

# 遺伝的アルゴリズムと深層強化学習に基づく ジョブスケジュール最適化問題 —PCB加工プロセスをモデルとして—

汪 永豪<sup>†</sup> 山田 孝子<sup>‡</sup>  
関西学院大学<sup>†</sup> 関西学院大学<sup>‡</sup>

## 1 はじめに

本研究では、プリント基板(PCB) 製造工場を対象に、同一加工機械を複数回使用する生産プロセスをモデル化しスケジューリング最適化をした。静的な環境では遺伝的アルゴリズム(GA)と深層強化学習に基づく Deep Q-Network(DQN)を用い手法を性能比較した。さらに機械故障等が発生する動的な環境に柔軟に再スケジュールできる DQN の適応性能を検討した。

## 2 生産管理問題のモデル

本研究では受注生産する工場の製造工程、受注型ジョブショップスケジュール問題(Job Shop Scheduler problem:JSSP)を対象とし、JSSPを制約付き最適化問題として定式化する。目標関数としては:

$$\text{opt. } JSSP = \min \left\{ \max_s T J_s \right\} \quad (1)$$

ここで,  $T J_s$  は第  $s$  種別製品の最終加工工程の完了時刻を表す。

本最適化問題では各ジョブセットに含まれる全製品種別の最終加工完了時刻の最大値をスケジューリング案の終了時刻として求め、各スケジューリング案の終了時刻の中で最小終了時刻(makespan)のスケジューリング案を最適とする。

この JSSP は待機している加工機械に対して加工対象となるジョブを割り当てる逐次意思決定問題とみることができる。この意思決定過程は次の図 1 に示します。

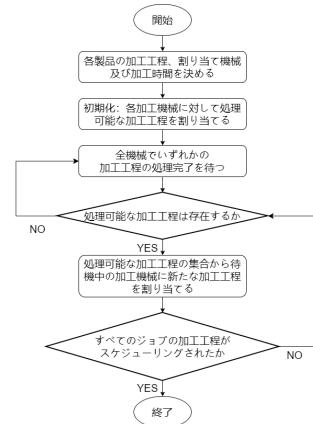


図 1: ジョブスケジューリング問題フローチャート

## 3 実験用データ自動生成

JSSP の検討には、クライアント側から提供される注文書が不可欠である。ただし実データの入手は困難であるので、本研究では先行研究 [3] の PCB 製造工程を参考に、各加工機械に対して加工役割を割り当てる人工的なデータを生成する。データ生成では注文書のばらつきやロット数を確率的に変動させた。

## 4 遺伝的アルゴリズムによる解の改善

本研究では、JSSP に対して GA を用いた最適解探索手法を採用する。

各染色体は、全ジョブに属する加工工程を一度ずつ含む遺伝子列として表現する。各遺伝子は「ジョブ  $J_s$  の第  $q$  工程」とし、制約を満たすようスケジュールを構成する。各ジョブは処理可能なジョブからランダムに工程を選択配置して初期染色体を生成する。各染色体デコードで得られた総加工時間 makespan を適応度として評価する。トーナメント選択により最良個体を選択し、親染色体を決定して制約を保持する交差手法として PPS(Precedence Preservative Crossover) という交差手法で子個体を生成する。以

Job Scheduling Optimization Based on Genetic Algorithms and Deep Reinforcement Learning  
—A Model of the PCB Manufacturing Process—

<sup>†</sup> Wang Yonghao, Kwansei Gakuin University  
<sup>‡</sup> Takako Yamada, Kwansei Gakuin University

上の進化計算を繰り返し、収束後に得られる最良染色体を最終解とする。最終解に対応するカントチャートと世代ごとの *makespan* の折れ線グラフで視覚的に評価する。

