Transactions of the Korean Nuclear Society Autumn Meeting
Goyang, Korea, October 24-25, 2019

A Power Increase Autonomous Algorithm based on Deep Reinforcement Learning

Daeil Lee, Awwal Mohammed Arigi and Jonghyun Kim*
Department of Nuclear Engineering, Chosun University, 309 pilmun-daero, Dong-gu, Gwangju, 501-709, Republic of Korea
dleodlf1004@Chosun.kr, awwal@Chosun.kr, Jonghyun.kim@Chosun.ac.kr

1. Introduction

An autonomous operating system can manage the overall system without the intervention of a human [1]. For handling the whole system without human, the autonomous operating system needs some abilities, such as defining the operational strategy and flexible operation according to dynamic system parameters. Artificial Intelligence (AI) techniques are one of method that can give these abilities to the autonomous system.

Deep reinforcement learning method, which is a kind of AI method, has been used as a way to achieve the given operational goal of systems. Deep reinforcement learning methodology is a combination of deep neural network method and reinforcement learning. It is known that deep learning applications can overcome limitations of classic controllers that cannot reflect human experience and control the non-linear parameters. On the other hand, deep reinforcement learning method can find the trajectory to achieve operational goal though interaction with a given environment.

Nuclear power plants (NPPs) are operating with analog and digital instrumentation and control (I&C), which is controlled by the combination of the manual control by operators and the automatic control by automatic algorithms [2]. However, the operator's intervention is still essential for controlling systems in NPPs.

Power increase operation is a part of the start-up operation in the NPPs. During the start-up operation, operator's roles are to control components to increase reactor power and generate electricity. This operation requires a higher cognitive demands than the full power operation in which most of the controls are performed by automatic algorithms. This operation mode need to check, monitor, control the plant based on operating procedures. Therefore, manual controls may cause high operator's burden. It is known that the power start-up/shutdown operation modes have a high probability of human errors due to operator's burden and workload [3].

This study suggests an autonomous algorithm for increasing power from 2% to 100%. First, this study analyzes general operating procedures and time line during power increase operation. Then, it suggests a power increase algorithm by using an asynchronous advantage actor-critic (A3C) and Long Short-term Memory (LSTM). The A3C is a kind of deep reinforcement learning method that can train the neural network by using the parallel actor-learners based on CPU threads and the asynchronous network update [4]. As neural network applied in A3C, this study used LSTM network. LSTM has an ability to extract features to the given domain's problem from non-linear variable. In addition, LSTM can consider the past experience from the time series data by using long term memory cell. Validation and demonstration of the algorithm are also presented in this study.

2. Analysis of Power Increase Operation

This section analyzes the existing operational strategy for developing an autonomous algorithm. General operating procedures are reviewed to identify main operator's activities in the operation. Then, the time line was also analyzed to determine the sequence of activities.

This study used a compact nuclear simulator (CNS) with a Westinghouse 900 MWe, 3 loops pressurized water reactor (PWR) as the reference plant.

2.1. Analysis of general operating procedure

General Operating Procedure (GOP) provides instructions for the operator's tasks during the start-up operation. The GOP during the start-up operation usually contains manual tasks, compared with other operational modes, e.g., the full power mode. This is because the operation at this mode is less automated.

This study identifies major tasks in GOP and the conditions of manual operations to perform the task. In addition, control types, i.e., on/off or regulating, are also identified for conditions. Table I shows the operator's tasks and conditions.

Table I: Operational rule for increasing power.

Step	Control type	Condition
1	Regulating	Increase reactor power from 2% to 6% ~ 10%
2	On/Off	If reactor power is 4% then the main feed-water pump 1 start.
3	On/Off	If reactor power is 10% then click intermediate and power range block buttons.
4	Regulating	If reactor power is 10% then the turbine RPM set-point is1800RPM.
5	Regulating	Increase reactor power to 14% ~ 16%
6	Regulating	If reactor power is between 10% and 20% then load set-point is 200MWe.
7	Regulating	If reactor power is over 10% then acceleration set-point is 50MWe/min.
8	On/Off	If the turbine RPM is 1800RPM and reactor power is over 15% then push the network breaker.
..	..	..

2.2. Timeline Analysis

A timeline analysis has been carried out to develop a normative sequence of start-up operations, which is required for training the algorithm. The timeline analysis includes systems that should be operated at a specific power level, the system's mode (i.e., manual or automatic), and the power trend over time. This analysis is based on the operation conditions in Table I.

Fig. 1 shows the result of timeline analysis. Five control functions are included such as steam generator (S/G) level control, rod control, synchronous control, turbine load control, and pump control. Main feedwater pumps and condenser pumps are included in pump control during the power increase operation.

