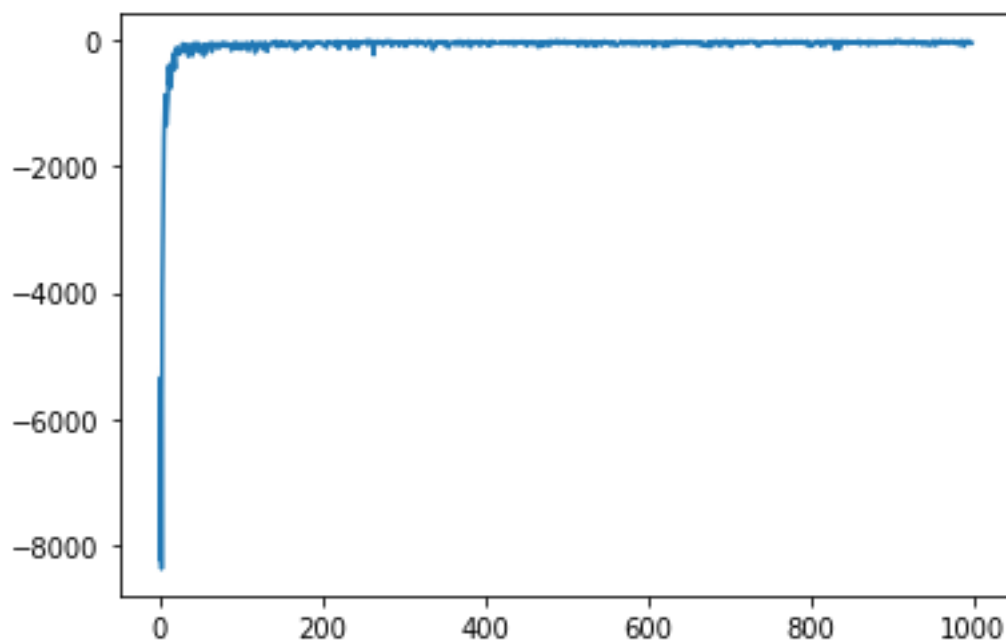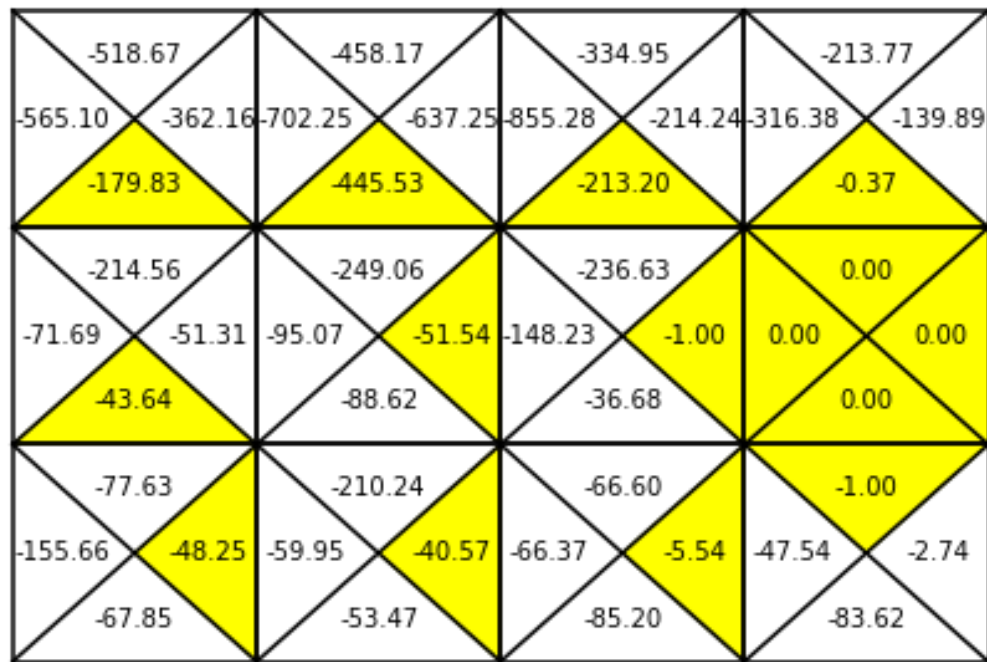# Experiments and Analysis(40%)

