Decision-making of Online Rescheduling Procedures Using Neuroevolution of Augmenting Topologies

Teemu Ikonen, Keijo Heljanko and Iiro Harjunkoski

Aalto University
School of Chemical
Engineering

UNIVERSITY OF HELSINKI

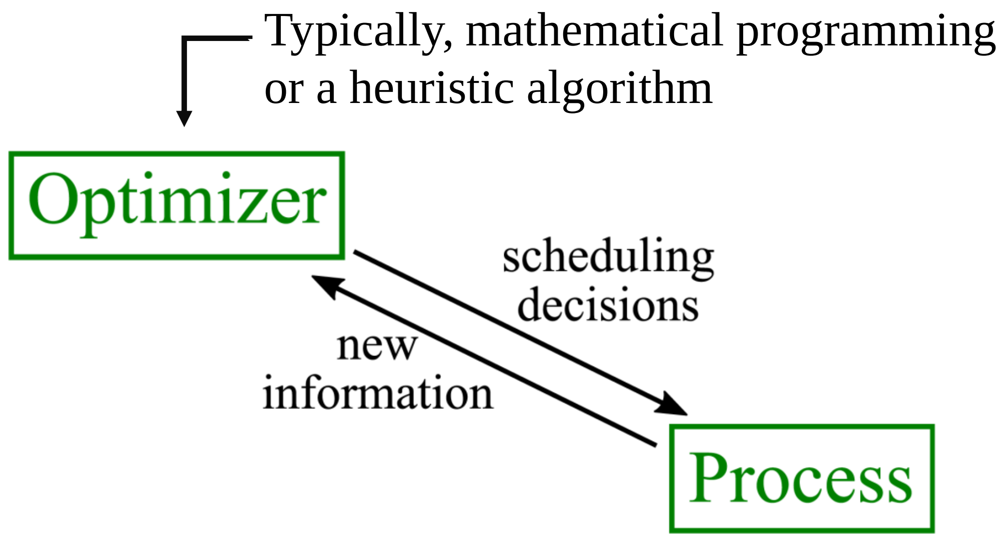
singpro

2019 AIChE Annual Meeting

Orlando, FL, United States

November 12, 2019

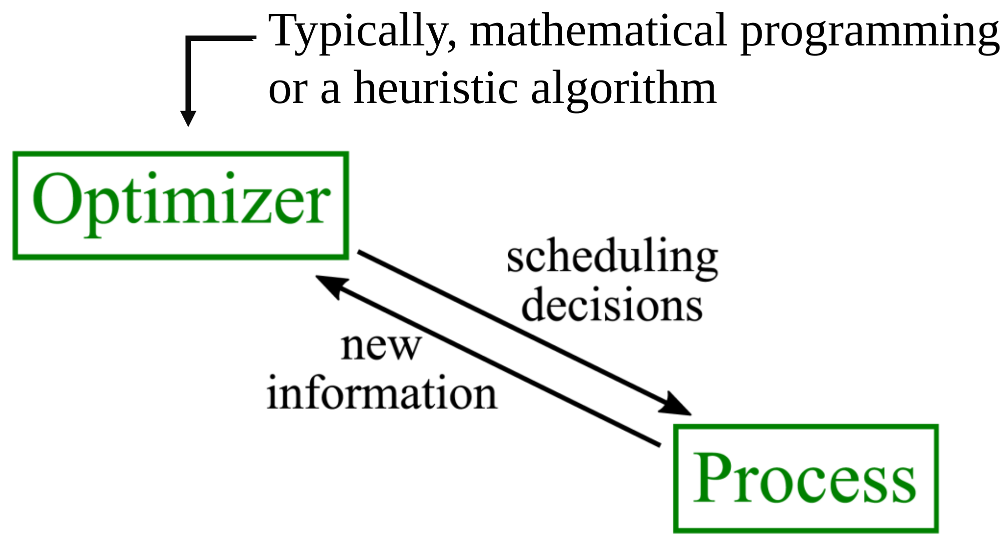
Introduction: Online Rescheduling Decisions



Introduction: Online Rescheduling Decisions

Questions:

- When to trigger a new rescheduling procedure?
- Mathematical programming or a heuristic algorithm?
- How far ahead to schedule (i.e. horizon length)?
- How much computing resource to allocate?



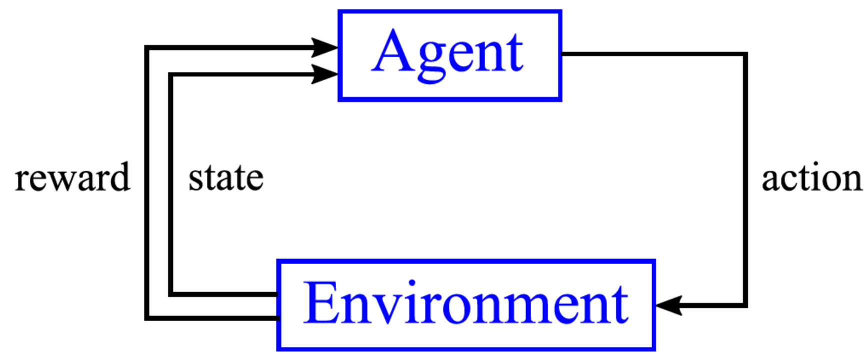
Outline

- **Reinforcement learning (RL)**
  - Introduction
  - Neuroevolution of Augmenting Topologies (NEAT)
- **Proposed approach**
- **Test case**
  - Periodic rescheduling (benchmark)
  - NEAT agent
- **Conclusions**

Reinforcement Learning

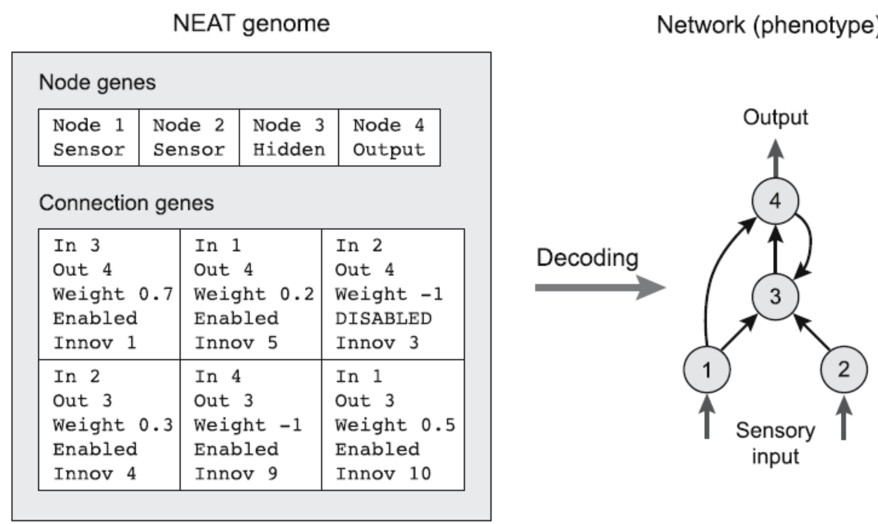
Reinforcement Learning

- **One of the three main branches of machine learning**
  - Along with supervised and unsupervised learning
- **A goal-seeking agent learns while interacting with an environment**
- **Exploration and exploitation**
- **A wide variety of different algorithms**
  - Q-learning (Watkins & Dayan, 1992)
  - Deep Q-Network (Mnih et al., 2015)
  - NEAT (Stanley & Miikkulainen, 2002)
  - trust region policy optimization (Schulman et al., 2015), etc.



