KNS2020 Autumn Meeting

Re-exploration of pretrained artificial intelligence model for a nuclear power plant autonomous operation

Dec 17, 2020

Jae Min Kim

Hong Jun Yang

Seung Jun Lee*

Ulsan National Institute of Science and Technology
Address 50 UNIST-gil, Ulju-gun, Ulsan, 44919, Korea
Tel. +82 52 217 2371 **Web.** psa.unist.ac.kr

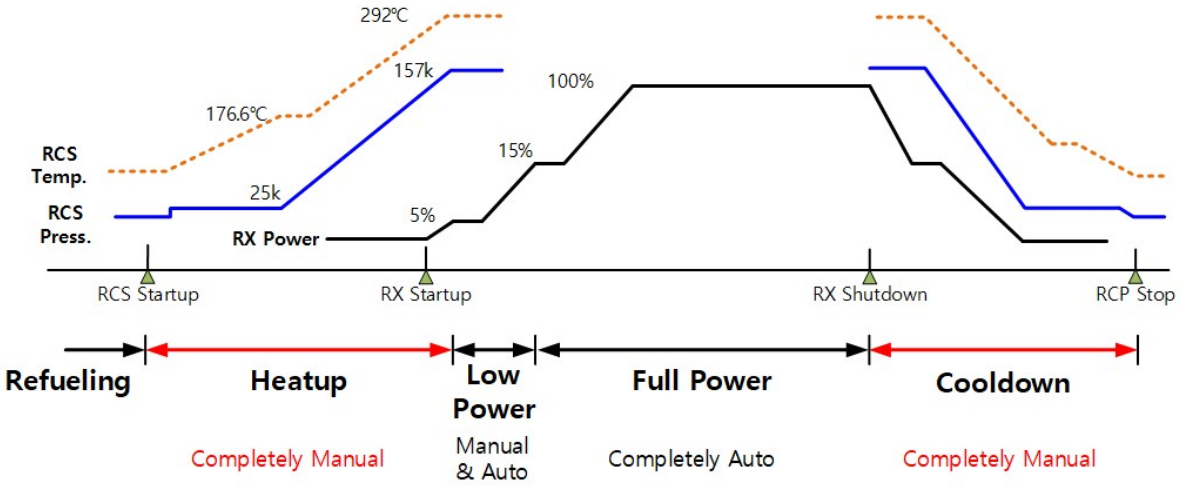
CONTENTS

1. Introduction
2. Purpose
3. Framework of autonomous operation
4. NPP environment
5. Feasibility study
6. Application
7. Summary
8. Future work

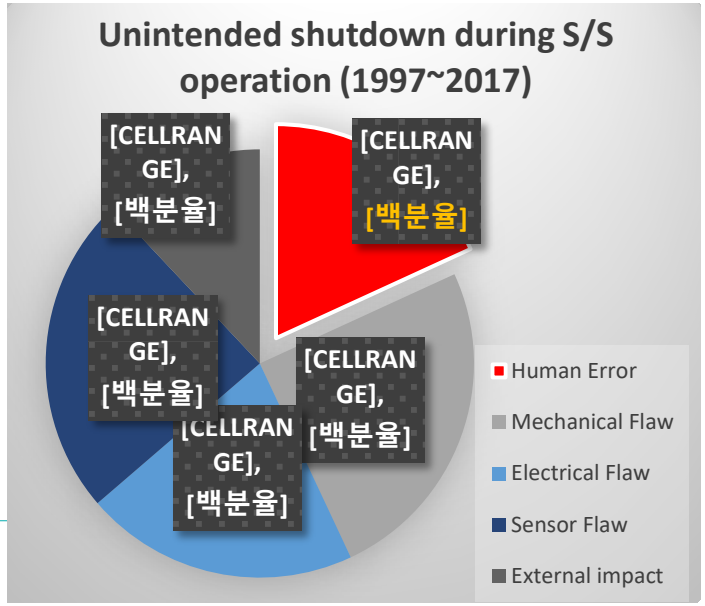
Introduction

❖ Characteristics of startup and shutdown operation

- Cold-shutdown mode ↔ Hot-standby mode : completely manual
- More than 20 hours to finish startup and shutdown operation
- High burden for operators – 18% of unintended shutdown during startup and shutdown operation by human error



<Operating mode during an NPP operation>

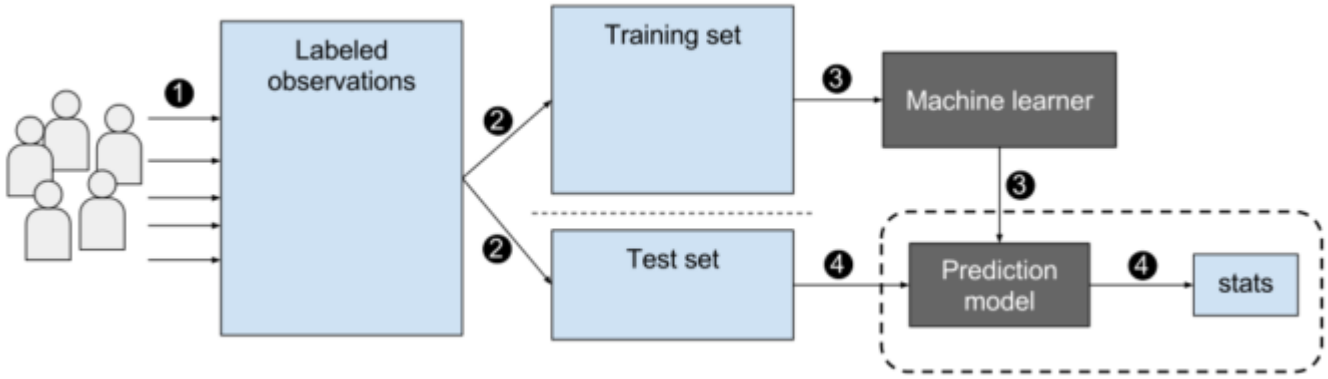


Ref. Advanced MMIS toward substantial reduction in human errors in NPPs (2013)
Human error

