1 Overview

1.1 Environment

- The game is a single-agent environment.
- The game is a complicated optimization where the agent must decide incoming jobs to accept and which to taken jobs to work on.
- The goal of the game earn the most possible money from completing jobs

1.2 State space

- The state space is the current day, as well as data for each job offer and currently taken job.
- Each job is represented by:
    1. number of workers required
    2. complication probability (this determines how long the job takes to complete)
    3. soft deadline
    4. hard deadline
    5. payment
    6. expected length
    7. rate of pay (payment / (workers * expected length)

1.3 Action space

The action space is 27 binary outputs:

- 7 binary outputs for which of the up to 7 incoming jobs to accept
- 20 binary outputs for the up to 20 jobs that can be accepted at a time.
- Note: you may want to use an action mask so superfluous actions are ignored when updating

1.4 Rewards

- You receive the payment of any jobs completed on time
- Payment is reduced as the job passes the soft deadline
- You lose 0.3*payment for any jobs not completed by the hard deadline

1.5 Terminal States

- The episode ends after 100 days

2 Instructions

Full details of the rules can be found in the instructions folder

3 Software

Software you can play (to get ideas for reward shaping and such) can be found in the software folder

4 Heuristic Pretraining

If you wish to do pretraining, let me know and I will add instructions on how to generate data from heuristics