Neuroevolution of Augmenting Topologies (NEAT)

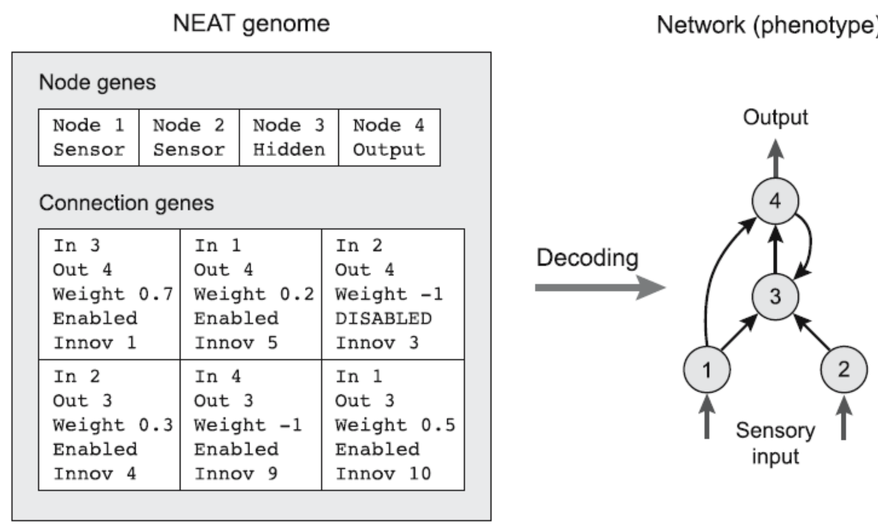
- **First proposed by Stanley and Miikkulainen (2002)**
- **A genetic algorithm that simultaneously evolves the topology and weighting parameters of a neural network**



Graph: Floreano et al., (2005)

Neuroevolution of Augmenting Topologies (NEAT)

- **Complexity of neural network (NN) is minimized**
  - Initiated from very simple NNs
  - The complexity of NNs is incrementally increased during the evolution
- **The performance is reported to compare well against gradient-based backpropagation algorithms (Such et al., 2017)**



Graph: Floreano et al., (2005)

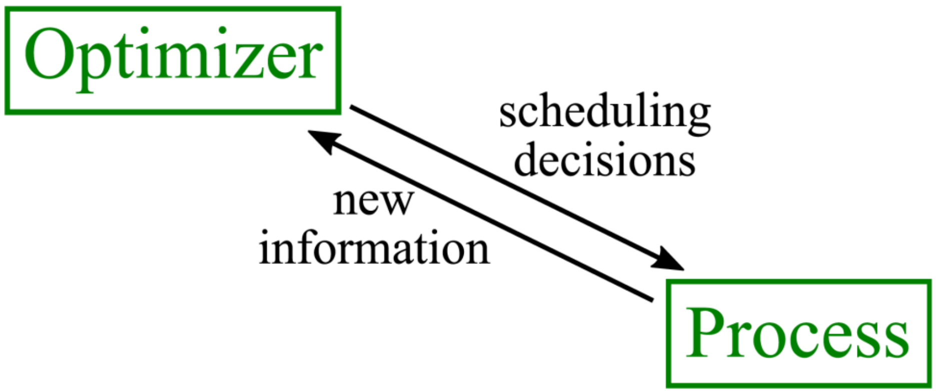
Aalto University
School of Chemical
Engineering

RL in Scheduling

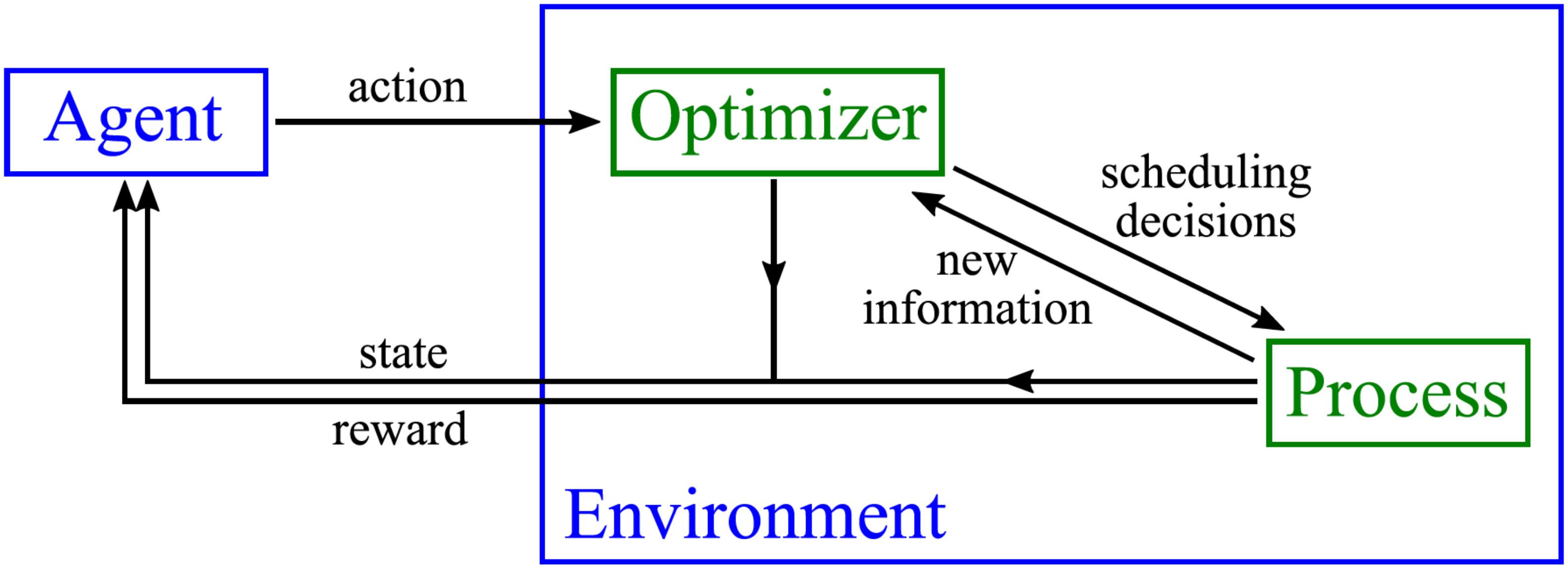
- **RL has been used to**
  - make explicit scheduling decisions (Semrov et al., 2016, Atallah et al., 2018)
  - repair outdated schedules (Palombarini & Martinez, 2012)
  - define dispatching rules from historical scheduling data (Priore et al., 2014, Aydin et al. 2000)

Proposed Approach

Proposed Approach



Proposed Approach



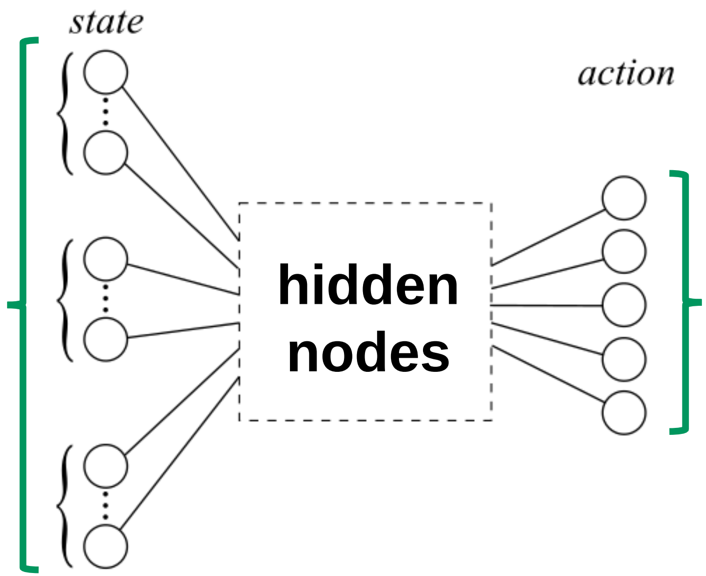
Proposed Approach

Agent's decisions (i.e., actions on the optimizer):

- When to trigger a new rescheduling procedure? — Action 1
- Mathematical programming or a heuristic algorithm? — Action 2
- How far ahead to schedule (i.e. horizon length)? — Action 3
- How much computing resource to allocate? — Action 4

