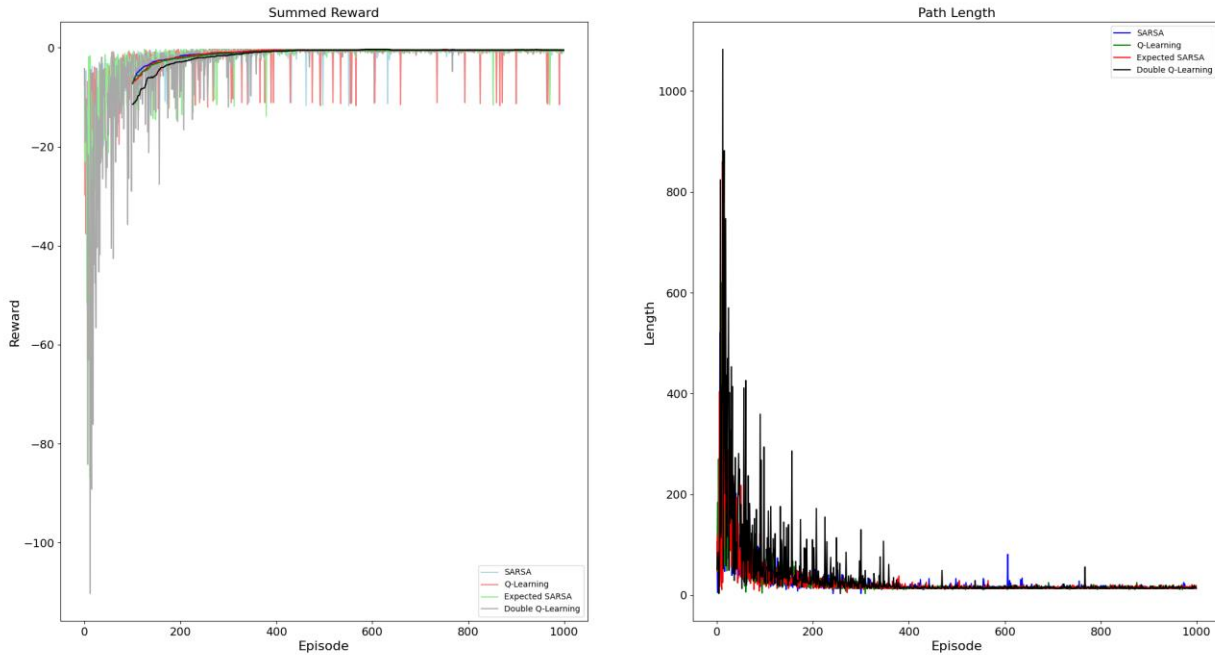# Q3: Analysis of Results

Quantitative Analysis of Results:

Task 1:


```
All experiments complete
Experiment Setup:
 - episodes:1000
 - sim_speed:0.1

[SARSA on Task 1] : SARSA : max-rew=-0.300 med-100=-0.500 var-100=0.050 max-episode-len=521.0
[Q Learning on Task 1] : Q-Learning : max-rew=-0.300 med-100=-0.500 var-100=5.565 max-episode-len=620.0
[Expected SARSA on Task 1] : Expected SARSA : max-rew=-0.300 med-100=-0.500 var-100=1.248 max-episode-len=860.0
[Double Q Learning on Task 1] : Double Q-Learning : max-rew=-0.300 med-100=-0.500 var-100=0.052 max-episode-len=1083.0
```
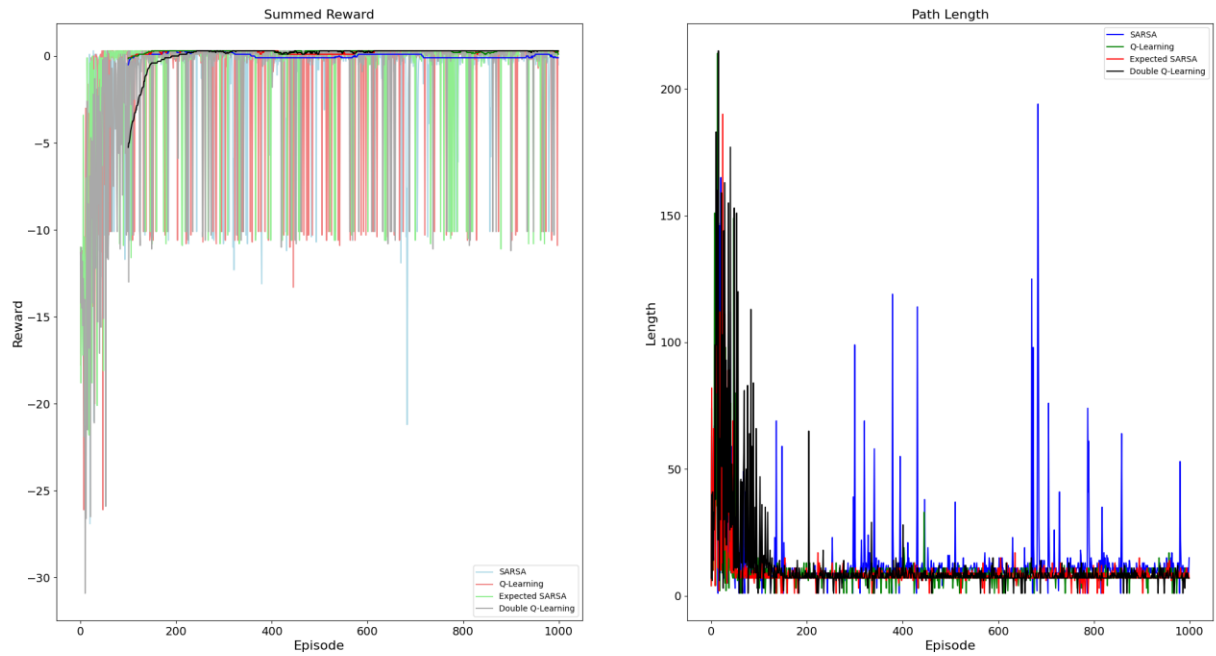
Algorithm	Max Reward	Med-100	Var-100	Max Episode Length
SARSA	-0.300	-0.500	0.050	521
Q-Learning	-0.300	-0.500	5.565	620
Expected SARSA	-0.300	-0.500	1.248	860
Double Q Learning	-0.300	-0.500	0.052	1083

