

深層強化学習を用いた生産計画技術の研究

Production Scheduling Based on Dispatching Rules with Deep Reinforcement Learning

多品種少量生産の製造業では、経営者と製造現場の両者を満足させるスケジュールの立案が求められている。しかし、そのようなスケジュールを立案するためのロジックの構築には、長年の経験やノウハウに基づく試行錯誤が必要であり、計画担当者の大きな負担となっている。そこで本研究では、経営者と製造現場の両者を考慮した評価指標を定め、深層強化学習により評価値の高いスケジュールを立案できる割付けルールを自動で構築する手法を提案した。小規模なベンチマーク問題を用いた実験により、提案手法は従来のルールよりも評価値の高いスケジュールを立案できることを確認した。今後は、研究成果の実務適用に向けて、計画規模の拡大や汎化性能の改善、機能の拡張を進める。

清藤 駿成 Seito Takanari
宗形 聡 Munakata Satoshi

1. はじめに

近年、顧客ニーズの多様化が進み、かつ製品のライフサイクルも短期化している。こうした状況の下、多品種少量生産の製造業では、資材投入から製品出荷までのリードタイム(LT)を短縮し、短納期への対応力を強化することが経営上の課題となっている。一方、製造現場では、製品や製造ラインごとの個別の制約を考慮して業務を効率的に実施することが課題である。製造業では、経営上の要求を満たし、かつ現場の作業を効率的に実行する生産スケジュールの立案が求められている。

生産スケジュールの立案には、割付けルールにしたがい作業を1つずつ設備に割付ける、ディスパッチング法が広く用いられている。既存の割付けルールは汎用的なものであるため、工場の計画担当者は、経営的な観点と現場的な観点を両立するスケジュールを立案できるように、複数のルールを組合せたり、製造ラインごとの個別制約をルールに加えたりしている。しかし、このような独自の割付けルールの構築には、経験やノウハウに基づく試行錯誤が必要であり、計画担当者の大きな負担となっている。人手不足でノウハウの伝承が困難となる中、計画担当者の負担を軽減するため、自社に適した割付けルールを自動で構築する技術が必要である。

そこで本研究では、経営的観点と現場的観点の両方を考慮したスケジュールの評価指標をもとに、高評価のスケジュールを立案するための割付けルールをAIに学習させる生産計画技術を提案する。

2. 従来手法

生産スケジューリングとは、有限の設備を用いて複数の作業を処理する際に、所与の制約を満たすように各作業の設備への割り当てを決める問題である。どの作業を、どの設備で、どの期間に処理するのかを決めることを作業の割付けという。制約には、各作業の処理順序の遵守や、設備で同時に処理できる作業は1つだけ、など多くの工場で共通する基本的な条件に加えて、段取り替えや設備の稼働時間、代替設備の有無など製造ラインごとの個別の条件がある。

生産スケジューリングの一般的な手法に、ディスパッチング法がある。この手法は、事前に構築した割付けルールにしたがい、作業を設備に1つずつ割付けていくことでスケジュールを立案する手法である。割付けルールとは、次に割付け可能な作業と、その作業を処理可能な設備のペア(以下、割付け候補と呼ぶ)の集合から、1つの割付け候補を選択するための規則である。実務で使用される割付けルールには、最早納期順、最小処理時間順、最大処理時間順、納期余裕時間順などがある。

しかし、これらの割付けルールは、どの工場でも使用される作業の納期、処理時間、余裕時間などを比較するだけの汎用的なものである。したがって、そのままでは工場ごとに異なる経営的観点や現場的観点を満足させるスケジュールの立案は難しい。そのため計画担当者は、複数のルールを組合せたり、製造ラインごとの個別制約をルールに反映させたりして、自社に適した独自のルー

ルを構築する。例えば、段取り替えを減らして作業の効率化を図る場合、同じ材料や治工を使うロットの作業を同一設備で連続して処理する個別制約を計画担当者はルール化する。これにより製造ラインでの効率的な作業スケジュールを立案できるが、LTの短縮などの経営的課題が改善するとは限らない。段取り替えが増えたとしても、いくつかのロットを別の設備で同時に処理する方が、全体のLTは短くなる場合もある。計画担当者は、経営的観点と現場的観点をバランスさせるために、ロットの属性に応じて作業を別の設備に割付けるルールとの組み合わせを検討する。しかし、どのような属性で作業を別の設備に割付けると良いのかのルール化は、計画担当者の経験やノウハウに基づく試行錯誤が必要であり、計画業務の大きな負荷となっている。また、一部のノウハウは担当者の暗黙知となっており、割付けルールとして表現すること自体が困難な場合もある。