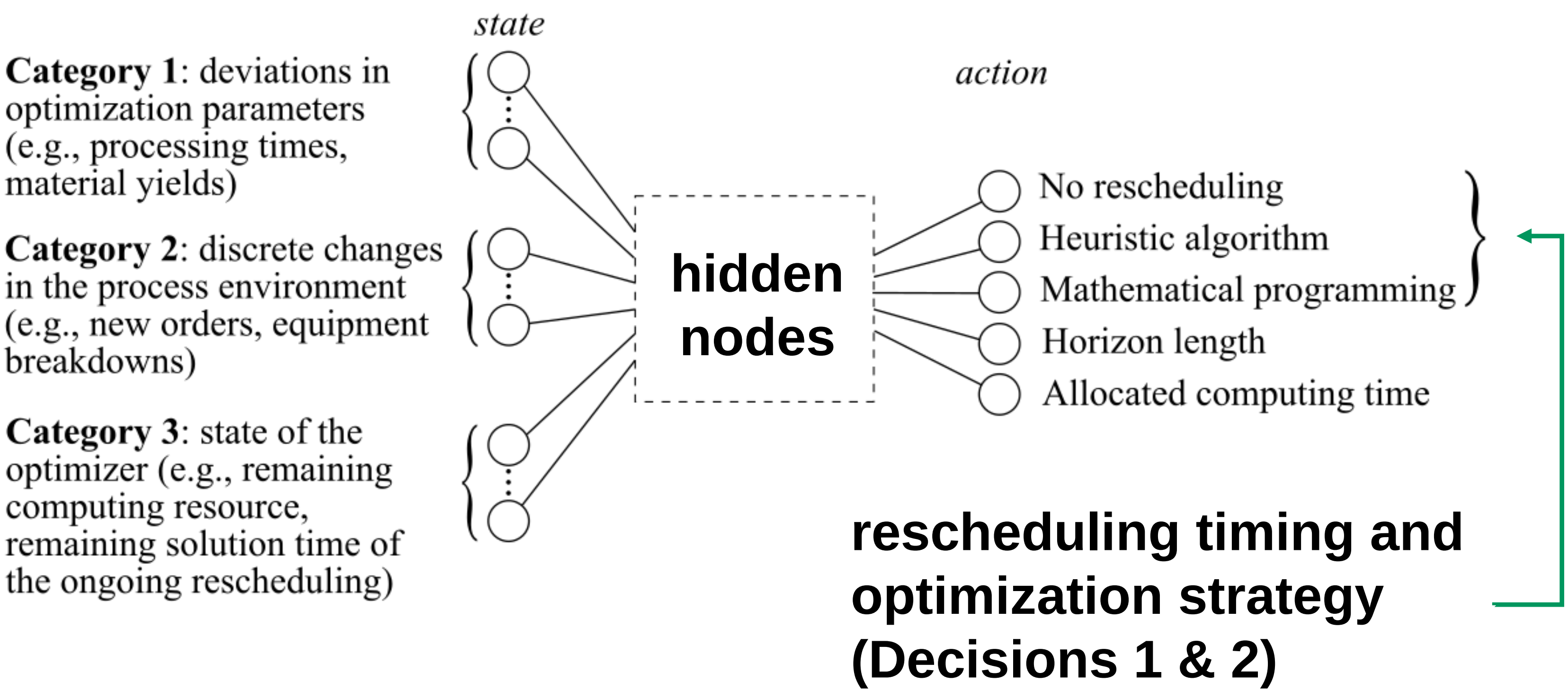
Proposed Approach

changes in the environment (i.e. the process and the optimizer



rescheduling decisions

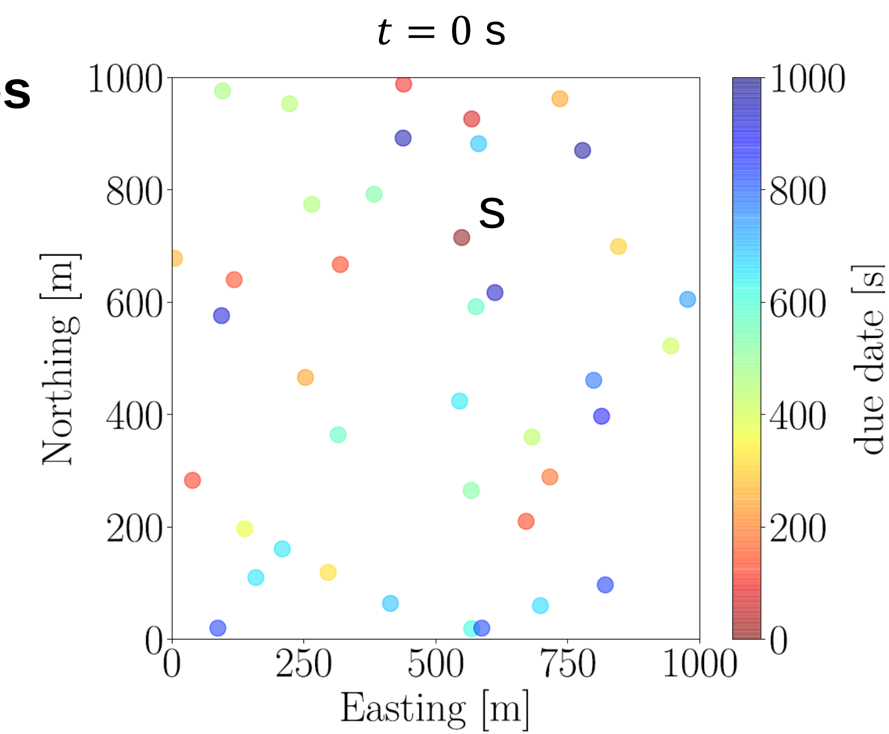
Proposed Approach



Test Case

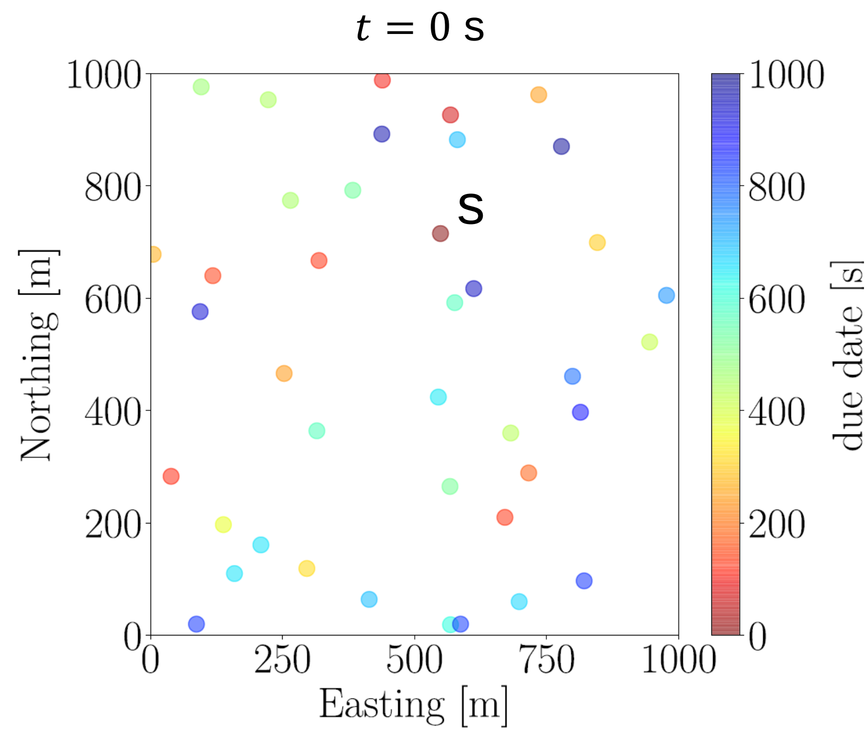
Test Case: Optimization Problem

- **A vehicle visits sites with due dates**
- **The objective is to minimize the delay sum of all visits**
- **Optimizer: ant colony optimization (ACO)**
- **Computing budget for all rescheduling procedures: 50 s**
- **Sites**
  - 40 sites are known at $t = 0$ s
  - 10 sites are received during the process



Test Case: Optimization Problem

- **Randomly generated locations, order and due dates**
- **5 training instances**
- **10 test instances**

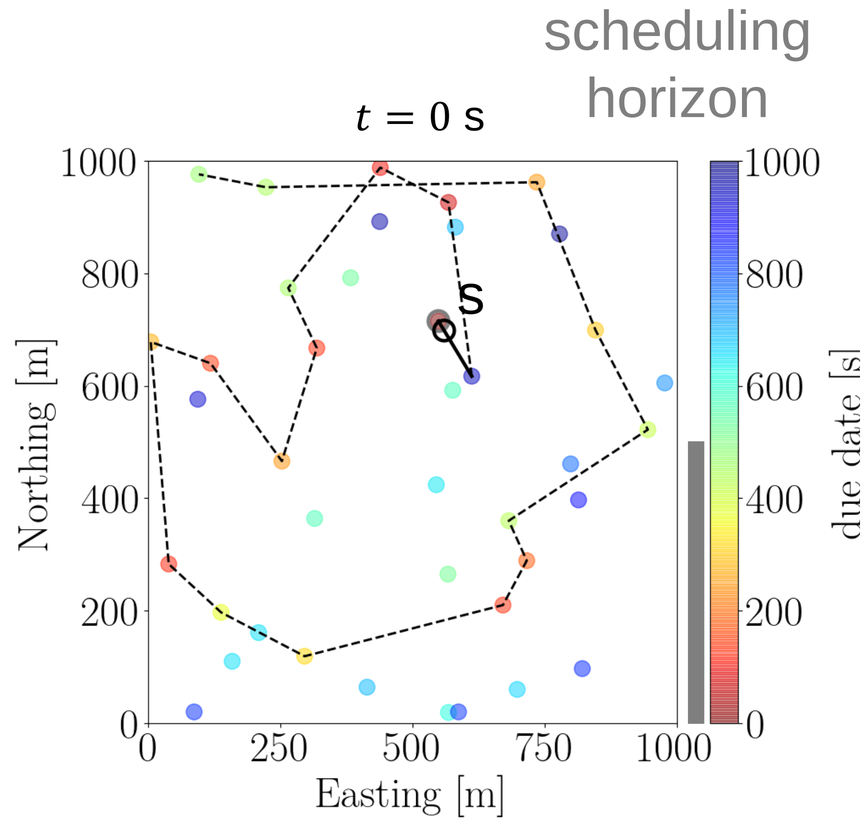


Test Case: Periodic Rescheduling

Location of the vehicle…
- at the **start**
- at the **end**

… of the optimization procedure.

