Monthly Report 4 - Analysis and Results

My experimental results aren't ready to share yet. I can share the an overview of what results will be included and what they will show.

My research question is: To what extent can Reinforcement Learning optimise personnel scheduling problems?

When a model completes a scheduling problem it will receive a reward and count of constraints violated (working hours, consecutive shifts etc).

The key metric is average reward (the accumulated reward per time step). For each experiment average reward will be plotted against episode number. An episode is one run through of a problem so these plots will ideally show that an RL agent is learning.

There are many possible variants for the design of this system (an RL agent for personnel scheduling) including feature engineering, reward shaping and model architecture.

The following experiments will be run to determine the optimal system:

Reward shaping

The design of the reward influences learning. I will test several different reward formulations.

Algorithms

There a many RL algorithms. I will test the following:

A2C

PPO

DQN

Max length of episodes

Try different episode lengths.

Number of message passing steps

The graph neural network is used to choose actions. Number of message passing steps can be optimised.

Take best model from above experiements run on several random seeds. This is due to randomness inherent in RL.

Compare with Benchmark method; Linear Programming.