



Figure 1: Basic agent-environment relationship in a Markov decision process. The agent chooses an action A_t and the environment returns a new state S_{t+1} and a reward R_{t+1} . The dotted line represents the transition from step t to step $t + 1$ [?].