

ロボットとAIによる 新生産システムの開発と量産適用

谷川 兼一
山田 圭祐

石川 琢磨
宮井 敦司

高橋 千優

労働人口の減少による人手不足を背景に、ロボットを活用した生産の自動化は、より活発化している。OKIデータでも、内製自動装置やロボットを用いた自動生産システムを開発し、生産に対する「作業の自動化」に取り組んできた。しかし、従来の自動生産システムでは、技術者が逐次動作を指定することによりロボットの行動シーケンスを作成しなくてはならなかったため、複数の自動装置からなる工程に複数の製品が同時に流れる生産ラインの場合は、考えられる全ての組合せを技術者が指定する必要があった。すなわち、技術者の負担が増大してしまうため、現実的ではなかった。

このような課題を解決すべく、実際の工場を模した仮想工場をコンピューター上に構築し、その仮想工場で想定されるあらゆる工程の組み合わせをもとに最適なロボットの行動シーケンスを自動で学習する「思考の自動化」を実現する生産システムを開発し、実際の生産ラインに導入した。「思考の自動化」を実現したことで、生産システムが自動で学習することができるため、技術者の負担を軽減することができるようになった。本稿では、このロボットとAIによる生産システム（以下、本生産システム）を紹介する。

従来自動生産システムの課題

図1に、本生産システムを導入した実工場での自動生産工程の概略図を示す。形成、組立、固着、特性検査を担当する4台の個別自動装置と1台のロボットにより構成される。ただし、組立工程では、一部の生産をロボットが担当する。各工程間の搬送はすべて1台のロボットが行う。各個別自動装置には、最大で5個のキャリア置き場がある。ここでキャリアとは、工程を流れる製品を複数固定することができる治具である。工程入口にあるキャリアがすべての生産工程を経由して出口へ移動すれば、1バッチ分の製品が生産完了となる。図1に示すように、1台のロボットが4台の自動装置へのキャリアの搬送を担当するため、ロボットから見ると常に複数の行動選択肢が存在している。

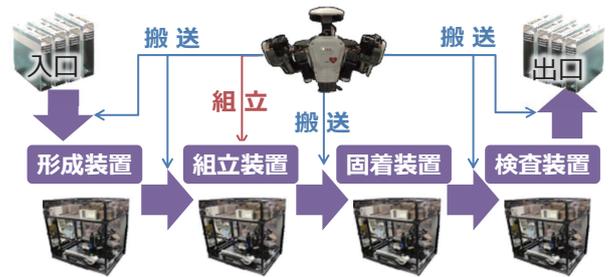


図1 実工場での自動生産工程の概略図

図2を用いて、技術者がロボットに教示すべき行動数を考える。工程の入口には80カ所のキャリア置き場があり、形成装置には5カ所のキャリア置き場がある。各装置間を結ぶ矢印はロボットによるキャリアの搬送行動を示す。入口のキャリア置き場「1」からは、形成装置のキャリア置き場「1」～「5」の5カ所へ搬送できる。入口のキャリア置き場は80カ所あるので、入口と形成装置間のロボットによるキャリア搬送に対応する行動数は400ステップとなる。同様に考えると、実工場全体でロボットの総行動数は835行動となる。

実際にはロボットは、刻々と変化するキャリアの配置状況に応じ、その時点で最適な行動を835通りの中から選択する必要がある。各キャリア置き場には、搬送ができるか否かの2状態が存在する。例えば、前の工程が工程完了後の状態であり、その後の工程のキャリア置き場にキャリアがない状態の場合、搬送できる状態である。図2に示すように、本生産システムではキャリア置き場の総数は176であるため、キャリア置き場の状況は 2^{176} パターン存在する。

したがって技術者は、ロボットに対して $2^{176} \times 835$ パターンの中から状況に応じた最適な行動シーケンスを指定する必要があり、さらにエラー停止や復旧した場合の行動なども考える必要があるため、考える全ての行動を技術者が指定して「作業の自動化」を実現することは困難であった。

