

Implementation of Example 8.3 from Sutton-Barto

Alexander Lyzhov

Skolkovo Institute of Science and Technology

E-mail: nikkou.postbox@gmail.com

1 Experimental Results

This figure compares the cumulative reward of several RL algorithms with incorporated planning. They are solving an episodic problem of getting from point A to point B on a grid. The trick is that when half of epochs are completed, a new, more efficient pathway is opened. Versions with advanced exploration with various parameters perform much better than simple eps-greedy exploration.

