてコンピュータシミュレーションによる実験を行った。いずれのロボットも、直進した先に目的地があるものとする。実験は500エピソード行い、1エピソードはいずれかのロボットが衝突した時点で終了した。また、500ステップを過ぎても目的地に到達しない場合もエピソードを終了した。図6にロボットCの各エピソードにおける報酬和とエピソード終了までに要したステップ数を示す。初期は明らかに衝突のために少ないステップ数で終了しているが、学習が進むにつれて方策が改善されていき、400エピソードあたりで報酬和・ステップ数が収束していることがわかる。また、図7は最終エピソードにおける各ロボットの軌跡を示している。ロボットAは目的地に向かって直進しているが、他の3台のロボットはいずれも左迂回をすることで目的地に到達している。

また、5台のロボットで同様の実験を行ったが、1000エピソードまで学習させても衝突回避行動は発現しなかった。

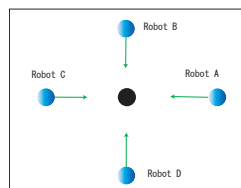


Fig.5 Initial setting of experiment 2

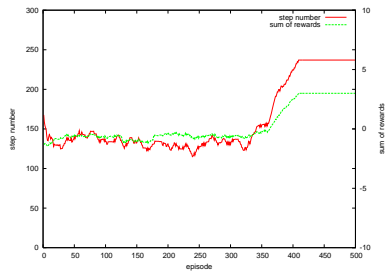


Fig.6 Machine C's step number and sum of rewards

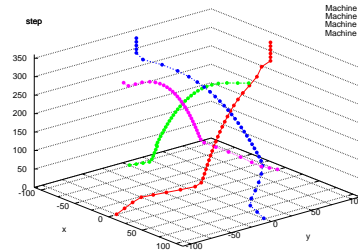


Fig.7 The trajectory of each robots at the last step

4 おわりに

自律移動ロボット群の衝突回避問題に着目して、実環境における競合解消問題に対して機械学習手法の提案を行い、環境がある程度単純な場合には良好な結果が得られた。

今後は、行動にかかるコストの削減、デッドロック状態の解消やロボット形状の多様化といった問題を解決しなければならない。このためには、コストの見積もりを反映させた報酬関数の設計、行動集合の拡張、そして面積の表現方法の変更などが必要となる。現在は実機実験の準備を進めている。

参考文献

- [1] J Borenstein and J Koren. Real-time obstacle avoidance for fast mobile robots. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 19, No. 5, pp. 1179–1187, 1989.
- [2] Pieter Jan't Hoen and Sander M. Bohte. Collective intelligence with sequences of actions Coordinating actions in Multi-Agent Systems. *Machine Learning:ECML 2003*, pp. 181–192, 2003.
- [3] 太田順, 新井民夫, 倉林大輔. 知能ロボット入門. コロナ社, 2000.
- [4] 三上貞芳. 強化学習のマルチエージェント系への応用. *人工知能学会誌*, Vol. 12, No. 6, pp. 37–41, 1997.
- [5] R. S. Sutton and A. G. Bart. *Reinforcement Learning—An Introduction—*. The MIT Press, 1998. (三上 貞芳, 皆川 雅章 訳 (2000) 『強化学習』. 森北出版).