Engineering Risk Analysis Group
Prof. Dr Daniel Straub

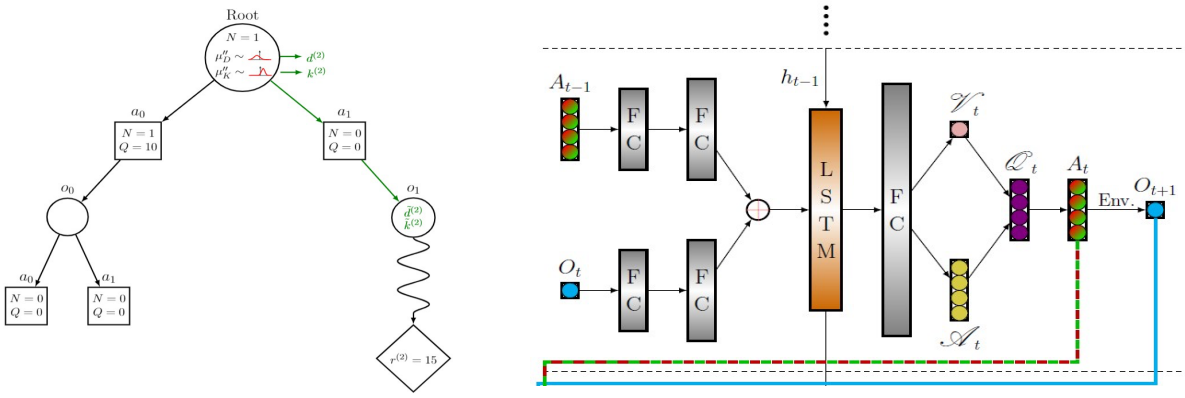
Master's thesis

Reinforcement Learning for Inspection and Maintenance Planning – A Comparison of Neural Networks and Monte Carlo Tree Search

Daniel Koutas, July 2022

Background

Infrastructure maintenance is extremely challenging but has the potential of saving a lot of time, effort and money if optimal decisions are made based on the information at hand. This work investigates the optimal maintenance planning problem of a one-component system subject to deterioration, which is modeled as a partially observable Markov decision problem (POMDP). The goal is to find a strategy which minimizes the expected life cycle cost (LCC), corresponding to the sum of action and failure costs incurred throughout the life cycle of the system.



The two proposed methods to be compared for the solution of the POMDP reference problem are the Monte Carlo Tree Search (left) and a Deep Recurrent Q-network (right).

Methodology

We compare two reinforcement learning methods to solve the optimal maintenance strategy subject to uncertainty, namely deep neural networks (NNs) and Monte Carlo tree search (MCTS). Both methods approximate the value of taking a certain action given an observation history. While deep NNs use whole deterioration trajectories for training, MCTS sequentially progresses until the end of the system lifetime. We investigate different method performance metrics, including the achieved expected LCC of the resulting maintenance strategies, the effect of the observation error, and a time-dependent mapping from belief state to action.

Results

Both methods produce lower expected LCCs than some given benchmark strategies. However, the NNs are computationally faster and result in lower expected LCCs compared to MCTS. In addition, they can handle high noise in the measurements, whereas MCTS performs well for low and medium measurement noises but fails to adapt to the high noise regime. By contrast, the visualization and the interpretation of the NNs' trained strategies are difficult, as custom test scenarios cannot be imposed. Here, the MCTS has their key advantages, as any starting point can be inferred which greatly facilitates the interpretation of the resulting strategies.

Supervised by MSc Elizabeth Bismut (TUM)
Prof Dr Daniel Straub (TUM)