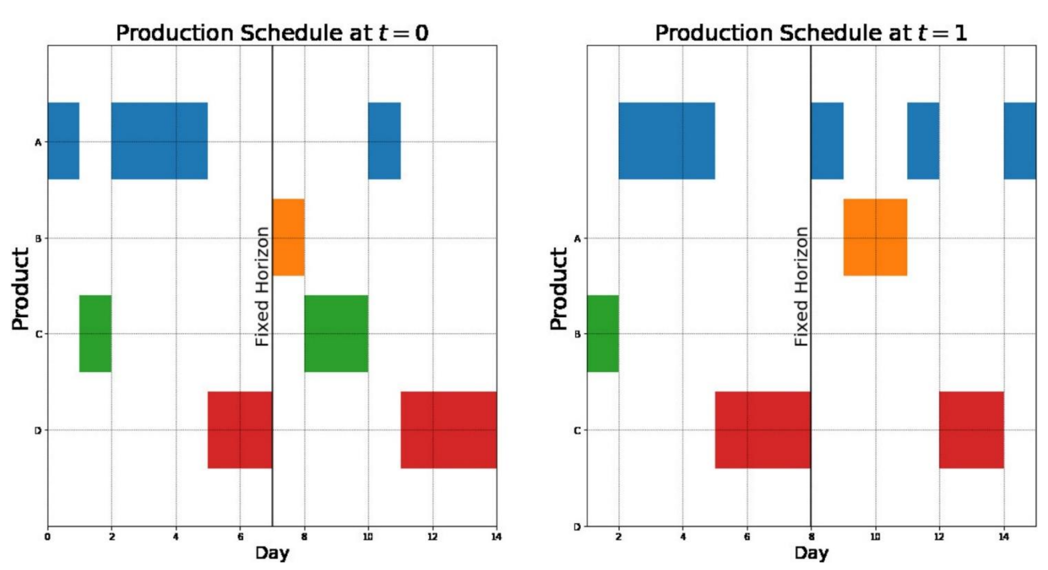
An Industrial Application of Deep Reinforcement Learning for Chemical Production Scheduling

Christian D. Hubbs[*,†]
Adam Kelloway[*]
Nikolaos V. Sahinidis[‡]
Ignacio E. Grossmann[†]
John M. Wassick[*]

Motivation

- **Goal**
  Develop a model to schedule a continuous chemical reactor owned and operated by Dow
- **Motivation**
  Scheduling is a difficult, stochastic combinatorial optimization problem
  The role is critical for maintaining operational stability

Chemical Reactor Scheduling



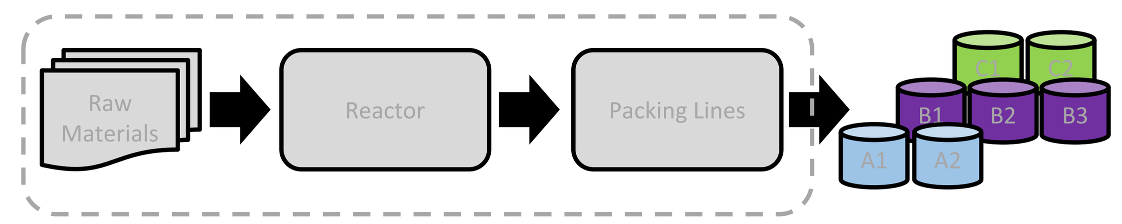
Gannt chart showing plant with four products. Production plants often impose additional constraints on schedules such as not making changes within the next week to allow employees to have stable shifts and avoiding certain product transitions which may incur high costs or place the equipment under large stresses. [1]