Introduction

❖ Characteristics of startup and shutdown operation

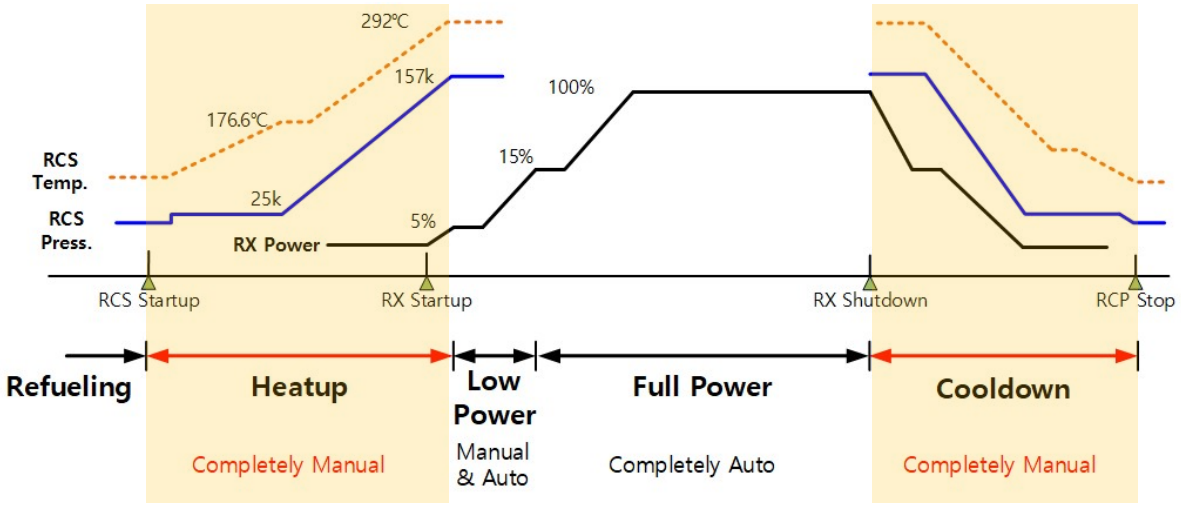
- Cold-shutdown mode ↔ Hot-standby mode : completely manual
- More than 20 hours to finish startup and shutdown operation
- Supervisory learning requires huge amount of training datasets
- Lack of operation datasets



<Example of supervisory learning>

Purpose

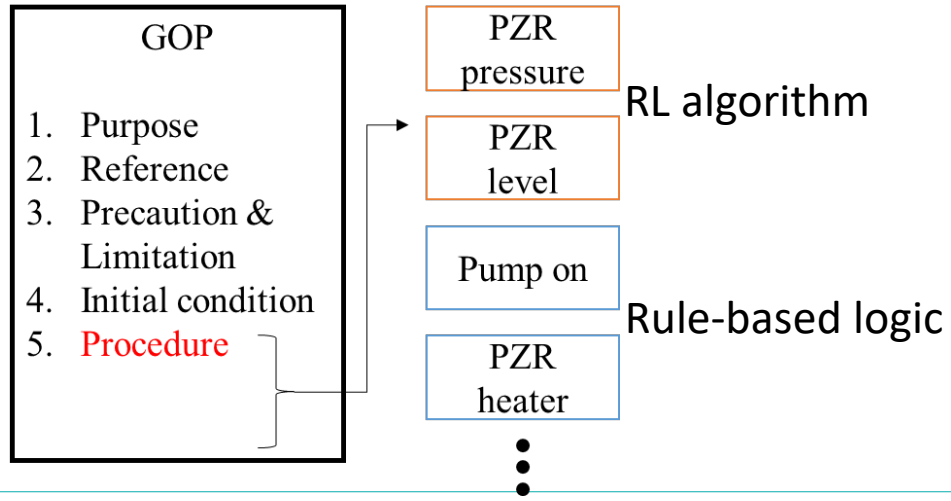
- **To reduce the human error by eliminating human intervention**
- **Develop autonomous operation system for startup and shutdown operation**
  - Framework for autonomous operation is suggested
- **Determine what algorithms can be applied to NPP operation**
  - Agents are expected to learn operating policy for each block

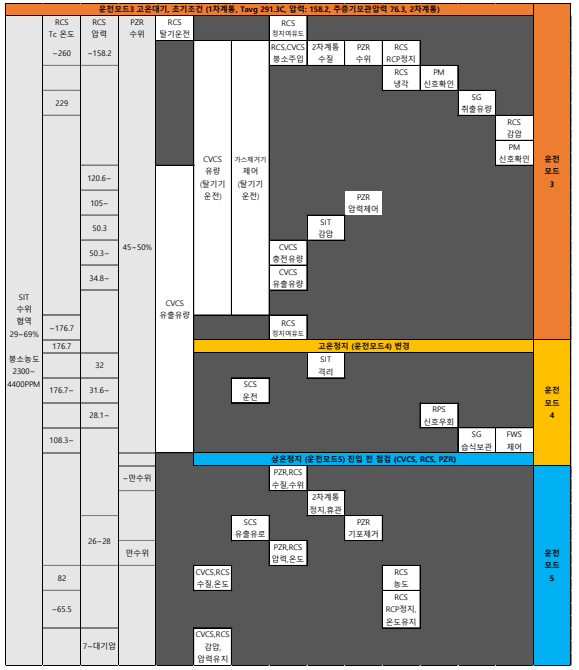


Framework of autonomous operation

❖ General operating procedure (GOP)

- Including refueling, **startup**, **shutdown** and power operation
  - Limiting condition for operation (LCO),
  - Components control
  - Check the plant status
- GOP analysis → Operating blocks