The control mode of the systems is divided into automatic and manual modes. Furthermore, the manual mode is categorized into Type A and Type B manual control. Type A is a manual control for which the GOP provides specific instructions on the control target. For instance, if the GOP instructs operators to start a pump or open a valve, this is Type A control because the component and the state are clearly defined. Type B is a manual control for which the GOP does not provide specific guidelines on the target. A representative example is a rod control. Generally, the GOP does not describe how many steps of control rod should be withdrawn or inserted, but the target of reactor power, e.g., increase the reactor power to 20%. In this type of control, the determination of the number of steps to withdraw or insert relies on the operator's experience by monitoring plant parameters in situation.

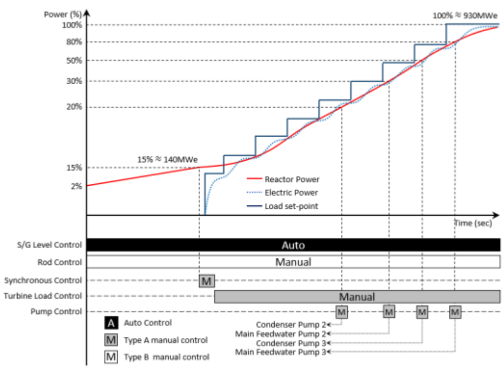

Fig. 1. Timeline for increasing the reactor power from 2% to 100%

3. Development of power increase algorithm

An algorithm for the autonomous power increase in the start-up operation has been suggested by using AI techniques. This study applied different techniques according to the types of manual control modes, as shown in Fig. 2. For the Type A control, a rule-based systems has been applied, while the Type B control was implemented by a deep reinforcement learning.

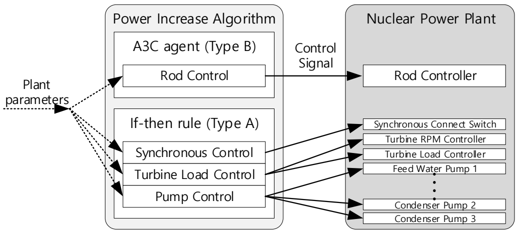

Fig. 2. Structure of power increase algorithm

3.1. Algorithm for Type A control

For implementing the Type A control, a rule-based system based on if-then rules was suggested. Fig. 3 shows the rules for Steps 2 and 3 in Table I. The conditions and operator's actions in the step are converted to if-then rules in Table II.

Table II: Converted if-then rule

Rule number	If-then rule
1	IF reactor power = 4% THEN: start main feed-water pump 1 AND go rule 2
2	IF reactor power = 10% THEN: go rule 3
3	IF the state of intermediate range signal and power range signal is not block THEN: click intermediate and power range block buttons AND go rule 4
4	IF reactor power = 10% THEN: go rule 5
..	..

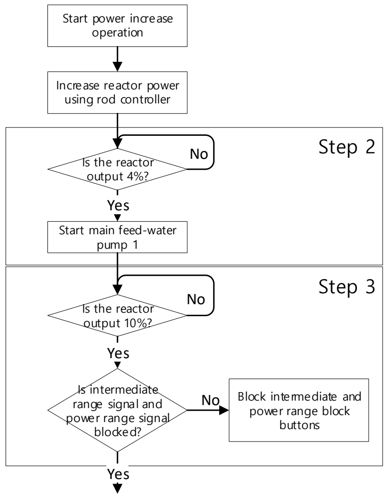

Fig. 3. Example of if-then rule from step 2 to 3

3.2. Algorithm for Type B control

3.2.1. Background of A3C

Type B control is implemented using the deep reinforcement learning algorithm, i.e., A3C method. The concept of A3C algorithm is to train an agent to work in multiple environments. The agent can gain a lot of experiences from the interaction with multiple environments at the same time. The agent uses the accumulated experiences to train how to control the components to achieve the goal or solve problems.

Fig. 4 shows the training method of A3C. A3C consists of main agent and local agent. The local agent collects experiences in a series of independent episodes, and each episode comprises of a sequence of turns. At the end of one episode, the local agent is trained based on the collected experience during the episode. Then, the main agent is asynchronously updated with a trained local agent. After training, the main agent can find a path to achieve the operational goal through a trial-and-error.

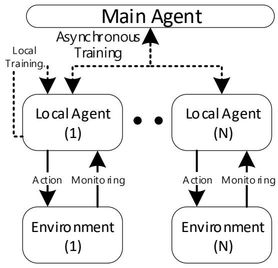

Fig. 4 A3C training method

3.2.2. Reward design

The goal of deep reinforcement learning is generally assigned by a reward(s). Then, the agent is trained to maximize the cumulative reward. The agent can be compensated by rules designed to solve problems in a given domain. As a training guideline, this study uses the operation boundary. Fig. 5 shows the success criteria to receive a reward in the A3C algorithm.

