

機械学習による AGV の自律搬送スケジューリング

北見工業大学：○宮田慎太郎， 渡辺美知子

要旨

本研究では，FMS(Flexible Manufacturing System)工場に予め設定された複数の工作機械を複数の無人搬送車(Automatically Guided Vehicle, AGV)が周回する問題を対象とする．AGV が工場内で自律的な走行を行うために機械学習を採用し，その運行スケジュールの数値実験を行いその有用性を確認する．

1. はじめに

FMS 工場において，AGV を用いた多品種少量生産システムが加工機械の高い稼働率を上げるためには，AGV がどのような走行をするかを定めるスケジューリングが大きなウェイトを占めている．このような問題は，一般に NP 完全問題に属する大規模複雑な離散事象の組合せ最適化問題となり，厳密に解くことは殆ど不可能である．

本研究では，一方通行の走行レーン上を走行する複数 AGV に関して，指定された全加工機械に立ち寄り総 AGV の周回完了時間を最小とするようなスケジューリングを機械学習<sup>1)</sup>に基づいて各 AGV が自律的に決定する手法を取り扱う．工場内を走行する AGV の自律的意思決定には，確率的学習オートマトン(Stochastic Learning Automaton, SLA)を採用し，総 AGV の周回完了時間を最小とするような運行スケジュールを行う事を目的とする．最後に，本手法の有効性を確認するために数値実験を行い検証する．

2. 問題の記述

本問題では，図 1 に示す様な一方向走行レーンを持つ仮想 FMS 工場モデル<sup>2)</sup>を採用する．工場内は，加工機械が走行レーンに沿って並べられている．各工作機械を

$$M = \{M_{xy} : x=1, 2, \dots, u, y=1, 2, \dots, v\} \quad (1)$$

とおく．走行レーンを走る AGV を

$$A = \{A_i : i=1, 2, \dots, n\} \quad (2)$$

とする．今，AGV,  $A_i$  が行き先を指定された加工機械の集合を

$$M(A_j) = \{A_{xy} : x \in X(A_j), y \in Y(A_j)\} \quad (3)$$

とおく．ここで， $X(A_i)$ ,  $Y(A_i)$  は AGV,  $A_i$  の行き先の加工機械の添字集合である．全 AGV が図 1 に示す操車場から出発し，指定されたすべての加工機械を回り，操車場に戻る全 AGV の周回完了時間を  $T(A_i)$ ，加工機械を周回する順序を  $O(M(A_i))$  とすれば，

$$\min_{O(M(A_i))} F_1(A_i) = T(M(A_i)) \quad (4)$$

を満たすような AGV の巡回スケジュール  $O(M(A_i))$  を決定する問題となる．また，各 AGV,  $A_i$  の操車場の出発順序を  $S(A_i)$  とすると，全 AGV の周回完了時間  $T(A_i)$  は，出発順序にも影響されるから，

$$\min_{S(M_i)} F_2(A_i) = T(S(A_i)) \quad (5)$$

を満たす必要がある．このような評価を行う AGV には，以下の制約を設定する．

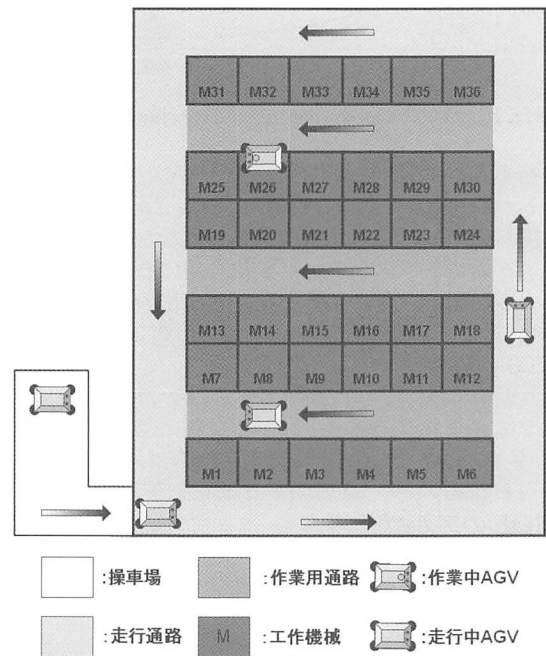


図 1. 一方向走行レーによる FMS 工場モデル [制約条件]