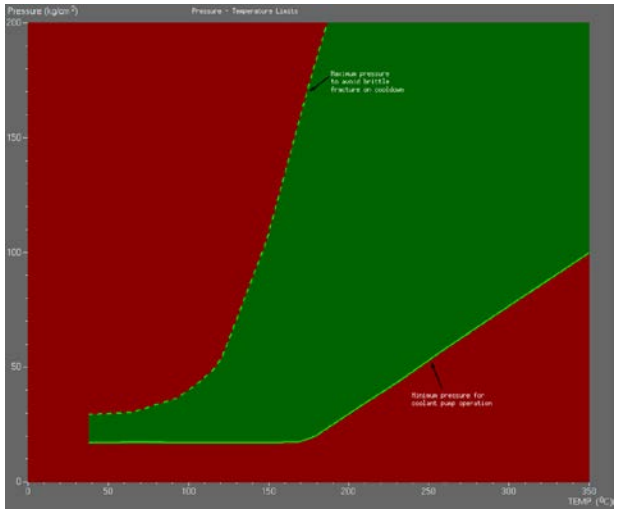
Framework of autonomous operation

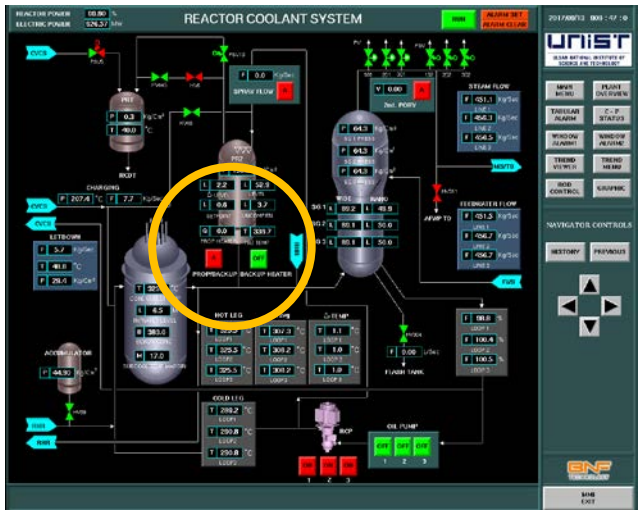
❖ General operating procedure (GOP)

- GOP analysis → Block operation
- Operating blocks have own entry and terminal conditions
- Algorithms can be either reinforcement learning or rule-based logic

RCS pressure and temperature must be maintained within the limitation curve.

→ **A.I. agent** (complicated)



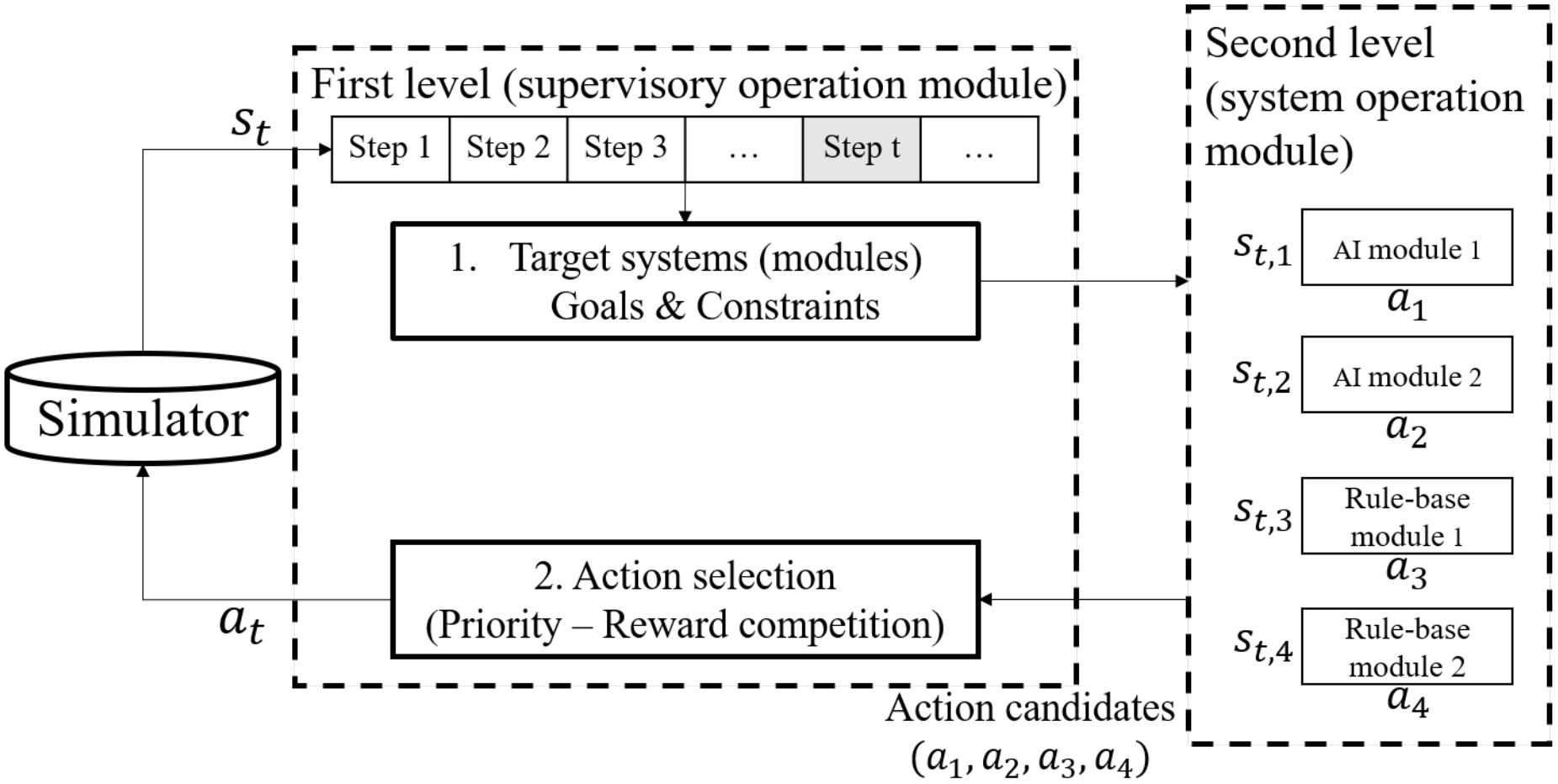


Energize all pressurizer heater groups and start increasing the pressurizer temperature.