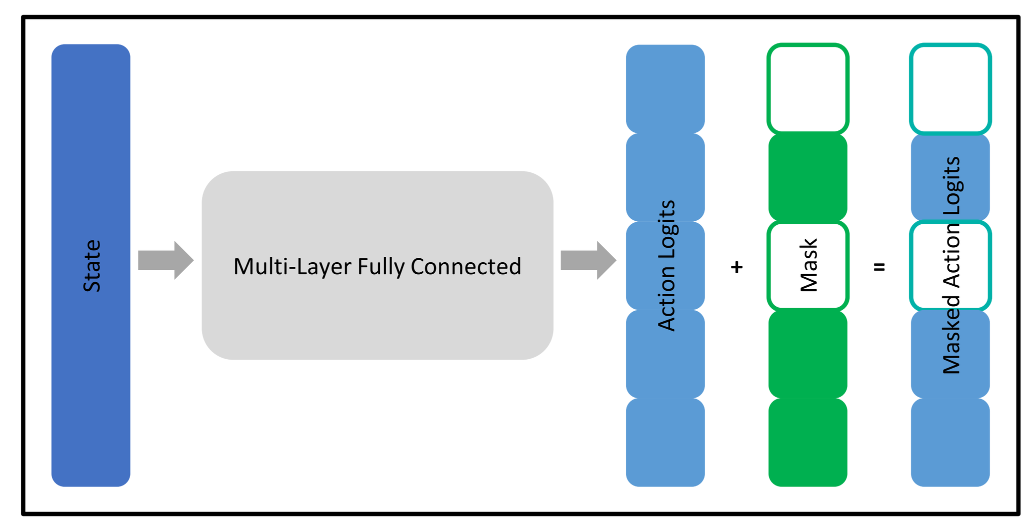
- Scheduling is critical for maintaining steady and successful operations. Good schedules lead to lower inventory, higher customer service levels, less off-spec material, and better margins.
- The job is very stressful and leads to high turnover rates for schedulers due to constant changes from pricing, customer requests, and production interruptions.
- Classic solutions such as stochastic programming models are not used due to computational costs, issues with scalability, and difficulty maintaining such models.
- Reinforcement learning provides a potential avenue for addressing the uncertainty in the domain and developing good schedules quickly to support our scheduling staff.

* Digital Fulfillment Center, Dow Chemical, Midland, MI
† Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA
‡ H. Milton Stewart School of Industrial & Systems Engineering and School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA

