

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 11 日現在

機関番号：24403

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25820024

研究課題名(和文)5軸加工機のためのマルチエージェント強化学習を用いた自律分散型スケジューリング

研究課題名(英文)Autonomous Distributed Scheduling by Using Multi-agent Reinforcement Learning for 5-axis Machine Tools

研究代表者

岩村 幸治 (Iwamura, Koji)

大阪府立大学・工学(系)研究科(研究院)・准教授

研究者番号：40332001

交付決定額(研究期間全体)：(直接経費) 3,200,000円

研究成果の概要(和文)：機械加工の分野で生産リードタイムを短縮するための方法として、5軸加工機による工程集約型の高能率生産が注目されている。本研究では、種類の異なる5軸加工機およびジョブから構成される加工システムを対象として、マルチエージェント強化学習を用いた自律分散型のスケジューリングシステムの開発を行った。このシステムでは、各5軸加工機の仕様にカスタマイズされた5軸加工用CAMを用いて、加工可能なすべてのジョブのNCデータを作成し加工時間を求め、その情報に基づいて各5軸加工機におけるスケジュールを作成する。計算機実験を行った結果、リードタイムをはじめとする生産に関する目的関数を改善できることを示した。

研究成果の概要(英文)：Process integration achieved by 5-axis machine tools have received attention in recent years from the viewpoint of reduction of production lead time. New autonomous distributed scheduling systems by using multi-agent reinforcement learning are proposed in this research. The proposed scheduling systems consist of 5-axis machine tools, 3DCAD database, customized 3DCAM corresponding to types of 5-axis machine tools, and the jobs. The proposed scheduling systems carry out following 3 steps. At first, the individual 5-axis machine tools refer the 3DCAD database to obtain the 3DCAD data of jobs which can be carried out next machining process. Next, NC data and processing time of all jobs are created by 3DCAM. Finally, the individual machine tools and jobs carry out the autonomous distributed scheduling process by using multi-agent reinforcement learning. It is shown that objective functions of manufacturing systems are improved by using proposed system through the case studies.

研究分野：機械生産工学

キーワード：5軸加工機 マルチエージェント強化学習 自律分散型生産システム スケジューリング

1. 研究開始当初の背景

ものづくり産業の経営環境は厳しく、各企業はユーザーニーズに迅速に反応し市場に製品を投入するため、短いリードタイムで多品種少量生産を行う必要がある。機械加工の分野で生産リードタイムを短縮するための方法として、5軸加工機による工程集約型の高効率生産が注目されている。

5軸加工機は、左右、前後、上下の3軸に加えて傾く2軸も制御して、ジョブを加工する加工機であり、取付面を除いた5面を段取り無しで一度に加工できるため、工程集約による段取り時間の削減、同時多面加工による高精度維持、加工治具の省略・簡略化による治具費用の削減などに効果がある。しかし、各軸を工具側またはジョブ側のどちら側に配置するかは組合せにより、多くの種類の5軸加工機が存在する。

一方、5軸加工機で加工を行うためのNCデータの作成には、5軸加工用のCAMが用いられるが、同じ製品を作る場合でも5軸加工機の種類が異なる場合、または5軸加工機の種類が同じでも使用する工具の種類が異なる場合は、NCデータが異なるため加工時間は異なる。

5軸加工機に関する研究として、例えば、回転軸のアライメント誤差を正確に測定し、NCで補正することによって運動精度を向上させる研究(文献1)、5軸加工用CAMにおいて機械運動を考慮して工具経路を生成する研究(文献2)など活発な研究が行われている。しかし、従来の研究は5軸加工機的设计および単体の運用に注目した研究であり、複数の5軸加工機に注目して、全体として効率的に運用するための研究は十分行われていない。

2. 研究の目的

本研究課題では、図1に示すように、複数の種類の異なる5軸加工機およびジョブから構成される加工システムを対象として、5軸加工機を有効に活用することができるマルチエージェント強化学習を用いた自律分散型のスケジューリングシステムの開発を行う。

自律分散型生産システム概念は、多品種少量生産のための生産システムとして、1990年代から活発に議論されており、RFID(Radio Frequency Identification)など工場の情報・通信技術の発展に伴い近年実用化も進められている。本研究課題では、図1に示すように、各ジョブのRFIDを参照することでデータベースから得られる3DCADデータを用いて、各5軸加工機に対応するCAMから作成されたNCデータおよびその加工時間に基づいて、ジョブの生産リードタイムをはじめとする生産システムの目的関数を最適化するスケジューリングシステムの開発を行う。

