

タブーサーチ Q 学習による ジョブショップスケジューリング問題の解法

山田 俊生[†] 安達 雅春^{††}

[†] 東京電機大学工学研究科 ^{††} 東京電機大学工学部電気電子工学科

1. はじめに

近年、製造業において、製品の開発・生産期間の短縮化の傾向が強まっている。そのため、生産者側はスピードに富んだ生産システムの構築が要求されている。その中でスケジューリングは非常に重要な問題であり、本研究ではジョブショップスケジューリング問題 (Job-shop Scheduling Problem: JSP) について考える。JSP は先行研究[1]において、強化学習が用いられており、本研究では強化の改良版となるタブーサーチ Q (Tabu Search-Q: TS-Q) 学習法を用いた。

2. ジョブショップスケジューリング問題

JSP は順序関係がある複数の仕事を仕事ごとに複数の機械で処理する場合、全仕事が終了する時間を最短にする機械の稼働スケジュールを決める問題である。また、いくつかの制約が存在し、本研究では以下の 3 つの制約を有する問題を対象とする。

- (a) ある仕事の処理中に機械の稼働順序を変えてはいけない
- (b) ある仕事の処理を複数の機械が同時に行わない
- (c) ある仕事の処理中の機械が他の仕事の処理を行わない

JSP における稼働スケジュールモデルとしてガントチャートが多く使用されており、本研究においても、解の表現として、ガントチャートを作成する。

3. タブーサーチ Q 学習

TS-Q 学習法における Q 学習は強化学習の学習法の 1 つである。強化学習はエージェントと呼ばれる個体がある環境で行動を選び、その行動により環境が変化していく中でエージェントが報酬を獲得する。Q 学習ではある環境の状態である行動を選択した時の行動価値 Q 値から最良値を選び、行動を選択する。そして、Q 値はエージェントの行動後に報酬によって更新される。この一連の動作を 1 回の学習として指定した学習回数まで行い、指定した学習回数までを 1 エピソードとして指定したエピソード数まで行なう。

強化学習には最適解を求める際に局所最適解に陥ることがある。それを防ぐために TS を用いる。TS は最良値を選ぶ際にその遷移情報をタブーリスト (Tabu List: TL) に記録し、TL に存在する遷移を行わないようにする手法である。そして、この TS と Q 学習を組み合わせた手法が TS-Q 学習法である。

Q 学習の更新式を下式に示す。

$$Q(s_t, a_t) = (1 - \alpha) \times Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_a Q(s_{t+1}, a)]$$

ここで、 s は状態、 a は行動、 t は時間、 $Q(s_t, a_t)$ は Q 値、 α は学習率、 γ は割引率、 r は報酬である。

4. シミュレーション実験方法

本研究では JSP のベンチマーク問題となっている la01 (仕事数 10 × 機械数 5) を用いた。ガントチャートの最大時間は 1500 とした。エージェントと Q 値は各仕事に用意する。行動の選択は各機械の稼働時間の前進、後退、及び全機械の待機を含めた (機械数) × 2 + 1 である。報酬を獲得するための目的関数は状態値である。状態値は (仕事数) × (機械数) の行列に、制約 (c) に反する処理が行われてしまった場合、対応する機械に 1 を入れ、それ以外の機械に 0 を入れる。報酬は状態値の行列の数値の合計値が減少した場合、 $r = 5$ を与え、連続で行動した時に合計値が更に減少した場合、 $r = 10$ を与える。また、前進後退時に最大・最小時間 (0) から逸脱した場合と制約 (b) に反する処理が行われた場合は $r = -10$ を与える。

5. シミュレーション結果

TS-Q 学習によって得られた解の一例をガントチャートとして図 2 に示す。縦軸が仕事番号、横軸が時間、機械ごとに色分けして、無色は機械を稼働させていないことを示す。図 2 の仕事 1 の場合、左から機械番号 2, 1, 5, 4, 3 の順番で処理を行うことを示している。

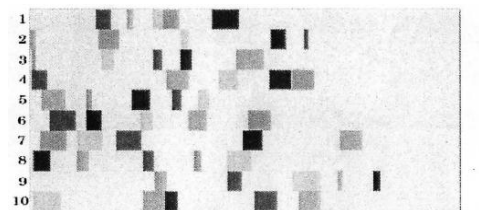


図 2. TS-Q 学習法のガントチャート

6. まとめ

本研究では TS-Q 学習の JSP への適用を試みた。この解法と ϵ -greedy 法をシミュレーションによって比較した結果、TS-Q 学習法の方が良好な解を得られる頻度が高いことを確認した。

参考文献

- [1] 岩村 幸治「自立分散型リアルタイムスケジューリングへのマルチエージェント強化学習の適用」システム制御情報学会論文誌, vol.26, No.4, pp.129-137, (2013).