- Realized route —
- Scheduled route ---



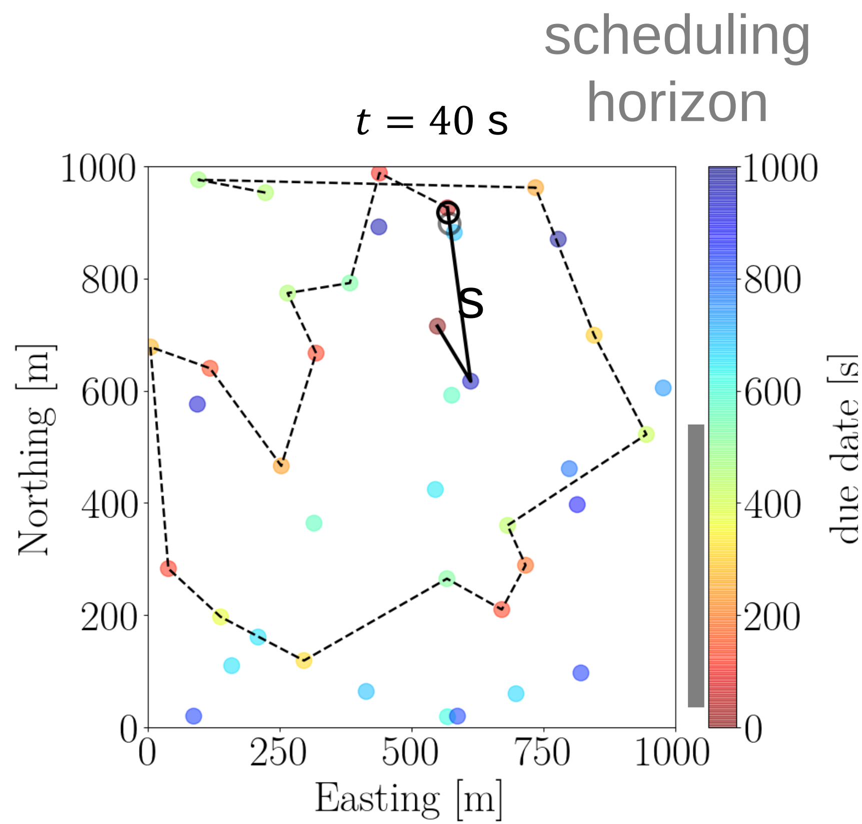
Aalto University School of Chemical Engineering

Test Case: Periodic Rescheduling

Location of the vehicle…
- at the **start** ○
- at the **end** ○

… of the optimization procedure.

- Realized route —
- Scheduled route ---



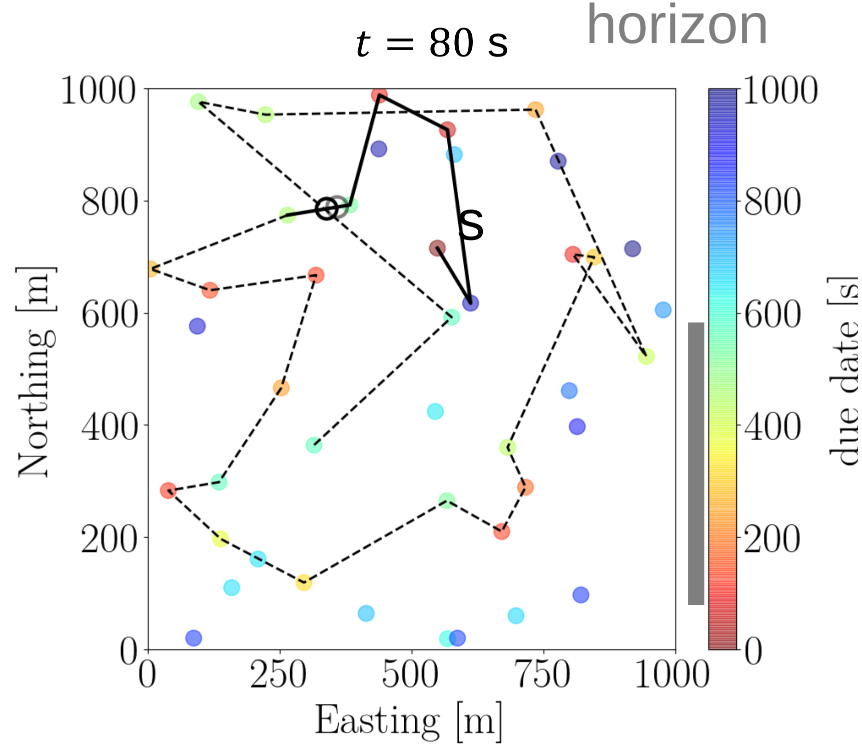
Test Case: Periodic Rescheduling

Location of the vehicle…
- at the **start**
- at the **end**

… of the optimization procedure.

- Realized route —
- Scheduled route ---

scheduling horizon



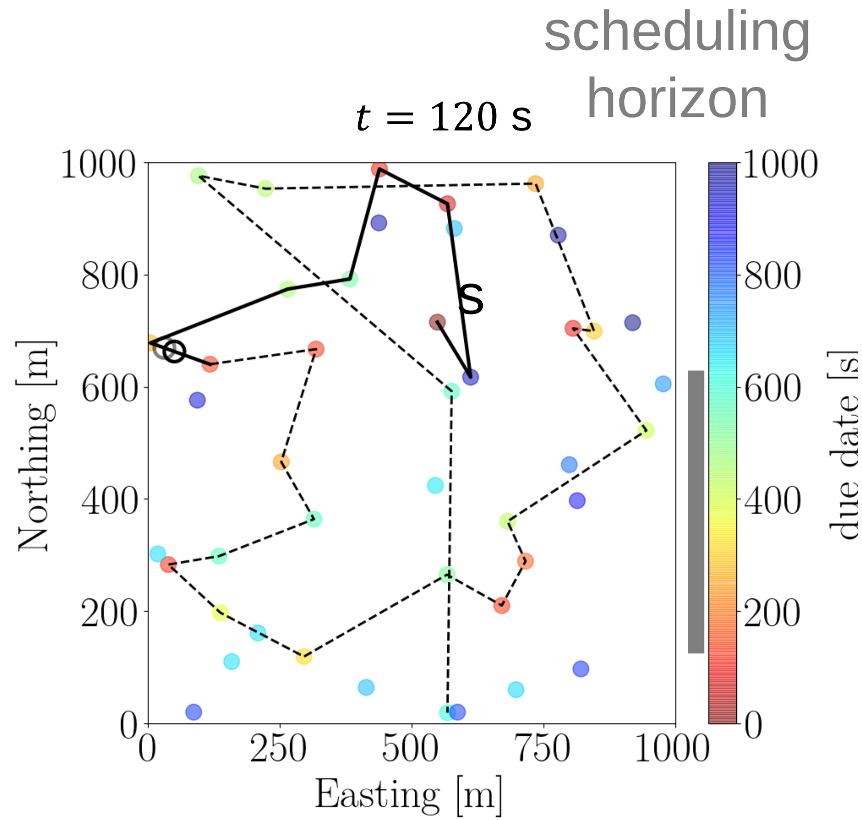
Test Case: Periodic Rescheduling

Location of the vehicle…

- at the **start**
- at the **end**

… of the optimization procedure.