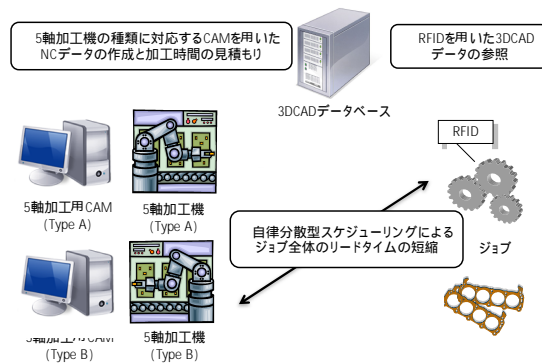


図1 5軸加工機のための自律分散型スケジューリングシステム

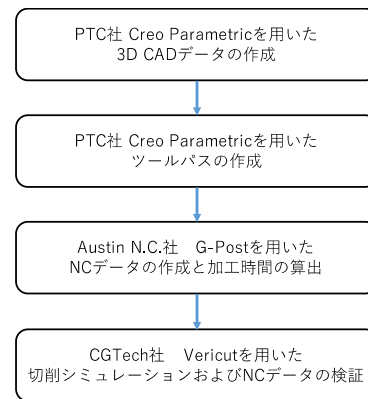


図2 NCデータの作成手順

3. 研究の方法

(1) 5軸加工機のための自律分散型スケジューリングシステムの開発

本研究課題では、図1に示す5軸加工機、5軸加工用CAM、ジョブ、RFIDおよび3DCADデータベースから構成する5軸加工機のための自律分散型スケジューリングシステムを開発する。ただし、5軸加工機に関しては、実際のハードウェアの代わりに、CGTech社の切削シミュレーションソフトであるVERICUTを用いた。システム開発にあたり、特に重要となるNCデータの作成と加工時間の算出方法について以下にまとめる。

3DCAMのカスタマイズ

5軸加工機で加工を行うためのNCデータを、3DCADデータから作成するためには、3DCAMおよびポストプロセッサが用いられるが、同じ製品を作る場合でも5軸加工機の種類が異なる場合、または5軸加工機の種類が同じでも使用する工具の種類が異なる場合は、NCデータが異なるため加工時間は異なる。そのため、5軸加工機の種類および所持する工具に応じて3DCAMおよびポストプロセッサをカスタマイズし、対象とするすべてのジョブのNCデータの作成と加工時間の算出を行う。

3DCADデータを用いたNCデータの作成と加工時間の算出

本研究課題では図2に示す手順で、3DCADデータからNCデータの作成と加工時間の算出を行っている。はじめに、PTC社の3DCAD・

CAM である Creo Parametric を用いて対象とするジョブの 3DCAD データの作成を行う。次に同ソフトを用いて、ツールパスの導出を行う。その後、Austin N.C. 社のポストプロセッサである G-Post を用いて、各工作機械に対応した NC データの作成と加工時間の算出を行う。最後に、VERICUT を用いて切削シミュレーションを行い、NC データの検証を行う。

(2) スケジューリングプロセス

研究代表者らの従来の研究(文献 3)では、ジョブエージェントおよびリソースエージェントが、自身の目的関数であるジョブの滞留時間および加工コスト、リソースの加工精度、稼働率を最適化するための次の加工プロセスのスケジュールを、効用値に基づいて決定する手法を提案している。この手法では、待機状態にあるジョブエージェントとリソースエージェントが、ジョブエージェントの次の加工プロセスを実行するリソースエージェントを決定する。以下、単にエージェントとよぶ場合は、ジョブエージェントおよびリソースエージェントを示す。

次に実行する加工プロセスを決定するために以下の四つのプロセスを実行する。

(プロセス 1) エージェントの状態の収集

各エージェントは、時刻 t における他のエージェントの状態に関する情報を収集する。すなわち、ジョブエージェントは各リソースエージェントが加工可能な加工プロセスを、リソースエージェントは各ジョブエージェントのまだ実行されていない加工プロセスを求める。

(プロセス 2) 候補エージェントの選択

各エージェントは、次の加工プロセスを実行するために候補となるエージェントを求める。すなわち、ジョブエージェントは、次の加工プロセスを加工することができるリソースエージェントを求める。リソースエージェントは、まだ実行されていない加工プロセスの中で実行することができる加工プロセスを求め、その加工プロセスを持つジョブエージェントを候補とする。