以上のように、製造業では経営的観点と現場的観点を両方を満足させる割付けルールの構築が求められているが、そのようなルールの構築は高度なノウハウが要求される負荷の大きい業務となっている。計画担当者の負担を軽減するために、自社に適した割付けルールを自動で構築する技術が必要である。

3. 提案手法

そこで本研究では、強化学習²⁾を用いてAIに割付けルールを学習させる手法を提案する。強化学習とは、AIが自ら試行錯誤を行い、その経験の中から良い行動を見出し、学習していく仕組みである。これを生産スケジューリングに適用することで、スケジュール立案の試行錯誤の中から、経営的観点と現場的観点を両立したスケジュールを立案するための割付けルールを自動で学習できる。

3.1 深層強化学習による割付けルールの学習

まず、強化学習での割付けルールの表現方法を説明する。強化学習では、深層ニューラルネットワーク(以下、DNNと略す)を用いて割付け候補を選択する。DNNは、パラメータ θ を持ち、スケジュール sch の入力に対して方策 p を出力する関数である： $p = f(sch; \theta)$ 。方策 p とは、スケジュール sch での割付け候補 i の選択確率 π_i を並べたベクトルである。割付けルールは、割付け候補の集合から選択確率の最も高い割付け候補を選ぶことに該当する。したがって、割付けルールの学習とは、評価値の高いスケジュールを立案する割付け候補の選択確率を高くするように、DNNのパラメータを調整することに相当する。

```

input: 作業のリスト oprs
1.  $\epsilon = 1.0$  // 探索係数を初期化
2. for training_count = 1, ..., T:
3.   schs = [] // スケジュールのリストを初期化
4.   for schedule_count = 1, ..., S:
5.     sch = [] // スケジュールを初期化
6.     while is_not_completed(sch): // 完成かを判定
7.       r = get_random() // 一様乱数[0, 1]を取得
8.       if  $r \leq \epsilon$ : // ランダムな割付け
9.         {opr, rsc, period} = allocate_random(sch)
10.      else: // DNNを用いた割付け
11.        {opr, rsc, period} = allocate_dnn(sch)
12.        sch = sch + [{opr, rsc, period}]
13.      schs = schs + [sch] // スケジュールを追加
14.    targets = create_targets(schs) // 教師を作成
15.    update_dnn(targets) // DNNを更新
16.     $\epsilon = \max(\epsilon - 0.1, 0.1)$  // 探索係数を更新
    
```

図 1 強化学習のアルゴリズム

次に、強化学習のアルゴリズムを図1により説明する。全体としては、スケジュール立案を繰り返しながら、経営的観点と現場的観点を両方を考慮して定義した評価指標に基づき、高評価のスケジュールを立案した割付けの選択確率が高くなるようにDNNのパラメータを更新する流れとなっている。特にステップ14では、完成スケジュールの立案過程である割付けから方策の教師 π を作成する。評価値の高いスケジュールの割付けを学習するために、評価値が上位P個のスケジュールを用いた。教師 π は、実際に選択した割付け候補 i の選択確率 π_i を1とする、one-hotベクトルとした。ただし、同じ立案状況に対して異なる割付けを選択したスケジュールが複数存在する場合、評価値の順位 r に対して、 $\pi_i = (2/3)(1/3)^{r-1}$ とした。またステップ15では、DNNの出力した方策 p と教師 π の平均二乗誤差が小さくなるように、確率的勾配降下法であるAdamを用いて、DNNのパラメータ θ を更新する。

3.2 深層ニューラルネットワークの構成

本節ではDNNの構成を説明する。DNNが経営的観点と現場的観点を両方を考慮できるようになるには、入力されたスケジュールから、経営指標である製造KPIなどの全体情報と、個別の制約の充足状況を判断できる詳細情報の両方を特徴量として抽出できる必要がある。そのような特徴量を抽出するためには、作業の開始終了時間などの数値情報だけでなく、作業順序や割付け設備などの非数値情報を考慮する必要がある。しかし、従来のDNNは画像などの数値情報の入力が前提であり、非数値情報の処理はできない。したがって、数値情報と非数値情報の両方を取り扱うことのできるDNNが必要である。

そこで本研究では、入力となるスケジュールを頂点と辺から構成されるグラフとしてモデル化し、グラフを処理できるDNNであるグラフ畳み込みニューラルネットワーク³⁾(GCN)を用いた。これにより、スケジュールの数値情報だけでなく非数値情報も1つのモデルの中で表現できる。また、数値情報と非数値情報を用いた畳み込み処理により、スケジュールの全体情報と詳細情報の両方を考慮した特徴量を抽出できる。

