

自動搬送ロボットのための分散オンライン行動プランニング

割田 智†

藤田 桂英‡

†東京農工大学 工学部 情報工学科

‡東京農工大学大学院 工学研究院 先端情報科学部門

1 はじめに

近年、社会におけるマルチエージェントシステムの活用が広がっている。マルチエージェントシステムの重要な問題の1つに、倉庫におけるピッキング作業の自動化がある。ピッキング作業とは指定された品物を倉庫の中から探し出して指定された場所まで配達する作業のことである。従来は人間が作業を行っていたが、ロボットによる作業の自動化が進んでいる。このような背景から複数のロボットが協調してピッキング作業を行う問題が注目を集めている。一方、これらの研究の多くはロボットを集中管理して最適化問題を解くというアプローチを取っており、各ロボットが自律的に行動を決定する分散型の手法に関する研究はあまり多くない。各ロボットが自律的に行動を決定するアプローチを取っている先行研究としては、環境をマルコフ決定過程としてモデル化しモンテカルロ木探索を適用する手法に関する研究が存在する [1]。しかし、この研究においてはロボットの衝突が考慮されておらず、現実世界に適用する上ではモデル化が不十分であるという問題があった。

そこで本研究では、ロボット同士が衝突するような移動を禁止する制約の下で、複数のロボットが協調して品物を自動搬送する環境、及びこの環境において効率的にピッキング作業を行えるような分散型の手法を提案する。また、このような環境のシミュレーションを行えるシミュレータを開発し、シミュレータを用いて提案手法の評価を行う。時間内に配達できた品物の個数を比較することで、既存手法と比べて効率的に品物の配達ができることを示す。

2 問題設定

倉庫がグリッドグラフ G として表されており、 G にはデポ（駐車場）と呼ばれるノードが1つ決まっている。倉庫内には N 体のエージェントが存在し、1 タイムステップごとに各エージェントが独立に行動を決定し実行する。エージェントの取りうる行動はその場に留まる、隣接ノードへの移動、現在のノードにある品物の回

収、品物の配達（デポでのみ可）の4種類であり、品物の回収または配達を行うと報酬を得られる。品物を持っているエージェントはデポに品物を配達するまで新たに品物を回収することはできない。また、1 タイムステップごとにある確率 p で G のノード集合からランダムに選ばれたノード v に品物が出現する。エージェントの目的はこのような環境の下で T ステップ以内に得られる報酬の合計（累積報酬）を最大化することである。

この問題の環境においては、エージェントが「衝突」（すれ違いを含む）する移動は許されず、そのような移動をしようとしたエージェントはその場に留まるとする。また、エージェント a が品物を回収または配達したとき、全エージェントが一律に報酬 $reward$ を獲得し、エージェント a は追加の報酬 $bonus$ を獲得する。

3 提案手法

エージェント同士の衝突を考慮した環境において、効率的にピッキング作業を行うための手法を2つ提案する。

3.1 衝突回避付き Iterative Greedy

モンテカルロ木探索で行動を決定するためには他のエージェントの行動を予測する必要がある。この行動予測の結果に基づいてシミュレーションを行うため、行動予測の精度は十分高くなければならない。一方、行動予測は繰り返し実行されるため高速であることも求められる。ある程度精度が高く高速な手法として全エージェントが貪欲法に基づくアルゴリズムに従うとして行動予測を行う手法が提案されており [1]、その中で Iterative Greedy と呼ばれるアルゴリズムが最も優れていることが明らかになっている。しかし、本研究で扱う問題にモンテカルロ木探索を適用する上では、Iterative Greedy をそのまま行動予測に用いるのは精度の観点から不十分である。Iterative Greedy は「衝突禁止制約」を考慮したアルゴリズムではないため他のエージェントと衝突するような行動であっても回避行動をとらない。一方、エージェントがモンテカルロ木探索で行動を決定している場合そのような行動は回避されるので、予測される行動と現実の行動が乖離し精度が下がってしまう。

そこで、本研究では「衝突回避付き Iterative Greedy」というアルゴリズムを行動予測に用いることを提案する。衝突回避付き Iterative Greedy は、Iterative Greedy に

Decentralized Online Planning for Autonomous Mobile Robots
†Department of Computer and Information Sciences, Faculty of Engineering, Tokyo University of Agriculture and Technology
‡Division of Advanced Information Technology and Computer Science, Institute of Engineering, Tokyo University of Agriculture and Technology

「既に行動を決定した他のエージェントと同じノードに移動することになる, またはすれ違うことになる場合はその行動を選択できない」という制約を追加したアルゴリズムである.