→ **Rule-based logic** (simple)

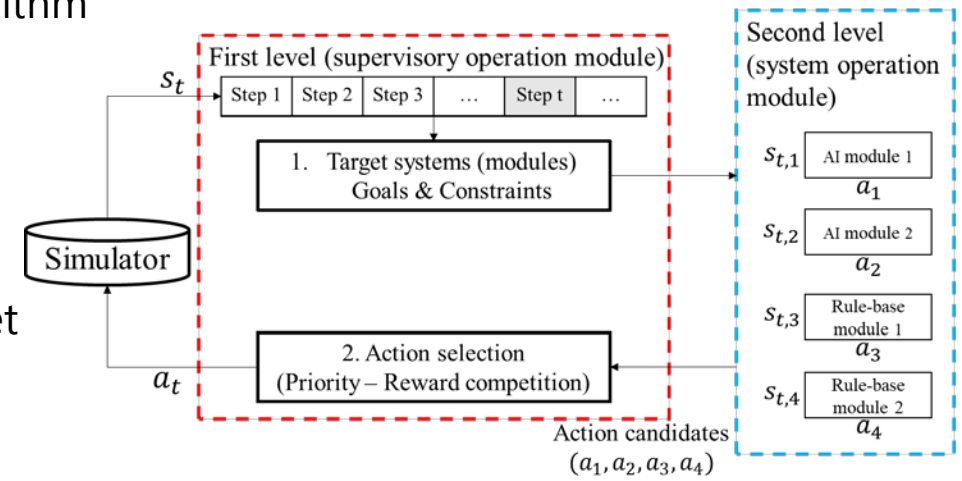
Framework of autonomous operation

❖ **Framework of autonomous operation system for NPPs**



Framework of autonomous operation

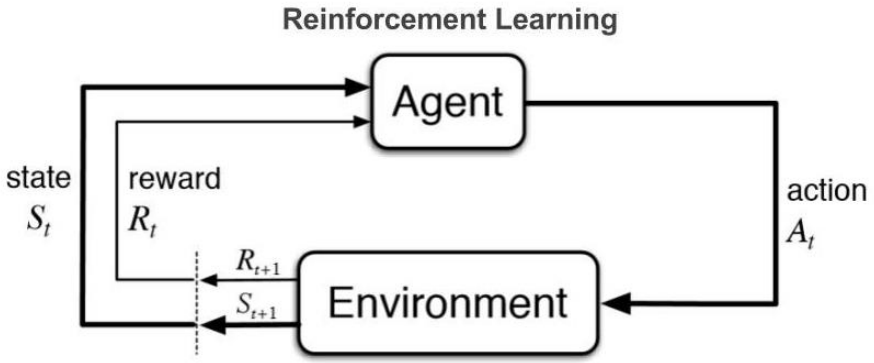
1. **Supervisory operation module (1st level)**
   - Input: states from an NPP environment
   - Monitoring NPP status based on RCS temperature and pressure
   - Activation of operating blocks in the 2nd level
2. **System operation modules (2nd level)**
   - Module for system and component operation
   - Blocks including AI or rule-based algorithm
   - Output: actions
3. **Action selection (1st level)**
   - Collect action candidates
   - If there is a conflict, the final action set is selected based on criteria

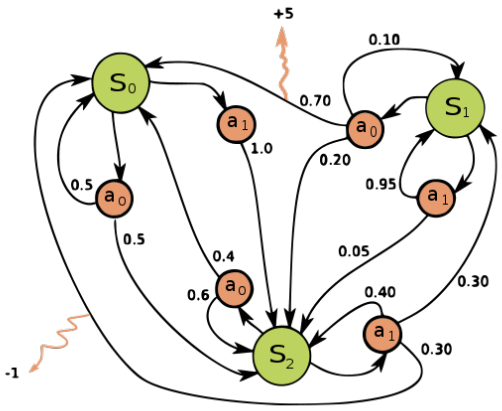


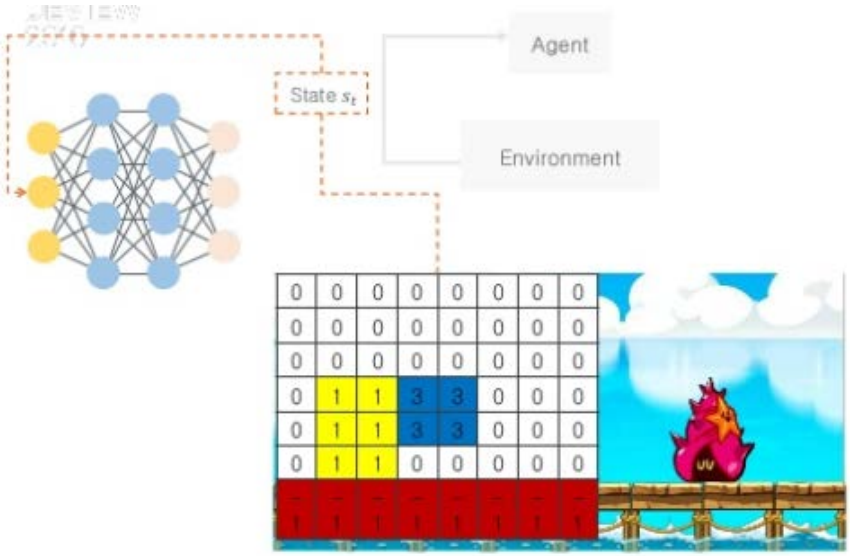
NPP environment

❖ NPP to training environment

- Define an NPP as an environment where agents can interact with
- State, action, reward, state transition probability, discount factor