(プロセス 3) 効用値の算出

各エージェントは、(プロセス 2) で求めた候補エージェントについて、候補エージェントを次の加工プロセスとして実行した場合の目的関数の値に基づいて、自身の目的関数に対応する効用値を求める。目的関数の性質および単位が異なるため、各エージェントは効用値を 0 から 1 の範囲で求める。

(プロセス 4) 協調

各エージェントは、候補エージェントの効用値をコーディネーションエージェントに送る。コーディネーションエージェントは、各エージェントからの効用値に基づいて、次の加工プロセスを実行するためのリソースエージェントとジョブエージェ

ントの組合せを決定する。この時、コーディネーションエージェントは、効用値の和が最大となるリソースエージェントとジョブエージェントの組合せを求める。

ジョブエージェントおよびリソースエージェントは、以上の四つのプロセスをすべての加工プロセスが終了するまで繰り返し実行し、スケジュールを作成する。

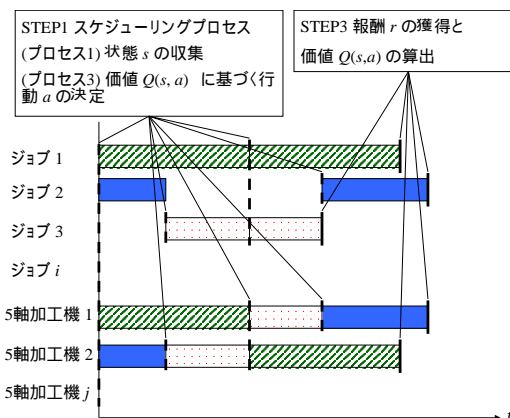


図 3 マルチエージェント強化学習の適用

(3) マルチエージェント強化学習を用いたスケジューリングプロセス

本研究課題では、(2) 節に示した効用値に基づくスケジューリングプロセスにおいて、マルチエージェント強化学習を適用することにより、エージェントが生産システムの状態に応じて、適切に効用値の算出方法を切り替えることができる手法を提案する。なお、本研究課題では、リソースエージェントは 5 軸加工機エージェントに相当する。

図 3 は、本研究課題で提案するジョブエージェントおよび 5 軸加工機エージェントへのマルチエージェント強化学習の適用方法を、ジョブエージェントが三つ、5 軸加工機エージェントが二つの場合の例におけるガントチャートを用いて示している。ガントチャートの横軸は時間軸を、縦軸はジョブエージェントおよびリソースエージェントを示し、加工プロセスを実行するエージェントの組合せの時間的流れを棒図表で示している。また、棒図表における色の違いはジョブエージェントの種類をそれぞれ示している。

各エージェントは、以下の 6 つの STEP を実行することにより、強化学習に基づいて適切な効用値の算出基準を獲得することができる。

STEP0 この STEP は、強化学習を行う場合の報酬 r を算出する基準を作成するために行う。ジョブエージェントおよび 5 軸加工機エージェントは、(2) 節に示した効用値に基づくスケジューリング手法を適用しスケジュールを作成する。ここで得られたスケジュールを基準とし、STEP1 から STEP3 の強化学習プロセスを実行する。

また、各エージェントは、強化学習に用い

る状態 s と行動 a の組合せを示すルール (s, a) に対する価値 $Q(s, a)$, 報酬の累積値 $SumReward(s, a)$ および報酬を獲得した回数 $RewardCount$ を 0 に初期化する。

STEP1 ジョブエージェントおよび5軸加工機エージェントは、(2)節に示した効用値に基づくスケジューリングプロセスの(プロセス1)から(プロセス4)を実行する。ただし、強化学習の適用のため、(プロセス1)および(プロセス3)を、それぞれ以下のように置き換える。

(プロセス1) エージェントの状態の収集

各エージェントは、時刻 t における他のエージェントの状態に関する情報を収集する。また、強化学習に用いる生産システムの状態 s を観測するため、目的関数ごとに待機中のエージェントの数を収集する。

(プロセス3) 効用値の算出

各エージェントは、候補エージェントの効用値を決定するために、状態 s に対する行動 a の価値 $Q(s, a)$ に基づいて、 ε グリーディ手法(文献4)を用いて行動 a を決定する。 ε グリーディ手法は、確率 $1-\varepsilon$ で、一番大きい価値を持つ行動を選択し、確率 ε ($0 < \varepsilon < 1$) でランダムな行動を選択する。また、 ε グリーディ手法により決定された行動 a に対応するルール (s_t, a_t) を保存する。

