

Title	フローショップ・スケジューリング問題への強化学習の適用
Author(s)	田中, 雄介
Citation	
Issue Date	1999-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1320
Rights	
Description	Supervisor:平石 邦彦, 情報科学研究科, 修士

フローショップ・スケジューリング問題への 強化学習の適用

田中雄介

北陸先端科学技術大学院大学 情報科学研究科

1999 年 8 月 13 日

キーワード： フローショップ・スケジューリング, 強化学習, 人工ニューラル・ネットワーク, 遺伝アルゴリズム.

本研究は, 強化学習の生産スケジューリング問題に対する, 適用可能性の考察を行う. そのために, フローショップ・スケジューリングに関して, ジョブの総処理時間最小化問題 ($F||C_{max}$ 問題と表記する) を考える.

先行研究では, 実用的観点から, ジョブショップ・スケジューリングに関して, ジョブの総処理時間最小化問題に対して強化学習を適用し, 従来法に比べてより良い実行可能解をより速い計算時間で獲得したという報告がある. 先行研究で成功した要因のひとつとして注目されるのは, 問題を強化学習問題として定式化する際, 種々のヒューリスティクスを参考にして, 探索空間の大幅な削減を図った点である.

他方, 本研究で対象にする $F||C_{max}$ 問題は, 2 機械問題の場合, ジョンソン則による最適解獲得が証明されている. しかし, 3 機械以上の場合, 総処理時間を最小化するスケジューリング方策は, 小規模な問題を除いては, 計算量の観点から求めるのが困難である. ジョブショップの特殊な問題であるフローショップを対象にするのは, 最適化則を参考にした定式化が可能になることと, その定式化を土台にして, 全数探索以外最適解が獲得できない問題領域への有効性も評価できる利点があるからである.

以上より, 本研究における目的を, 次のように設定する.

- いまだ全数探索以外最適解が獲得できない本問題に対し, 強化学習問題としての定式化を考案する.
- 実装評価により, その有効性を吟味し, 強化学習の適用可能性を評価考察する.

まず, ジョンソン則を参考に, 3 機械問題への適用を前提にして, 強化学習問題としての定式化を考案する. 強化学習の構成要素は, 環境, エージェント, 政策, 報酬, 価値関数の 5 つに分けられる. 環境は, 機械やジョブ集合である. エージェントは, 第 1 機械前の待ち行列からのジョブの選択を意思決定する.

ジョブ仕掛順の意思決定が, 強化学習の逐次意思決定へ定式化されるとき, 2 つの考え方がある. ひとつは, 全ジョブの仕掛順を決めたスケジュールを, 漸次改善していく過程を逐次意思決定とする考え方がある. しかし, 本研究では, 待ち行列前のジョブから次に仕掛るジョブを選択して, 全ジョブの処理が終了するまでの過程を逐次意思決定とする考え方を採った.

政策は, エージェントが知覚する状態に対して, その行為を決定する. エージェントの学習とは, 累積報酬を最大化する政策を獲得することだが, 状態と行為の定式化次第で, 学習結果が大きく左右される. 行為は, 第 1 機械前の待ち行列から, 仕掛けるジョブ 1 つを選択することになる. 問題を解く手掛かりが予め分かっている場合, 選択対象を絞り込むことがルール獲得に役立つ. そこで, 3 種類の選択候補となるジョブ集合を作った.

- a1) ジョンソン則から得られる仕掛順の規則性をそのまま利用したジョブ集合
- a2) ジョンソン則から得られる仕掛順の規則性を一部利用したジョブ集合
- a3) 待ち行列内のジョブが全て含まれるジョブ集合

次に, 状態記述として, 指標を 7 つ考案した. 前工程とは第 1, 第 2 工程を, 後工程とは第 2, 第 3 工程をさす.

- s1) 全ジョブ中, 待ち行列内の残りジョブ数
- s2) 待ち行列内のジョブのうち, 前工程時間が後工程時間より長いものの数
- s3) 前工程で, 全ジョブに対する, 待ち行列内ジョブの残り負荷
- s4) 後工程について s3 と同様
- s5) 待ち行列内のジョブで, 前工程処理時間が全ジョブ中最小のジョブが残っているかどうか
- s6) 後工程について s5 と同様
- s7) 後工程の機械稼働率

総処理時間最小化問題であるから, 報酬として, 最終状態における総処理時間の負値を与える. 価値関数の学習には $TD(\lambda)$ 法を用い, 人工ニューラル・ネットワークによる関数近似を行う. また, 強化学習単独では最適政策獲得までに多くの訓練回数を必要とするため, 遺伝アルゴリズムを組合せ, その優良解を価値関数に学習させることで, 訓練回数の短縮を図る強化学習を考案した.

予備的に 2 機械問題で, ジョンソン則により最適解が獲得される問題で解いた場合, s2 と a1 の状態と行為の定式化の組み合わせでは, ジョンソン則と同一解を得られることが確

認できた。また、それ以外にも複数の状態と行為の組合せで実験を行ったが、最適解の7%増程度の総処理時間を実現する学習結果も観測された。このように定式化が不完全でも学習できることは、全数探索以外最適解獲得ができない問題領域でも学習が有効となる可能性を示唆している。

さらに、3機械問題でも、2機械問題と同等の学習効果をあげることができた。しかし、2機械問題への読み替えを行ってジョンソン則により求めた解を上回ることができなかった。

また、遺伝アルゴリズムによる優良解を学習させた強化学習では、より少ない訓練回数で同程度の安定した解を得ることができた。

以上をまとめると、

- フローショップ・スケジューリング問題に対し強化学習問題としての定式化手法を述べた。
- ジョンソン則で最適解獲得ができる2機械問題と、全数探索以外最適解獲得ができない3機械問題を実装評価した結果、双方において上記定式化の有効性が確認できた。

今後の課題としては、強化学習の定式化の改善を図る努力を続けることと、他の確率的探索手法との比較を行うことなどが挙げられる。また、対象をフローショップからジョブショップへ、目的関数についてもジョブの総処理時間最小化以外へと拡張し、問題がより複雑になる中でも、学習能力のあるエージェントに関する研究が必要であるものと考えられる。