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

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

PLot for expected reward vs True reward with learning rate 1.125 and discount factor of .49