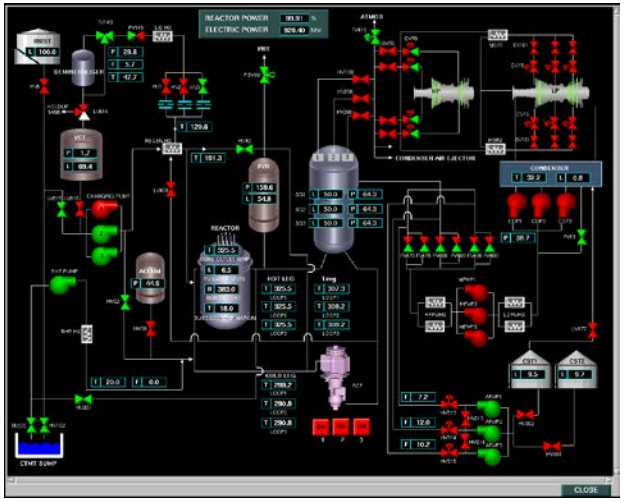


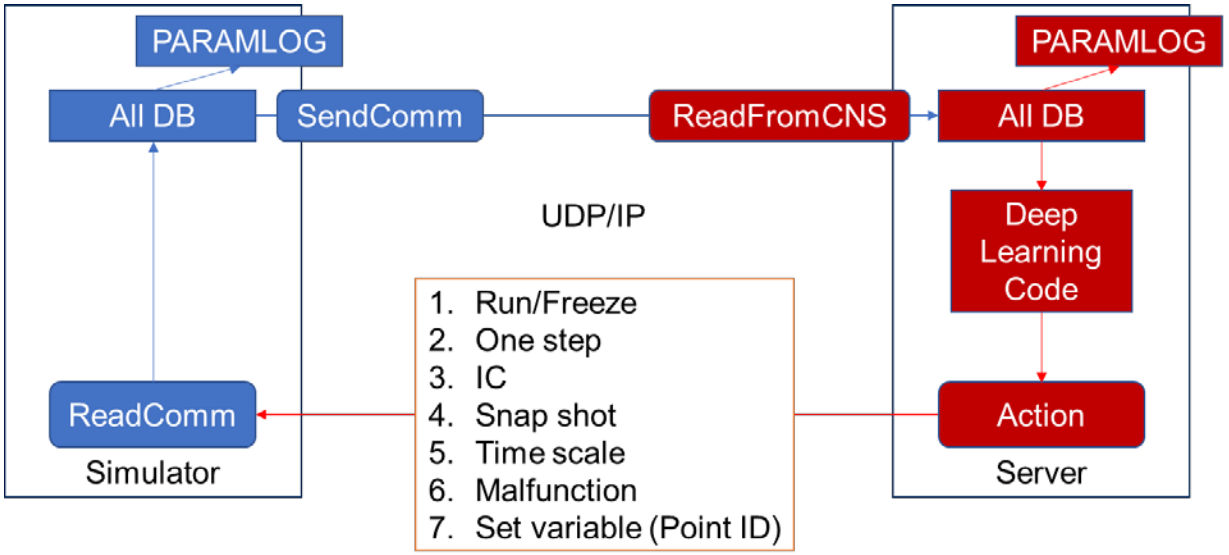


NPP environment

❖ Training environment (Compact Nuclear Simulator, 993MWe WH 3-loop PWR)

- State: physical features representing an NPP status
  - RCS temperature, PZR pressure, PZR level, ...
- Action: controllable components
  - Valves, pumps, PZR heater, ...

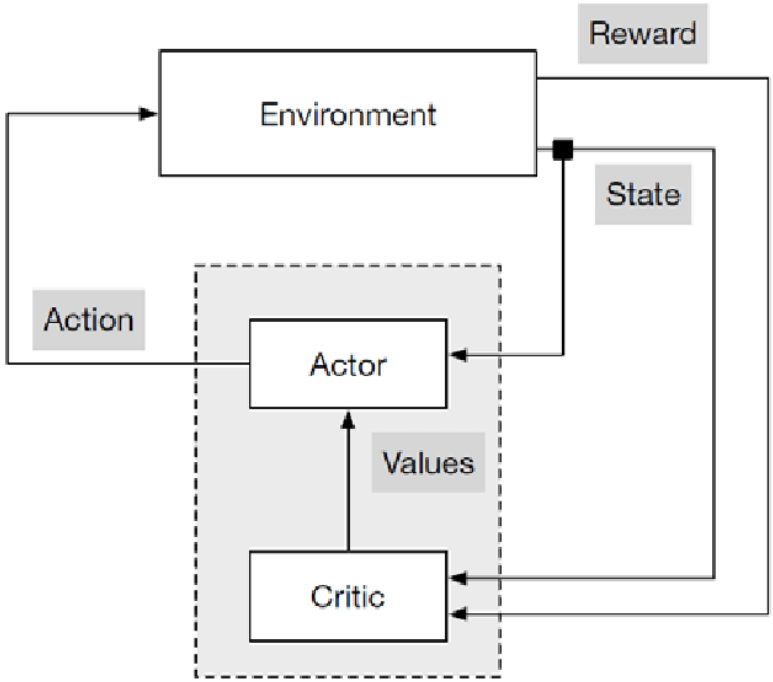




Feasibility study

❖ Advantage Actor-Critic (A2C) algorithm

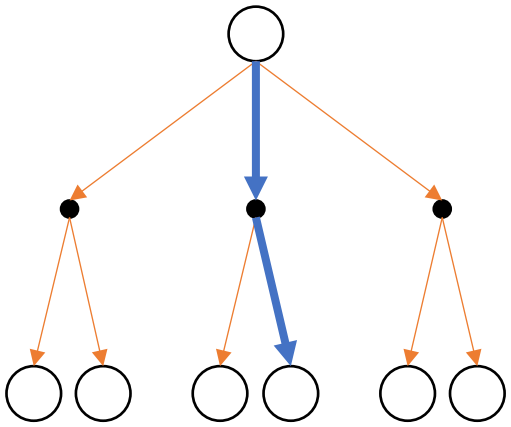
- One of policy gradient algorithms
- Actor determines action
  - Output shape: [action size]
  - Action selected stochastically
- Critic evaluates state value
  - Output shape: [1,]
  - Advantage → Actor



Feasibility study

❖ Exploration vs. Exploitation

- RL agent can undergo enormous pathways for episodes
- Exploitation sticks to the best decision based on current information
- Exploration tends to take more pathways
- Exploration and exploitation are trade-off relationship

