Reinforcement Learning for Job Shop Scheduling with Maintenance Operations and Partially Observable Deteriorating Machines

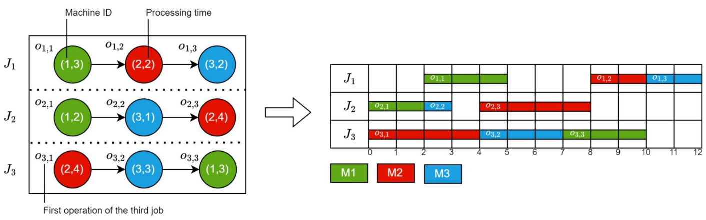
Christoph Schmidl[1], Merlijn Krale[1], Nils Jansen[1,2]

[1]Radboud University, Nijmegen, The Netherlands
{christoph.schmidl,merlijn.krale}@ru.nl
[2]Ruhr-University Bochum, Germany
n.jansen@rub.de

How to learn optimal maintenance schedules, despite machine state uncertainties?

- **Reinforcement Learning (RL):** learning a scheduler for **job shop scheduling problems** with **maintenance operations** and **partially observable machine states.**
- **Active measuring:** act-then-measure technique [1] to reduce amount of **maintenance operations** by revealing the **actual machine state** with a **cost** involved.
- **Classifier:** approximates the machine state based on **deteriorating processing times.**
- **Applications** include **predictive maintenance** and shop scheduling problems.

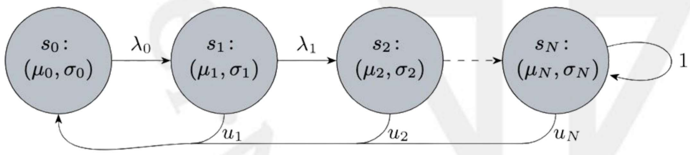
The Job Shop Scheduling Problem



- Every job $J_i \in \mathcal{J}$ must be processed by all m machines in $\mathcal{M}$ in the order given by the operation indices of job J_i

- **Constraints**: Precedence, No-overlap, No-preemption
- **Objective**: Minimizing makespan

Machine State Deterioration



- Deterioration of each machine M_j is modelled by a continuous-time Markov chain (CTMC)
- At each state $s_{i<N}$ the machine has a chance to deteriorate to state s_{i+1} given by the exponential distribution $\lambda_i e^{\lambda_i t}$
- For each deterioration state, we assume processing speeds follow a normal distribution given as $\mathcal{N}(\mu_i, \sigma_i^2)$

Setting: Job Shop Environment

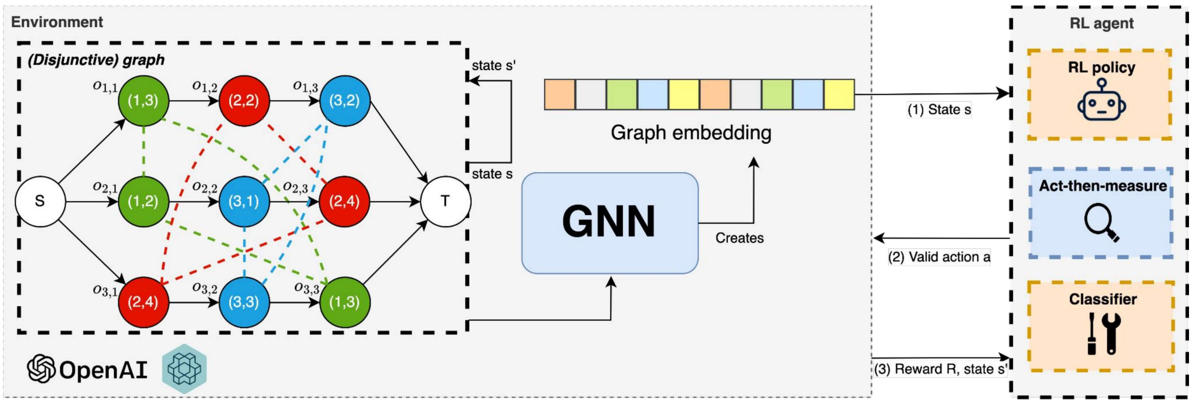


Figure 1: The Reinforcement Learning agent takes the pre-processed graph embedding as its state input. The maintenance classifier detects deteriorating processing times. The act-then-measure component reveals the machine state while reducing the number of costly maintenance operations.

Research outlook

- Modelling the problem as a **partially observable Markov decision process (POMDP).**
- Implementing the RL pipeline including the **act-then-measure** [1] technique and **maintenance classifier.**
- Running experiments on various **benchmark instances** or create **own test suite** with maintenance feature.

References

[1] Krale, M., Simão, T. D., & Jansen, N. (2023). Act-Then-Measure: Reinforcement Learning for Partially Observable Environments with Active Measuring. *arXiv preprint arXiv:2303.08271*.

Acknowledgements

DEUCE. Data-Driven Verification and Learning under Uncertainty.

This research has been partially funded by the NWO project 17931 (SAM-FMS) in the MasCot program and the ERC Starting Grant 101077178 (DEUCE).