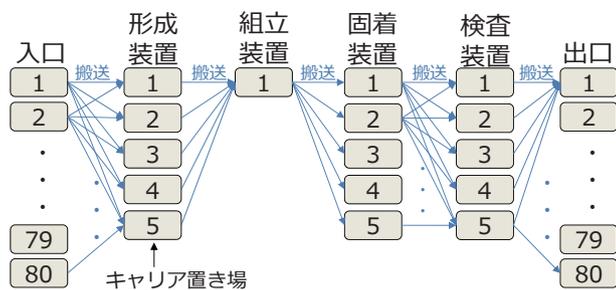


図2 自動生産工程におけるキャリア配置

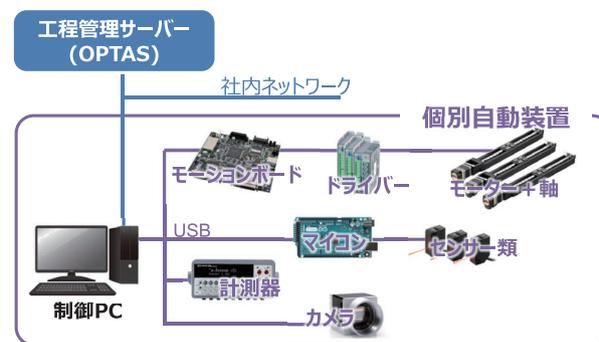


図4 個別自動装置のシステム構成図

本生産システムの構成

図3を用いて本生産システムの構成を説明する。本生産システムは、実工場、仮想工場、判断部、工程管理サーバー（以下、OPTAS^{*1)}）で構成され、生産の「作業の自動化」を達成する際に生じる技術者の負担を、仮想工場「思考の自動化」により大幅に軽減できる。

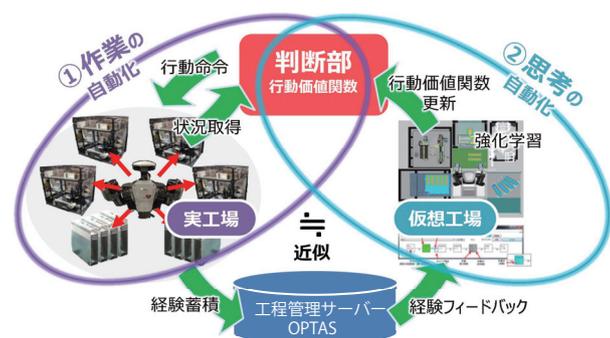


図3 ロボットとAIによる生産システムの概略図

実工場は、すでに図1に示したとおり、複数の個別自動装置とロボットにより構成されている。各工程の生産は個別自動装置が行い、ロボットは一部の組立て作業と搬送を行うことで、複数工程の無人化を実現した。ロボットには人型の次世代産業用ロボットであるNEXTAGE（カワダロボティクス製）¹⁾を用いた。

図4に個別自動装置のシステム構成図の一例を示す。必要な動作はモーターと駆動軸の構成により実現される。すべての個別自動装置にはカメラが搭載され、画像連動動作制御、製品の位置合わせ、検査、判別、認識など多岐にわたる機能を実現している。また、計測器やセンサーを搭載し、必要に応じて検査することもできる。これらすべての要素は、制御PCにより制御される。制御PCは社内ネットワークに接続され、OPTASや判断部と通信できる。本個別自動装置を社内内で内製開発したことにより、各工程を自動化することができた。

*1) OKI data Production control and Total Analysis Systemの略。

判断部は、実工場からキャリアの配置などの状況を取得し、状況に応じた最適な行動命令を出すことで、実工場全体を統括的に制御する。

仮想工場はコンピューター上で実工場の作業をシミュレーションすることで、多くの技術者と時間を要するさまざまな生産工程の組合せをソフトウェア的に実現し、短時間でロボットの動作を学習させるための環境である。最新の実工場のデータを仮想工場にフィードバックすることにより現実に近い学習ができる。