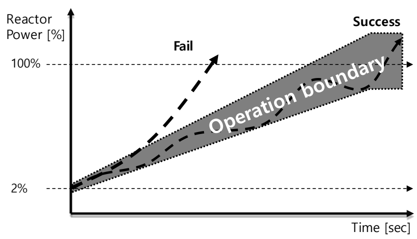

Fig. 5. Operation boundary of A3C

3.2.3. LSTM network modeling

Long Short-Term Memory (LSTM) network, which is a type of recurrent neural network, is applied in A3C's agents for handling nuclear power plant data that are non-linear and time-serial. LSTM network constitutes the main body of agent.

Fig. 6 shows the structure of the LSTM network applied in A3C. The LSTM network consists of 3 layers. The input layer has time-length = 10 sec and the number of input parameters = 8. Input parameters include parameters used in rod controller (i.e., current power, average temperature, rod position, and electric power, based on the GOP) as well as operational boundary (upper / lower boundary value, the distance between power and boundary values). The LSTM also has 2 layers and 8 LSTM cells for each layer. The output layer consists of actor (for calculating policy) and critic (for calculating value) layers which is parts of the A3C algorithm. The A3C algorithm gets the probability of actions according to policy and value parameters.

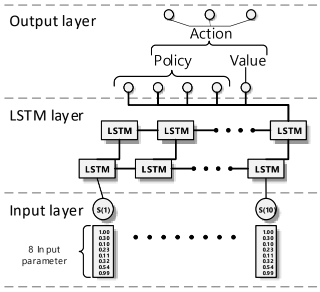

Fig. 6. Structure of LSTM network in an A3C agent

4. Experimental

4.1. Training environment

For training and validating the algorithm for the autonomous power increase, the CNS is used as a real-time testbed. In this study, the CNS is installed on three sub-computers. Then, each sub-computer can run 20 separate CNS simulations, so the total 60 CNSs can be simulated at the same time.

4.2. Experimental result

Fig. 7 and Fig. 8 show the experimental results with the rates of 1%/min. The A3C agent in the algorithm for the autonomous power increase has been trained by 8800 episodes in the simulation for 43 hours. Fig. 7 shows that the autonomous power increase algorithm can increase the reactor power within the operation boundary. In addition, it also shows that the algorithm successfully increase the electrical power from 0MWe to 900MWe.

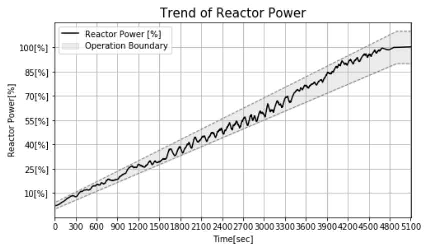

Fig. 7. Trend of reactor power from 2% to 100%

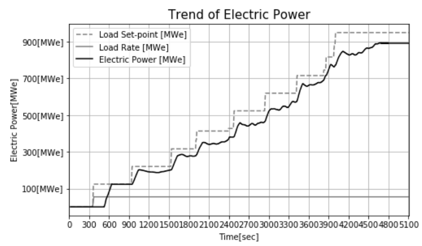

Fig. 8. Trend of electric power from 0 MWe to 900MWe

5. Conclusion

This study proposed an algorithm for autonomously increasing reactor power and electrical power from 2% to 100%. The algorithm for the autonomous power increase was developed based on the analysis of power increase operation. Then, the algorithm was implemented by using the A3C and the LSTM, which are AI techniques. The experimental results have shown that the algorithm can operate systems to increase power from 2% to 100%.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korean government(Ministry of Science and ICT) (2018M2B2B1065651) and Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning(N01190021-06).

REFERENCES

[1] H. Basher, and J. S. Neal, Autonomous Control of Nuclear Power Plant, Citeseer, 2003.
[2] J. Y. Kim, K. C. Kown, I. K. Hwang, D. Y. Lee, W. M. Park, J. S. Kim, and S. J. Lee, Development of Advanced I&C in Nuclear Power Plants: ADIOS and ASICS, Nuclear Engineering and Design, Vol. 207, No. 1, pp. 105-119, 2001.
[3] D. I. Lee, and J. H. Kim, Autonomous Algorithm for Start-Up Operation of Nuclear Power Plants by Using LSTM, International Conference on Applied Human Factors and Ergonomic, Springer, 2018.
[4] M. Aggarwal, A. Arora, S Sodhani, and B. Krishnamurthy, Improving Search Through A3C Reinforcement Learning Based Conversational Agent, International Conference on Computational Science, Vol. 3, pp. 273-286, 2018.