- 1) 走行レーンは一方向にのみ通行可能である．図 1 に示すように周回が可能である．
- 2) AGV は走行レーン上での追い越しを禁止する．
- 3) 各工作機械は AGV の待避バッファを持ち，一台のみ収容可能である．被加工物は AGV からバッファにおいて加工機械に渡され，指定された加工が終了するまで待機する．
- 4) AGV は待避バッファ上の AGV の追い越しが可能である．
- 5) 各 AGV は AGV 間の衝突を回避する．
- 6) AGV の移動時間は図 1 の 1 マスを移動する時間を 1 単位時間とする．被加工物の加工時間は予め単位時間を用いて設定する．
- 7) AGV は指定された加工機械すべてを巡回すると，速やかに操車場に戻り，移動を中止する．

3. 確率的学習オートマトンによる AGV 自律走行

3.1 SLA による AGV の定式化

各 AGV,  $A_i$  を以下のようなオートマトンに定式化する．

$$A_i = \{S_i, O_i, I_i, \Theta_i, \Phi_i\} \quad , i=1, 2, \dots, n \quad (6)$$

$n$  は AGV の台数を表し，各項目  $S_i$ ,  $O_i$ ,  $I_i$ ,  $\Theta_i$ ,  $\Phi_i$  はそれぞれ状態，出力，入力，状態遷移関数，出力関数を表す．状態遷移関数及び出力関数の両方，または一方が確率的で

あるとき、確率的オートマトンと呼ぶ。確率的オートマトンを学習オートマトンとして AGV に適用するため、各項目を次のように設定する。

### 3.2 状態 S

各 AGV が認識できる環境の状態を以下 8 通りに定める。

S0: AGV の前方 4 単位時間内に他の AGV が存在する。

S1: 1AGV の前方 4 単位時間内に指定の加工機械が存在する。

S2: 左側に走行経路が存在する。

S3: すぐ前方に壁が存在する。

S4: AGV の右側 4 単位時間内に他の AGV が存在する。

S5: AGV の右側 4 単位時間内に指定の加工機械が存在する。

S6: すべての加工機械を巡回。

S7: AGV が走行中またはバッファで待機中。

S0~S7 は状態が真のとき 1, 偽のとき 0 とする

$$S = \{S_j = S_j \in s_0 \times s_1 \times \dots \times s_7\} \quad (7)$$

### 3.3 出力 O

AGV の各状態 Si における行動を O とし、以下 5 つとする。

O<sub>0</sub>: 直進      O<sub>1</sub>: 左折

O<sub>2</sub>: 待機      O<sub>3</sub>: 加工機バッファに入る, 出る

O<sub>4</sub>: 操車場に戻る

$$O = \{O_j; j = 0, 1, \dots, 4\} \quad (8)$$

ただし AGV が待機中の時には任意の加工時間が終了するまで強制的に待機する。また、右折、後退は禁止する。

### 3.4 入力 I

入力はオートマトンの行動に対する評価とする。

評価は以下のスカラー評価 b を与える。

$$b = d(t-1) - d(t) \quad (9)$$

d(t) は現在の位置から次の目的地までの移動時間である。

### 3.5 状態遷移関数 Θ, 出力関数 Φ

状態遷移関数 Θ は一般に状態 S(t-1) から S(t) への遷移を定義する。ここでは AGV の学習を実行するため、以下の確率遷移行列を採用する。状態 S(t) から S(t+1) への状態遷移関数は、状態 S(t) から行動 O(t) への関数 θ(t) と行動 O(t) から O(t+1) への関数 φ(t) に分離可能であれば、二つの合成 θ, φ で表現できる。従って、θ(t) と φ(t) を状態遷移関数 Θ と出力関数 Φ の代わりに採用してもよい。SLA では θ(t) として、状態 S<sub>j</sub> と行動 O<sub>k</sub> の対を確率で表現し、これを状態 - 行動確率遷移行列 P = [P<sub>jk</sub>] (j = 0, 1, ..., 127, k = 0, 1, ..., 4) で表す。また、φ(t) を時刻 t の状態が S<sub>j</sub> であるときの行動を決定する関数とし、以下に示す式で確率的に決定する。

