AI-based Plant Control

Go Takami *1

Expectations for machine learning and artificial intelligence (AI) are growing globally and related investment has been increasing in a diverse range of businesses. Machine learning is used for autonomous car driving and robot control, and the use of its applications is rapidly increasing in factory automation (FA). On the other hand, practical process control techniques based on machine learning and AI have not yet been developed for process automation (PA) although data analysis using process data is becoming more common. PID control still remains the main technique and advanced control techniques by experts are used when complicated control is required.

Not to simulate but to apply machine learning technology to actual equipment, we performed AI-based control of a three-tank-level control system, which is a popular educational kit for process control. This paper introduces the details of the experiment and the machine learning technology used to control this system.

INTRODUCTION

In recent years, there are high expectations for artificial intelligence (AI) and the machine learning technology that supports it, with investments in this area increasing rapidly. Yokogawa Electric Corporation and Yokogawa Solution Service Corporation have also been working to apply machine learning to business, and have published various achievements related to machine learning (1)-(6). Previously, the focus was on using machine learning to detect and predict abnormalities in facilities. Many analyses based on machine learning are performed using process data. In the field of process control, although there are a few examples of applying machine learning to simulators, practical technology applicable to actual facilities has not yet been developed. PID control technology is mainly used for control in the PA field while advanced control technology is used in complex process control. However, field workers who use these control methods often complain that the results are not always satisfactory, and so experts manually operate facilities and change set points depending on the conditions.

This paper introduces an experiment of applying machine learning to process control, which uses a three-tank-level control system in the control education kit in Yokogawa’s training center. We also discuss the possibility and feasibility of control with machine learning.

REINFORCEMENT LEARNING

In this control experiment with machine learning, we used reinforcement learning, which will be outlined in this section.

Machine learning technologies can be divided into the following three categories by type of study method and input data:

- Supervised learning
- Unsupervised learning
- Reinforcement learning

In supervised learning, “label” information is assigned to data in advance and the data are then given to the learner (learning algorithm) to create models. For example, images labeled “dog” or “cat” are given to the learner, which creates a model that can then distinguish any images of dogs and cats. Deep learning, which is often used for image recognition, is a

*1 AI Business Development Department, Information Technology Center, IA Products and Service Business Headquarters
type of supervised learning.

In unsupervised learning, data without label information are given to a learner to create models. For example, in the case of Internet shopping, multiple buyers’ purchase history is given to a learner, which then creates a model of the purchase trends of buyers and categorizes customers by their history. However, since there is no label information, human operators need to consider how the data have been categorized.

Although both supervised learning and unsupervised learning differ in terms of the existence of labeled data, they both require data to be prepared for learning in advance. In contrast, reinforcement learning does not necessarily need any data for learning; it learns based on the unique concept of “reward.” Figure 1 shows a typical configuration of reinforcement learning.

![Figure 1](image1.png)

**Figure 1** Configuration of reinforcement learning

The agent and the environment are the two major factors of reinforcement learning. The agent autonomously takes action and the environment changes accordingly. The agent then receives the changed environment as the status and samples data. At this time, a reward value is calculated according to the change in the status caused by the action. By sampling the action, changed status, and reward value, the agent learns as it pursues higher reward values. This process is similar to children piling up blocks, as explained with reference to Figure 2.

When a child tries piling blocks, he repeatedly fails but gradually learns how to pile them high. In reinforcement learning, the child is the agent and the blocks are the environment. The child continues trial and error considering the status of the blocks. When he can pile them high, he feels happy, and when the blocks collapse, he feels sad. In reinforcement learning, piled blocks are “status,” piling is “action,” and the feelings of joy and sadness are “reward.” The child does not consider Newton’s laws or calculate the location of the center of gravity; he empirically learns how to pile blocks high. Similarly, in reinforcement learning the agent learns how to earn better rewards based on the status obtained by its action and resulting rewards. Following its own strategies, the agent searches for better approaches to earn higher rewards, and it occasionally finds solutions that humans cannot come up with.

As described earlier, in reinforcement learning the agent samples data. Therefore, different from supervised learning and unsupervised learning there is no need to prepare data in advance. However, reinforcement learning has one problem: the need for numerous trials. Even in the case of blocks, repeated trials are needed to pile them high. Well-known examples of reinforcement learning include AlphaGo, which beat professional human Go players, and the deep Q-network, which mastered several dozen Atari games. These programs performed thousands or millions of trials in cyber space to achieve the goal, but so many trials is not practical in the real world. Therefore, reinforcement learning has rarely been applied to problems in the real world.

To solve this problem, Yokogawa developed an innovative method by using Omega Simulation’s plant simulator for vinyl acetate manufacturing plants. This method is called the Factorial Kernel Dynamic Policy Programming and was announced at the 2018 IEEE International Conference (2). Trials usually need to be repeated at least tens of thousands of times, whereas our method can boost production only with 30 trials under specific conditions. If this programming is applicable to other conditions, reinforcement learning will be widely used in actual facilities instead of simulations. To verify its feasibility, we conducted an experiment with a three-tank-level control system.

**CONTROL OF THREE-TANK-LEVEL SYSTEM**

**System Configuration**

A three-tank-level control system is a training device for control in Yokogawa’s training center (Figure 3). Figure 4 shows the system configuration.