STEP2 5軸加工機エージェントが、ジョブエージェントのすべての加工プロセスを終了するまでSTEP1のスケジューリングプロセスを実行する。

STEP3 ジョブエージェントおよび5軸加工機エージェントは、スケジューリングが終了した時点における自身の目的関数の値から、報酬 r を得る。また、得られた報酬 r を用いて、モンテカルロ法(文献4)に基づいて価値 $Q(s, a)$ を算出する。

モンテカルロ法では、エージェントは、報酬を得た時点から、つぎに新しい報酬を得る時点までの g 個のルール (s_t, a_t) ($t=0, 1, \dots, g-1$) を保存する。ここで、ルール (s, a) は状態 s と行動 a の組合せを示す。この時、価値 $Q(s, a)$ は、以下の式から算出する。

$$SumReward(s_t, a_t) = SumReward(s_t, a_t) + r \quad (1)$$

$$Q(s, a) = SumReward(s, a) / RewardCount \quad (2)$$

ここで、 $SumReward(s, a)$ および $RewardCount$ は、ルール (s, a) における報酬の累積値、およびルール (s, a) において報酬を獲得した回数である。

STEP4 新しく生産システムに投入されたジョブエージェントに対してSTEP0からSTEP3を実行し、スケジューリングを行う。
STEP5 STEP4を繰り返し、各エージェントの状態 s に対する行動 a の実行と報酬 r の算出を繰り返し行うことで価値 $Q(s, a)$ を収

束させる。

(4) エージェントの状態・行動・報酬

(3)節に示した手順において、STEP1の(プロセス1)における状態 s , STEP1の(プロセス3)における行動 a , およびSTEP3における報酬 r の与え方が、強化学習の観点から重要となる。これらの与え方を以下に示す。

状態 s

ジョブエージェントおよび5軸加工機エージェントが観測する生産システムの状態 s を、以下の式に示す。

$$s = (s_1, s_2, s_3, s_4) \quad (3)$$

ここで、 s_{type} ($type=1, 2, 3, 4$) は、それぞれ、稼働率、加工精度、滞留時間、加工コストの目的関数を持つ待機中のエージェントの数である。

行動 a

ジョブエージェントおよび5軸加工機エージェントは、行動 a を実行するために、以下の式における n ($n=1/5, 1/3, 1, 3, \text{または} 5$) の値を決定する。

$$UV' = (UV)^n \quad (4)$$

ここで、 UV は、ジョブエージェントおよび5軸加工機エージェントが算出した効用値である。すなわち、(4)式を適用することにより、待機中の各目的関数のエージェントの数に基づいて、効用値を修正することができる。

報酬 r

STEP3におけるスケジューリングが終了した時点で、ジョブエージェントおよび5軸加工機エージェントは、目的関数の値から報酬 r を算出する。

STEP0で求めた基準スケジュールの結果とSTEP1からSTEP3で求めた強化学習を適用したスケジュールの結果を比較し、目的関数の値が改善された場合に正の報酬を、そうでない場合に負の報酬を与える。報酬 r は、以下の式から求める。

$$r = \sum_{h=1}^r r_h / \quad (5)$$

ここで、 r はエージェントの総数を表す。また、 r_h は目的関数の種類ごとに算出されたエージェント h の報酬の値を表す。稼働率は最大化、加工精度、滞留時間および加工コストは最小化の目的関数であるため、ジョブエージェントおよび5軸加工機エージェント h は、自身の目的関数の種類に基づいて、以下の式から報酬 r_h を算出する。

(a) 目的関数が稼働率の場合

$$r_h = (proOFV_h - preOFV_h) / preOFV_h \quad (6)$$

(b) 目的関数が加工精度、滞留時間、または、加工コストの場合

$$r_h = (preOFV_h - proOFV_h) / preOFV_h \quad (7)$$

ここで、 $proOFV_h$ および $preOFV_h$ は、(3) 節に示した STEP1 から STEP3 を実行し、強化学習を適用することにより得られた目的関数の値、および STEP0 を実行し強化学習を行わない場合により得られた目的関数の値を表す。

