

Description

Consider a 5x10 gridworld, as shown in Figure 1. The agents start at the same random location and have four actions (move in each of the four directions). There is a reward of 30 to catch T, a target that starts at the red square and moves randomly by one square at each time step. In addition, there is reward of -1 for every time step the agent is in this gridworld.