![Figure 2](image2.png)

**Figure 2** Example of reinforcement learning

As described earlier, in reinforcement learning the agent samples data. Therefore, different from supervised learning and unsupervised learning there is no need to prepare data in advance. However, reinforcement learning has one problem: the need for numerous trials. Even in the case of blocks, repeated trials are needed to pile them high. Well-known examples of reinforcement learning include AlphaGo, which beat professional human Go players, and the deep Q-network, which mastered several dozen Atari games. These programs performed thousands or millions of trials in cyber space to achieve the goal, but so many trials is not practical in the real world. Therefore, reinforcement learning has rarely been applied to problems in the real world.

To solve this problem, Yokogawa developed an innovative method by using Omega Simulation’s plant simulator for vinyl acetate manufacturing plants. This method is called the Factorial Kernel Dynamic Policy Programming and was announced at the 2018 IEEE International Conference (2). Trials usually need to be repeated at least tens of thousands of times, whereas our method can boost production only with 30 trials under specific conditions. If this programming is applicable to other conditions, reinforcement learning will be widely used in actual facilities instead of simulations. To verify its feasibility, we conducted an experiment with a three-tank-level control system.

**CONTROL OF THREE-TANK-LEVEL SYSTEM**

**System Configuration**

A three-tank-level control system is a training device for control in Yokogawa’s training center (Figure 3). Figure 4 shows the system configuration.

![Figure 3](image3.png)

**Figure 3** Three-tank-level system
The three-tank-level control system used in this experiment has a control path and a disturbance path that simulates disturbances. Pump 1 pumps up water from the storage tank and discharges it into water tank 3, which is located on the top. Water flows into water tank 2 and water tank 1 and then returns to the storage tank. In this system, the goal is to control the level of water tank 1 by adjusting the aperture of a valve for the output to water tank 3. Since there is a time lag until the adjustment of the valve aperture is reflected in the level of water tank 1, it is difficult to maintain the level of water tank 1 at the target by manual operation.

Figure 5 shows the results of PID control with the level of water tank 1 set as the observation point. The horizontal axis of the graph expresses time (s) and the vertical axis shows the level of water tank 1 (%). An ultimate sensitivity method was used for tuning PID control and the control target was 30%. However, overshoot occurred as shown as (1), and it took more than 200 seconds to reach the target range within ±5% of the set value. We applied reinforcement learning to control this three-tank-level control system.

The goal of this experiment was to control the level of water tank 1 at 30%. The following formula was used to calculate the reward value.

\[
\text{Reward value} = 0.1 \times (30 - |L1001 - 30|)
\]

### Table 1 Input and output of the three-tank-level system

<table>
<thead>
<tr>
<th>Observation target</th>
<th>Tag name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of water tank 1 (%)</td>
<td>L1001</td>
</tr>
<tr>
<td>Level of water tank 2 (%)</td>
<td>L1002</td>
</tr>
<tr>
<td>Level of water tank 3 (%)</td>
<td>L1003</td>
</tr>
<tr>
<td>Flowrate of the control path (%)</td>
<td>F1001</td>
</tr>
<tr>
<td>Flowrate of the disturbance path (%)</td>
<td>F1002</td>
</tr>
<tr>
<td>Aperture of the valve (%)</td>
<td>V001</td>
</tr>
</tbody>
</table>

Thirty trials were performed in the experiment. Each trial consisted of 200 steps (operations) and each step was performed at 2-second intervals. We observed the status.

### Learning Process and Results

Each trial started with water tank 1 empty. Before the start of a trial, the valve was closed and all the water in water tank 1 was drained. The agent adjusted the aperture of the valve 200 times in each trial. Operations were chosen from among the five options. The agent observed the resulting status and calculated the reward value. The model was updated each time the trial was terminated after 200 steps. The learning was repeated until 30 trials were completed.

Figure 6 shows the results of controlling the level of water tank 1 with the model after 30 trials. As in Figure 5, the horizontal axis of the graph expresses time and the vertical axis shows the level of water tank 1. The graph shows that the control successfully converged to the set point of 30%. The remarkable feature is the time taken for control to stabilize. In PID control, overshoot occurred and it took more than 200 seconds to stabilize. In contrast, there was no overshoot with the control after reinforcement learning, and it rose smoothly to the set point and stabilized in about 100 seconds, halving the time to startup. In batch or other processes, this is expected to shorten the waiting time until control is stabilized when starting up a facility.
In addition, we carried out another experiment by adding a disturbance to the model after reinforcement learning. Furthermore, the experiment started with the level of water tank 1 raised. Figure 7 shows the results. During the learning process, trials were always started with water tank 1 empty. Although this full-tank condition had not been learned, undershoot did not occur and stable control was achieved until reaching the set point. In the original experiment, the level of water tank 1 became high in some trials. The agent seems to have used these data to achieve control in a different condition. Depending on the learning process, it may be possible to create a model that can respond to disturbances.

![Figure 7](image)

**Figure 7** Experiment was started with the level of water tank 1 raised

**CONCLUSION**

This paper introduced an example of applying machine learning as a practical technology in a new field, which is significant progress. Leveraging the results, we are planning to accumulate application cases not only in experiment kits but also in actual facilities. This will require developing technology that ensures safety and versatility. We intend to bring innovations to the control field by continuously developing technologies for the following:

- Safe learning method
- Robustness to disturbances
- Online learning
- Applicability of models to other facilities
- Response to changes in set points

**REFERENCES**

4. “Yokogawa and NAIST apply reinforcement learning to chemical plants and achieve advanced control with a small number of trials,” Nikkei Robotics (March), No. 44, 2019 (in Japanese)