3.2 楽観的行動選択

各エージェントがモンテカルロ木探索で行動を決定することになると, 他のエージェントと衝突してしまうような行動の評価値 (累積報酬の平均値) は高くなりづらいため選択されにくくなる. 結果としてある程度エージェント同士の衝突を回避させることができる. 一方で衝突を回避しようとしてエージェントがお互いに遠回りしてしまい, 効率が下がってしまうという課題がある. このことから「衝突を回避するためにあえて遠回りする」行動と「他のエージェントが回避してくれると仮定して最短経路を選ぶ」行動を状況に応じて使い分ける必要があることがわかる.

そこで本研究では「楽観的行動選択」という手法を提案する. 楽観的行動選択は「楽観度」という値に基づいて貪欲な行動の評価値を補正することで, 貪欲な行動が選択されやすいようにする手法である. 現在の状態を s , 状態 s において Iterative Greedy を用いたときに選択される行動を a_{greedy} とする. また, 状態 s における行動 a の補正前の評価値を $eval(s, a)$, 補正後の評価値を $eval'(s, a)$ と表す. 楽観度を $opt(0 \leq opt \leq 0.5)$ としたとき, $eval'(s, a)$ は以下の式を満たす.

$$eval'(s, a) = \begin{cases} eval(s, a) \times (1 + opt) & (a = a_{greedy}) \\ eval(s, a) & (otherwise) \end{cases}$$

4 実験

提案手法の有効性を評価するため, 表 1 に示す 6 種類の手法でエージェントの行動を決定したときのシミュレーション結果を比較する. 評価指標としては, 時間内に配達できた品物の個数を用いる.

エージェント数は 3, グラフのノード数は 214, シミュレーションを行うステップ数 T は 100 であり, 品物が出現する確率 p は 0.3 である. また, $reward$ は 100, $bonus$ は 70 としている.

実験結果を図 1 に示す. 6 種類の手法を, 時間内に配達できた品物の個数の平均値の昇順に並べると Greedy, GreedyCA, MCTS, MCTS OPT, MCTS GreedyCA, MCTS OPT GreedyCA の順になった. 提案手法を用いない場合 (MCTS) は平均 7.9 個しか配達できていないのに対し, 行動予測に衝突回避付き Iterative Greedy を用いた場合 (MCTS GreedyCA) は平均 9.7 個, 楽観的行動選択を用いた場合 (MCTS OPT) は平均 8.5 個配達することが

名前	説明
Greedy	Iterative Greedy
GreedyCA	衝突回避付き Iterative Greedy
MCTS	モンテカルロ木探索
	Greedy で行動予測
MCTS GreedyCA	モンテカルロ木探索
	GreedyCA で行動予測
MCTS OPT	モンテカルロ木探索
	Greedy で行動予測
	楽観的行動選択を行う
MCTS OPT GreedyCA	モンテカルロ木探索
	GreedyCA で行動予測
	楽観的行動選択を行う

表 1: 6 種類の手法

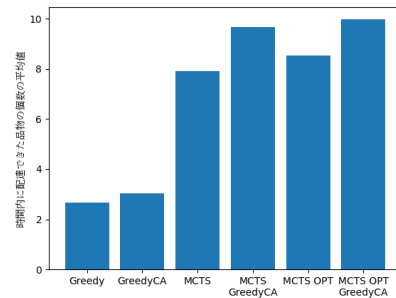


図 1: warehouse-large での実験結果

できており, 提案手法がピッキング作業の効率化に有効であることが確認できる. また, 2 つの手法を組み合わせた場合 (MCTS OPT GreedyCA) は平均 10.0 個配達することができており, 最も高い性能を示した. これは, 「衝突を回避する行動」と「貪欲な行動」を状況に応じて使い分けることができるためである.

5 まとめ

本研究では, エージェント同士の衝突を考慮した環境において, 効率的にピッキング作業を行うための手法を 2 つ提案した. また, シミュレータを用いた実験により, 提案手法の有効性を示した.

参考文献

- [1] Daniel Claes, Frans Oliehoek, Hendrik Baier, and Karl Tuyls. Decentralised online planning for multi-robot warehouse commissioning. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, AAMAS '17*, page 492–500, Richland, SC, 2017. International Foundation for Autonomous Agents and Multiagent Systems.