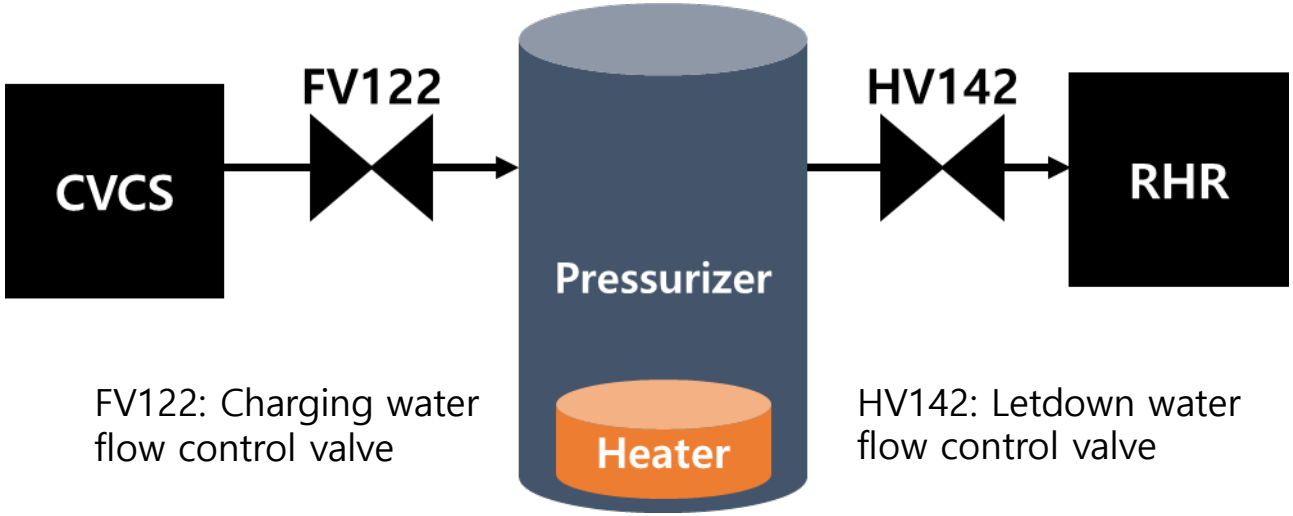


Feasibility study

❖ Pressurizer pressure control

Parameter	Unit
PZR pressure	kg/cm^2
RCS temperature	°C
FV122 status	-
HV142 status	-
Trend value	kg/cm^2

Action list (total 5 actions)
No action
FV122 Close for 2 seconds (-1.5%)
FV122 Open for 2 seconds (+1.5%)
HV142 Close for 2 seconds (-1.5%)
HV142 Open for 2 seconds (+1.5%)
FV122 Close & HV142 Close for 2 seconds
FV122 Close & HV142 Open for 2 seconds
FV122 Open & HV142 Close for 2 seconds
FV122 Open & HV142 Open for 2 seconds



FV122: Charging water flow control valve

HV142: Letdown water flow control valve

CVCS: Chemical Volume Control System
RHR: Residual Heat Removal system
PZR: Pressurizer

Feasibility study

❖ Pressurizer pressure control

- Reward function is designed to get a positive reward as PZR pressure approaches to the target value
- The agent gets a negative reward if PZR pressure moves away from the target value
- The agent gets large penalty if the episode ends up early

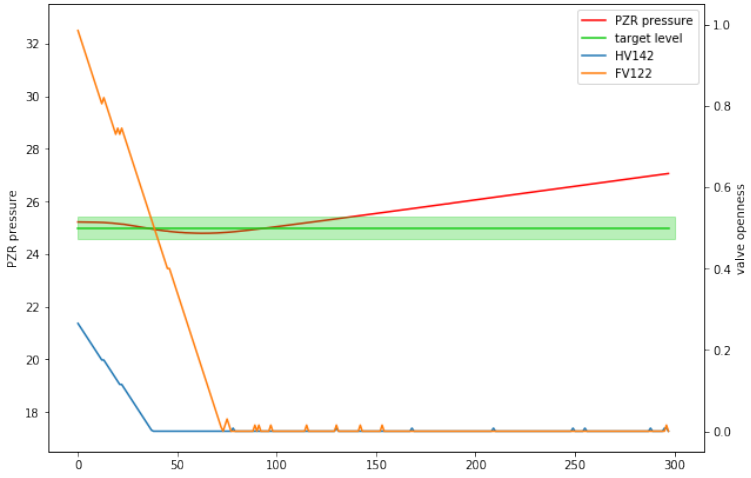
$$R = \begin{cases} 0.5 * (BC_u - BC_d) - (abs(P_t - P_c)) \\ (-1) * (step_{max} - step_c)\ (if\ done) \end{cases}$$

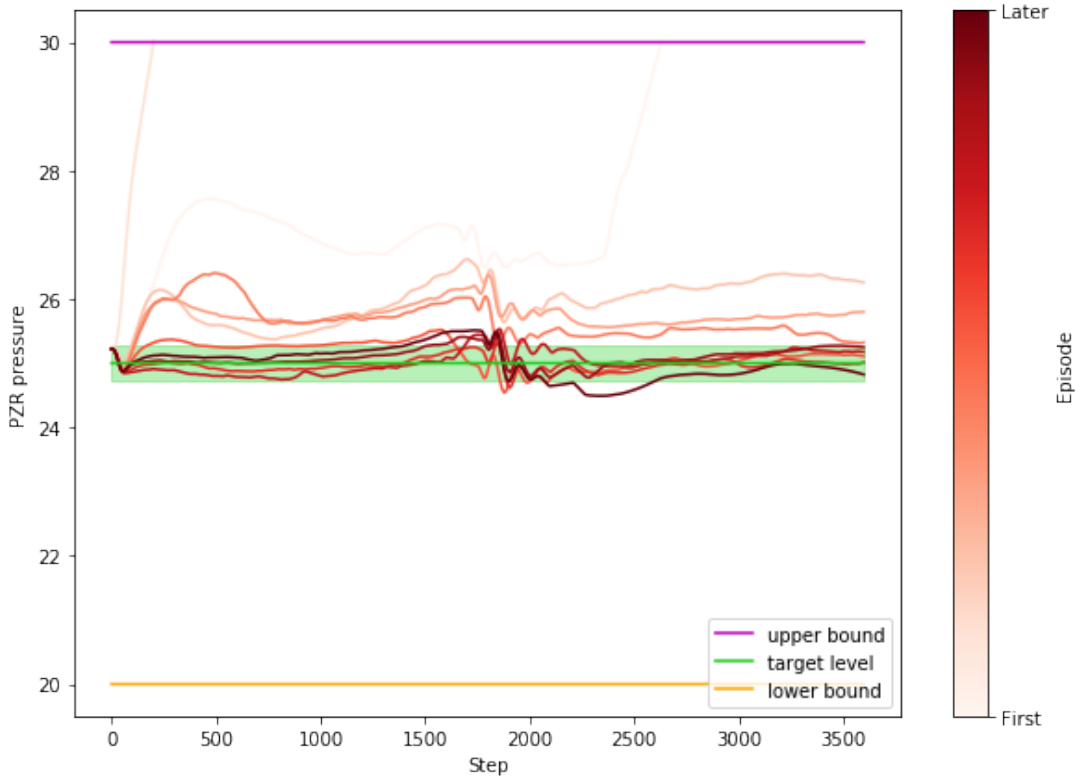
BC_u: upper boundary condition
BC_d: lower boundary condition
P_t: target pressure
P_c: current pressure