本生産システムでは、実工場でのロボット以外のセンサーやカメラなどから得られる位置情報を外部変数としてロボットにフィードバックすることができる。この外部変数を参照し、ロボットに教示した基本動作の位置座標を現在の状況に応じて変化させることができる。そのため、835通りの行動は、5通りの基本行動の教示で再現でき、ロボットの動作教示にかかる技術者の負担を大幅に削減することができた。

しかし、キャリア置き場の状況が 2^{176} 通りであることには変わりがない。これら膨大な組合せに対応した最適行動を仮想工場により自動的に学習する「思考の自動化」によって、「作業の自動化」を実現している。

「思考の自動化」：行動価値関数と強化学習

判断部が状況に応じて最適な行動命令を下すためには、複数の行動選択肢の中から、ある状況に応じた最適な行動が一つ選択できれば良い。本生産システムでは判断根拠の指標として、表1に例を示す行動価値関数を用いた。

ここで行動価値関数 $Q(S, A)$ とは、キャリア置き場の状況 S_n における複数の行動選択肢 $A_1 \sim A_m$ に対して、それぞれの行動価値を数値化したもので、最も行動価値の高い行動を選択すれば、その状況に応じて最適な行動を選択することができる。すでに述べたように、行動数 A_m は835通りあり、状況数 S_n は 2^{176} 通りある。

表1 行動価値関数 Q(S,A) の例

		← 835個 →				
		行動	A1	A2	...	Am
2176通り	状況	S1	3.0	1.0		0.5
	S2	2.0	8.0			10.0
	...					
	Sn	3.5	1.0			-6.0

状況に応じて最適な行動を選択するためには、行動価値関数 $Q(S, A)$ の最適化が必要である。本生産システムでは、行動価値関数の自動獲得を強化学習（用語解説参照）²⁾ の一種である Q 学習²⁾ により行った。

生産システムでは最適化すべきは生産性である。今回の場合、あるバッチに対応した合計処理時間を短くするように、表1の各行動価値 $Q_t(S_t, A_t)$ を最適化することにした。まず、1バッチ分の搬送行動を行い、各行動に要した処理時間 T_t 及び各行動により状況が遷移した結果に応じた報酬 R_{t+1} の時系列データを保存する。ここで報酬 R_{t+1} は、行動時間の逆数であり、 $R_{t+1} = p / T_t$ である。p は係数である。

次に、1バッチ分の行動完了後にバッチ処理にかかった合計処理時間を算出し、報酬を再計算する。今回のバッチ合計処理時間と前回までのバッチ合計処理時間とを比較して、今回のバッチ合計処理時間のほうが長ければ各行動報酬 R_{t+1} を-1倍とし、逆に短ければそのままの報酬値とした。

この再計算された報酬 R_{t+1} を用いて、1バッチ分のすべての行動価値 $Q_t(S_t, A_t)$ を更新する。この更新は、1バッチ分の行動価値関数を時系列上の順番に辿る形で、次式に基づいて行われる。

$$Q_t(S_t, A_t) \leftarrow Q_t(S_t, A_t) + \alpha \left[R_{t+1} + \gamma \max_A Q_{t+1}(S_{t+1}, A) - Q_t(S_t, A_t) \right]$$

ここで、 $\alpha (0 \leq \alpha \leq 1)$ は学習率であり、 $\gamma (0 \leq \gamma \leq 1)$ は割引率である。

また、(t+1) 回目の行動選択には、 ϵ -Greedy法を用いた。ある確率 ϵ でランダムに行動を選択し、それ以外では、

$$\max_A Q_{t+1}(S_{t+1}, A)$$

により、対応する全行動 A の中から行動価値最大の行動を選択する。

この Q 学習を所望のバッチ数分繰り返すことにより行動価値関数 $Q(S, A)$ が最適化され、「任意の状況に対してどの行動を選択することが最適であるか？」が自動的に学習される。

「思考の自動化」：仮想工場の役割

現実世界である実工場で、最適化に必要な1000バッチ分を学習すると、2年以上かかってしまうが、仮想工場ではオフラインで Q 学習すると、1000バッチ分の Q 学習を約5分で行うことができる。