- Realized route —
- Scheduled route ---



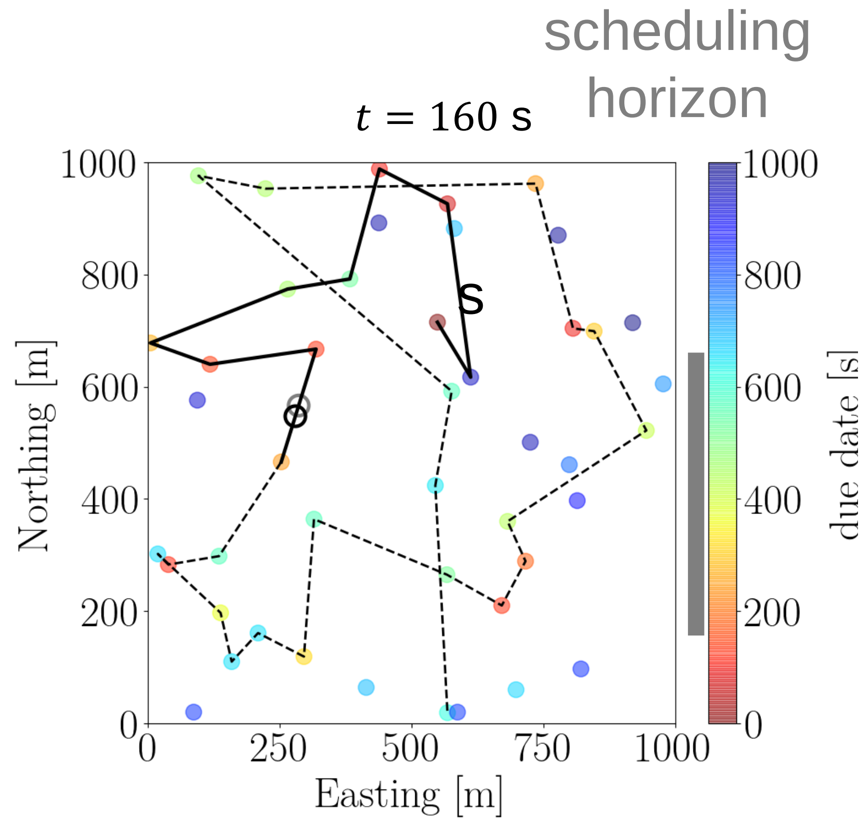
Test Case: Periodic Rescheduling

Location of the vehicle…

- at the **start** ○
- at the **end** ○

… of the optimization procedure.

- Realized route —
- Scheduled route ---



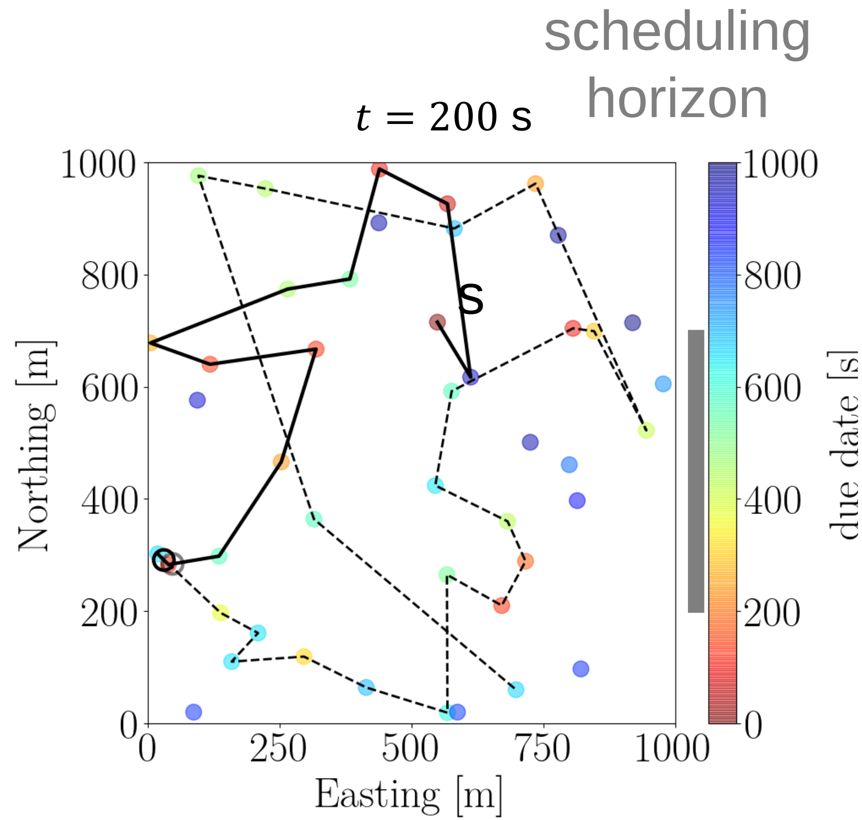
Test Case: Periodic Rescheduling

Location of the vehicle…

- at the **start** ○
- at the **end** **○**

… of the optimization procedure.

- Realized route —
- Scheduled route ---



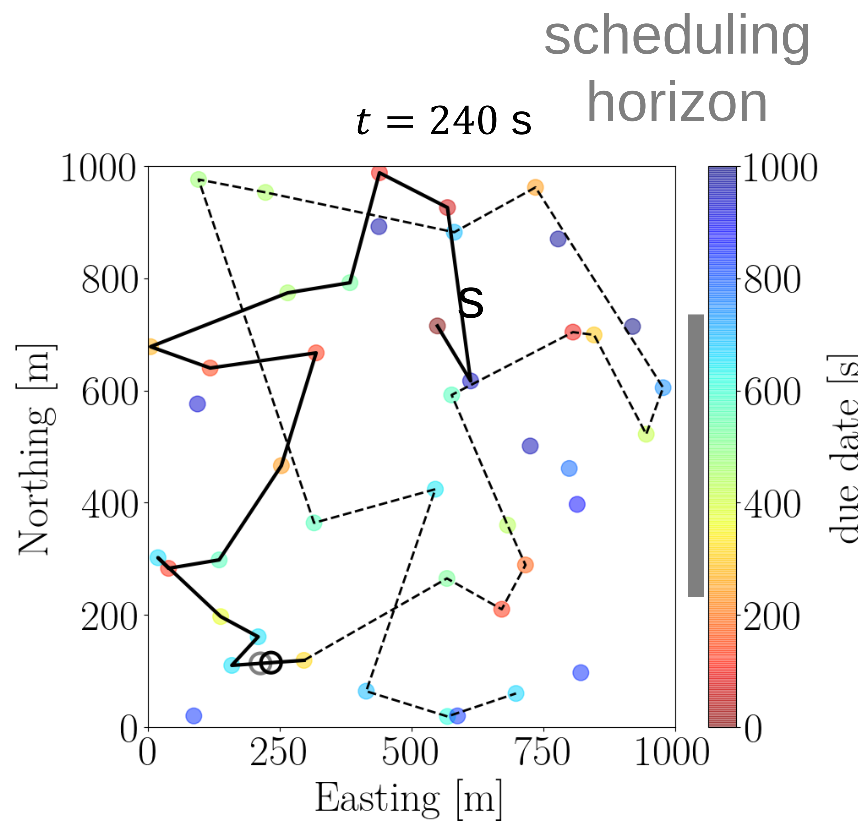
Test Case: Periodic Rescheduling

Location of the vehicle…

- at the **start**
- at the **end**

… of the optimization procedure.

- Realized route —
- Scheduled route ---



Aalto University
School of Chemical
Engineering

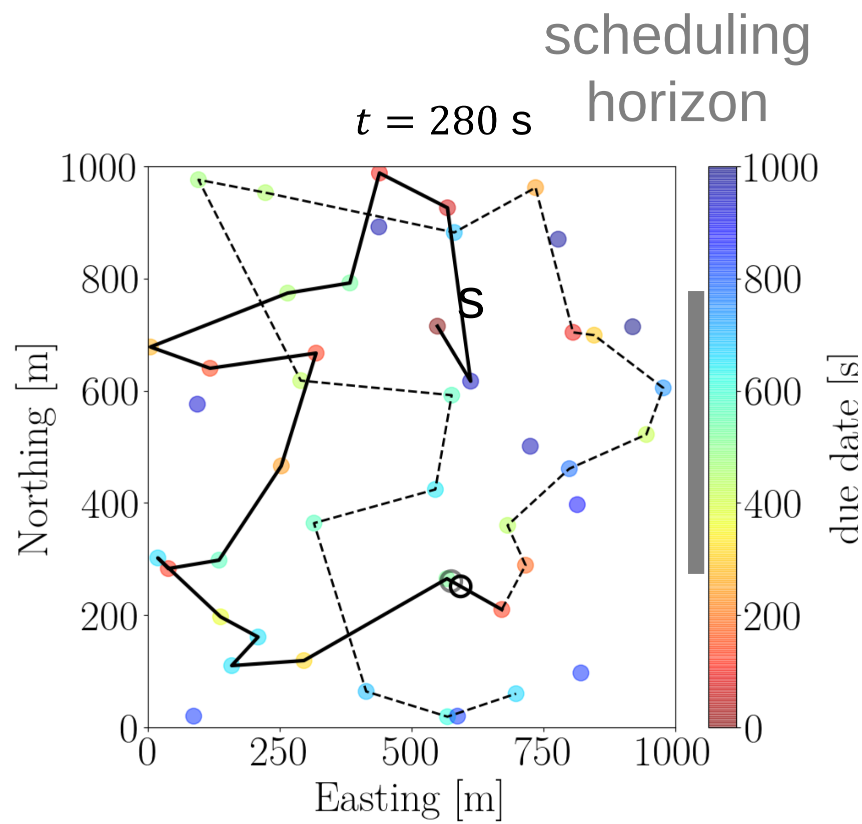
Test Case: Periodic Rescheduling

Location of the vehicle…

- at the **start**
- at the **end**

… of the optimization procedure.