Names	Values
Target value	25 kg/cm^2
Terminal condition	20 ~ 30 kg/cm^2
Actor layers	3 (24 hidden nodes)
Actor activation function	Softmax
Critic layers	3 (24 hidden nodes)
Critic activation functions	Relu
Learning rate	0.001
Max episodes	(1) 500 (2) 2000
Max time step	(1) 300 (2) 3,500
Discounting factor	0.99

Feasibility study

- The results of pre-training were outside the target range
- After 300 pre-training, the agent restarted longer episodes (2000)
- The final agent successfully controls PZR pressure staying the target range



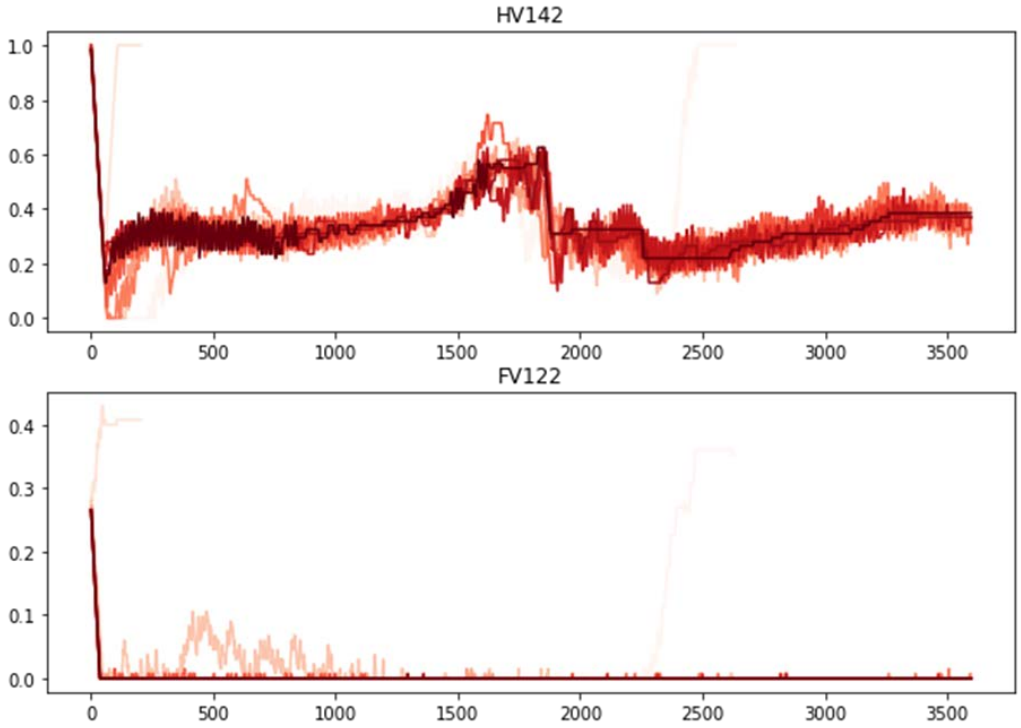


Feasibility study

- When we checked the last episode, PZR pressure successfully stayed in the target area and scored high
- The charging water flow control valve was closed, and operation is performed only with the letdown water flow control valve.