(1) スケジュールグラフ

まず、スケジュールからグラフを構成する方法を説明する。入力されたスケジュールに対して、表1の頂点を生成し、表2の有向辺を頂点間に張ることでグラフを構成する。これにより、作業の開始終了時間などの数値情報は頂点の属性値、作業順序や割付け設備などの非数値情報は有向辺として表現できる。スケジュールをグラフに変換した例を図2に示す。例えば、辺A→Xは作業Aが設備Xで処理可能で実際に割付けられたこと、辺A→Fは設備Xでは作業Aの次に作業Fが処理されること、辺A→Bは工程の作業順序の遵守をそれぞれ表現している。

(2) グラフ畳み込みニューラルネットワーク

次に、スケジュールグラフを入力とする GCN の構成を説明する。GCN の処理を式(1)~(8)に示す。なお、パラメータ $\theta = (A, a, B)$ 、および GRU と LSTM のパラメータは、各添え字 (i, j, e, t, s) に対して異なるものを用いた。

【入力層】

$$h_i^0 = x_i A_i + a_i \quad (1)$$

【畳み込み層】

$$m_i^t(e) = \sum_{j \in \text{Adj}(i,e)} [h_i^{t-1}, h_j^{t-1}] B_e^t(i,j) \quad (2)$$

$$h_i^t = \text{RU}_i^t([m_i^t(e_1), \dots, m_i^t(e_{\#edge_i})], h_i^{t-1}) \quad (3)$$

【出力層】

$$r^0, q_k^0, c_k^0 = \mathbf{0} \quad (4)$$

$$q_k^s, c_k^s = \text{LSTM}_k^s(r^{s-1}, q_k^{s-1}, c_k^{s-1}) \quad (5)$$

$$e_k^s = h_k^{\text{cnv}} \cdot q_k^s \quad (6)$$

$$p^s = \text{Softmax}([e_1^s, \dots, e_{\#dsp}^s]) \quad (7)$$

$$r^s = \sum_{k \in \text{Dsp}} p_k^s h_k^{\text{cnv}} \quad (8)$$

入力層では、スケジュールの数値情報である各頂点 i の k 次元の属性値 x_i を用いて、 d 次元の特徴量 h_i^0 を計算する：式(1)。畳み込み層では、グラフの有向辺として表現された非数値情報を、各頂点 i の特徴量 h_i^t に反映する。 t は処理番号、 $\#edge_i$ は頂点 i に接続する辺の種類数である。

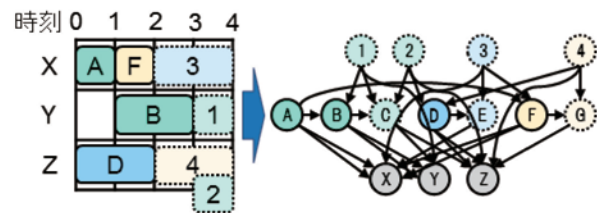
表 1 スケジュールグラフの頂点の種類

名称	属性値
未割付け作業	最早開始時間, 最早終了時間, 最早作業時間
割付け済み作業	開始時間, 終了時間, 作業時間
設備	処理効率*
割付け候補	最早開始時間, 最早終了時間, 最早作業時間

* 同一の機能を持つ設備の性能差の表現に使用する。

表 2 スケジュールグラフの有向辺の種類

名称	元頂点種類→先頂点種類
作業順序	作業→作業
前後作業	割付け済作業→割付け済作業
割付け設備	割付け済作業→設備
処理可能設備	作業→設備
割付け候補作業	割付け候補→未割付け作業
割付け候補設備	割付け候補→設備
割付け候補前後作業	割付け候補→割付け済作業



A,B,D,F : 割付け済み作業, C,E,G : 未割付け作業,
X,Y,Z : 設備, 1,2,3,4 : 割付け候補

図 2 スケジュールグラフの例

式(2)では、頂点 i から辺の種類 e を介して隣接する頂点 j の特徴量 h_j^{t-1} を、メッセージ $m_i^t(e)$ として集計する。式(3)では、各辺の種類 e のメッセージ $m_i^t(e)$ を用いて、頂点 i の特徴量 h_i^t を更新する。有向辺として表現された非数値情報は、隣接する頂点からのメッセージを受け取る処理に変換され、各頂点の特徴量に反映される。これにより、作業順序などの詳細情報が特徴量に抽出される。また、式(2)(3)を N_{cnv} 回繰り返すことで、直接隣接していない頂点の情報も伝播され、スケジュール全体の情報が特徴量に反映されるようになる。出力層では、式(4)~(8)を N_{out} 回繰り返す。割付け候補 k の特徴量 h_k^{cnv} から方策 p^{Nout} を計算する。 s は処理番号、 $\#dsp$ は割付け候補数である。本研究では、割付け候補 k の選択確率 π_k を、対応する特徴量 h_k^{cnv} だけでなく、他の割付け候補の特徴量も考慮して計算できるように、Set2Set⁴⁾ と呼ばれる手法を用いた。