4. 研究成果

ケーススタディ

提案した手法の有効性を検証するために、スケジューリングのケーススタディを行った。生産対象として、二つの加工プロセスを実行することで完成品となるジョブエージェントが 16 個の場合、20 個の場合、30 個の場合の 3 通りを考える。これらのジョブは、いずれもスケジューリングの開始時点で加工可能な状態であるものとする。本研究では、同じジョブを繰り返し生産する場合を想定し、学習を繰り返すことによる目的関数の値の変化で学習の効果を評価する。

また、対象とする生産設備として、5 軸加工機 10 台を持つ生産システムを考える。また、各 5 軸加工機エージェントはすべてのジョブの加工プロセスを実行することができるが、加工プロセスを実行する場合の加工時間は、(1) 節に示した方法により得られた NC データに基づく加工時間を用いる。また、加工コストについては加工時間に比例する値を、加工精度はランダムでそれぞれ与えた。

種類および所持する工具の異なる 5 軸加工機エージェントに対して、12 通りのケースを作成し実験を行った。なお、 ε グリーディ手法における ε は、0.2 とした。

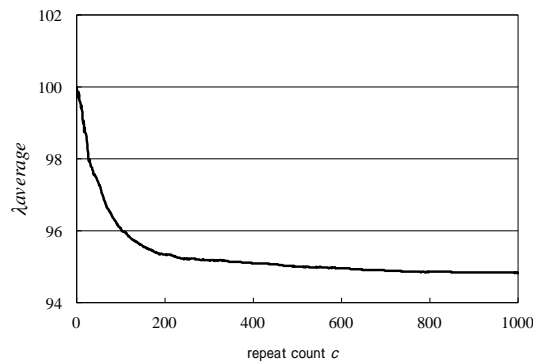


図4 実験の結果

図4に結果を示す。同図において、横軸および縦軸は、それぞれスケジューリングの繰り返し回数 c および全 12 ケースにおける平均改善率 $\lambda_{average}$ の平均値を示している。平均改善率 $\lambda_{average}$ は、スケジューリングの繰り返し回数 c が c^* に到達する時点の改善率 r_h の平均値を示し、以下の式から求める。

$$\lambda_{average} = \sum_{c=1}^{\omega} r_h \quad (8)$$

ここで、 c は c 回目のスケジューリングに

おける改善率 r_h の値である。

改善率 r_h は、すべてのエージェントの目的関数の値について、提案手法により得られた値と、学習を行わない従来手法により得られた値との比率を表し、以下の式から求める。

$$r_h = \sum_{h=1}^v \mu_h / \mu_h \quad (9)$$

ここで、 μ_h および μ_h は、エージェント h の目的関数の値の改善率、およびエージェントの総数を表す。 μ_h はエージェントの目的関数の種類に基づいて以下の式から求める。

(a) 目的関数が稼働率の場合

$$\mu_h = preOFV_h / proOFV_h \quad (10)$$

(b) 目的関数が加工精度、滞留時間、または、加工コストの場合

$$\mu_h = proOFV_h / preOFV_h \quad (11)$$

ここで、 $proOFV_h$ および $preOFV_h$ は、提案手法および従来手法を用いた場合のエージェント h の目的関数の値を示す。

同図からわかるように、従来手法と比較して目的関数の値を改善できていることがわかる。

まとめ

本研究課題では、5 軸加工機を活用するためのマルチエージェント強化学習を用いた自律分散型スケジューリングシステムの開発を行った。研究成果と今後の課題を以下にまとめる。

- 5 軸加工機、5 軸加工用 CAM、ジョブ、RFID および 3DCAD データベースから構成する 5 軸加工機のための自律分散型スケジューリングシステムを開発した。5 軸加工機の種類および所持する工具の情報に基づいてカスタマイズされた 3DCAM およびポストプロセッサを用いて、NC データと加工時間を算出した。
- 5 軸加工機のための自律分散型スケジューリングシステムとして、効用値に基づくスケジューリングプロセスを拡張し強化学習を適用することにより、ジョブエージェントおよび 5 軸加工機エージェントが、生産システムの状態に応じて、適切に効用値の算出方法を切り替えることができる手法を提案した。
- ケーススタディを行い提案手法の有効性を検証した。この結果、マルチエージェント強化学習を行うことにより、従来手法と比較して適切な効用値の算出基準を獲得できることを示した。
- 以上から、開発した自律分散型スケジューリングシステムを用いることで 5 軸加工機を有効に利用し、生産に関する目的関数を最適化するためのスケジュールを作成することができることを示した。

