Development and Introduction of New Production System with Robot and AI

Kenichi Tanigawa Takuma Ishikawa Chihiro Takahashi
Keisuke Yamada Atsushi Miyai

Against the backdrop of labor shortage brought on by the decrease in the labor force, the move to automate production with robots is becoming increasingly active. Even OKI Data has been working on "automated work" of production through the development of an automated production system that utilizes in-house manufactured automatic devices and robots. However, with a conventional automatic production system, the engineer would need to create action sequences for the robots by programming sequential motions. If multiple products simultaneously flow down a production line that handles processes involving several automatic devices, it would be necessary for the engineers to specify every possible combination of processes. Since this would greatly increase the burden on the engineers, it was not realistic.

To solve such a problem, a virtual factory imitating a real factory was built on a computer. The virtual factory was used to develop a production system with "automated thinking" that will self-learn the robot's most optimal action sequences based on every possible combination of processes. The developed system was then implemented on an actual production line, and as a result, the burden on the engineers were significantly reduced. This article introduces the new production system that utilizes robot and AI.

Issues with Conventional Automatic Production System

Figure 1 shows a schematic of an automated production process at a real factory where the new production system was implemented. It consists of one robot and four individual automatic devices that are responsible for forming, assembling, fixing, and characteristic inspection. However, in the assembly process, the robot partially handles production. Transfers between processes are performed by the one robot. The individual automatic devices have, at a maximum, five carrier positions. The carrier is a jig capable of securely holding multiple products flowing along the process. Production of one batch of products is completed when the carrier at the entrance of the process moves to the exit after passing through all the production processes. As **Figure 1** shows, since one robot is responsible for transferring carriers to the four automatic devices, there are always multiple action choices when viewed from the robot.

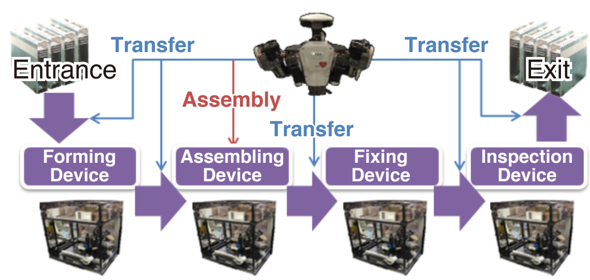

Figure 1. Schematic of an Automated Production Process

The number of actions that an engineer should teach the robot is considered using **Figure 2**. There are eighty carrier positions at the entrance of the process and five at the forming device. The arrows connecting each device indicate the carrier transfer actions of the robot. From the entrance position "1," the carrier can be transferred to one of the five positions, "1" thru "5," of the forming device. Since there are eighty positions at the entrance, the number of robot actions corresponding to the carrier transfer between the entrance and the forming device is 400 actions. In a similar matter, the total number of robot actions at the entire real factory is 835.

In actuality, the robot needs to select the optimal action from among the 835 patterns according to the ever-changing situation of the carrier position at that time. Each position has two states of whether transfer is possible or not. For example, if a process reaches completion and there is no carrier in the position of the subsequent process, it is a transferable state. As can be seen in **Figure 2**, since the total number of positions in this new production system is 176, there are 2^{176} position states.

Therefore, the engineer needs to specify an optimal action sequence for the robot from the $2^{176} \times 835$ patterns according to the situation. Furthermore, it is necessary to consider the actions for cases of error stops and recoveries. Thus, it is difficult for the engineer to realize "automated work" through the specification of all possible actions.

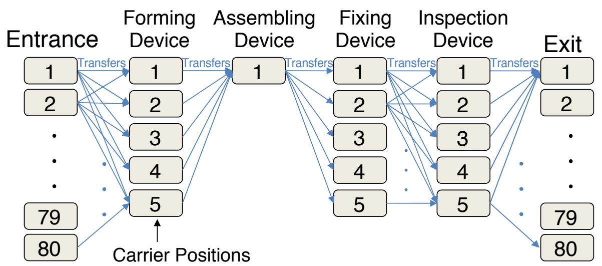

Figure 2. Carrier Positions in an Automated Production Process

Configuration of New Production System

The configuration of the new production system is described using **Figure 3**. The new production system consists of a real factory, a virtual factory, a decision unit and a process management server (OPTAS*1)). The "automated thinking" of the virtual factory drastically reduces the burden on the engineers that arises when trying to achieve "automated work."

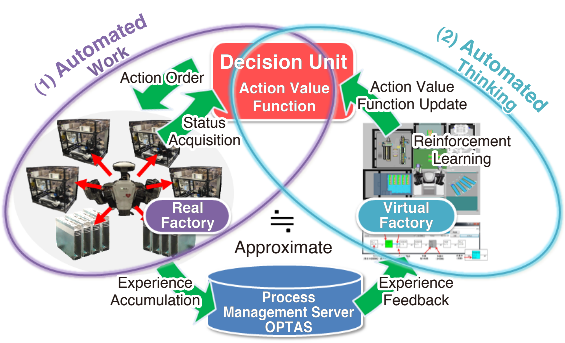

Figure 3. Schematic of Production System with Robot and AI

As was previously shown in **Figure 1**, the real factory consists of a robot and multiple individual automatic devices. Production at each process is carried out by an individual automatic device. The robot handles part of the assembly work and the transfer between processes. The result is an achievement of multiple unmanned processes. The robot that was used is NEXTAGE[1)], a humanoid next generation industrial robot manufactured by KAWADA ROBOTICS.