$$O_k = \text{Prob}(P_{jk}) \quad (10)$$

ここで、Prob() は確率 P<sub>jk</sub> に従って行動 O<sub>k</sub> を選択する関数を示す。

### 3.6 SLA の学習アルゴリズム

SLA の学習は、P の初期値の和が 1 となるように均等に設定する。式(10)に従い行動を確率的に決定し、評価 b に

基づき P を更新する。その手法として以下に示す L<sub>R-εP</sub> (Liner Reward - ε Penalty in S - model)<sup>3)</sup> のアルゴリズムを採用した。現在の状況を S<sub>c</sub> とする。

if b > 0 then α = 1, β = 0 and b = 1

if b = 0 then α = 0, β = -1 and b = 0

if b < 0 then α = 0, β = -1 and b = -1

ΔP<sub>jk</sub> = λab + μβ(1 - b)

if j = c then ΔP<sub>jk</sub>(t + 1) = P<sub>jk</sub>(t) + ΔP<sub>jk</sub>

Else

ΔP<sub>jk</sub>(t + 1) = P<sub>jk</sub>(t + 1) - ΔP<sub>jk</sub>/4

Endif

ここで λ, μ は学習係数と呼ばれ、1 > λ ≥ μ > 0 である。

## 4. 数値計算実験と考察

図 1 の様な FMS 工場モデルにおいて、AGV が 5 台、各 AGV に与えられる作業数を 5 つとして数値計算実験を行った。学習係数 λ, μ は、それぞれ 0.001, 0.00001 とし、学習回数を 500 回行った結果を図 2 に示す。この図からは学習を繰り返すことで、全体的に作業時間が減少していることがわかる。しかしながら、学習回数が 400 回を越してもなお、作業時間が 1000 を超えている周回が多数あるなど、周回完了時間にばらつきが非常に多いことがわかる。この原因としては、SLA の学習回数が 500 回と少ないことが考えられ、この学習回数を数万~数十万と増やすことで良い結果が得られると予想される。

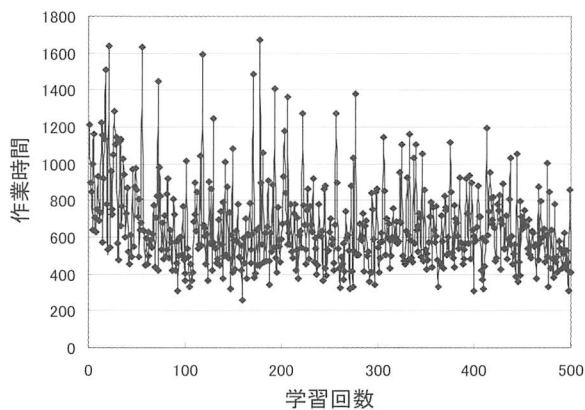


図 2. AGV の学習状況

## 5. おわりに

本研究では、一方通行の走行レーン上を走行する複数 AGV に関して、指定された全加工機械を巡回する総 AGV の周回完了時間を最小とするような運行スケジュールの数値実験を行った。AGV の自律的行動獲得のために SLA を採用し、周回完了時間が減少するのが確認された。

### 参考文献

- 1) 三上貞芳, 皆川雅章, 強化学習, 森北出版, (2000)
- 2) 古川正志, 渡辺美知子, 旭川工業高等専門学校研究報 文大 4 号, 旭川工業高等専門学校, pp99-109, (1998)
- 3) K. S.Naren dra and M. A. L Thathacher ; Leaning Automata An introduction, Prentice Hall (1989)