<引用文献>

高山直士,他3名,5軸制御マシニングセンタの精度向上に関する研究(旋回軸のアライメント誤差の測定とその補正),日本機械学会論文集(C編),Vol.77, No.782, pp.3951-3960,(2011).
金子誠,他1名,回転2軸の動きを考慮した5軸制御加工用工具経路生成法,精密工学会誌,Vol.76, No.1, pp.101-105,(2010)
岩村幸治,他2名,ホロニック生産システムにおけるリアルタイムスケジューリングに関する研究:第1報 効用値に基づく意思決定プロセス,日本機械学会論文集C編, Vol.69, No.682, pp.1733-1738(2003)
高玉圭樹,マルチエージェント学習 相互作用の謎に迫る,コロナ社(2003)

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計1件)

岩村幸治,真弓宗久,谷水義隆,杉村延広,自律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, 査読有,
<http://doi.org/10.5687/iscie.26.129>

[学会発表](計20件)

田宮隆宏,岩村幸治,平林直樹,杉村延広,3Dプリンタおよび5軸加工機の有効利用のための作業配分およびスケジューリングに関する研究,生産システム部門研究発表講演会2016,2016年3月14日,東京理科大学(千葉県・野田市).
菅野翼,岩村幸治,谷水義隆,杉村延広,総納期ずれ最小化のためのバックワードおよびマルチエージェント強化学習を用いたスケジューリング手法,第58回自動制御連合講演会,2015年11月14日,神戸大学(兵庫県・神戸市).
Jie CHEN, Koji IWAMURA, Yoshitaka TANIMIZU, Nobuhiro SUGIMURA, Development of Simulation Model for Transportation Processes of Autonomous Distributed AGV Systems by using AnyLogic, AnyLogic Conference 2015, November 5, 2015, Philadelphia (USA).
Jie Chen, Koji Iwamura, Yoshitaka Tanimizu, Nobuhiro Sugimura, Development of Simulation Model for Autonomous Distributed AGV Systems by Applying Social Force Model, The 34th JSST Annual Conference International Conference on Simulation Technology,

October 14, 2015, Toyama International Conference Center (Toyama・Toyama).

岩村幸治,石井大皓,田宮隆宏,谷水義隆,杉村延広,5軸加工機のための3DCAMを用いたスケジューリングシステムの開発,スケジューリング・シンポジウム2015,2015年9月27日,青山学院大学(東京都・渋谷区).

陳杰,岩村幸治,谷水義隆,杉村延広, Social force モデルを用いた自律分散型 AGV システムの搬送シミュレーション,生産システム部門研究発表講演会2015,2015年3月16日,慶應義塾大学(神奈川県・横浜市).

石井大皓,岩村幸治,谷水義隆,杉村延広,5軸加工機の効率的な運用のための3DCAMを用いたスケジューリングシステム,生産システム部門研究発表講演会2015,2015年3月16日,慶應義塾大学(神奈川県・横浜市).

Koji Iwamura, Takayuki Yokote, Yoshitaka Tanimizu, Nobuhiro Sugimura, Autonomous Distributed Scheduling by Using Multi-Agent Reinforcement Learning for Minimizing Sum of Earliness and Tardiness, ISCIE/ASME International Symposium on Flexible Automation 2014, July 14, 2014, Awaji Yumebutai International Conference Center (Hyogo・Awaji).

Kentaro Tanaka, Koji Iwamura, Yoshitaka Tanimizu, Nobuhiro Sugimura, A Study on Tool Allocation Method Considering Load Balancing and Optimization of Transportation, ISCIE/ASME International Symposium on Flexible Automation 2014, July 14, 2014, Awaji Yumebutai International Conference Center (Hyogo・Awaji).

岩村幸治,森田浩平,谷水義隆,杉村延広,歩留まりと作業時間のばらつきを考慮した作業者の工作機械への配置とスケジューリング,生産システム部門研究発表講演会2014,2014年3月17日,電気通信大学(東京都・調布市).

森田浩平,岩村幸治,谷水義隆,杉村延広,加工時間の確率分布を考慮した工作機械への作業配置とスケジューリング,日本機械学会2013年度年次大会,2013年9月10日,岡山大学(岡山県・岡山市).

6. 研究組織

(1)研究代表者

岩村 幸治 (IWAMURA, Koji)
大阪府立大学・工学研究科・准教授
研究者番号:40332001