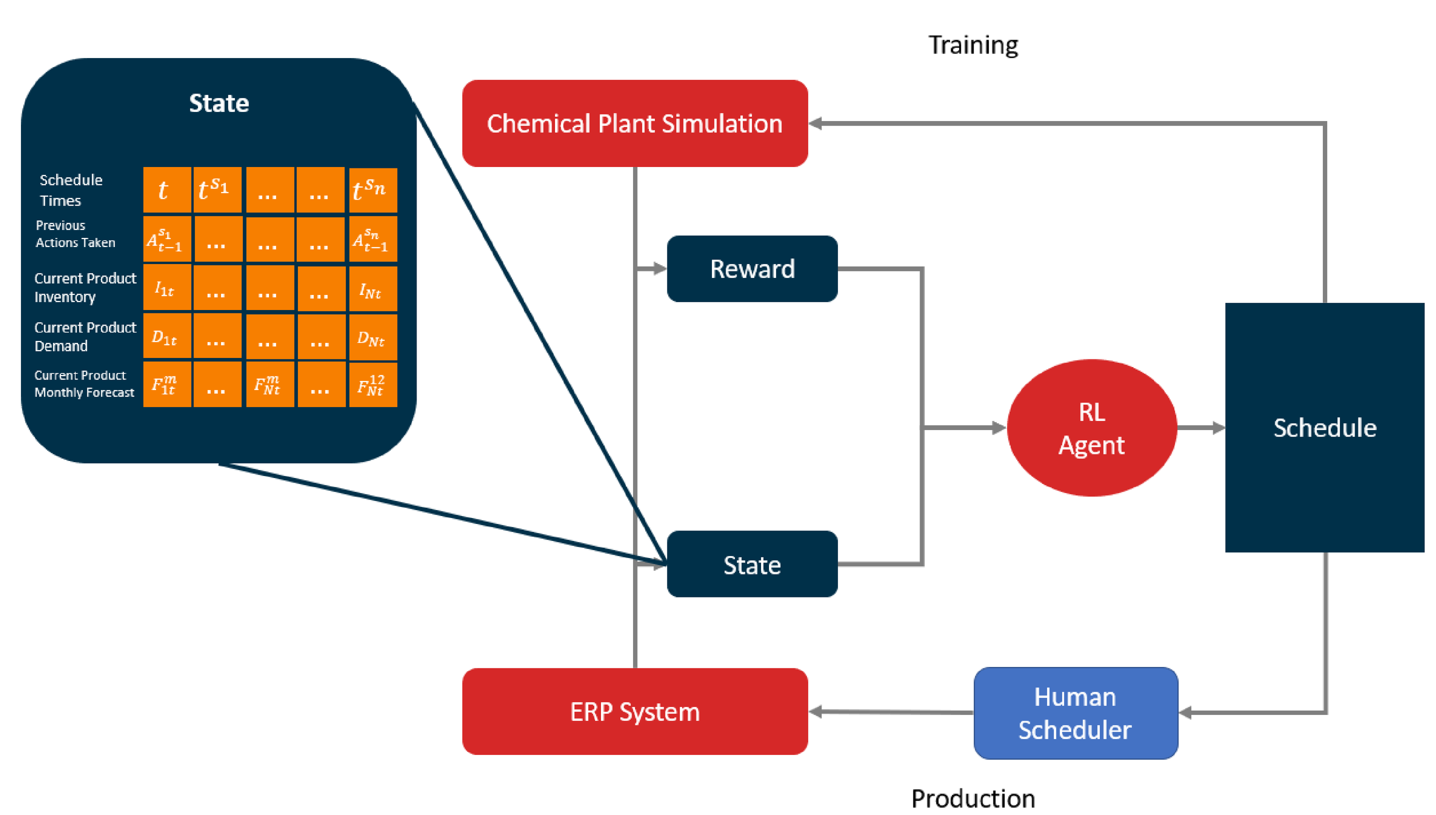
Scheduling Environment

- Environment built to simulate the scheduling and production process of a Dow facility.
- Simulated demand and pricing based on historical data
- Multi-stage model with a reactor followed by a bagging line
- Type-changes create off-grade that eats into margins and reduces the amount of prime product available
- Minimum production constraints maintained
- Trained to optimize trade-off between customer service levels, production costs and inventory to maximize profitability