Figure 4 shows an example system configuration of an individual automatic device. The required motions are achieved with the use of motors and drive shafts. All the individual automatic devices are equipped with a camera to enable various functions such as image linked motion control, product alignment, inspection, differentiation and recognition. The devices can also be equipped with measuring instruments and sensors to perform other inspections as necessary. All these elements are controlled using a control PC, which is connected to the in-house network for communication with OPTAS and the decision unit. Each process was efficiently automated through the in-house development of these individual automatic devices.

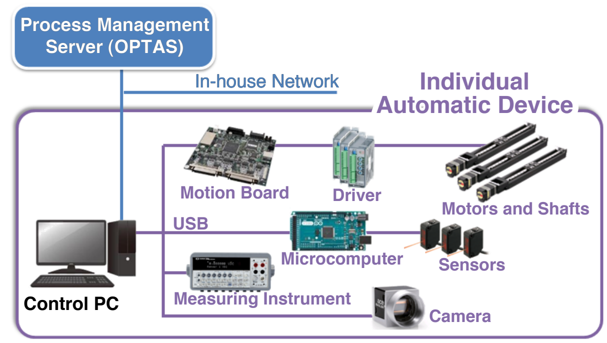

Figure 4. System Configuration of Individual Automatic Device

The decision unit gathers the situational status such as carrier position from the real factory and gives out optimum action orders according to the situation, thereby comprehensively controlling the entire factory.

A virtual factory is a computer environment for simulating the work of a real factory and realizes various combinations of production processes with software, which otherwise would require numerous engineers and time to perform. Thus, it enables the robot's motions to be learned in a short period. Feeding back data from the real factory to the virtual factory brings learning closer to reality.

In the new production system, position information obtained at the real factory from sensors and cameras other than those equipped on the robot can be fed back to the robot as an external variable. Referencing this external variable, the position coordinates of the basic motion taught to the robot can be changed according to the current situation. Therefore, the 835 actions can be reproduced with the teaching of five basic actions, and the engineers' burden in teaching the robot's motion is drastically reduced.

However, there are still 2^{176} different situations for the carrier position. Through "automated thinking," the optimum actions corresponding to the massive number of these combinations are learned automatically at the virtual factory, thus allowing the realization of "automated work."

"Automated Thinking": Action Value Function and Reinforcement Learning

For the decision unit to issue an optimum action command for a situation, it is sufficient if one optimum

*1) OPTAS: OKI data Production control and Total Analysis System action corresponding to the situation can be selected from the multiple action options. In the new production system, the action value function shown as an example in Table 1 served as an indicator for the decision basis.

The action value function Q (S, A) is the action values represented numerically for each of the multiple action options A1 thru Am available for the carrier position situation Sn. The optimal action for a situation is selected when the action with the highest action value is selected. As mentioned previously, the number of actions Am is 835 and the number of situations Sn is 2^{176}.

Table 1. Example of Action Value Function Q (S,A)

835 Actions

2^{176} Situations

Actions / Situations	A1	A2	・・・	Am
S1	3.0	1.0		0.5
S2	2.0	8.0		10.0
・・・				
Sn	3.5	1.0		-6.0

Selection of an optimal action according to the situation requires optimization of the action value function Q (S, A). In the new production system, automatic acquisition of action value function was carried out using Q-learning[2], which is a type of reinforcement learning (refer to glossary)[2].

Productivity is what should be optimized in a production system. In this case, it was decided that each action value Q (S_t, A_i) in **Table 1** would be optimized to shorten the total processing time for a certain batch. First, transfer actions for one batch is performed, and the time-series data of the processing time T_t required for each action and reward R_{t+1} that corresponds to the result of the situation transition due to each action are saved. Here, the reward R_{t+1} is the reciprocal of the action time, and $R_{t+1}=p/T_t$. p is a coefficient.

Next, after the action completion of one batch, the total processing time taken for the batch processing is calculated, and reward is recalculated. This total batch processing time is compared with the total batch processing time up to the previous time. If the latest total batch processing time is longer, each action reward R_{t+1} is multiplied by -1, and conversely, if it is shorter, the recalculated reward value is used as is.

Using the recalculated reward R_{t+1}, all the action values Q_t (S_t, A_i) for one batch are updated. The updates are performed based on the following equation while tracing the action value functions of one batch in chronological order.

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

Here, $\alpha(0 \leq \alpha \leq 1)$ is the learning rate, and $\gamma(0 \leq \gamma \leq 1)$ is the discount rate.

