

要 旨

本報告では、自動倉庫に在庫している製品割り当て問題の各評価に乗じているウェイトを各評価に基づいて学習と更新を繰り返して、任意の環境に応じたウェイト値の自動調整を行う。また、製品割り当てと強化学習のための基礎研究としてウェイトの自動調整をGAに適用し、その有効性を数値実験で確かめる。

1. はじめに

立体自動倉庫に在庫している製品割り当て問題は、数理的に解が得ることが困難な組み合わせ最適化問題であるが、GAによる解法で最良解が得られることを報告してきた¹⁾。このGAに用いた評価は、多目的関数にウェイトを乗じたスカラー関数で表現されている。しかし、一般的にウェイト決定は、設計者の経験と勘に頼ることが多く、環境に応じたウェイトの自動調整が望ましい。

本研究では、自動倉庫に在庫している製品割り当て問題にGAを適用し、あるウェイト値で各評価の最良解を求める。次に、ウェイト調整として再度GAを適用して各評価の最良解を求め、これらの各評価に基づいて各ウェイトを更新し、最良ウェイトを求めることを目的とする。また、各ウェイト値調整の有効性を検証するために数値実験を行う。

2. 自動倉庫の製品割り当て問題

自動倉庫Gにおける製品割り当て問題は、ラック $G_{x,y}(x=1,2,3,\dots,m \quad y=1,2,3,\dots,n)$ で構成される立体自動倉庫とする。このラック内のパレット $X_{x,y}(x=1,2,3,\dots,m \quad y=1,2,3,\dots,n)$ 上に積載されている製品 K を割り当てる場合、以下の評価と制約条件で記述される。

- 最短経路出庫時間: $T_{i,j}$

$$\text{minimize } F_1 = \sum_{i=1}^m \sum_{j=1}^n 2 T_{i,j} X_{i,j} \quad (1)$$

$$T_{i,j} = l_{i,j} / v, \quad l_{i,j} = \sqrt{|x_i^2 + |y_j^2} / v$$

- 最短ピッキング作業時間: $Y_{i,j}$

$$\text{minimize } F_2 = \sum_{i=1}^m \sum_{j=1}^n Y_{i,j} X_{i,j} \quad (2)$$

$$Y_{i,j} = \begin{cases} P_{ick}^{all} \\ Q^* \times P_{ick} \end{cases}$$

- 最小出庫パレット数: $X_{i,j}$

$$\text{minimize } F_3 = \sum_{i=1}^m \sum_{j=1}^n X_{i,j} \quad (3)$$

- 出庫残最小: $Q_{i,j} - Q^*_{i,j}$

$$\text{minimize } F_4 = \sum_{i=1}^m \sum_{j=1}^n (Q_{i,j} - Q^*_{i,j}) X_{i,j} \quad (4)$$

- 製品 k の出庫要求数: S^k

$$\sum_{i=1}^m \sum_{j=1}^n S^k_{i,j} X_{i,j} \geq S^k \quad (5)$$

- 自動倉庫内の空ラック率: M

$$\sum_{i=1}^m \sum_{j=1}^n (1 - X_{i,j}) \geq M \quad (6)$$

3. ウェイトの自動調整

製品割り当て問題の評価 F は、上記の評価 $F_1 \sim F_4$ にウェイト w_d を乗じて任意の環境に応じた割り当てを行う。これを以下の式で示す。

$$F = \sum_{d=1}^4 F_d w_d \quad (7)$$

この評価に用いるウェイトは、人間の経験と勘に依存することが多く、自動的にウェイトを調整することが望まれる。以下にウェイトの自動調整を行う更新法とアルゴリズムを示す。

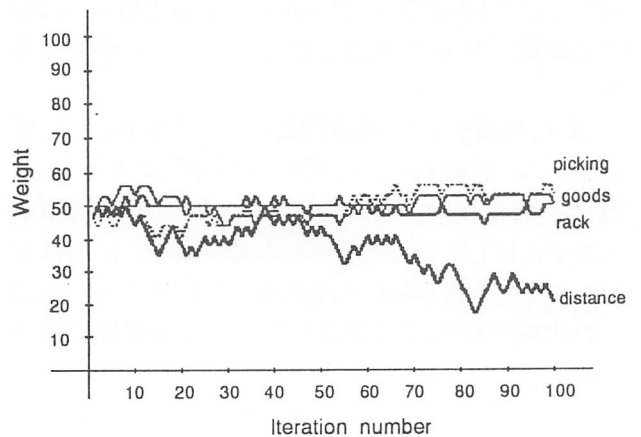


図 2. ウェイトの学習状況

[ウェイトの更新方法]

