

# RL-Based Control of Benchmark CSTR Processes with Safety Considerations

Seunghyun Cho, Hain Lee, Jaehyun Oh

May 2, 2025

## 1 Backgrounds

### 1.1 CSTR

Continuous Stirred Tank Reactor (CSTR) is a representative chemical reactor in which fluid is continuously fed, and reaction and mixing occur simultaneously. It is assumed that the reactants are perfectly mixed within the reactor, and the system is characterized by dynamic features such as reaction rate, concentration, and temperature. In particular, due to nonlinearity, even small input variations or noises can significantly affect the system behavior. Thus, this system requires precise control strategies.

### 1.2 Process Control

Process Control is a field of engineering that aims to ensure stable operation and optimal performance of industrial processes. It is utilized in various industries such as chemical, petrochemical, and pharmaceutical to control variables (e.g., temperature, pressure, flow rate) of various equipment including reactors, distillation columns, and heat exchangers. Due to complex process characteristics such as nonlinearity and time delay, control methods beyond traditional Proportional-Integral-Derivative (PID) control—such as Model Predictive Control (MPC) and reinforcement learning-based approaches—are being actively studied.

## 2 Project Summary

The failure of Process Control can directly lead to safety issues. Due to the nonlinear characteristics of the CSTR process, it poses challenges in control, requiring more flexible and adaptive control strategies. Despite the importance of safety, only a few studies have focused on utilizing reinforcement learning methods to CSTR processes in the context of safety. This project aims to address this gap by proposing an RL-based control framework that explicitly considers safety constraints in the operation of CSTR processes. Accordingly, this study applies deep reinforcement learning algorithms such as Deep Q-Network (DQN), Proximal Policy Optimization (PPO), and Deep

Deterministic Policy Gradient (DDPG) to perform dynamic control of the process, and seeks to validate the variation in control performance and potential for ensuring process stability depending on reward function design through the Van de Vusse reaction, a benchmark for CSTR processes.

## 3 Related Works

### 3.1 Library

Bloor et al. increased the accessibility of RL research in Process Control by proposing a standardized environment for the field.[1]

### 3.2 Previous Researchs

Park et al. improved control performance by 12% compared to MPC by optimizing the reward function considering the nonlinearity of the CSTR.[2] Yu et al. successfully achieved stable control of the CSTR using a PPO-based control framework, even under disturbances.[3] Bloor et al. proposed a standardized environment for comparing the performance of reinforcement learning and nonlinear MPC.[4] Chen et al. (1995) provided an initial benchmark for the nonlinear control problem of CSTR, laying the foundation for subsequent studies.[5] Additionally, Park et al. provided a comprehensive review of RL applications in Process Control, confirming its effectiveness in controlling the complex dynamics of the CSTR.[6]

## 4 Project Milestones

1. Modify the reaction equation in the PC-Gym library to describe the Van de Vusse reaction, and add the Tank Level used in [3] as an additional consideration. Redefine the corresponding States and Actions to suit this new configuration.
2. Implement baseline models for each algorithm using DQN, PPO, and DDPG.
3. Adjust the reward function and explore methods to enhance the benchmark performance of each baseline model.

## References

- [1] MaximilianB2, “pc-gym: Process control environments for reinforcement learning,” <https://github.com/MaximilianB2/pc-gym/tree/main>, accessed: 2025-04-29.
- [2] 박준수, “강화학습 기반 화학 공정 제어 성능 향상을 위한 보상 함수 시뮬레이션 연구,” *제어 로봇 시스템학회 논문지*, 2024.
- [3] M. Yu, B. Li, S. Zhao, N. Roy, and B. Zhang, “Ppo-based resilient control framework for safer operation of exothermic cstr,” *Process Safety and Environmental Protection*, vol. 193, pp. 558–576, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957582024014721>
- [4] M. Bloor, J. Torracca, and C. Tsay, “Pc-gym: Benchmark environments for process control problems,” *arXiv preprint arXiv:2410.22093*, 2024. [Online]. Available: <https://arxiv.org/abs/2410.22093>
- [5] H. Chen, H. Kremling, and F. Allgöwer, “Nonlinear predictive control of a benchmark cstr,” *Proceedings of the 3rd European Control Conference, Rome-Italy.*, pp. 3247–3252, 01 1995.
- [6] J. Park, H. Jung, J. W. Kim, and J. M. Lee, “Reinforcement learning for process control: Review and benchmark problems,” *International Journal of Control, Automation and Systems*, vol. 23, pp. 1–40, 2025.