For the (t+1)th action selection, the ε-Greedy method was used. An action is randomly selected with a certain probability ε, and for others, according to

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

From all the corresponding actions A, the action with the highest action value is selected.

Repeating this Q-learning for the desired number of batches optimizes the action value function Q (S, A), and “Which action is optimal for an arbitrary situation?” is automatically learned.

“Automated Thinking”: Role of Virtual Factory

In a real world factory, it would take more than two years to learn the 1000 batches necessary for optimization, but with the virtual factory, off-line Q-learning will enable the same 1000 batches to be learned in five minutes.

In addition to the learning and simulation functions, a function was developed to link the real and virtual factories in real-time when the real factory is in operation. With this function, the operating status can be displayed in real time. When an error occurs, the condition, location and handling method are also displayed to allow error isolation and recovery work to be performed by the operator alone.

Initial parameters of the virtual factory are set manually. Depending on the state of the initial parameters, the virtual factory optimization may not be sufficient. In the new production system, actual operation data accumulated in OPTAS can be fed back to the virtual factory and used for learning. By repeating this process, the virtual factory’s degree of optimization can be improved. When there are no changes in processing time, etc. of each process, it has been confirmed that by repeating the learn ⇒ actual operation cycle two or three times will bring the virtual factory to a near optimal state. If there are changes in processing time, etc. of each process, learning for the new situation is performed, and the virtual factory can be re-optimized.

Effectiveness of New Production System

The effectiveness of the new production system is shown in the great workload reduction that the “automated thinking” brings for the engineers when they “automate work.”

Since the action value function is automatically generated through learning at the virtual factory, it becomes unnecessary to implement the action sequences for all the number of batches on the robot side as with a conventional case. Furthermore, there was a 15% productivity improvement with the automatically generated sequences when compared against human-generated sequences (when maximum number is transferred to the carrier positions of individual automatic devices). Therefore, it also eliminates the technical man-hours needed for optimizing the sequences. Moreover, at the study stage it was assumed that the fixing and inspection devices would each have five carrier positions as shown in **Figure 2**, but after the learning, it was found that one position each was optimal. As a result, the fixing and inspection devices were successfully downsized, and it was possible to suppress the cost of facility investment.

With the new production system, the engineers' workload for building a production system using a robot and individual automatic devices was reduced to 1/10 of a conventional system. **Photo 1** shows the implemented production system in operation.

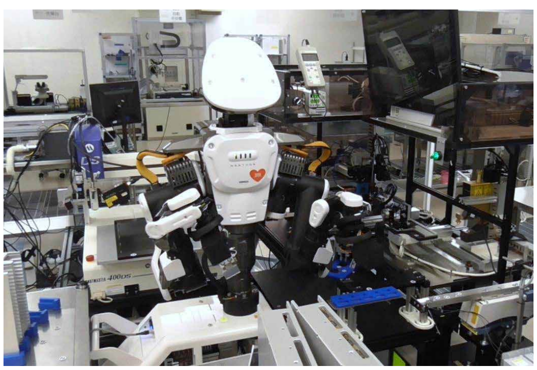

Photo 1. Implemented Production System

Future Development

In the production line that was implemented, the operation rates of individual automatic devices are high, but the operation rate of the robot is only about 30%. This is because the robot is used in a fixed position, and it can only handle individual automatic devices that are within range of the robot's motions. If the robot's position can be moved such as with an unmanned carrier, a single robot will be able to handle about three times the process and further improve efficiency.

The virtual factory of the new production system can be constructed to model a factory where the same robot handles part of the production line for another product. Accordingly, it is also possible to cope with automation of complicated high-mix low-volume production lines in which different products are mixed.

The decision unit of the new production system instructs the robot with optimal actions. However, the optimal actions can be directed not only to robots but also to "humans."

In this way, the new production system can handle various processes, and since productivity improvements can be expected with fewer engineer man-hours, development is continuing for lateral deployment within the company. ◆◆

References

1) KAWADA ROBOTICS CORPORATION: NEXTAGE, https://www.kawadarobot.co.jp/index_en.html
2) Richard S. Sutton, Andrew G. Barto, Sadayoshi Mikami, Masaaki Minagawa: Reinforcement Learning, First Edition, pp.159-161, 2001/12/27, Morikita Publishing Co., Ltd.

Authors

Kenichi Tanigawa, Core Technology Center, HW Engineering Division, Oki Data Corporation

Takuma Ishikawa, Core Technology Center, HW Engineering Division, Oki Data Corporation

Chihiro Takahashi, Core Technology Center, HW Engineering Division, Oki Data Corporation

Keisuke Yamada, Production Engineering Dept., Production Management Division, Oki Data Corporation

Atsushi Miyai, Production Engineering Dept., Production Management Division, Oki Data Corporation

TIPS

【Glossary】

Reinforcement learning

Type of machine learning using a computer. In order to reach a more correct result for a problem to be solved, learning is conducted on a trial-and-error basis to maximize the self-obtained reward.