Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 18-19, 2023

Simulation Data Selection Method for the Development of Effective Artificial Intelligence Application for Nuclear Power Plants

Hyun-Chul Lee[a*]
[a]*Korea Atomic Energy Research Institute, 989-111 Daedeok-daero, Yuseong-gu, Daejeon, South Korea 34057*
[*]*Corresponding author: leehc@kaeri.re.kr*

1. Introduction

There are many research and development worldwide to improve safety and availability by using artificial intelligence technology to operate nuclear power plants. Most R&D goals are system development for preemptive diagnosis and decision-making support, and the main means are based on machine learning or deep learning methods.

Supporting post-accident mitigation measures is a great economic effect because it allows time-consuming and capital-intensive accident handling tasks to be completed as early as possible, but preventing the occurrence of accidents or diagnosing them early has a foundation effect to minimize the need for accident mitigation. Therefore, it is clear that diagnosis is the application field that can expect the most benefits in terms of effectiveness compared to input resources.

In order to utilize artificial intelligence technology in the field of diagnosis, decision criteria are required to distinguish between normal and abnormal, and for this, process profiles for each normal and abnormal are required. However, abnormal or dangerous conditions do not occur frequently in general engineering systems, and thus it is very difficult in reality to obtain data about abnormal conditions from past histories.

To solve this problem, it is often tried to create a virtual environment and implement abnormal conditions to obtain data, and even this is usually hard in most industries.

In the nuclear domain, the situation is better. Due to the great fundamental requirements for nuclear power plant safety, measures have long been adopted to visually confirm the ability to cope with abnormal conditions. Computer codes for interpreting the safety of nuclear power plant processes for various abnormal conditions have been developed and used, and education and training simulators have been installed and continuously operated to secure operator's ability to cope with accidents, which is a bridgehead for nuclear power plant safety. In other words, it is possible to simulate various abnormal states in a virtual environment without actual risk.

This provides an excellent foundation for the use of artificial intelligence technology in the field of nuclear power plant operation, and is also a means to check accuracy and effectiveness when the results of R&D are used for actual nuclear power plant operation.

The simulator usually uses both variables that can be obtained from nuclear power plant measurement control facilities and virtual variables that cannot be measured in actual situations to simulate real situations. However, artificial intelligence developers often use virtual variables that cannot be acquired in a real environment for research and development. This study will discuss the selection of variables and data that should be careful when developing applications or systems using simulators as a source of nuclear power plant process data.

2. Methods and Results

In this section simulation variables and an effective selection method are discussed.

2.1 Simulator Variables

Simulators define and use a large number of variables to simulate various situation. According to usage characteristics, simulation variables can be classified as follows.

- Variable for analytic models (Vm)
- Variable for human interaction (Vh)
- Variable for program (Vp)

Vm includes input/output variables for implementing mathematical and physical analytic models required for describing nuclear installations and process behavior of systems. Vp includes variables for configuring computer programs.

Vh used for interaction with human operators or trainees can again be subdivided into the following.

- Alarm variables
- Indicator variables
- Controller variables

Developers can obtain detail values of all variables from simulation computers at any time. In the real situation, nuclear plant operators can monitor alarms and indicators and mostly the states of indicators are recorded in the plant computer. It means a new system can usually get indication information only with a very high possibility.

Therefore, it is very important to separate simulation variables into measurable and non-measurable variables from the variable availability point of view. Measurable variables are available from both the simulator and the real plant, but non-measurable variables are available from the simulator only.

2.2 Variable Selection Method

In order to secure applicable process data, the variables that can be acquired at the nuclear power plant where the application or system is operated must first be identified. The process information recorders and communication means installed at the nuclear power plant should be identified, and the variables available in the environment in which the developed application or system is installed should be accurately defined. A defined variable is called an preliminary effective variable.

It should be noted that the available variables vary depending on the characteristics of the nuclear power plant. Therefore, OPR1000 and APR1400 nuclear power plants differ in the range and content of the effective variables provided.

Second, it is necessary to identify the simulator variables used for development. The number of simulator variables is larger than the preliminary effective variables and usually ranges from 10,000 to hundreds of thousands. There are many internal variables defined by simulator developers for their development work convenience in Vm, and there are many integrated variables in Vh that have been processed several variables to provide high-level information. Reducing the number of variables is inevitable because it is difficult to effectively handle all simulation variables in artificial intelligence algorithms. The operation of selecting important variables is called feature selection, and eventually a variable representing the movement of many variables is selected, and in this case, non-measurable variables can be selected. In order to prevent the development of artificial intelligence results that cannot be applied, it is necessary to adjust the preliminary effective variable to become a representative variable. It is common for developers to perform this task themselves. Sometimes there are variables that are available in real nuclear power plants, but that the simulator cannot provide. The reason is that in most cases the introduction of a new system has not yet been reflected in the simulator, and variable selection should be attempted in consideration of the scope of use of the system to be developed.

Finally, it is necessary to adjust (mainly reduce) the number of selected measurable representative variables. This is because it acts as a realistic constraint. The period of data transmission and reception, the memory capacity of the environment in which the developed system is installed, and the communication capacity must be considered. For this, there will be cases where data loss must be endured.

3. Conclusions

It is practically inevitable to use a simulator to develop a operation support system that uses artificial intelligence technology applicable to nuclear power plants.

This study discusses how to select representative variables that are actually applicable from many variables provided by the simulator. Representative variables should be selected from the effective variables in consideration of the source of information on the nuclear power plant to be applied, the characteristics of the simulator, and practical restriction of I&C systems.

REFERENCES

[1] KAERI, ITF-OPR1000 Simulator, 2020.
[2] KAERI, CNS Manual, 2000.
[3] M. S. Raza, Understanding and Using Rough Set Based Feature Selection: Concepts, Techniques and Applications, Springer Singapore, 2017.