B4M36SMU

Reinforcement Learning 3 - Convergence of Q-learning and SARSA

Monday 21$^{\text{st}}$ May, 2018

GLIE

- Off-policy Q-learning needs to satisfy the infinite exploration condition;
- SARSA needs to satisfy both conditions of GLIE in order to converge.
- For further details, please visit: `https://doi.org/10.1023/A:1007678930559`
- If you use ε-greedy policy, it is more practical to use different ε for each state.

Learning rate

- In TD-learning and Q-learning, learning rate should decrease so that sum of α_t diverges and sum of α_t^2 converges.
- For further details, please visit: https://doi.org/10.1023/A:1022676722315
- Again, it is more practical to use different learning rate for each $Q(s, a)$ value.