

Mean reward over last 100 episodes

- Reinforcement agent - penalty,  $c=0.05$
- Reinforcement agent - penalty,  $c=0.075$
- - Reinforcement agent - penalty,  $c=0.1$

Mean reward over last 100 episodes