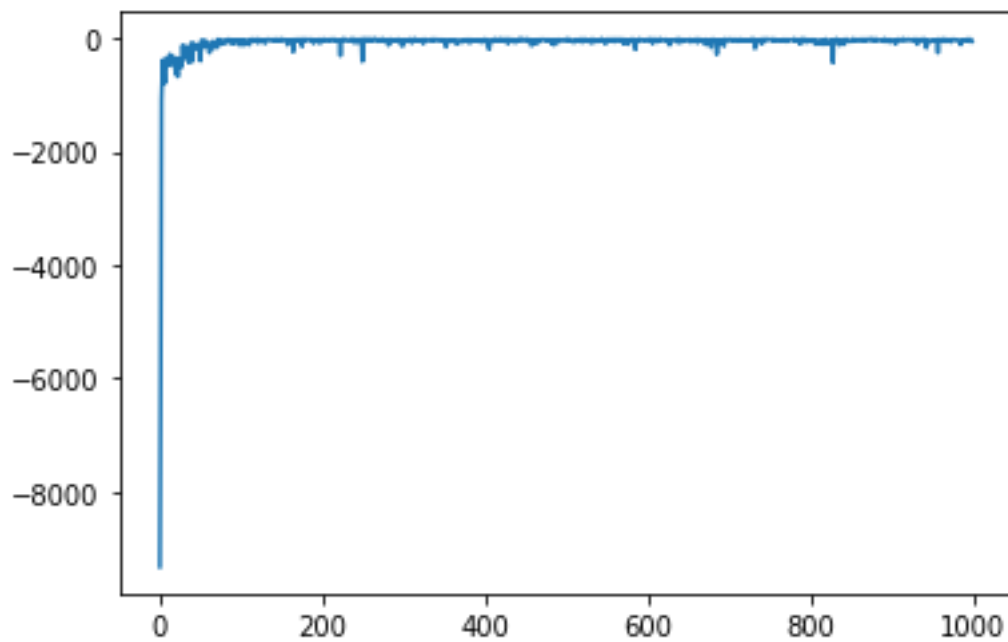1. Plot the q_values in your result.(20%)

   State-action values：

   ● First-visit MC control

-518.67	-458.17	-334.95	-213.77
-565.10 -362.16	-702.25 -637.25	-855.28 -214.24	-316.38 -139.89
-179.83	-445.53	-213.20	-0.37
-214.56	-249.06	-236.63	0.00
-71.69 -51.31	-95.07 -51.54	-148.23 -1.00	0.00 0.00
-43.64	-88.62	-36.68	0.00
-77.63	-210.24	-66.60	-1.00
-155.66 -48.25	-59.95 -40.57	-66.37 -5.54	-47.54 -2.74
-67.85	-53.47	-85.20	-83.62

- Every-visit MC control

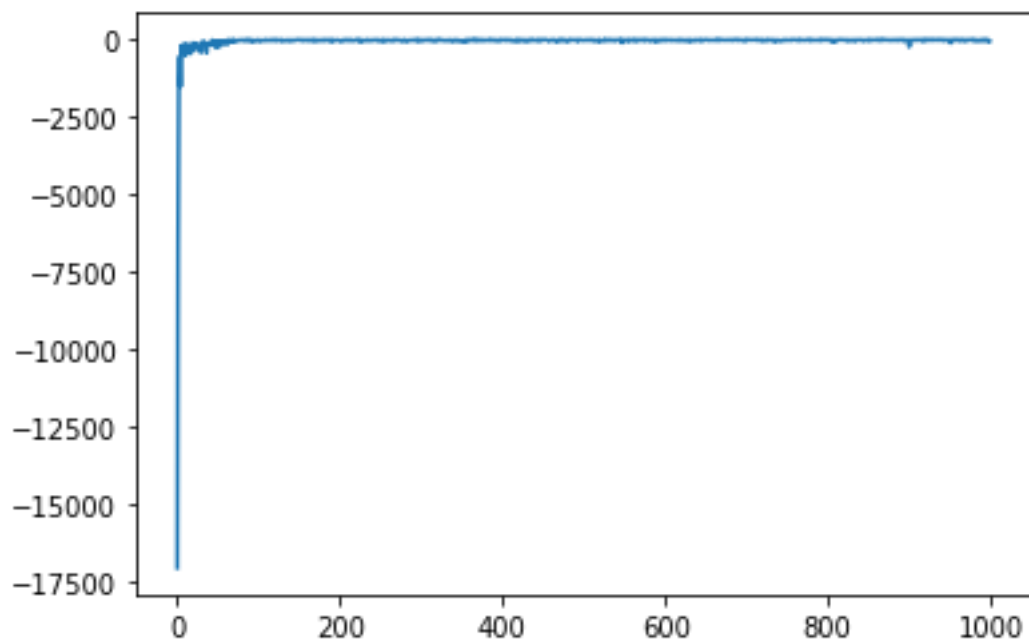2. Whether q_values are reasonable? Explain your result.(10%)

Reasonable。因為是使用 on-policy 的蒙地卡羅，所以會選擇最遠的避開 Swamp 的路徑移動，且在 Swamp 中，會直接向下離開 Swamp。

3. Transfer state-action values to state values and plot it.(10%)

State values：

- First-visit MC predict

	1	2	3	4
1	-81.87	-114.83	-121.76	-329.43
2	-36.49	-29.45	-19.64	0.0
3	-20.02	-14.1	-10.19	-1.52

- Every-visit MC predict

	1	2	3	4
1	-139.97	-202.58	-105.85	-196.03
2	-42.08	-44.15	-15.31	0.0
3	-61.51	-52.86	-7.44	-3.25