以上により、グラフからスケジュールの全体情報と詳細情報を反映した割付け候補の特徴量を求め、経営的観点と現場的観点の両方を考慮した方策を計算できる。

4. ベンチマーク問題での検証

経営的観点と現場的観点の両方を考慮して学習した割付けルールが、従来の割付けルールよりも高評価値のスケジュールを立案できるかどうかを、ベンチマーク問題で検証した。ベンチマーク問題にはORLib⁵⁾のabz5(100作業, 10設備)を用いた。設備の処理効率は一律1.0とした。現場的観点の評価には、基本制約である作業の順序制約と設備制約を用いた。ただし本評価では、割付け候補の選出過程で基本制約は常に満たされるようにしている。経営的観点の評価には、LTの短縮化と関連するメイクスパンを用いた。メイクスパンとは、すべての作業が終了するまでの時間である。比較対象には、メイクスパンの最も短いスケジュールを作成した、最早着手可能時間順の割付けルールを用いた。図1のアルゴリズムにより割付けルールを学習し、各学習の終了時の割付けルールを用いたディスパッチング法によりスケジュールを立案し、メイクスパンを評価した。使用したハイパーパラメータを表3、および実験環境を表4に示す。

各学習回数でのスケジュールの評価値を図3に示す。学習全体に約40時間を要した。また、DNNを用いた割付けルールでの1回のスケジュールの立案には約6.2秒を要した。学習回数が増えるにしたがって、スケジュールの評価値が改善され、従来手法よりも評価値の高いスケジュールを立案できるようになっている。スケジュール立案の試行錯誤の中から、評価値の高いスケジュールを立案できる割付けを見出し、その選択確率が高くなるようにDNNを学習できたためであると考えられる。

5. おわりに

本研究では、経営的観点と現場的観点の両方を満足させるスケジュールを立案する割付けルールを自動で構築するために、深層強化学習を用いて割付けルールを学習する手法を提案した。100作業のベンチマーク問題を用いた検証の結果、学習した割付けルールは、従来の割付けルールよりも評価値の高いスケジュールを立案できた。

しかし、本検証は小規模なベンチマーク問題に対して提案手法の有効性を示しただけである。今後は、①実務レベルである1,000作業規模の問題に対しても、評価値の高いスケジュールの立案が実用的な時間でできること、②学習時と異なる問題に対しても評価値の高いスケジュールを立案できること、③現場的観点として製造ラインの個別制約を考慮して学習できること、が求められる。実用化に向けて引き続き研究開発を進めていく。

表3 ハイパーパラメータの設定

T	12	d	8
S	1,024	N_{cnv}	16
P	4	N_{out}	8

表4 実験環境

OS	Windows 7 Professional
CPU	Intel Core i7-6700, 3.40 GHz
メモリ	16.0 GB
論理コア数	8
DNNフレームワーク	PyTorch(1.0.1)

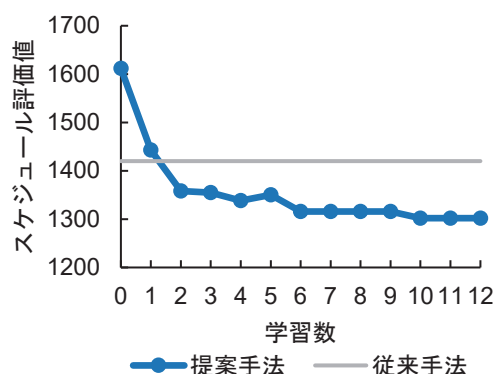


図3 各学習回数でのスケジュールの評価値

参考文献

- 1) M. L. Pinedo, Scheduling: Theory, Algorithms, and System (4th Edition), Springer (2016)
- 2) W. Zhang, et al., A Reinforcement Learning Approach to Job-Shop Scheduling, Proceedings of the 14th IJCAI, Vol. 2, pp. 1114-1120 (1995)
- 3) J. Gilmer, et al., Neural Message Passing for Quantum Chemistry, Proceedings of the 34th ICML, Vol. 70, 1263-1272 (2017)
- 4) O. Vinyals, et al., Order Matters: Sequence to Sequence for Sets, arXiv preprint arXiv:1511.06391 (2015)
- 5) ORLib (2019/7/18 閲覧), <http://people.brunel.ac.uk/~mastjbj/jeb/info.html>



清藤 駿成 2017年入社
 ビジネスインキュベーション部
 機械学習・AIを用いた意思決定支援
 技術、生産計画技術の研究開発
 takanari.seito.cw@hitachi-
 solutions.com



宗形 聡 2003年入社
 ビジネスインキュベーション部
 機械学習・AIを用いた意思決定支援
 技術、データ分析技術の研究開発
 satoshi.munakata.tu@hitachi-
 solutions.com