Task 2:


```
All experiments complete
Experiment Setup:
 - episodes:1000
 - sim_speed:0.001

[SARSA on Task 2] : SARSA : max-rew=0.300 med-100=-0.100 var-100=1.231 max-episode-len=194.0
[Q Learning on Task 2] : Q-Learning : max-rew=0.300 med-100=0.250 var-100=7.279 max-episode-len=214.0
[Expected SARSA on Task 2] : Expected SARSA : max-rew=0.300 med-100=0.200 var-100=5.555 max-episode-len=190.0
[Double Q Learning on Task 2] : Double Q-Learning : max-rew=0.300 med-100=0.300 var-100=7.166 max-episode-len=215.0
```
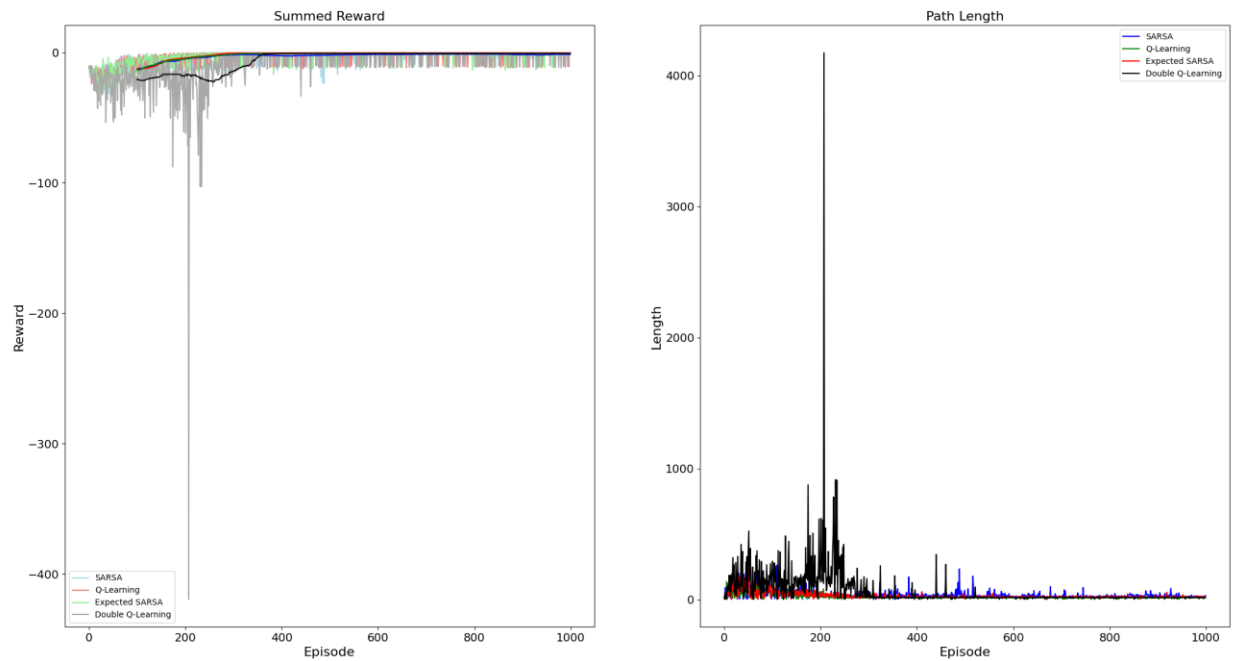
Algorithm	Max Reward	Med-100	Var-100	Max Episode Length
SARSA	0.300	-0.100	1.231	194
Q-Learning	0.300	0.250	7.279	214
Expected SARSA	0.300	0.200	5.555	190
Double Q Learning	0.300	0.300	7.166	215

Task 3:

Task 3

```
All experiments complete
Experiment Setup:
 - episodes:1000
 - sim_speed:0.001

[SARSA on Task 3] : SARSA : max-rew=-0.400 med-100=-1.400 var-100=2.139 max-episode-len=367.0
[Q Learning on Task 3] : Q-Learning : max-rew=-0.200 med-100=-0.400 var-100=12.487 max-episode-len=274.0
[Expected SARSA on Task 3] : Expected SARSA : max-rew=-0.200 med-100=-0.800 var-100=9.196 max-episode-len=241.0
[Double Q Learning on Task 3] : Double Q-Learning : max-rew=-0.400 med-100=-0.600 var-100=14.694 max-episode-len=4174.0
```

Algorithm	Max Reward	Med-100	Var-100	Max Episode Length
SARSA	-0.400	-1.400	2.139	367
Q-Learning	-0.200	-0.400	12.487	274
Expected SARSA	-0.200	-0.800	9.196	241
Double Q Learning	-0.400	-0.600	14.694	4174

Qualitative Analysis of Results:

Q-Learning provides the best performance in terms of rewards and path efficiency, however due to it being off policy it is has more variance.

Double Q- Learning takes the longest to reach convergence, and the most instability in its path length. Double Q-learning showed the worst performance on all tasks.

SARSA showed steady and stable learning, however it was slower to reach convergence then Q-Learning in the tasks.

Expected SARSA's performance was in between Q-Learning and SARSA, as it is both efficient and stable in terms of path length and efficiency.