学習やシミュレーション機能のほかに、実工場稼働時に実工場と仮想工場をリアルタイムにリンクさせる機能も開発した。本機能により、リアルタイムに稼働状況を表示ができる。また、エラー状況やエラー箇所やエラー対処方法も表示されるため、エラー特定と復旧作業を作業者のみで行うことができる。

仮想工場の初期パラメーターは、人手で設定する。その初期パラメーターの状態によっては、仮想工場の最適化が十分ではない場合がある。本生産システムでは、OPTASに蓄積された実稼働時のデータを仮想工場にフィードバックし、そのデータを使って学習することができる。これを繰り返す行うことで、仮想工場の最適化の度合いを改善することができる。各工程の処理時間などの変化が無ければ、学習→実稼働を2~3回循環させることにより、ほぼ最適状態に収束することが確認されている。各工程の処理時間などの変化があった場合は、新しい状況に対する学習が行われて再度最適化することができる。

本生産システムによる効果

本生産システムの導入による効果は、「作業の自動化」を行う際の技術者工数を「思考の自動化」により大幅に削減したことにある。

仮想工場での学習により行動価値関数が自動生成されるため、従来のようにロボット側にバッチ数すべての行動シーケンスを実装する必要がなくなった。また、自動生成されたシーケンスは、従来どおり人が考えたシーケンス（個別自動装置のキャリア置き場に最大数搬送する場合）と比較すると、15%の生産性向上が得られた。したがって、シーケンスを最適化する技術対応工数も不要となった。さらに、検討段階では図2のように、固着装置及び検査装置も5カ所のキャリア置き場がある想定であった。しかし、学習の結果、固着装置及び検査装置におけるキャリア置き場は1カ所が最適であることが判明した。その結果、固着装置及び検査装置のサイズダウンに成功し、設備投資を抑制することができた。

本生産システムによって、ロボットと個別自動装置を用いた生産システムの構築の技術者工数を、従来の1/10に抑制できた。導入した生産システムの稼働状態を写真1に示す。



写真1 導入した生産システム

今後の展開

今回導入した生産ラインでは、個別自動装置の稼働率が高いが、ロボットの稼働率は30%程度である。これは、ロボットが位置固定状態で使用され、ロボットの可動範囲内に配置された個別自動装置しか担当できないためである。したがって、ロボットを無人搬送車などで移動させれば、一台のロボットで約3倍の工程を担当することができ、一層、効率化できる。

また本生産システムの仮想工場は、ほかの製品の生産ラインの一部を同一のロボットが担当するような工場モデルであっても構築できる。したがって、異なる製品が混在する少量多品種の複雑な生産ラインの自動化にも対応できる。

今回開発した生産システムの判断部は、ロボットへ最適な行動を指示した。しかし、最適な行動は、ロボットに限らず「人」へも指示できる。

このように、今回開発した生産システムはさまざまな工程にも対応でき、少ない技術者工数によって生産性を向上させることが期待されるため、社内での水平展開を目指して現在も開発を継続している。 ◆◆

参考文献

- 1) 株式会社カワダロボティクス：NEXTAGE、<http://nextage.kawada.jp/>
- 2) Richard S. Sutton、Andrew G .Barto、三上貞芳、皆川雅章：強化学習、1版、pp.159-161、2001/12/27、森北出版

筆者紹介

谷川兼一：Kenichi Tanigawa. 株式会社沖データ HW技術本部 要素技術センター

石川琢磨：Takuma Ishikawa. 株式会社沖データ HW技術本部 要素技術センター

高橋千優：Chihiro Takahashi. 株式会社沖データ HW技術本部 要素技術センター

山田圭祐：Keisuke Yamada. 株式会社沖データ 生産統括本部 生産技術部

宮井敦司：Atsushi Miyai. 株式会社沖データ 生産統括本部 生産技術部

TiPO【基本用語解説】

強化学習

コンピューターによる機械学習の一種。解決すべき課題に対し、より正しい結果を得るため、試行錯誤を通じて自ら得られる報酬が最大化するよう学習を進める。