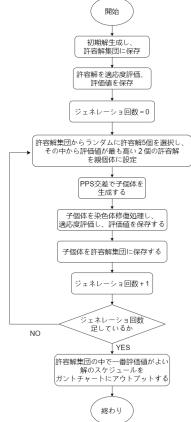


図 2: 遺伝的アルゴリズムフローチャート

## 5 深層強化学習による解の改善

本研究では価値関数に基づく強化学習手法である Deep Q-Network(DQN) を用いる。DQN では、各状態  $s_t$  における行動  $a_t$  の価値を表す行動価値関数  $Q(s_t, a_t)$  をニューラルネットワークにより近似し、各状態における行動選択を  $Q$  値に基づいて行う。

DQN のメインネットとターゲットネットとして、メインネットでは、状態  $s_t$  を入力として各行動の  $Q$  値を出力し、 $\varepsilon$ -greedy 法により実行行動  $a_t$  を決定するため用いる。また、経験再生でサンプルされた  $(s_t, a_t, r_t, s_{t+1})$  に対して、メインネット出力から  $Q(s_t, a_t)$  を算出し、誤差逆伝播によりメインネットのパラメータ更新を行う。一定ステップ間隔ごとにメインネットのパラメータをターゲットネットへコピーするハード更新を行う。固定テストデータ集合に対する定期評価では、評価値のばらつきを抑える目的からターゲットネットを用いて計算を行う。

## 6 実験結果

121 個のジョブセットに GA と DQN 両方により最適スケジューリングの *makespan* の比較は図 3 に示した通り、多くのケースで DQN の *makespan* は GA と同程度（比率 1.0 前後）で、一部では GA を上回る結果も見られるが、ジョブセットによっては性能が低下する場合も存在した。

散布図は 4 は横軸を *makespan* とし、縦軸をジ

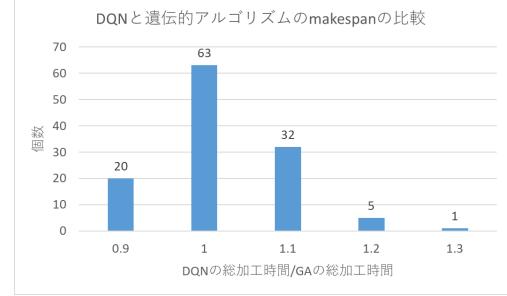


図 3: DQN と GA の *makespan* の比較

ヨブセットの縦軸とし、GA と DQN で得られた *makespan* の差を表現した。ジョブセットのプロセス処理時間の分散が小さいまたは大きいほど、DQN は GA に匹敵する性能が明らかにする。

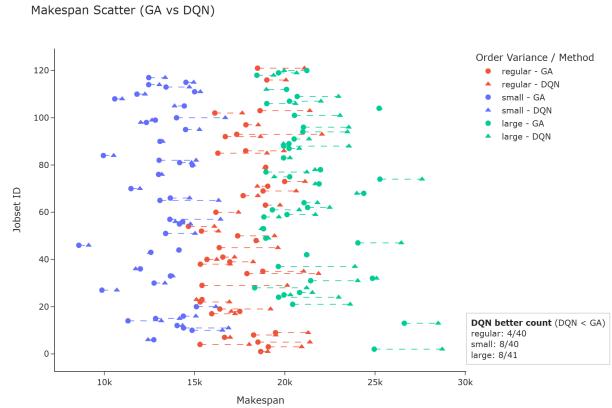


図 4: 同じジョブセットにより GA と DQN の *makespan* 比較

## 参考文献

- [1] S. Lee, Y. Cho and Y.H. Lee, "Injection Mold Production Sustainable Scheduling Using Deep Reinforcement Learning" *Sustainability*. vol. 12, 8718(2020).
- [2] L.B. Wang, X. Hu,..., "Dynamic job-shop scheduling in smart manufacturing using deep reinforcement learning" *Computer Networks*. vol. 190, 107937(2021).
- [3] Andrzej Kiernich, Jerzy Kalenik,..., "Impact of Particular Stages of the Manufacturing Process on the Reliability of Flexible Printed Circuits" *Sensors*.25, 140, (2025).