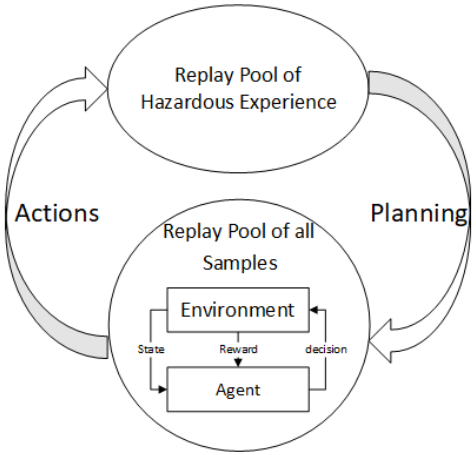
Improving Deep Reinforcement Learning by Safety Guarding Model via Hazardous Experience Planning

Pai PENG, Fei ZHU, XingHong LING, PeiYao ZHAO, Quan LIU

Frontiers of Computer Science, DOI: 10.1007/s11704-021-0250-y

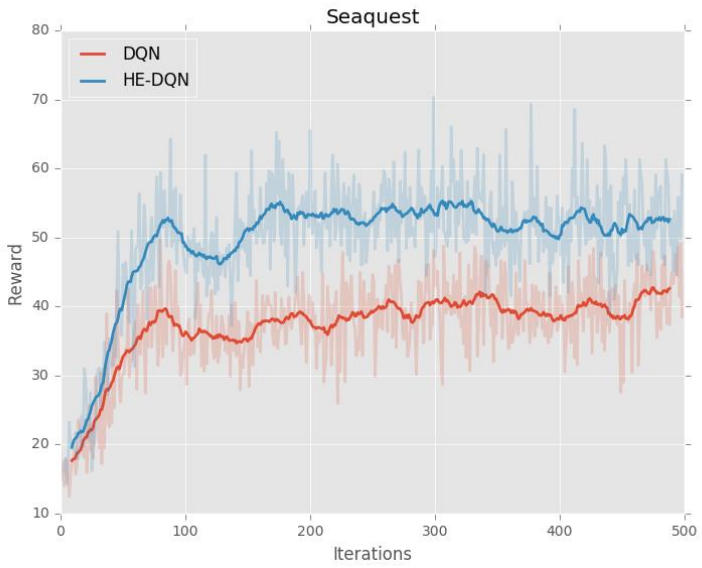
Problems & Ideas

- Problem of agent falling into hazardous state because of taking improper action
  - most traditional reinforcement learning algorithms are based on unconstrained stochastic policies, which is very likely to lead to failure
  - little concern has been put on taking advantages of valuable experiences that relate to the failure of the task
- Ideas: Build up a safety guarding model via hazardous experience planning
  - takes advantage of the hazardous experience replay buffer
  - build up a safety guarding model



Main Contributions

- Comparison of average rewards between DRL and DRL equipped with HE



- Comparison of game rounds between HE-DRL and other safe deep reinforcement learning