

Reinforcement Learning

■ Implementation

- parameters
 - Episode: every 30 seconds
 - Batch amount: 64 observations
 - discount rate: 0.99
 - replay memory size: 10000
- network design
 - input layer: 11 nodes
 - hidden layers: 2 layers, each 22 nodes
 - activator: Relu
 - output nodes: 4 nodes (= 4 action space)
 - activator: linear (approx q-function)
 - loss function: MSE

Implementation (Animation Image)

■ Result



episode 1



episode 50



episode 140



episode 240