- Realized route —
- Scheduled route ---



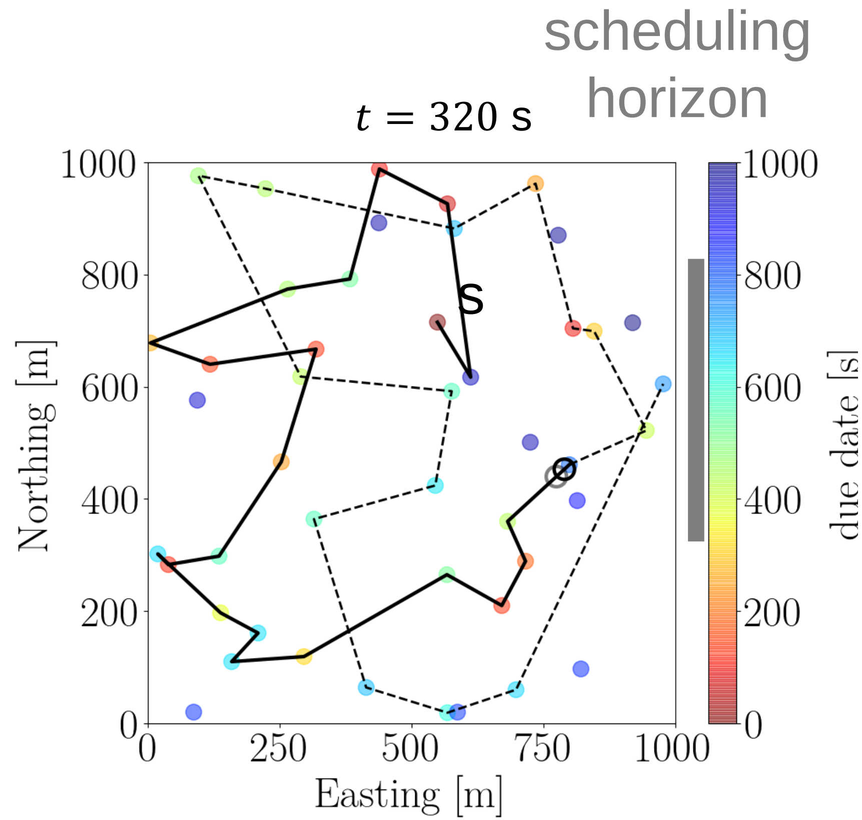
Aalto University
School of Chemical
Engineering

Test Case: Periodic Rescheduling

Location of the vehicle…
- at the **start**
- at the **end**

… of the optimization procedure.

- Realized route —
- Scheduled route ---



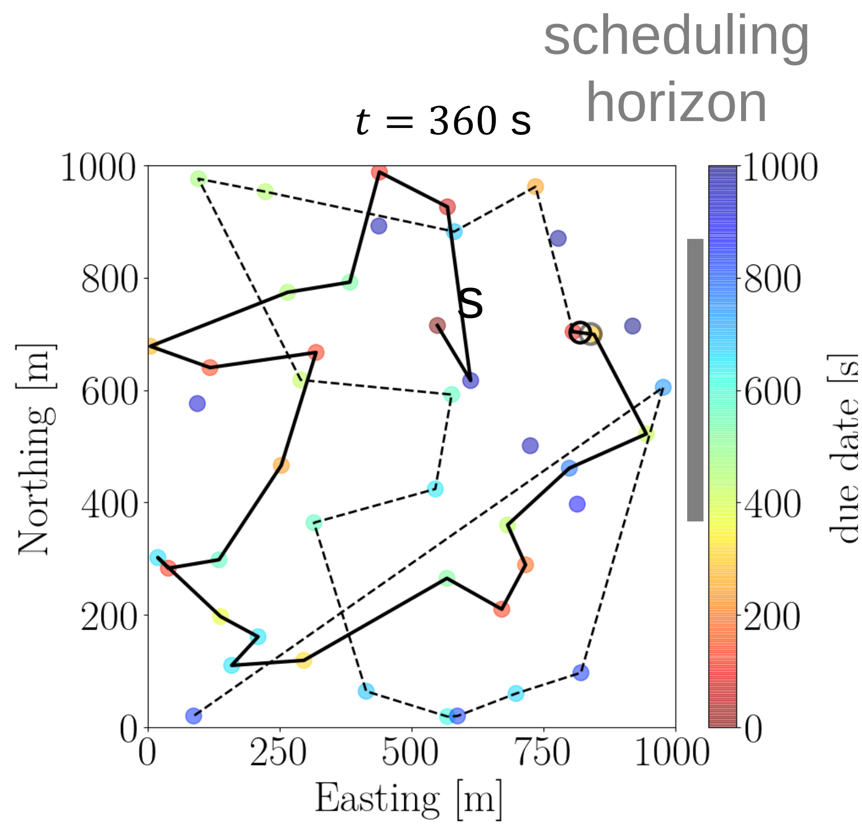
Test Case: Periodic Rescheduling

Location of the vehicle…

- at the **start** ○
- at the **end** ○

… of the optimization procedure.

- Realized route —
- Scheduled route ---



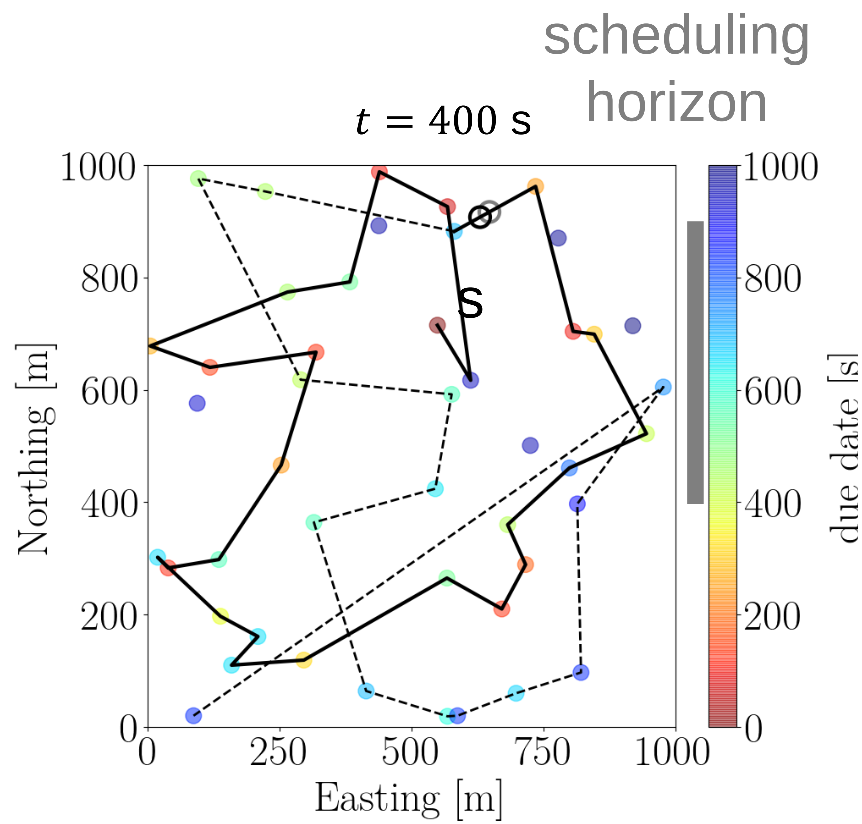
Aalto University School of Chemical Engineering

Test Case: Periodic Rescheduling

Location of the vehicle…
- at the **start** ○
- at the **end** ○

… of the optimization procedure.

