A picture containing chart

Description automatically generatedFollowing figures were obtained from the runs:

The deterministic run (is\_slippery=False) finds the goal almost every single time. Especially when the number of episodes is increased.

Chart, histogram

Description automatically generated

The non-deterministic run (is\_slippery = True) gives worse results as expected. The average reward does not seem to get noticeably better with more episodes run.

Chart

Description automatically generated

The non-deterministic run (is\_slippery = True) with Q-table update algorithm improved. I think the result is marginally better than with old update algorithm. I tried to play with alpha and gamma but could not find meaningful improvement with those tweaks.