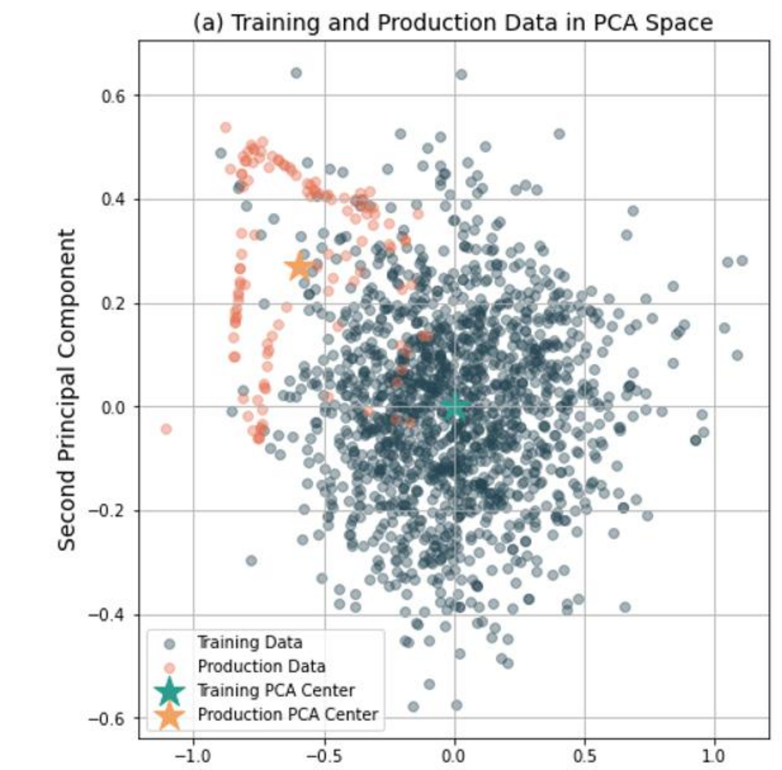


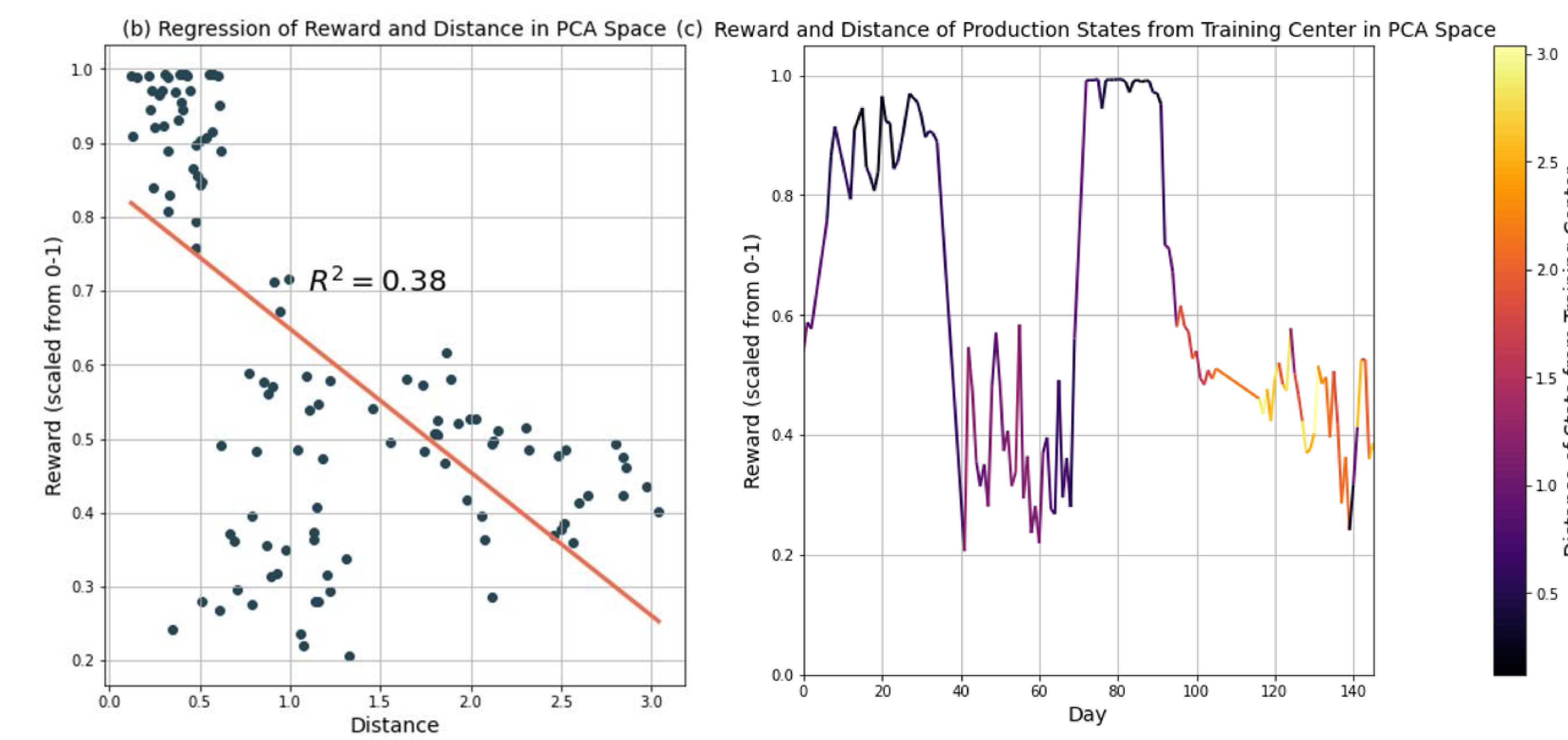
- Trained using Population Based Training (PBT)[2]
- Use PPO algorithm and action masking
- Deployed with human-in-the-loop
  - Allows feedback on schedules
  - Helps ease the business into relying on model-based schedule
  - Important for monitoring



Results

- Good feedback from schedulers
- Reduces time to decisions
- Promotes operational stability
- Use PCA of state space to understand where the system has trained and when it may need additional attention





- Rewards decrease as decision state moves away from center of training data
- Significant drift from training regime caused by upstream data issues stemming from ERP system upgrade which led to a prolonged period of unreliable data

Future Work

- Expand to multi-agent system
- Coordinate schedules across sites and geographical regions to optimize the production and supply network
- Build in planned maintenance periods (e.g. turnarounds) to assist with inventory control when plants are expected to be offline for extended periods of time

References

1. Hubbs *et al.* "A deep reinforcement learning approach for chemical production scheduling." *Computers and Chemical Engineering*, 141, 2020
2. Jaderberg *et al.* "Population Based Training of Neural Networks." 11 2017. URL http://arxiv.org/abs/1711.09846.