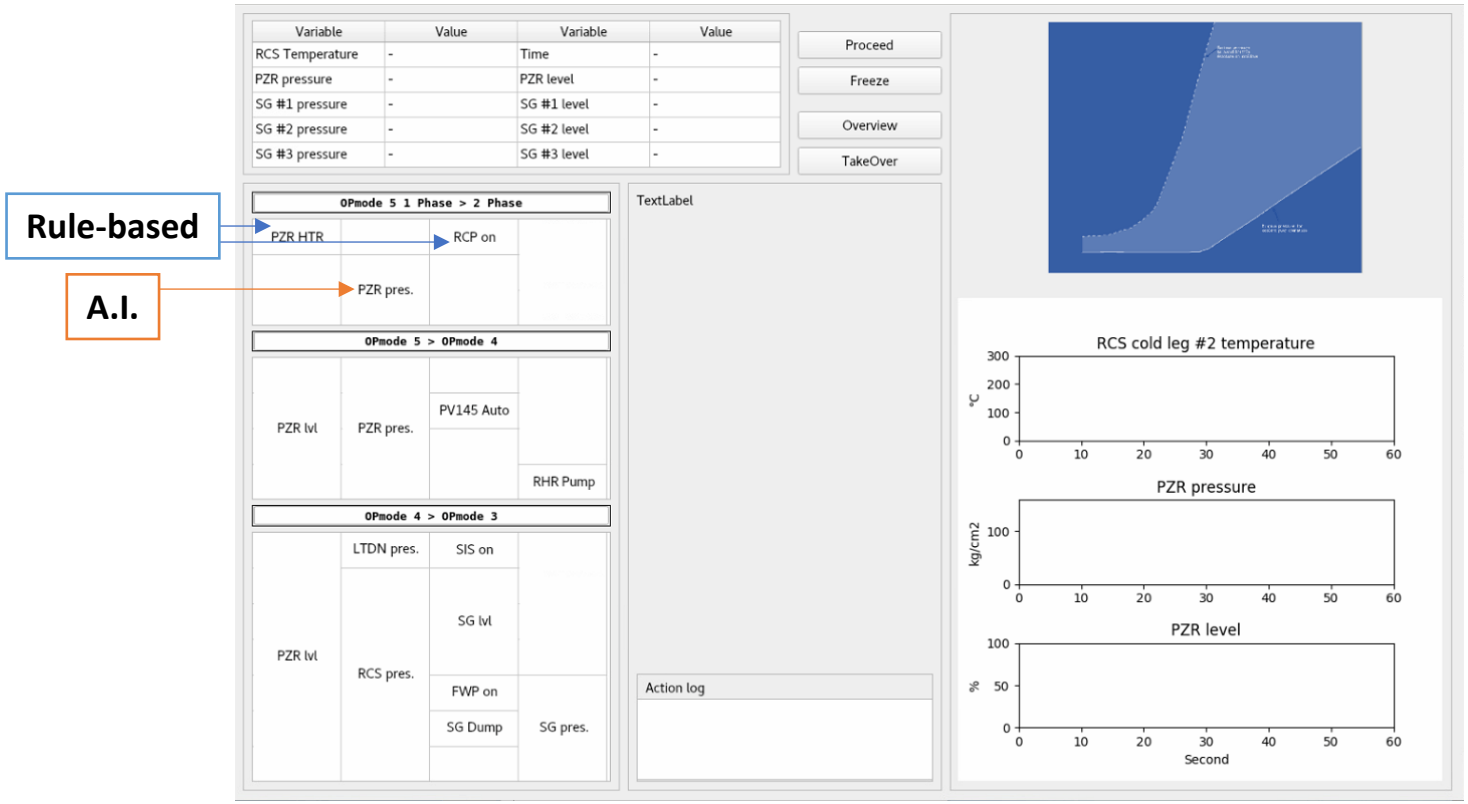
Feasibility study

- When we checked the last episode, PZR pressure successfully stayed in the target area and scored high
- The charging water flow control valve was closed, and operation is performed only with the letdown water flow control valve.



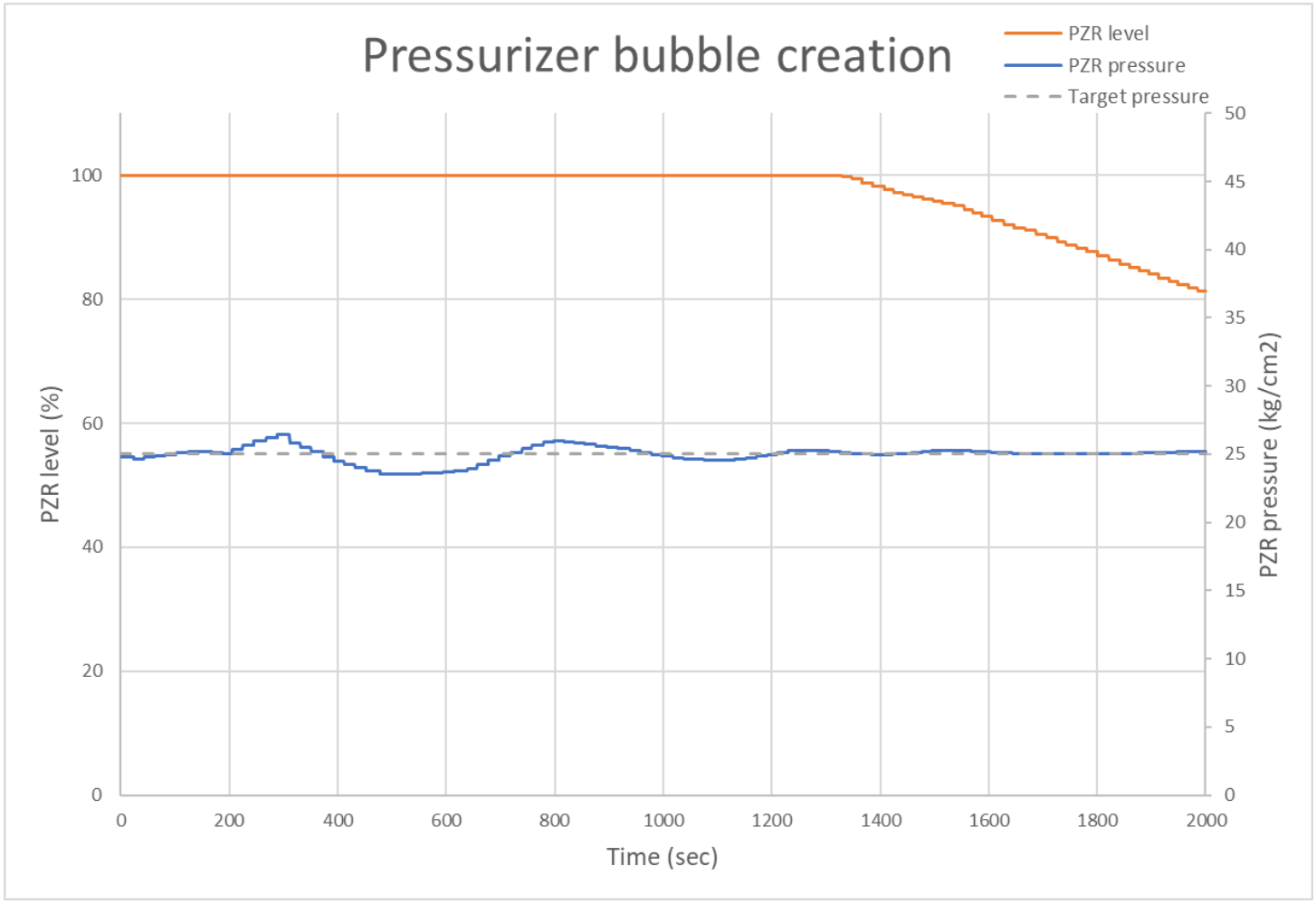
Application

- Combined operating blocks for startup operation
- Solid phase (before bubble creation)
  - FV122, HV142 control PZR pressure



Application

- PZR bubble creation - success



Summary

1. When considering the characteristics of startup/shutdown operation, it is advantageous to implement autonomous operation system using reinforcement learning.
2. Framework of autonomous operation system consisting of two modules
   - Monitoring an NPP status and activating necessary blocks
   - Operating blocks for system and component operation
   - Priority among outputs from the second level
3. After building a training environment using an NPP simulator, feasibility study was conducted using actor-critic algorithm.
4. Re-exploration of the pre-trained agent led to successful learning.
5. As an application, PZR bubble creation operation was implemented by integrating operating blocks.

Future work

- Application: training more operating blocks
  - Increase the number of A.I. blocks
  - Expand operating section of the integrated model
- When expanding an operating range, there may be conflicting actions in the supervisory module
- Future work is about preparing for possible action conflicts in supervisory operation module


THANK YOU
unist
ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY
2009
FIRST IN
CHANGE