Learning rate	reward	policy	Exploration Factor
.125	Reward: 7.82	00111110022	.99
0.0125	Reward: 5.87	00111210001	.99
0.00125	Reward: 5.87	00111210001	.99
1.125	Reward: 9.67	00100122222	.99
1.125	Reward: 9.62	00001022222	.88
1.125	Reward: 10.12	00000112222	.55
1.125	Reward: 10.16	00001112222	.49

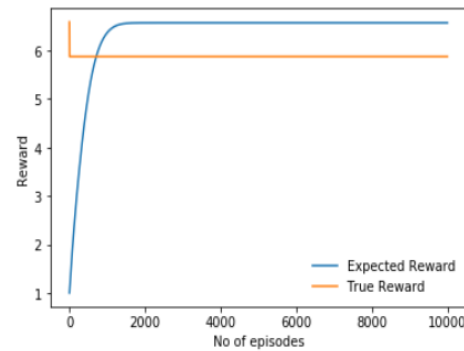*Table: Rewards at different learning rates and exploration factor*

So we can see that decreasing the learning rate decreases the final reward and decreasing the discount factor increases the reward.
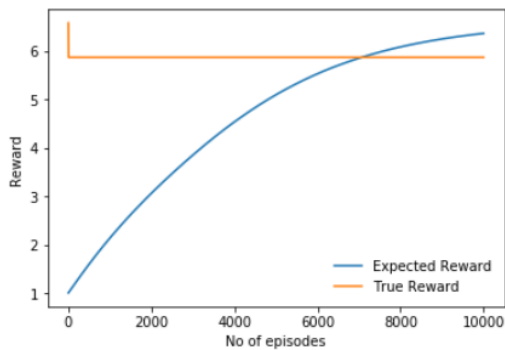
PLOTS

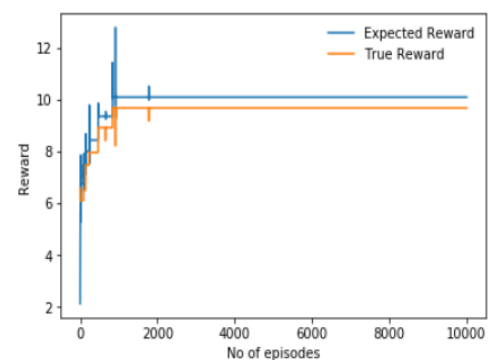True vs expected reward @ learning rate=0.125 & discount factor =0.99
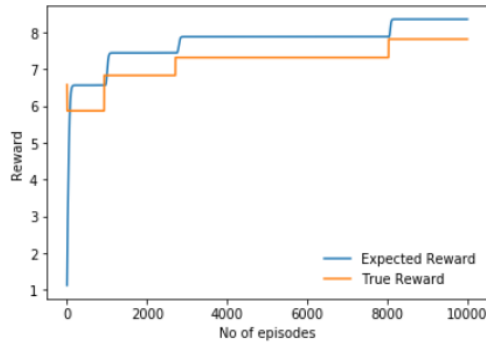
True vs expected reward @ learning rate=0.0125 & discount factor =0.99

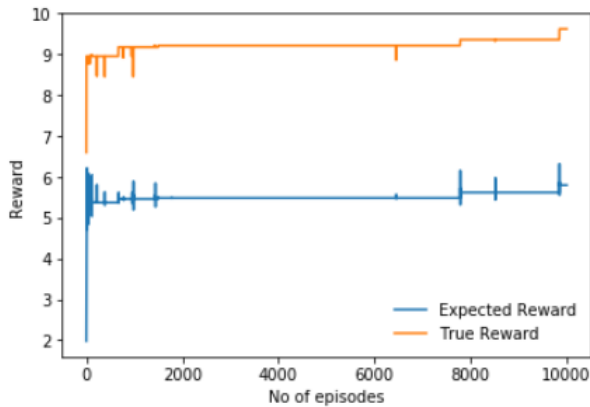True vs expected reward @ learning rate=0.00125 & discount factor =0.99

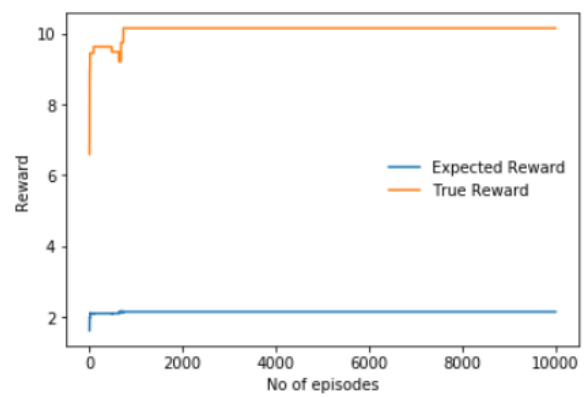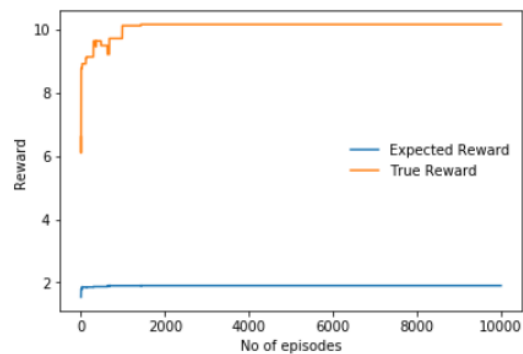True vs expected reward @ learning rate=1.125 & discount factor =0.99

PLOTS

True vs expected reward @ learning rate=1.125 & discount factor =0.88

True vs expected reward @ learning rate=1.125 & discount factor =0.55

True vs expected reward @ learning rate=1.125 & discount factor =0.49

FEEDBACK:

It took extensive search from google to understand what openai gym is? how does env works and what are spaces of type boxes.

It wasn't that great to have an exercise to directly get started with reinforcement learning.

Overall took almost 3 days of searching and reading. Basically figuring out what TO DO's mean rather than learning something new.

SUBMITTED BY:

Manav Madan

Stefan Mohrle

Vanshaj Taxali