- Realized route —
- Scheduled route ---



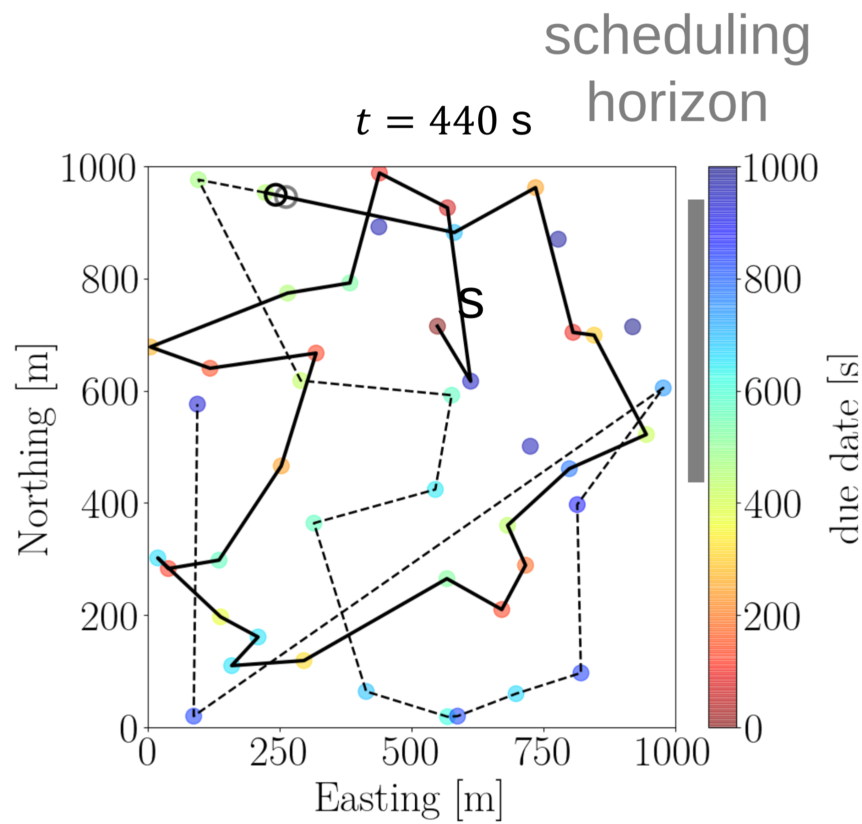
Test Case: Periodic Rescheduling

Location of the vehicle…

- at the **start** ○
- at the **end** **○**

… of the optimization procedure.

- Realized route —
- Scheduled route ---

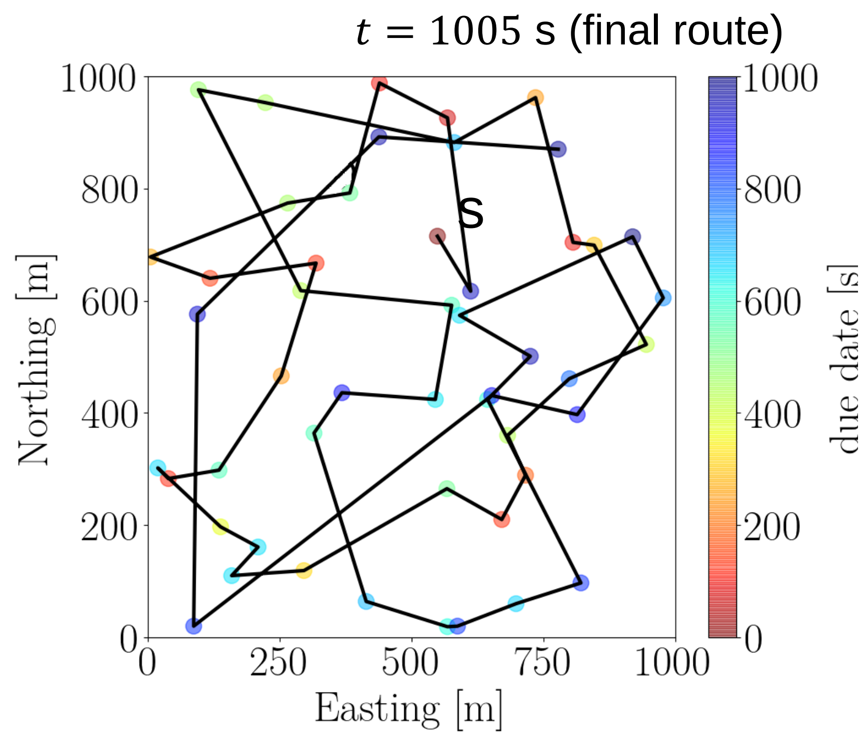


Test Case: Periodic Rescheduling

Location of the vehicle…
- at the **start** ○
- at the **end** **○**

… of the optimization procedure.

- Realized route —
- Scheduled route ---

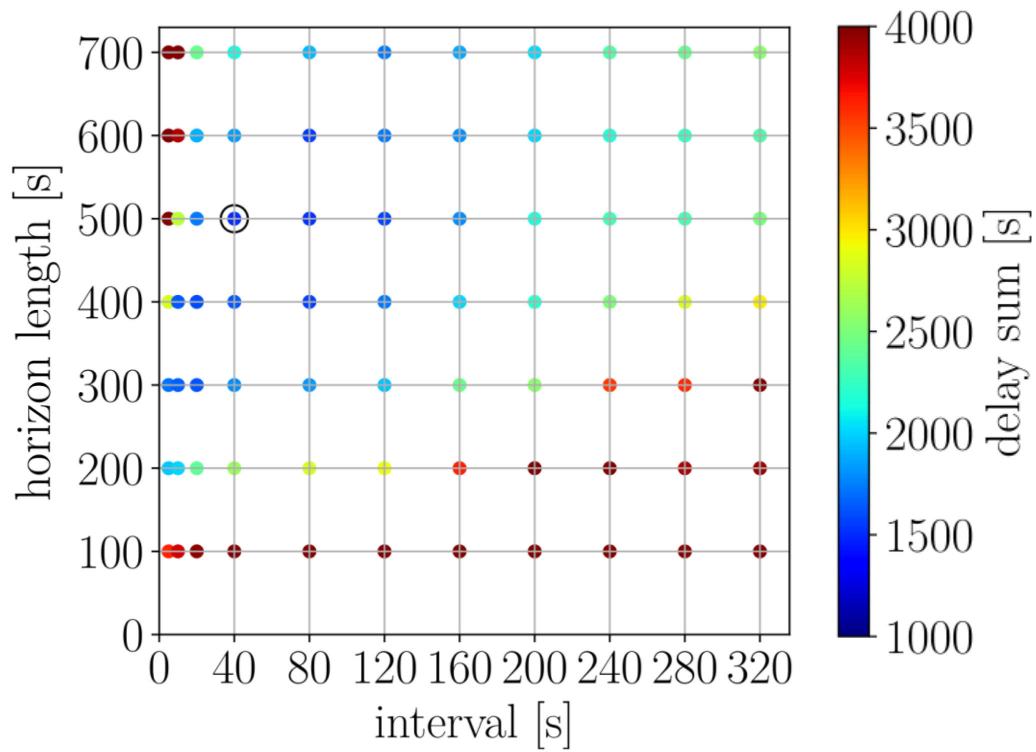


Final delay sum is 1504.9 s

Aalto University
School of Chemical
Engineering

Test Case: Parameter Tuning

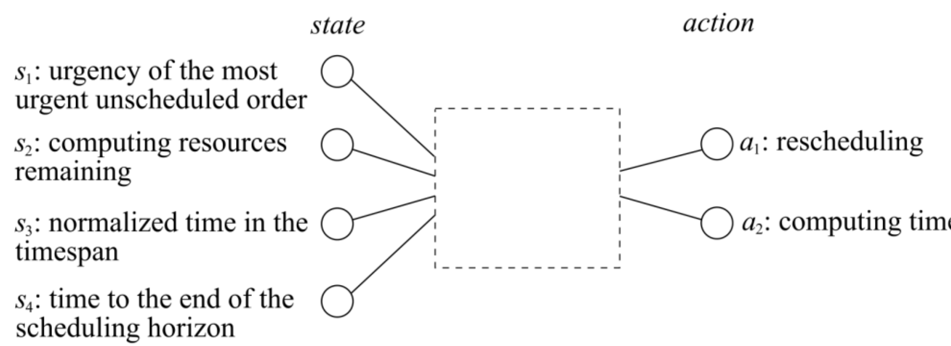
- Grid search in the space of horizon length and rescheduling interval
- Average delay sum of 5 randomized instances of the test case
- The optimized parameter combination is circled



(the darkest red points exceed the scale)

Test Case: State and Action Spaces, and Reward

- **A simplified version of the approach with two rescheduling decisions**
  - Rescheduling timing (a_1)
  - Computing time per procedure (a_2)
- **We use the NEAT algorithm**
  - Population size of 80
  - 80 generations
  - 20 hours of training on a cluster of 20 CPUs
- **The reward is the final delay sum multiplied by -1**
- **The agent can act at an interval of 1 s**