初期ウェイト値 w_d ($d=1,2,3,4$) と更新量を Δg_d ($d=1,2,3,4$) を与え、以下の更新を繰り返す。

$$\text{if then } F_{d(t)} < F_{d(t+1)} \\ w_d = w_d + \Delta g_d \quad (8)$$

$$\text{if then } F_{d(t)} < F_{d(t+1)} \\ w_d = w_d - \Delta g_d \quad (9)$$

このとき、 Δg_d は $F_{d(t+1)}$ が $F_{d(t)}$ より減少する w_d の符号を採用する。このフローチャートを図1に示す。

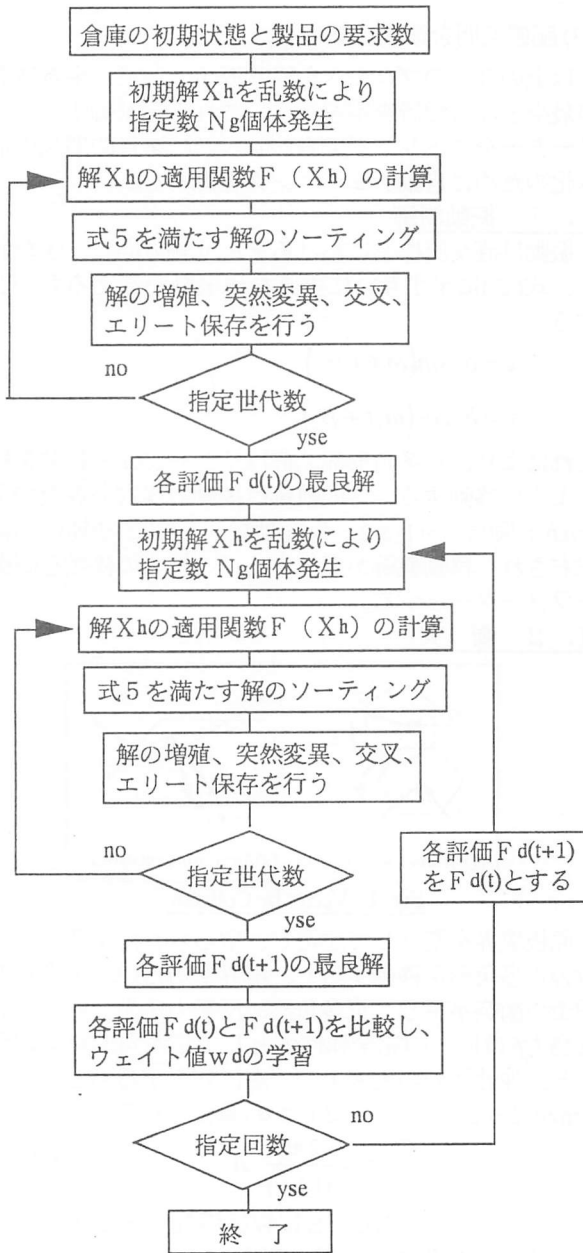


図1. ウェイト調整アルゴリズム

4. 数値実験

以下の条件で数値実験を行う。

- 自動倉庫ラック数：20 × 30
- 製品Kの在庫ラック：27

1ラックの製品量：1 ~ 30

製品要求数：154

初期個体数：50

突然変異：0.2

世代数：50

交叉：0.2

初期ウェイト値：50

学習回数：100

学習量：3

上記の条件で100回のウェイト更新を行った時の変化状況を図2のグラフに示す。このグラフで横軸は更新回数を示し、縦軸はウェイト値を示している。また、この図で各ウェイトは最短経路出庫時間をdistance、ピッキング作業時間をpicking、出庫パレット数をrack、在庫残goodsで示した。最短経路時間以外の評価は、各評価に基づいて成功の報酬と失敗の報酬が交互に現れ、初期ウェイト値の近傍を試行錯誤している。これは、常に最適値が得られているためと考えられる。最短経路出庫時間のウェイト値は、初期ウェイトより減少していくのが見られた。これは、この評価値が他のものよりかなり大きいため、それを補正しようとしているものと考えられる。

図3には、図2で求めた各ウェイト値を製品割り当て問題に適用したときの各評価値の変化状況を示す。この図では、コストが比較的小さいピッキング作業時間、出庫パレット数、在庫残のコストを20倍にして表示している。

5. おわりに

本報告では、自動倉庫に在庫している製品割り当て問題の評価に用いているウェイト値の自動調整を行う方法を示した。従来のウェイト値は、人間の経験と勘に頼るヒューリスティックな手法であったが、本手法を採用することにより自動的にウェイトの調整を行える可能性がある。

本研究では、強化学習によるウェイト調整のための予備的研究としてウェイトの調整が可能かどうかを調べた。今後、強化学習による方法を確立する予定である。

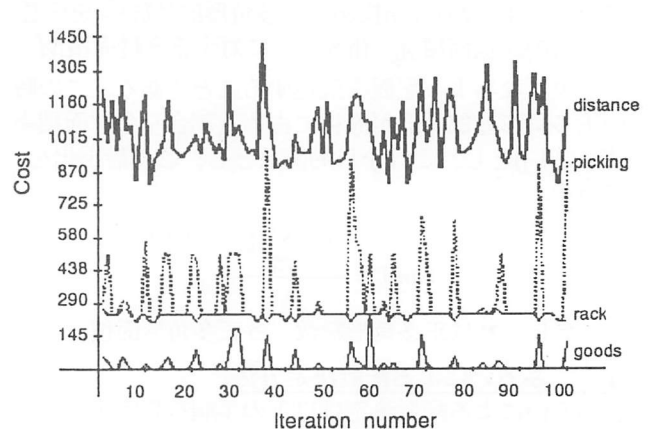


図3. 各評価のコスト状況

参考文献 1) 渡辺美知子他；流通C I Mシミュレータの基礎研究，1995年度精密工学会春季大会学術講演会講演論文集，1995