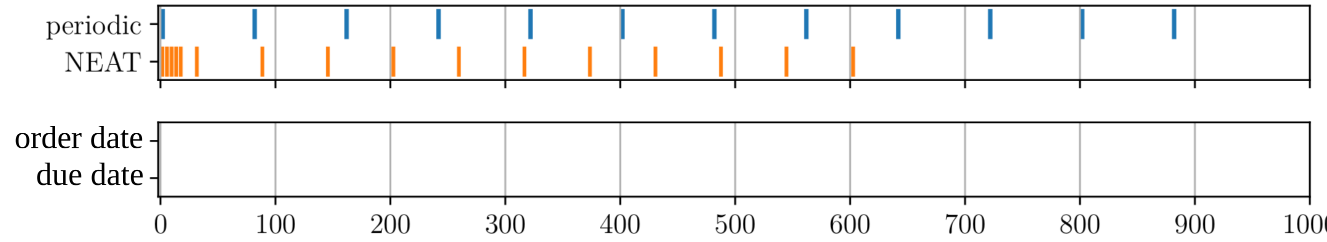


Aalto University
School of Chemical
Engineering

Test Case: Results

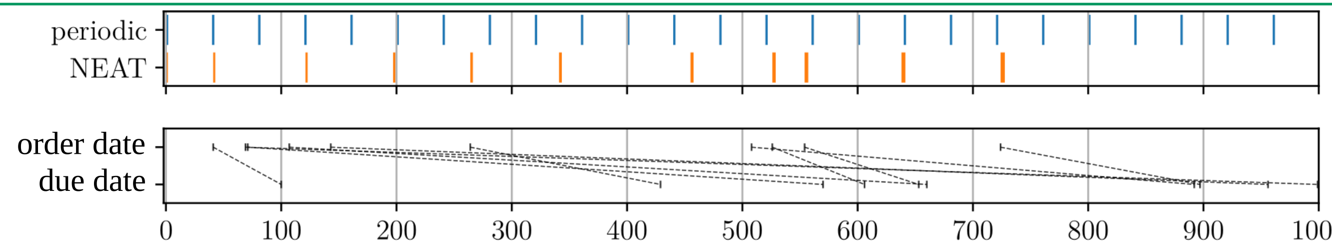
Test case 1:

- new orders 0/50



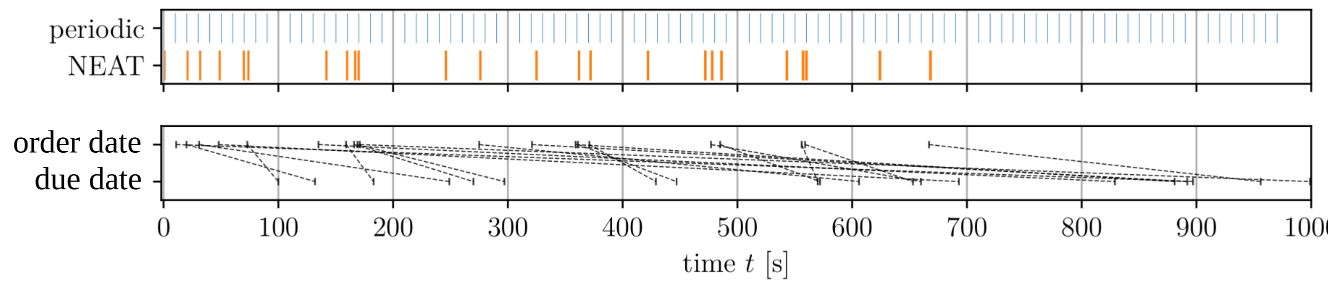
Test case 2:

- new orders 10/50



Test case 3:

- new orders 20/50

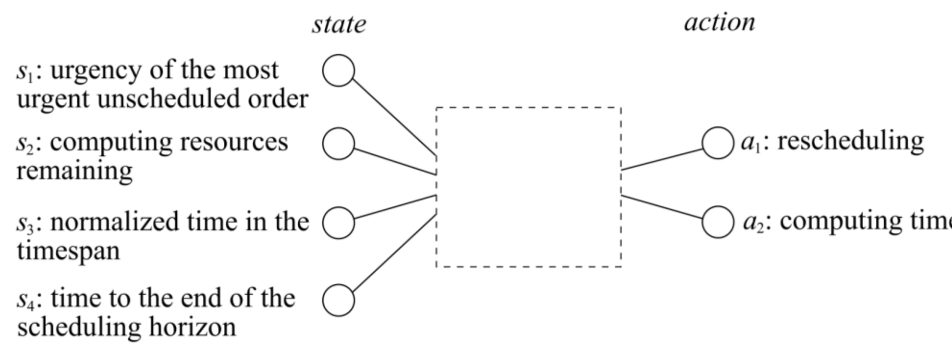


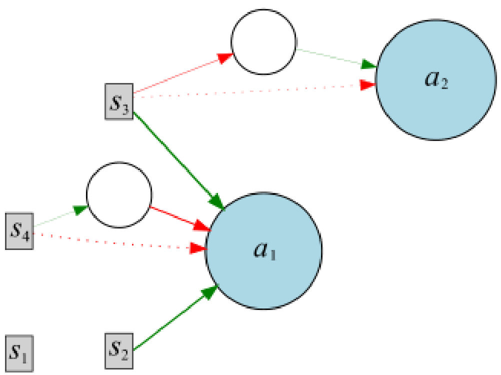
Aalto University
School of Chemical
Engineering

Test Case: Results

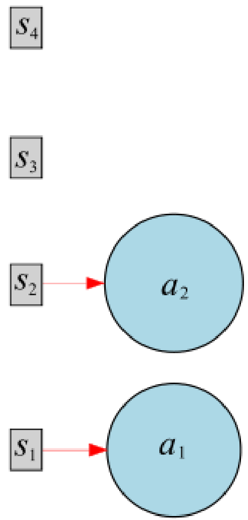
Test case	Rescheduling method	Average delay sum [s] training instances	Average delay sum [s] test instances	Difference [%] training instances	Difference [%] test instances
1	periodic	1120.7	1233.9		
	NEAT agent	1061.3	1103.5	-5.30	-10.57
2	periodic	1504.9	1524.4		
	NEAT agent	1365.9	1461.8	-9.24	-4.11
3	periodic	1793.0	2095.1		
	NEAT agent	1635.2	1775.0	-8.80	-15.28

Test Case: Neural Network Topologies

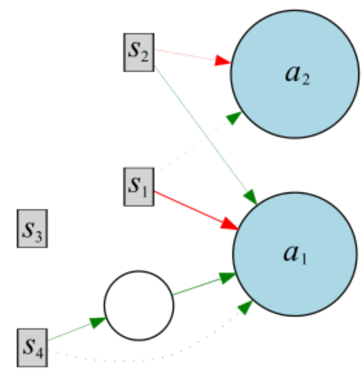




Test case 1:
- new orders 0/50



Test case 2:
- new orders 10/50



Test case 3:
- new orders 20/50

Conclusions

- **We propose an approach where**
    - The process and scheduling optimization together form the environment in reinforcement learning
    - An RL agent is trained to make four decisions on rescheduling procedures
- **In the three test cases, a simplified version of the approach yields, on average, better results than periodic rescheduling**
- **Future work investigates**
    - the approach with all four decisions
    - other RL algorithms for the purpose

References

Watkins, C. J., & Dayan, P. (1992). Q-learning. Machine learning, 8(3-4), pp. 279-292.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G. et al. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), pp. 529.

Stanley, K. O., & Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. Evolutionary computation, 10, pp. 99-127.

Schulman, J., Levine, S., Abbeel, P., Jordan, M., & Moritz, P. (2015). Trust region policy optimization. In International conference on machine learning, pp. 1889-1897.

Floreano, D., Dürr, P., & Mattiussi, C. (2008). Neuroevolution: from architectures to learning. *Evolutionary Intelligence*, 1(1), pp. 47-62.

Such, F. P., Madhavan, V., Conti, E., Lehman, J., Stanley, K. O., & Clune, J. (2017). Deep neuroevolution: Genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning. *arXiv preprint arXiv*:1712.06567.

Hausknecht, M., Lehman, J., Miikkulainen, R., & Stone, P. (2014). A neuroevolution approach to general Atari game playing. IEEE Transactions on Computational Intelligence and AI in Games, 6(4), 355-366.

Semrov, D., Marsetic, R., Zura, M., Todorovski, L., & Srdic, A. (2016). Reinforcement learning approach for train rescheduling on a single-track railway. Transportation Research Part B: Methodological, 86, 250-267.

References

Atallah, R. F., Assi, C. M., & Khabbaz, M. J. (2018). Scheduling the operation of a connected vehicular network using deep reinforcement learning. IEEE Transactions on Intelligent Transportation Systems, 20, pp. 1669-1682.

Palombarini, J., & Martinez, E. (2012). Smartgantt – an interactive system for generating and updating rescheduling knowledge using relational abstractions. Computers & Chemical Engineering, 47, pp. 202-216.

Priore, P., Gomez, A., Pino, R., & Rosillo, R. (2014). Dynamic scheduling of manufacturing systems using machine learning: An updated review. AI EDAM, 28, pp. 83-97.

Aydin, M. E., & Oztemel, E. (2000). Dynamic job-shop scheduling using reinforcement learning agents. Robotics and Autonomous Systems, 33, pp. 169-178.

Acknowledgement

The financial support from Academy of Finland, through project SINGPRO, is gratefully acknowledged.

ACADEMY OF FINLAND

Aalto University
School of Chemical
Engineering